



HAL
open science

A pedagogical resources recommender system for teachers

Nader Abdelmalek

► **To cite this version:**

Nader Abdelmalek. A pedagogical resources recommender system for teachers. Education. Université de Technologie de Compiègne, 2023. English. NNT : 2023COMP2771 . tel-04678292

HAL Id: tel-04678292

<https://theses.hal.science/tel-04678292v1>

Submitted on 27 Aug 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Par Nader ABDELMALEK

*Un système de recommandation de ressources
pédagogiques pour les enseignants*

Thèse présentée
pour l'obtention du grade
de Docteur de l'UTC



Soutenue le 10 novembre 2023

Spécialité : Informatique : Unité de recherche Heudyasic
(UMR-7253)

D2771

UNIVERSITÉ DE TECHNOLOGIE DE COMPIÈGNE



Doctoral Thesis

Un système de recommandation de ressources pédagogiques pour les enseignants

Par **Nader ABDELMALEK**

Directrices de thèse :

Marie-Hélène Abel & Christine Lahoud

Spécialité : Informatique

Laboratoire HEUDIASYC

(Heuristique et Diagnostic des Systèmes Complexes) UMR CNRS 7253

École doctorale ED71 - Sciences pour l'ingénieur

Thèse soutenue le 10 novembre 2023 devant le jury composé de :

M. Karim SEHABA (Président)

Mme Hajer BAAZAoui (Rapporteur)

M. Sébastien IKSAL (Rapporteur)

Mme Anne BOYER (Examinatrice)

M. Sylvain LAGRUE (Examinateur)

Mme Marie-Hélène ABEL (Directrice de thèse)

Mme Christine LAHOUD (Directrice de thèse)

UNIVERSITÉ DE TECHNOLOGIE DE COMPIÈGNE

Abstract

A pedagogical resources recommender system for teachers

by Nader Abdelmalek

The landscape of education is ever evolving and in recent years, it has been particularly impacted by the pandemic and associated technological developments. In this context, teachers struggle with diverse challenges that impact their professional development and their educational practices. These challenges may concern their own expertise of their courses, the motivation of their learners, and their awareness of the educational environment. They are generally linked to multiple contexts such as the social and cultural contexts of teachers as well as learners, the environmental context, and the educational context. These challenges vary in intensity according to different situations, for example an inexperienced teacher teaching a specialised course to a difficult audience. The socio-cultural settings of teachers can significantly impact on how he or she meets these challenges. The use of pedagogical resources plays a key role in facilitating teaching and learning. Teachers can use or produce such A resource for learners, but they can also use them to answer their own questions. A pedagogical resource can be considered as any teaching material that helps teachers to deliver effective teaching to their learners. In this thesis, we aim to help teachers meet their challenges by recommending pedagogical resources that consider of their complex and multifaceted context.

We focus on the methods of representing and integrating the multiple contexts of teachers. The followed methodology involves ontological modelling that enables semantic reasoning and has led to the development of the Teacher Sentimental and Collaboration Contexts ontology (TSCCO). This ontology encapsulates the multiple contexts that characterises teachers and aims to facilitate data integration.

Realising the recommendations, we have developed the Context-Aware Pedagogical Resource Recommender System (CAPRRS). CAPRRS combines collaborative and content-based filtering techniques with a hybrid contextual filtering approach, guided by the multiple contexts. It uses a combination of two feature selection techniques (variance-based and mutual information) to identify relevant features. Collaborative filtering takes advantage of recorded interaction traces, while content-based filtering considers data describing pedagogical resources, offering more personalised recommendations.

The main findings highlight the importance of feature selection and the effectiveness of CAPRRS in providing context-aware recommendations of pedagogical resources. These results contribute to a deeper understanding of the relationships that may exists between teacher preferences, contextual features, and recommender approaches. They contribute to the research domain of both education and recommender systems.

Un système de recommandation de ressources pédagogiques pour les enseignants

Le monde de l'éducation est en constante évolution et ces dernières années, il a particulièrement été impacté par la pandémie et les évolutions technologiques associées. Dans ce contexte, les enseignants doivent faire face à de nombreux défis qui ont des conséquences à la fois sur leur évolution professionnelle et leurs pratiques éducatives. Ces défis peuvent concerner leur propre expertise du domaine enseigné, la motivation des apprenants et la prise en compte de l'environnement éducatif. Ils sont généralement liés à différents contextes tels que le contexte social et culturel de l'enseignant et celui des apprenants, le contexte environnemental et le contexte éducatif. Ces défis varient en intensité selon différentes situations comme par exemple un enseignant sans expérience qui enseigne un cours spécialisé auprès d'un public récalcitrant. Le contexte socioculturel de l'enseignant peut impacter sa façon de relever ces défis. L'usage de ressources pédagogiques joue rôle majeur pour faciliter l'enseignement et l'apprentissage. Un enseignant peut se saisir ou produire de telles ressources destinées aux apprenants mais également s'en servir pour répondre à ses propres questions. Une ressource pédagogique peut être considérée comme tout matériel didactique qui aide les enseignants à dispenser un enseignement efficace à leurs apprenants. Dans ce travail de thèse, nous avons pour objectif d'aider les enseignants à relever leurs défis au moyen de recommandations de ressources pédagogiques tenant compte de leur contexte qui est complexe et multiple.

Nous nous sommes concentrés sur les moyens de représenter et d'intégrer les différents contextes des enseignants. La méthodologie suivie implique une modélisation ontologique permettant un raisonnement sémantique et a permis le développement de l'ontologie du Contexte Sentimental et de Collaboration de l'Enseignant (TSCCO). Cette ontologie encapsule les différents contextes caractérisant les enseignants et vise à faciliter l'intégration des données.

Pour effectuer les recommandations, nous avons développé un Système de Recommandation de Ressources Pédagogiques Contextuelles (CAPRRS). CAPRRS combine des techniques de filtrage collaboratif et de filtrage basé sur le contenu avec une approche hybride de filtrage contextuel, guidée par différents contextes. Il s'appuie sur l'association de deux méthodes de sélection de caractéristiques (la variance et l'information mutuelle) afin d'identifier les caractéristiques pertinentes. Le filtrage collaboratif tire parti des traces d'interaction enregistrées, tandis que le filtrage basé sur le contenu prend en compte les données décrivant les ressources pédagogiques, offrant ainsi des recommandations plus personnalisées.

Les principales conclusions soulignent l'importance de la sélection des caractéristiques et de l'efficacité de CAPRRS dans la fourniture de recommandations de ressources pédagogiques conscientes du contexte. Ces résultats contribuent à une compréhension plus approfondie des relations qui peuvent exister entre les préférences des enseignants, les caractéristiques contextuelles et les approches de recommandation. Ils contribuent à la fois aux recherches en éducation et en systèmes de recommandation.

Acknowledgements

I extend my deepest appreciation to Prof. Marie-Hélène ABEL and Mr. Christine LAHOUD, the thesis's supervisors, for their unwavering patience, invaluable guidance, and insightful feedback throughout the entirety of my doctoral journey. Their expertise and support have been fundamental in shaping this thesis.

Furthermore, I express my sincere gratitude to the French University in Egypt (l'Université Française d'Égypte) for their support, without which this research would not have been possible.

I would like to acknowledge my esteemed members of the Heudiasyc laboratory whose collaborative spirit, constructive critiques, and moral encouragement have been invaluable sources of inspiration and growth. Special gratitude is extended to my office mates and colleagues for their unwavering support.

Finally, I am profoundly grateful to my beloved family, especially my parents, sister, and spouse, whose emotional support and unwavering belief in my abilities have been a constant source of motivation and resilience throughout this academic journey.

Their collective support and encouragement have been indispensable in the completion of this doctoral thesis.

Contents

Abstract	iii
Acknowledgements	vii
I Introduction	1
1 Introduction	3
1.1 Background Information and Problem Statement	3
1.1.1 Social Problem	4
1.1.2 Scientific Problem	5
1.2 Thesis Organization	6
II State of the Art	9
2 State Of Art	11
2.1 Introduction	11
2.2 Educational Environment	11
2.2.1 Definitions	12
2.2.2 Teacher Profile	13
2.2.3 Educational Institution	14
2.2.4 Pedagogical Resource	14
2.2.5 Summary	15
2.3 Context	16
2.3.1 Definitions	17
2.3.2 Different Contexts	17
2.3.3 Teacher Contexts	18
2.3.4 Summary	19
2.4 Knowledge Representation	21
2.4.1 The way from Data to Knowledge	21
2.4.2 Knowledge and Its Representation	22
2.4.2.1 Knowledge Types	23
2.4.2.2 Knowledge Representation Techniques	23
2.4.2.3 Semantic Web Technologies and Semantic Reasoning	25
2.4.3 Knowledge Representation in Education	26
2.4.4 Summary	29
2.5 Feature Selection	29
2.5.1 Definitions	31
2.5.2 Types of Features	31
2.5.3 Feature Selection Approaches	32
2.5.4 Unsupervised Filter Approaches	33

2.5.4.1	Chi-squared	33
2.5.4.2	R-regression	34
2.5.4.3	F-regression	34
2.5.4.4	Analysis of Variance (ANOVA)	34
2.5.4.5	Mutual Information	35
2.5.4.6	Variance-based Threshold	35
2.5.5	Ontology-based Feature Selection	35
2.5.6	Feature Selection in Education	36
2.5.7	Summary	39
2.6	Recommender Systems	40
2.6.1	Types of Two-dimensional Recommender Systems	41
2.6.2	Context-Aware Recommender Systems	42
2.6.3	Recommender Systems in Education	42
2.6.4	Peer Recommendations in Education	46
2.6.5	Sentimental Context Importance in Recommender Systems	47
2.6.6	Summary	48
III	Contribution	51
3	Contribution Overview	53
3.1	Approach Overview	53
3.2	Approach Components	55
3.2.1	Multiple Contexts of Teacher and their Representation (T-Box)	55
3.2.2	Data Collection (A-Box)	57
3.2.3	Context-Aware Recommendations for Teachers	57
4	Teacher Contexts and Their Representation	59
4.1	Knowledge Engineering Methodology	59
4.2	Teacher Sentimental and Collaboration Contexts Ontology (TSCCO)	62
4.2.1	Integration of Three Ontologies	63
4.3	Teacher Context Ontology (TCO)	64
4.3.1	Teacher	65
4.3.2	Learner	67
4.3.3	Environment	68
4.3.4	Resource	68
4.4	Mood-Detection Ontology (MDO)	69
4.4.1	Mood Detection Model	69
4.4.2	Ontological Representation of Mood Concepts	70
4.5	MEMORAE Collaboration-Context Ontology (MCC)	71
4.6	Complete Contexts of Teacher	71
4.6.1	Teacher Profiling	72
4.6.2	Living Context	73
4.6.3	Working Context	73
4.6.4	Sentimental State	73
4.7	Semantic Reasoning	74
4.7.1	Preprocessing SWRL Rules	74
4.7.2	Resource and Peer Recommendation SWRL Rules	75

5	Data Collection	79
5.1	Data Sources Layer	79
5.2	Questionnaire - Teacher Profile	79
5.2.1	Questionnaire Methods	80
5.2.2	Elements Identification of Questionnaire	81
5.2.3	Teachers Questionnaire Trends	82
5.2.4	Questionnaire Adaptation	83
5.2.5	Questionnaire Realization	86
5.3	French Teacher Context Dataset (Multiple Contexts of Teacher)	96
5.3.1	Working Context Data	96
5.3.2	Living Context Data	99
5.4	Pedagogical Resources Dataset	102
5.4.1	Repository Data Retrieval	104
5.4.2	Application Programming Interface (API) Data Retrieval	109
5.5	Activity and Feedback	111
5.5.1	Activity Tracking	112
5.5.2	Sentimental Feedback	112
5.6	Data Ontological Mapping	113
6	Context-Aware Recommendations for Teachers	121
6.1	Variance-based Importance Feature Selection for Teacher Contexts (VIFSTC)	123
6.1.1	Features Data Retrieval and Vectorization	125
6.1.1.1	Features Data Retrieval	125
6.1.1.2	Features Data Vectorization	126
6.1.2	Mutual Information	127
6.1.3	Variance-based Feature Selection	128
6.2	Context-Aware Pedagogical Resources and Peer Teachers Recommender System	130
6.2.1	Hybrid 2D Pedagogical Resource Recommender Approach	131
6.2.1.1	User-based Collaborative Filtering Recommender Approach	132
6.2.1.2	Content-based Filtering Recommender Approach	132
6.2.2	Context-Aware Recommender Approaches	135
6.2.3	Context-Aware Peer Teacher Recommender Approach	139
IV	Experimentation and Results	143
7	Experimentation	145
7.1	Experimental Data	145
7.2	Experimental Setup for VIFSTC Feature Selection	147
7.3	Experimental Setup for Recommender System	148
8	Results	151
8.1	Experiment with Mood Detection	151
8.2	Experiment with VIFSTC Feature Selection	153
8.2.1	Experiments with different contexts	154
8.2.2	Discussion of Feature Selection Experiments	156
8.3	Experiments with Recommender System	161

8.3.1 Experiments with Different Number of Recommendations . . .	161
8.3.2 Experiments with Different Context Variations	163
8.3.3 Experimentation with Peer Scenarios	166
8.4 Results Overview	168
V Conclusion and Perspectives	171
9 Conclusion	173
10 Perspectives	177
10.1 Short-Term Objectives	177
10.2 Long-Term Objectives	178
Bibliography	181

List of Figures

1.1	Interconnections between the struggles of teachers.	5
2.1	Illustration of the inter-connected relations between teaching styles and learning ones.	13
2.2	The DIKW hierarchy according to (Ackoff, 1989) from the point-of-view of information systems and artificial intelligence.	22
2.3	Example of ontologies knowledge representation.	24
2.4	An overview of connections between the selected related articles to our work.	37
2.5	Percentage of recommender systems fulfilling the required characteristics.	44
3.1	An overview of the contributed approach of this thesis.	54
3.2	A partial T-Box snippet of TSCCO highlighting multiple contexts of teacher.	56
3.3	A partial A-Box snippet of mapping collected data into TSCCO highlighting multiple contexts of teacher.	58
4.1	Contextual factors of teachers.	61
4.2	Screenshot of Protégé view of TSCCO upper level concepts.	63
4.3	A partial T-Box representation of Teacher Sentimental and Collaboration Contexts Ontology (TSCCO)	64
4.4	T-Box of TCO and MDO integration.	65
4.5	A T-Box overview of TCO highlighting the main concepts.	65
4.6	A T-Box example of teacher and learner representation in TCO.	66
4.7	A T-Box example of resource representation in TCO.	68
4.8	An overview diagram of Moodflow@doubleYou platform.	69
4.9	Mood levels obtained by Moodflow@doubleYou.	70
4.10A	A T-Box example of mood detection ontology (MDO).	71
4.11	Different contexts partial T-Box representation of Teacher Sentimental and Collaboration Contexts Ontology (TSCCO)	72
5.1	An overview of the proposed approach's layers.	80
5.2	Data selection from datasets concerning geolocalization of educational establishments.	98
5.3	Data selection from datasets concerning Prioritized educational establishments (REP, REP+).	100
5.4	Overall data selection from datasets concerning working context of a teacher.	101
5.5	Overall data selection from datasets concerning living context of a teacher.	106

5.6	Activity and mood representation from TSCCO. (T-Box component)	113
5.7	Example of SPARQL query to retrieve all activities within an organization	115
5.8	Example of retrieved activity data	116
5.9	An example of time-series signal recording for sentimental state monitoring by MoodFlow@doubleYou	116
5.10	An example of time-series signal with respect to different activities and showing the neutral mood level of the person	117
5.11	An example of activity and mood assertion to TSCCO (A-Box assertion component)	117
5.12	Selected concepts representation from TSCCO. (T-Box component)	118
5.13	Resource concept representation in TSCCO. (T-Box component)	118
5.14	Example of D2RQ customized mapping file for TSCCO:Resource	119
6.1	An overview of the recommender approach.	123
6.2	An overview of the Variance-based Importance Feature Selection for Teacher Contexts (VIFSTC) approach.	124
6.3	An example of vectorized list of features.	127
6.4	Correlation between features.	129
6.5	Mutual information importance of features.	129
6.6	Overview of the 2D hybrid recommender approach.	131
6.7	Overview of user-based collaborative filtering recommender approach (2DCFR).	133
6.8	Overview of content-based filtering recommender approach (2DCBR).	134
6.9	Overview of different context-aware filtering recommender approach.	136
6.10	Representation of the multi-dimensional matrix (multipartite graph) of the multiple contexts of a teacher.	137
7.1	The hierarchy of the used data.	146
7.2	Example of SPARQL query to retrieve teacher-related data	147
7.3	Example of SPARQL query to retrieve pedagogical resource data within the organization "PRE103"	147
7.4	Experiment model for VIFSTC approach.	148
7.5	Flowchart of the experimental procedure for VIFSTC approach.	149
8.1	Moodflow@doubleYou output for Teacher A.	152
8.2	Moodflow@doubleYou output for Teacher B.	153
8.3	A-Box example of integrating the mood of Teacher A with TSCCO.	154
8.4	A-Box example of integrating the mood of Teacher B with TSCCO.	155
8.5	Correlation coefficient of all features.	155
8.6	Information gain importance for all features.	156
8.7	The selected features across all methods.	157
8.8	Classification accuracy of all contexts across all feature selection methods.	157
8.9	Classification accuracy without sentimental context across all feature selection methods.	158
8.10	Classification accuracy without living context across all feature selection methods.	158

8.11 Classification accuracy without working context across all feature selection methods.	159
8.12 Average classification accuracy across all feature selection methods.	159
8.13 FCA generated concepts lattice for the contexts of 4 teachers.	160
8.14 F1 scores obtained by the compared recommender approaches with respect to the different number of recommendations K	162
8.15 F1 scores obtained by the compared recommender approaches while neglecting the contextual features related to the living environment.	163
8.16 FF1 scores obtained by the compared recommender approaches while neglecting the contextual features related to the working environment.	164
8.17 F1 scores obtained by the compared recommender approaches while neglecting the contextual features related to the sentimental state.	165
8.18 Average overall F1 scores obtained by the compared recommender approaches across all experiments.	166
8.19 A-Box TSCCO instances of Teachers A, B and C used in peer scenario.	169

List of Tables

2.1	Aggregation of the ISCED educational levels.	14
2.2	Summary of related concepts to the educational environment.	16
2.3	Summary of types of contexts according to each definition	18
2.4	Multiple contexts where a teacher coexists	19
2.5	Summary of context definitions from the teacher perspective.	20
2.6	Summary of surveyed knowledge representation models.	30
2.7	Summary of surveyed feature selection approaches.	40
2.8	Summary of reviewed ERRSs	45
4.1	List of SWRL rules used in our approach	77
5.1	The sections of questions across all questionnaires	89
5.2	The types of questions across all questionnaires	92
5.3	Sources of working context datasets	97
5.4	Columns of working context datasets	102
5.5	Sources of living context datasets	103
5.6	Columns of living context datasets	104
5.7	Columns of resources repository datasets with description as provided by authors	107
5.8	Properties of Activity concept in MCC ontology	114
6.1	List of features	126
8.1	Experiment summary for Teacher A	152
8.2	Experiment summary for Teacher B	152

List of Abbreviations

TCO	T eacher- C ontext O ntology
MDO	M ood D etection O ntology
MCC	M EMORAe C ollaboration- C ontext O ntology
TSCCO	T eacher S entimental and C ollaboration C ontexts O ntology
VIFSTC	V ariance-based I mportance F eature S election for T eacher C ontexts
CAPRRS	C ontext- A ware P edagogical R esource R ecommender S ystem
ISCED	I nternational S tandard C lassification of E Ducation
REP	R éseau d'Éducation P rioritaire
REP+	R éseau d'Éducation P rioritaire R enforcé
DIKW	D ata, I nformation, K nowledge, and W isdom
RDF	R esource D escription F ramework
RDFS	R esource D escription F ramework S chema
OWL	W eb O ntology L anguage
SWRL	S emantic W eb R ule L anguage
LD	L inked D ata
URI	U niform R esource I dentifier
SPARQL	S PARQL P rotocol and R DF Q uery L anguage
TEL	T echnology- E nhanced L earning
LOM	I EEE L earning O bject M etadata
OER	O pen E ducational R esources
ITS	I ntelligent T utoring S ystem
FS	F eature S election
ANOVA	A Nalysis O f V ariance
RF	R andom F orest
RSFS	R andom S ubset F eature S election
TF-IDF	T erm F requency I nverse D ocument F requency
FCA	F ormal C oncept A nalysis
RS	R ecommender S ystem
CARS	C ontext- A ware R ecommender S ystem
ERS	E ducational R ecommender S ystem
ERRS	E ducational R esources R ecommender S ystem
AAD	A verage A bsolute D eviation
SoIS	S ystem of I nformation S ystems
TASCQ	T eacher A s S ocial C ontext Q uestionnaire
TIMSS	T rends in I nternational M athematics and S cience S tudy
PIRLS	P rogress in I nternational R eadng L iteracy S tudy
TALIS	T eaching A nd L earning I nternational S urvey
OECD	O rganization for E conomic C o-operation and D evelopment
CNIL	C ommission N ationale de l' I nformatique et des L ibertés
GDPR	G eneral D ata P rotection R egulation

CAPMAS	Central Agency for P ublic M obilization A nd S tatistics
FTCD	F rench T eacher C ontext D ataset
ONISEP	l' O ffice N ational d' I nformation S ur les E nseignements et les P rofessions
RAMSESE	le R épertoire A cadémique et M inistériel S ur les E tablisements du S ystème E ducatif
INSEE	I nstitut N ational de la S tatistique et des E tudes E conomiques
API	A pplication P rogramming I nterface
MI	M utual I nformation
JSON	J avaScript O bject N otation
2DHR	2D H ybrid R ecommender
2DCFR	2D C ollaborative F iltering R ecommender
2DCBR	2D C ontent- B ased R ecommender
k-NN	k - N earest N eighbor
SVD	S ingular V alue D ecomposition
CAPTRA	C ontext- A ware P eer T eacher R ecommender A pproach
PSO	P article S warm O ptimization
RMSE	R oot M ean S quared E rror
SVM	S upport V ector C lassification
MLP	M ulti L ayer P erceptron
RBM	B ernoulli R estricted B oltzmann M achine
MAE	M ean A bsolute E rror

Part I

Introduction

Chapter 1

Introduction

1.1 Background Information and Problem Statement

Teachers face significant challenges in their careers, which can impact their professional development and even lead to failure (Natale, 1993). Understanding and addressing these struggles requires considering various factors, such as expertise of teacher, learners, and working environment (Natale, 1993; Veenman, 1984). The definition of these challenges varies based on these factors and is often related to reluctance of learners to the learning process influenced by their social and cultural environment.

The level of struggle may vary depending on the uniqueness of these factors such as novice teachers who work in specialized areas. These teachers tend to struggle more compared to experienced teachers (Monk, 2007). Considering the socio-cultural context of environments of teachers can help mitigate these challenges (Morton et al., 2016). However, defining the multiple contexts of teachers is complex and requires a comprehensive understanding of their work and out-of-work settings (Varelas, 2012).

Personalized recommendations based on the multiple contexts of teachers can address these struggles. Such recommendations include training materials and various pedagogical resources which play a crucial role in teacher support (Bušljeta, 2013; Tuomi, 2013). However, the availability and organization of these resources are often missing. Additionally, resource recommendations need to consider the multiple contexts of each teacher, including their profile, social/living environment, and educational/working environment.

Information systems, specifically recommender systems, offer a formal approach for handling vast amounts of information and can be applied in the educational context to provide personalized resource recommendations (Maguire, 1991; Lü et al., 2012). Context descriptions within recommender systems can vary based on research perspectives and the specificity of the multiple contexts of a teacher (Xiao et al., 2018). In the educational field, context can be defined from different viewpoints, such as geography, educational institution structure, learners, or teachers.

The research challenge is focused on multiple contexts and context-aware resource recommendations for teachers. By addressing this challenge, we aim to provide valuable insights and potential solutions to support teachers in their professional development.

1.1.1 Social Problem

In the educational field, teachers play a vital role in the delivery of quality education to learners. However, teachers encounter various problems that can significantly impact their effectiveness and job satisfaction. This section aims to delve into the underlying factors contributing to these problems and propose cumulative solutions to address them.

The first identified struggle is the lack of learner motivation, which adversely affects the engagement of learners and accordingly educational outcomes (Natale, 1993). Regardless of the social conditions of learners, their relationship with the educational institution emerges as a critical factor in maintaining motivation. Consequently, teachers exert additional effort to sustain their interest, often leading to a blurred boundary between their professional and personal lives. The second struggle revolves around the professional-personal life balance of teachers (Natale, 1993). The demanding nature of the profession often results in work addiction and challenges in maintaining a healthy equilibrium between work responsibilities and personal well-being. The third struggle addresses the developing of self-evaluation skills for teachers. Such skills become crucial for teachers to effectively manage their workload and prioritize their professional and personal commitments (Natale, 1993).

Furthermore, teachers working in specialized areas face additional distinct struggles (Monk, 2007). First, learner discipline and behavior in educational institutions pose significant challenges, diverting valuable teaching time. Additionally, the cultural context and traditions of these areas can affect the teaching-learning process. Both learners and teachers encounter difficulties in adapting to the cultural nuances, potentially impacting the retention and comprehension of new information (Pence and Macgillivray, 2008). Moreover, learner mobility within specialized areas disrupts educational continuity which is driven by factors such as motivation loss and family movements. The identified struggles faced by teachers, including the lack of learner motivation, professional-personal life balance, self-evaluation, discipline and behavior in educational institutions, cultural differences, and learner mobility, exhibit interconnectedness. These interconnections imply that addressing one problem can have a positive impact on other related challenges (Roeser et al., 2012). As shown in Figure 1.1, a comprehensive approach can be adopted to propose cumulative solutions that simultaneously tackle multiple struggles. For instance, novice teachers can benefit from tailored recommendations for suitable educational resources to strengthen course content, addressing the lack of learner motivation. Additionally, providing appropriate training courses during the early stages of their careers can equip teachers with effective instructional strategies and classroom management techniques, helping to alleviate discipline and behavior issues. Moreover, promoting cultural sensitivity and diversity training for teachers can enhance their understanding of different cultural backgrounds, fostering a more inclusive learning environment and mitigating the impact of cultural differences on the teaching-learning process.

By addressing these interrelated problems through cumulative solutions, we can better support teachers with diverse backgrounds and profiles. Such an approach acknowledges the complex nature of the challenges faced by teachers and seeks to provide comprehensive support that considers the interconnectedness of

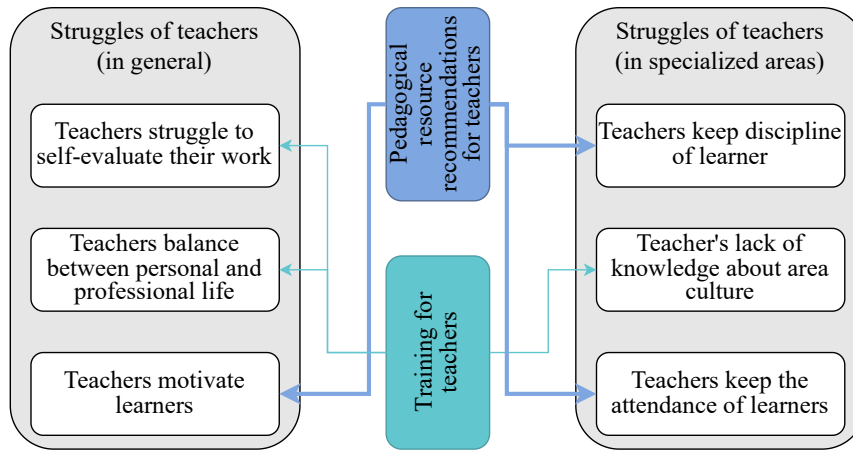


Figure 1.1: Interconnections between the struggles of teachers.

these struggles. Overall, by offering suitable resources and training courses, and considering the individual needs of teachers with different backgrounds, we can help teachers during the early stage of their career and make a positive impact in their classrooms and the education system.

1.1.2 Scientific Problem

In view of the stated struggles and challenges faced by teachers, there is a lack of comprehensive solutions that provide an extensive description of the contexts of teachers as well as personalized and context-aware recommendations (Manouselis et al., 2012). Existing educational descriptive data models target learners in a learner-centered approach (Papamitsiou and Economides, 2014). As a result, most of the educational data models can be considered learner-centered models (Romero and Ventura, 2020). Accordingly, there is a deficiency in the representation of different contexts of teachers and especially, the semantic representation (Jensen, 2019; Tarus, Niu, and Mustafa, 2018). Moreover, the integration of data from multiple contexts presents significant challenges, including variations in formats and schema (Golshan et al., 2017). A comprehensive approach is required to ensure the absence of data redundancy and facilitating data insertion and retrieval.

Existing recommender systems often overlook the diverse contexts in which a teacher coexists, including living environment, working environment, and sentimental state (Silva et al., 2023). These contextual factors play a crucial role in tailoring recommendations to personalized preferences and contexts of teachers. The subproblem of selecting the relevant factors representing each teacher is inherited from the previous problem (Adomavicius and Tuzhilin, 2010). This problem exists due to the diverse contexts of teachers which requires a precious selection of relevant factors from each context to enhance the recommendation performance and reduce its complexity (Lakiotaki, Matsatsinis, and Tsoukias, 2011). As a result, there is no definitive solution for this problem due to the lack of data organization of the mentioned contexts. Lastly, to evaluate the quality of recommendations, the problem of user evaluation and feedback arises which requires more than one method of feedback recording, such as sentimental state

detection and user activity tracking (Jannach et al., 2021). Such feedback can be used to enhance the upcoming recommendations.

After stating the various scientific problems, we deduce the research questions to be addressed in this thesis as follows:

1. How to represent the multiple contexts of teachers? And how to integrate data from these different contexts?
2. How taking account the multiple contexts of teachers to recommend suitable pedagogical resources to teachers using the most descriptive features?

These questions address various aspects of the stated problem, including understanding the struggles of teachers, defining their contexts, implementing a recommender system approach, considering contextual factors, addressing the multi-criteria problem, utilizing ontology representation and data integration, and ultimately generating context-aware resource recommendations for teachers.

1.2 Thesis Organization

The rest of this thesis is structured as follows:

PartII State of the Art encloses one chapter:

Chapter2 State of the Art begins with an introduction and establishes the research context and theoretical foundation. It then explores the educational environment, including definitions, teacher profiles, educational institutions, and resource descriptions. The concept of context is defined, different types of contexts are discussed, and teacher contexts are highlighted. Knowledge information representation and the various feature selection methods are explored, along with their relevance in education. Different types of recommender systems, particularly context-aware ones, are surveyed in the context of education.

PartIII Contribution comprises of four chapters:

Chapter3 Contribution Overview presents an overview of the recommender approach that adopting an ontology-driven approach that specifically caters to the perspective of teachers. It introduces the approach components: Teacher Sentimental and Collaboration Contexts Ontology (TSCCO), teacher data collection and the various recommender approaches which are explicitly discussed in the later chapters.

Chapter4 Teacher Contexts and Their Representation provides a comprehensive explication of diverse teacher contexts. Subsequently, leveraging the existing MEMORAe approach's ontology, an ontology-driven model is formulated to represent teacher contexts, referred to as the Teacher Sentimental and Collaboration Contexts Ontology (TSCCO).

Chapter5 Data Collection describes the data collection process to integrate relevant data for the multiple contexts which results in the creation of a French Teacher Context Dataset for experimentation. Interaction trace data and its importance in understanding teacher context

are discussed within the MEMORAe platform. A questionnaire is introduced as another data collection method, and its methods, elements, trends, adaptation, and realization are detailed.

Chapter6 Context-Aware Recommendations for Teachers commences with elucidating the process of extracting pertinent features from teacher contexts using the variance-based importance feature selection for teacher contexts (VIFSTC) approach. Subsequently, the feature selection technique is employed in context-aware pedagogical resource recommendation approach (CAPRRA), utilizing three distinct methods: PreFiltering, PostFiltering, and contextual modeling. Moreover, an innovative peer recommender approach is introduced, which aims to provide collaborators recommendations to teachers.

PartIV Experimentation is composed of two chapters:

Chapter7 Experimentation Protocol presents the variations of the dataset utilized for evaluating and testing the diverse components.

Chapter8 Results presents the subjective analysis of results, including sentiment feedback from teachers and its efficacy, as a means to assess the performance of the recommender approach.

PartV Conclusion and Perspectives contains two chapters:

Chapter9 Conclusion provides a comprehensive summary of the main findings of the thesis, while revisiting its objectives and research questions.

Chapter10 Perspectives delves into potential future directions and improvements for the proposed system, highlighting avenues for further research and development.

Part II
State of the Art

Chapter 2

State Of Art

2.1 Introduction

In the digital age, the realm of education is undergoing a transformative shift, propelled by technological advancements that facilitate personalized and adaptive learning experiences (Johnson et al., 2016). Central to this transformation is the integration of recommender systems, which are responsible for tailoring content and resources to different actors, thus enhancing the educational process (Bodily and Verbert, 2017). This section serves as a compass, guiding our exploration into the theoretical foundations that underlie the solutions of the research questions.

Navigating the complex educational landscape requires a profound understanding of its multifaceted dimensions, encompassing diverse online learning platforms and vast educational content repositories (Johnson et al., 2016). To address the challenge of curating content that resonates with teachers, our investigation starts by dissecting the different factors of educational environment that collectively shape the backdrop against which the representation of teachers in multiple contexts is based (Bodily and Verbert, 2017).

Our research objectives underpinning this inquiry revolve around the complexity of the recommendation process in educational contexts, driven by the dynamic interplay of multiple contexts that teachers engage with (Verbert et al., 2012). Knowledge representation emerges as a critical component, bridging raw data to meaningful knowledge, and enabling semantic reasoning to uncover deeper connections (Levesque, 1986). Additionally, feature selection techniques play a pivotal role in prioritizing attributes that contribute significantly to the recommendation process. We seek strategies to effectively capture and fuse the diverse contextual features that shape teachers' experiences (Liu et al., 2009). Finally, we explore the trends of recommender systems that offer highly personalized pedagogical resources aligned with multiple contexts of teachers (Dudek, 2000).

2.2 Educational Environment

The educational environment encompasses the physical, social, and psychological conditions of the learning and teaching process (Osguthorpe and Graham, 2003). It has a profound influence on learners' engagement, motivation, and learning outcomes and serves as the workspace for teachers' interactions with learners and peers (Hutchinson, 2003; Soemantri, Herrera, and Riquelme, 2010;

Goddard, O'Brien, and Goddard, 2006). In the context of recommender systems, the educational environment is defined as a set of contextual factors surrounding teaching and learning, including physical features, teacher profile, available resources, and socio-cultural context (Knoblauch and Hoy, 2008; George and Lal, 2019). This thesis focuses on the teaching process point-of-view and links it with recommender systems to provide contextually relevant recommendations, considering both teacher preferences and situational aspects in the evaluation of these recommendations (George and Lal, 2019).

This section contains a highlight of important definitions that provide a comprehensive understanding of the educational environment, in Section 2.2.1. Then, three main concepts are discussed from the teacher perspectives: teacher profile (Section 2.2.2), educational institution (Section 2.2.3), and pedagogical resources (Section 2.2.4).

2.2.1 Definitions

This section establishes a foundational understanding of key components within the educational environment, including teachers, teacher context, pedagogical resources, and educational institutions.

A teacher is an individual responsible for facilitating learning, designing lessons, assessing learner progress, and creating a supportive learning environment (Baylor and Ritchie, 2002; Grabinger and Dunlap, 1995; Horn and Little, 2010). Each teacher possesses a unique profile comprising professional experience, teaching style, and professional development (Ananiadou and Claro, 2009; Beijaard, Verloop, and Vermunt, 2000).

Teacher context encompasses the specific circumstances and factors influencing teachers' work and experiences, including living and working environment, educational institution, and sentimental state (Skaalvik and Skaalvik, 2011; Kelchtermans, 2005). It significantly shapes teaching practices and professional development needs.

A learner engages in knowledge, skill, or competency acquisition through education or training (Manuti et al., 2015). In the context of a recommender system, learners are students guided by teachers, and the system may consider learner profiles, learning styles, and preferences to enhance learning experiences (George and Lal, 2019).

Pedagogical resources encompass materials and tools supporting teaching and learning, such as textbooks, online courses, multimedia content, software, research articles, and interactive resources (Davis and Krajcik, 2005; Robin, 2016).

An educational institution refers to an organization providing formal education, including schools, colleges, universities, or other structured learning environments (Dib, 1988; O'Banion, 1997).

These definitions provide directional paths to explore more elements within the educational environment that are explicitly explored in the following subsection.

2.2.2 Teacher Profile

Teacher profiling is a crucial phase for understanding the educational environment, comprising various facets meticulously defined in literature (Mumtaz, 2000; Ananiadoui and Claro, 2009). The teacher profile is a comprehensive compilation of factors that intricately shape a teacher's teaching identity as follows:

Professional experience includes a teacher's previous employment history in education, detailing the duration of employment, institutions worked at, and roles and responsibilities held during that time. This aspect can be used to classify teachers into the different levels of experience.

Field of science or area of expertise is the specific area or discipline that teachers often specialize in. This factor highlights their expertise, aiding in matching them with the needs of educational institutions or other learning requirements.

Teaching style is the pedagogical approaches and teaching philosophy that are followed by teachers. Teaching styles can be categorized into five main types: authority, demonstrator, facilitator, delegator, and hybrid styles, each aligning with different learning styles as shown in Figure 2.1 (Opdenakker and Van Damme, 2006).

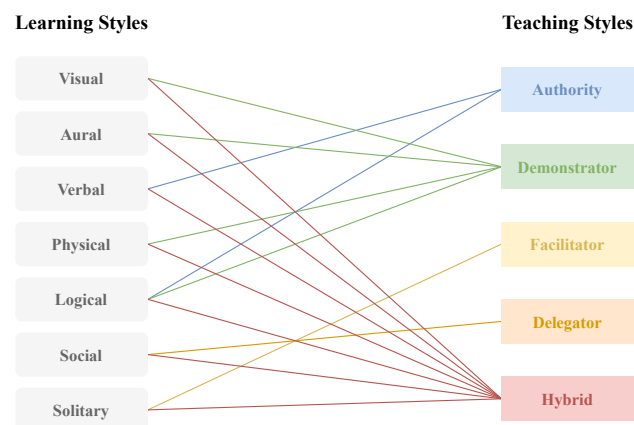


Figure 2.1: Illustration of the inter-connected relations between teaching styles and learning ones.

Professional development reflects a teacher's commitment to ongoing learning and staying current in their field through workshops, seminars, conferences, or courses attended. This factor is not considered as important as the previous ones but it can be of useful need to specify the weakness in a teacher profile.

These factors collectively provide a comprehensive portrait of a teacher's professional journey and pedagogical approach, contributing to an in-depth understanding of their role within the educational context.

2.2.3 Educational Institution

Educational institutions play a crucial role in society, providing structured education and instruction across various age groups and educational levels (Dib, 1988). These institutions encompass a wide spectrum of organizations committed to fostering learning and growth among students (O'Banion, 1997). More comprehensive definitions are achieved through additional elucidations of the different concepts of educational institutions, as follows:

Educational levels categorize educational institutions using frameworks like the International Standard Classification of Education (ISCED) (Darling-Hammond, 2007). This classification ranges from early childhood education and primary education to post-secondary non-tertiary education, short-cycle tertiary education, bachelor's or equivalent level, master's or equivalent level, and doctoral or equivalent level. Aggregating these levels, they can be broadly categorized into low education, medium education 1, medium education 2, and high education, as shown in Table 2.1.

Table 2.1: Aggregation of the ISCED educational levels.

Aggregated Category	ISCED 2011	Synonym
Low education	Levels 0-1	Primary levels
Medium education 1	Levels 2	Preparatory levels
Medium education 2	Levels 3-4	Secondary levels
High education	Levels 5-8	Undergraduate and postgraduate university levels

Specialized education programs aim to address educational inequalities and improve outcomes for students in disadvantaged areas such as REP (Réseau d'Éducation Prioritaire) and REP+ (Réseau d'Éducation Prioritaire Renforcé) in France (Stéfanou, 2017). These programs provide additional resources and support to educational institutions facing socio-economic challenges, with REP+ offering even more intensive reinforcement.

In summary, educational institutions serve as pillars of education across diverse levels and specializations, aiming to provide equitable opportunities for different learning contexts.

2.2.4 Pedagogical Resource

Pedagogical resources encompass a diverse array of materials, tools, and sources specifically designed or chosen to support teaching, learning, and educational activities (Davis and Krajcik, 2005). These resources play a pivotal role in enhancing the educational experience for both teachers and learners.

In the context of recommender systems, pedagogical resources take various forms according to (Davis and Krajcik, 2005; Ali et al., 2017), including:

- **Documents** These traditional resources provide structured content and information on specific subjects or topics in the form of e-books, textbooks and reference materials.
- **Digital and multimedia** have become increasingly prevalent in education with the advancement of technology. These can include interactive learning modules, online courses, educational websites, educational software, interactive games, and multimedia content (such as videos, animations, simulations) that enhance the learning experience and cater to different learning styles.
- **Research articles and journals** provide access to the latest research findings, studies, and academic literature. They are valuable for teachers and learners engaged in research, critical analysis, and deeper exploration of specific topics.

Pedagogical resources can be classified into simple and composite categories according to (Saleh, 2018). Simple pedagogical resources serve specific purposes in the learning process and can take the form of physical (traditional media and objects) or digital (multimedia files, programs, texts, URLs) resources. Composite pedagogical resources, on the other hand, consist of collections of multiple simple resources or more complex combinations of pedagogical content.

Access to pedagogical resources can be achieved through prerequisites, categorized as device (physical) and software (digital) requirements (Saleh, 2018). These prerequisites enable individuals to access and utilize the corresponding pedagogical resources effectively.

The integration and effective utilization of pedagogical resources continue to be a crucial aspect of modern pedagogy as educational technologies and methodologies evolve. These resources contribute to creating dynamic and enriching learning environments, fostering engagement, understanding, and overall educational enhancement.

2.2.5 Summary

In conclusion, the exploration of the educational environment in this section sheds light on the intricate interplay of factors that shape the teaching experience, as summarized in Table 2.2. The educational environment encompasses various dimensions, including the roles of teachers, the institutions that facilitate education, and the resources that support the learning process. By delving into definitions, teacher profiles, educational institutions, and pedagogical resources, a comprehensive understanding of the educational landscape emerges.

The essence of the educational environment aligns seamlessly with the focus of this thesis to describe the essential context of teachers' work environment which happens to be the educational environment. As teachers navigate the challenges of modern education, they require tailored recommendations that not only match their professional profiles and preferences but also consider the contextual factors in which these recommendations will be implemented.

The diverse pedagogical resources available play a pivotal role in enriching the teaching process and catering to the diverse needs and learning styles of students. Through this section, we incorporate a nuanced understanding of teacher

Table 2.2: Summary of related concepts to the educational environment.

Articles	Concept	Definitions
(Mumtaz, 2000; Ananiadoui and Claro, 2009; Beijaard, Verloop, and Vermunt, 2000)	Teacher Profile	<ul style="list-style-type: none"> • Professional experience • Field of science • Teaching style • Professional development
(Skaalvik and Skaalvik, 2011; Kelchtermans, 2005)	Teacher Context	<ul style="list-style-type: none"> • Living environment • Working environment • Educational institution • Sentimental state
(Dib, 1988; O'Banion, 1997; Stéfanou, 2017)	Educational Institution	<ul style="list-style-type: none"> • Educational levels • Specialized education
(Davis and Krajcik, 2005; Robin, 2016; Ali et al., 2017; Saleh, 2018)	Pedagogical Resources	<ul style="list-style-type: none"> • Form • Type • Prerequisite

profiles, the attributes of educational institutions, and the varied forms of pedagogical resources that contribute to a more robust and effective representation of the educational environment. Through the intricate interplay of these elements, this thesis endeavors to create a bridge between the theoretical understanding of pedagogical resources and their practical implementation within the multiple contexts of teachers.

2.3 Context

Context is a complex concept that plays a crucial role in enabling information systems, including user interactions (Adomavicius and Tuzhilin, 2010; Verbert et al., 2012). The main role of the context is to offer an understanding and adaptivity to the surrounding environment according to each user. It serves as the core of comprehending how external events influence user's behavior and consequently, decision-making processes (Carver and Turoff, 2007). Context allows information systems to deduce possible activities and information needs, facilitating appropriate explanation and enhancing the effectiveness of interactions (Avgerou, 2001). Additionally, context-aware approaches leverage the concept of

context to deliver personalized and appropriate information or services to users based on their specific needs (Adomavicius and Tuzhilin, 2010).

However, a unified definition of context remains vague, leading to various definitions and implementations of contextual information across different context-aware applications (Bazire and Brézillon, 2005). This diversity poses challenges in finding a standardized description of context. Additionally, in the realm of education, context plays a crucial role in shaping the teaching and learning processes, as well as the overall educational experience (Cranton and Carusetta, 2002; Verbert et al., 2012). Therefore, this section aims to delve into the definition of context and explore its various definitions and types, especially within the educational setting.

2.3.1 Definitions

Context is a multifaceted concept defined in various fields and disciplines. Linguistically, it refers to the surrounding circumstances that help explain a specific situation or activity¹. In social sciences, it encompasses broader social, cultural, and historical factors that influence individuals, groups, or events, including the social environment, cultural background, and economic conditions (Helliwell and Putnam, 2004). In psychology, context refers to environmental cues, stimuli, or conditions that provide additional information and influence perception, interpretation, and cognitive processing, incorporating factors like the physical environment, social cues, temporal information, and personal experiences (Schwarz and Sudman, 2012).

In the realm of recommender systems, context is defined as the set of environmental factors and situational information surrounding a user, capable of influencing their preferences, needs, and behaviors. Several key definitions have emerged in this context, emphasizing its role in characterizing the user's situation and improving the accuracy of recommendations. These definitions are summarized in Table 2.3 which mentions four main definitions from (Dey, 2001), (Adomavicius and Tuzhilin, 2005a), (Baltrunas, Ludwig, and Ricci, 2011), and (Sieg, Mobasher, and Burke, 2007). These definitions agree upon three main types: physical, social, and interactivity but the modal/sentimental type is mentioned by only two definitions.

2.3.2 Different Contexts

Context is further characterized by contextual dimensions and factors (Adomavicius and Tuzhilin, 2010). Contextual dimensions represent distinct aspects of context used for description or categorization, such as time, location, social relationships, and user preferences. Contextual factors are specific elements or components within these dimensions, contributing to the overall context and impacting the recommendation process. There are four broad types of contexts from the point of view of recommender systems: physical, social, modal/sentimental, and interactivity (Adomavicius and Tuzhilin, 2010; Raza and Ding, 2019). The physical type captures the time, location, and activity of the user during the

¹https://www.oxfordlearnersdictionaries.com/definition/american_english/context
- <https://dictionary.cambridge.org/dictionary/english/context>

Table 2.3: Summary of types of contexts according to each definition

Definition	Context Types
(Dey, 2001)	– Physical – Social – Interactivity
(Adomavicius and Tuzhilin, 2005a)	Situation-based – Physical – Social – Interactivity
(Baltrunas, Ludwig, and Ricci, 2011)	– Physical – Social – Modal/Sentimental
(Sieg, Mobasher, and Burke, 2007)	– Physical – Modal/Sentimental – Interactivity

recommendation intention to be used. The social type incorporates social relationships, connections, and influences in the recommendation process. The modal/sentimental one portrays the user's current sentimental state, including their objectives, mood, experience, and cognitive abilities. Interactivity type refers to the interaction activity between a user and different types of dimensions.

In the educational context and within recommender systems, context pertains to information characterizing the situation of a user or teacher, encompassing environmental factors, situational details, location, time, social context, activity, and the dynamic characteristics describing the current situation, including sentimental state. The primary objective of considering context is to enhance the accuracy and relevance of recommendations in various educational contexts.

In order to consider all these contexts for teachers, it makes the task of recommender systems more challenging by taking into account the different types of contextual dimensions such as physical dimension (work environment), social dimension (living environment), sentimental dimension (mood), and interactivity dimension (interaction activities). But incorporating these contextual dimensions helps understanding and addressing the diverse needs and preferences of teachers in different situations.

2.3.3 Teacher Contexts

The context can be classified differently when the educational/learning environment is considered (Raza and Ding, 2019). Additional types or extension of previously defined types arise with respect to the relevancy of some important factors that need to be considered within the educational/learning process (Leidner and Jarvenpaa, 1995). For example, social context can be extended by including economical and cultural dimensions. Socioeconomic context refers to the social and economic factors that impact education (Leidner and Jarvenpaa, 1995). It

includes factors such as the income levels, employment rates, and educational attainment of the community or region. Socioeconomic context can influence access to resources, funding for schools, and the overall educational opportunities available to students. Cultural context encompasses the sociocultural factors that shape education (Leidner and Jarvenpaa, 1995). It includes factors such as cultural norms, values, beliefs, and practices. Cultural context influences teaching approaches, curriculum design, and the inclusion of culturally relevant content in education. It also impacts student identity, motivation, and engagement in the learning process.

However, teacher identity is seen as involving multiple sub-identities, as being an ongoing process of construction, and as being influenced by multiple types of contexts such as social context and physical context (Akkerman and Meijer, 2011). These characterizations challenge the traditional representation of teacher identity as a possession of a defined set of assets (Beauchamp and Thomas, 2009). These multiple contexts indicate the different environments where a teacher co-exists including social/living environment, physical/working environment, and modal/sentimental state (Cooper and Olson, 2020; Gonzalez et al., 2007). Table 2.4 shows the different types of contexts where a teacher coexists in addition to their dimensions and factors. This representation illustrates the level of complexity that can be achieved by the problem of contexts for a single teacher. Therefore, we can include some dimensions of the socioeconomic context and cultural context with their corresponding teacher's contexts. For example, we can include location and area type dimensions that are represented in the cultural context as well the socioeconomic context through social and physical contexts which simplifies their representations.

Table 2.4: Multiple contexts where a teacher coexists

Context	Type	Dimensions
Living environment	Social context + Socioeconomic context	- Location - Area type
Working environment	Physical context + Cultural context	- Location - Area type - Educational institution type
Sentimental state	Modal/sentimental context	- Mood level - Activity

2.3.4 Summary

Context is a dynamic concept that significantly impacts information systems and recommender systems. It serves as a framework for understanding and adapting to the surrounding environment, ultimately enhancing user interaction and recommendation processes. In the realm of education, context plays a vital role in shaping teaching and learning experiences.

The diverse definitions of context across various disciplines and contexts highlight the intricate nature of this concept. In the context of recommender systems, several definitions emphasize the importance of considering various dimensions and factors that characterize a user's situation. These definitions include the well-known perspectives of Dey, Adomavicius and Tuzhilin, Baltrunas et al., and Sieg et al., all of which highlight context's role in improving recommendation accuracy and personalization. Context is characterized by contextual dimensions and factors. Contextual dimensions represent distinct aspects that can categorize or describe a context, while contextual factors are specific elements within those dimensions. This distinction allows for a deeper understanding of the factors influencing user behavior and preferences.

From the perspective of teachers, context becomes even more complex due to diversity of context types and dimensions, as summarized in Table 2.5. For instance, multiple contexts of teacher extend beyond the traditional physical and social settings to encompass socioeconomic and cultural types. These types shape teachers' interactions with pedagogical resources and peer teachers.

Table 2.5: Summary of context definitions from the teacher perspective.

Articles	Teacher Context	Context Type
(Dey, 2001; Adomavicius and Tuzhilin, 2005a; Baltrunas, Ludwig, and Ricci, 2011)	Living environment	Social context
(Leidner and Jarvenpaa, 1995; Raza and Ding, 2019)		Socioeconomic context
(Dey, 2001; Adomavicius and Tuzhilin, 2005a; Sieg, Mobasher, and Burke, 2007; Baltrunas, Ludwig, and Ricci, 2011)	Working environment	Physical context
(Dey, 2001; Adomavicius and Tuzhilin, 2005a; Sieg, Mobasher, and Burke, 2007)		Interactivity context
(Leidner and Jarvenpaa, 1995; Raza and Ding, 2019)		Cultural context
(Sieg, Mobasher, and Burke, 2007; Baltrunas, Ludwig, and Ricci, 2011)	Sentimental state	Modal/sentimental context

The representation of these multiple contexts is stated by three contexts: living environment, working environment, and sentimental state which covers five different types of contexts: social, physical, sentimental, socioeconomic, and cultural contextual types. However, the representation of these contexts must be carefully handled to ensure that the collected knowledge is efficiently utilized within any further process. Through the next section, the representation of such knowledge is discussed along with the educational environment.

2.4 Knowledge Representation

Knowledge representation is a broad concept that encompasses various aspects and dimensions (Brachman and Levesque, 1985). Knowledge representation involves the systematic organization, structuring, and encoding of knowledge and information in a manner that enables effective understanding, manipulation, and utilization. It goes beyond mere data storage and retrieval by capturing the semantics, relationships, and context of information, allowing for more sophisticated reasoning, inference, and analysis (Chandrasegaran et al., 2013).

In addition to representing facts, concepts, and relationships, knowledge representation often includes mechanisms for capturing uncertainty, ambiguity, and context (Van Harmelen, Lifschitz, and Porter, 2008). It aims to provide a comprehensive and coherent representation of knowledge that can be shared, communicated, and processed by both humans and machines (Brachman and Levesque, 1985). An essential aspect of knowledge representation is abstraction (Van Harmelen, Lifschitz, and Porter, 2008). It involves identifying and extracting the essential features and characteristics of objects, concepts, or phenomena, while omitting irrelevant or redundant details. Abstraction allows for the generalization of knowledge, enabling the application of learned principles to new situations and the extraction of insights from a large volume of data (Van Harmelen, Lifschitz, and Porter, 2008).

Furthermore, knowledge representation is closely tied to knowledge acquisition and knowledge engineering (Van Harmelen, Lifschitz, and Porter, 2008). It involves eliciting, organizing, and structuring knowledge from domain experts and other sources, and transforming it into a format suitable for storage, retrieval, and manipulation. This process often requires modeling tools, methodologies, and frameworks to capture and represent knowledge effectively. Knowledge representation also plays a vital role in knowledge-based systems, expert systems, and artificial intelligence applications (Van Harmelen, Lifschitz, and Porter, 2008). It enables the construction of knowledge bases, ontologies, or knowledge graphs that serve as the foundation for intelligent reasoning, decision-making, and problem-solving. By representing knowledge in a structured and machine-readable form, these systems can leverage automated reasoning algorithms, inference engines, and knowledge-driven processes.

2.4.1 The way from Data to Knowledge

The concepts of data, information, knowledge, and wisdom form a hierarchical structure known as the DIKW hierarchy Figure 2.2 (Ackoff, 1989). In this hierarchy, data is the foundational level in the DIKW hierarchy within informatics (Dalrymple, 2011). It comprises raw and unprocessed elements, such as facts, figures, symbols, or observations. These data elements are discrete and lack inherent context or meaning. Data can manifest in various forms, including numerical values, textual information, images, or any other representational format. It serves as the fundamental building block upon which the subsequent levels of the hierarchy are constructed.

Moving up the hierarchy, information represents the transformation of data into a more structured and meaningful state (Kebede, 2010; Dalrymple, 2011). This transformation process involves organizing, processing, and adding context

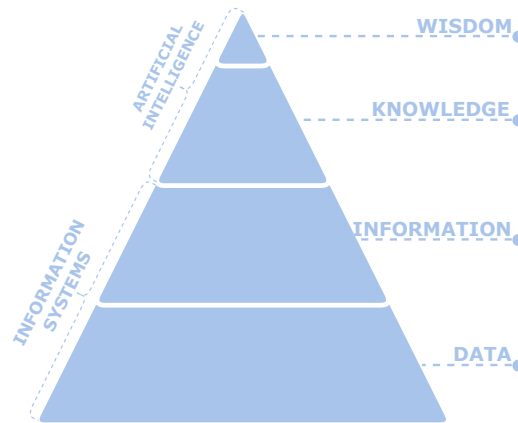


Figure 2.2: The DIKW hierarchy according to (Ackoff, 1989) from the point-of-view of information systems and artificial intelligence.

to raw data. Information adds clarity, interpretation, and relevance to data, making it more accessible and interpretable for individuals and systems. It is the bridge that connects raw data to knowledge, providing the necessary context for understanding and decision-making.

Knowledge advances beyond information as it encapsulates not only organized and contextualized information but also understanding, insights, and the capability to effectively apply this information (Kebede, 2010). Knowledge signifies a deeper level of awareness and familiarity with a specific subject or domain. It is acquired through experiences, learning, and the assimilation of information. Knowledge involves synthesizing and connecting pieces of information to form a coherent and actionable understanding.

At the pinnacle of the DIKW hierarchy lies wisdom which represents the highest level, characterized by individuals' ability to make judicious and sound decisions (Kebede, 2010). This capacity is based on their accumulated knowledge and extensive experience. Wisdom transcends mere knowledge; it involves discerning the most suitable course of action in complex and uncertain situations. In essence, wisdom bridges the gap between information and practical, informed action, embodying the culmination of the DIKW progression within informatics.

2.4.2 Knowledge and Its Representation

We can conclude that knowledge is the top of the informatics pyramid and hence, it represents the insights and application of information. Therefore, knowledge representation is indeed a crucial component of information systems (Levesque, 1986). By providing a structured framework for organizing and manipulating knowledge, it enables information systems to reason, make decisions, and perform complex tasks (Levesque, 1986).

By representing knowledge in a structured form, information systems can capture domain-specific information, facts, relationships, rules, and constraints (Hotz et al., 2014). This representation can be used to model real-world scenarios, understand complex situations, and generate intelligent behavior. It allows

these systems to apply logical reasoning, make inferences, and draw conclusions based on the available knowledge (Hotz et al., 2014).

Knowledge representation also plays a vital role in machine learning, where data is transformed into a structured format that allows algorithms to extract patterns, learn from examples, and make predictions (Carbonell, Michalski, and Mitchell, 1983). Furthermore, knowledge representation facilitates knowledge sharing, collaboration, and interoperability between different information systems. By adopting standardized representation languages and ontologies, systems can exchange and integrate knowledge (Nickel et al., 2015).

2.4.2.1 Knowledge Types

By representing the different elements of knowledge, information systems can acquire, organize, and apply knowledge to understand, interact with, and reason about the world in a manner that aligns with human-like intelligence. Each of the different types of knowledge serves a unique purpose according to the varieties that they target (Van Harmelen, Lifschitz, and Porter, 2008). The different types of knowledge are classified according to (Van Harmelen, Lifschitz, and Porter, 2008) into declarative, procedural, meta, heuristic, and structural knowledge.

Declarative knowledge constitutes objective information, such as facts, rules, and propositions. It can be expressed logically and is easily communicable and shareable. Procedural knowledge is centered on understanding how to execute specific tasks. It encompasses step-by-step procedures, rules, strategies, and algorithms necessary for achieving particular objectives. This knowledge is often represented using programming languages or algorithms. Meta knowledge delves into knowledge about knowledge itself, encompassing an understanding of the characteristics, limitations, and properties of knowledge. Heuristic knowledge is derived from past experiences and domain expertise. It includes practical rules, strategies, or techniques that guide decision-making and problem-solving processes. Structural knowledge pertains to comprehending the organization and relationships within complex problems or systems. It aids in breaking down intricate issues into more manageable sub-problems.

Each type of knowledge serves a unique purpose, and their representation enables information systems to understand, reason, and make informed decisions in various domains and tasks. By combining and utilizing different types of knowledge, information systems can exhibit intelligent behavior and perform complex cognitive tasks more effectively.

2.4.2.2 Knowledge Representation Techniques

Knowledge types are represented through different techniques of representation in information systems according to the definition of each type (Gregor, 2006). This knowledge representation is the process of creating structured models or frameworks to encode and store information and knowledge that can be used by information systems (Chandrasegaran et al., 2013). The purpose of knowledge representation is to enable information systems to understand, reason, and make informed decisions based on the encoded knowledge (Chandrasegaran et al., 2013). There are several approaches to knowledge representation, each with

its own knowledge types (Markman, 2013): logical representation, semantic networks, frames, ontologies.

Logical representation employs formal logic to express knowledge, employing logical rules, predicates, and symbols to convey relationships and facilitate logical reasoning (Levesque and Lakemeyer, 2001). Semantic networks utilize nodes and links to represent knowledge, with nodes denoting concepts or objects and links illustrating their relationships (Markman, 2013). This method is adept at capturing hierarchical structures and relationships (Borge-Holthoefer and Arenas, 2010). Frames structure knowledge as collections of attributes associated with objects or concepts, each containing slots for specific properties or links to other frames (Noy, Ferguson, and Musen, 2000; Rosse and Mejino Jr, 2003). This approach excels in representing structured knowledge and inheritance relationships (Lassila and McGuinness, 2001).

Ontologies provide a formal, explicit specification of concepts, properties, and relationships within a specific domain (Fensel and Fensel, 2001). They define a common vocabulary and a set of rules for reasoning about the domain knowledge (Noy, McGuinness, et al., 2001). Ontologies are commonly used in knowledge-based systems and semantic web applications (Davies, Studer, and Warren, 2006). In ontology-based knowledge representation, the T-Box (Terminology Box) and A-Box (Assertion Box) are used to organize and represent knowledge within a specific domain (Grimm, 2009), as shown in the example Figure 2.3 and as follows:

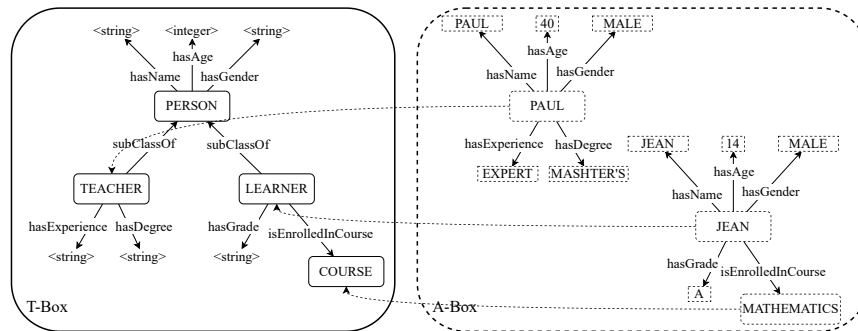


Figure 2.3: Example of ontologies knowledge representation.

T-Box represents the hierarchical structure of concepts, relationships, and properties within the domain. It defines the vocabulary and the schema of the ontology through general knowledge about the domain and is shared among multiple instances in the A-Box.

A-Box represents the specific instances or individuals in the domain and their relationships. It includes assertions or statements about the instances, specifying their properties, relationships, and values. A-Box provides concrete, instance-level knowledge about the domain. The methodology presented by ontology 101 guide (Noy, McGuinness, et al., 2001) offers a systematic approach to implement and construct new ontologies.

These representation methods vary in expressiveness, scalability, and suitability for different domains and problems. The choice depends on system requirements and the type of knowledge being represented. Declarative knowledge,

such as facts and rules, suits logical representation as well as ontologies (Patel and Jain, 2018). Procedural knowledge, involving task procedures, aligns with frames (Jakus et al., 2013). Meta-knowledge fits ontologies (Cristani and Cuel, 2005), heuristic knowledge suits semantic networks (Gopal, 2019), and structural knowledge finds its place in semantic networks (Jonassen, 2003).

It's important to note that this matching can be misfit if the choice of the specific requirements of the information system, the nature of the knowledge, and the problem domain oppose the nature of knowledge representation approach. Accordingly, a combination of different representation approaches may be used to capture and represent different types of knowledge within a system.

2.4.2.3 Semantic Web Technologies and Semantic Reasoning

The Semantic Web is a paradigm that extends the World Wide Web by adding a layer of meaning to the existing web content, allowing machines to understand and interpret the data (Fensel, 2005). Key components of the Semantic Web include ontologies, Resource Description Framework (RDF), and reasoning engines (Fensel, 2005). In the context of semantic web, ontologies define a formal vocabulary to represent knowledge and relationships within a specific domain and they are created using tools like Protégé framework². Protégé is a widely used, open-source software platform and ontology editor specifically designed for creating and managing ontologies. The platform, developed at Stanford University, is a key tool in the field of knowledge representation, semantic web development, and ontology engineering.

On the other hand, RDF³ serves as the backbone of the Semantic Web, allowing data to be expressed as subject-predicate-object triples (Fensel, 2005). RDF Schema (RDFS⁴) extends RDF by providing a basic vocabulary for defining ontologies (Fensel, 2005). Together, RDF and RDFS facilitate the creation of structured, interoperable data, which is crucial for representing and linking pedagogical resources and their metadata. OWL⁵ is a more expressive ontology language that enables the specification of complex relationships and constraints within ontologies (Fensel, 2005). This is particularly useful when modeling the intricate pedagogical concepts and their interconnections in educational contexts.

Semantic reasoning is fundamental in knowledge representation for information and recommender systems, enabling the inference of new knowledge and informed decision-making based on structured ontological relationships (Henze, Dolog, and Nejdl, 2004; Blanco-Fernández et al., 2008). It offers deductive capabilities by deducing insights from existing facts and rules, filling informational gaps. Additionally, it identifies inconsistencies, ensuring the integrity of knowledge representation. It extends beyond deduction to encompass abduction and induction, further enhancing domain understanding.

Semantic web rule language (SWRL⁶) rule is a formal statement within the Semantic Web framework that defines a condition-action pair, where specific conditions lead to certain actions or conclusions. SWRL is an extension of the

²<http://protege.stanford.edu>

³<https://www.w3.org/RDF/>

⁴<https://www.w3.org/wiki/RDFS>

⁵<https://www.w3.org/OWL/>

⁶<https://www.w3.org/submissions/SWRL/>

Web Ontology Language (OWL) and is used to represent semantic reasoning in terms of rules in the Semantic Web (Henze, Dolog, and Nejdl, 2004). SWRL rules are primarily designed to enhance the reasoning and inferencing capabilities of ontologies and knowledge bases. The SWRL rules are implicated between antecedents and consequent in the form *antecedent* \rightarrow *consequent*. Conjunctive atoms are used to express these antecedents and consequents. Conjunctive antecedents are represented as the following atoms: $a_1 \wedge a_2 \wedge \dots \wedge a_i \wedge \dots \wedge a_n$, where $1 \leq i \leq n$. In this work, atoms are classified into three forms: class $C(?x)$, property $P(?x, ?y)$, and "sameAs" relation *sameAs*($?x, ?y$). Class atom $C(?x)$ holds if x is considered an instance of class C , while property atom $P(?x, ?y)$ holds if x and y have a mutual relation P .

Incorporating Semantic Web techniques, such as Protégé, OWL, RDF, RDFS, and SWRL, holds the potential to enhance the effectiveness of resource discovery and recommendation in educational settings. These technologies provide a robust framework for modeling, representing, and semantically enriching the previously mentioned knowledge about the educational domain.

On the other hand, Linked Data (LD) refers to a set of principles and best practices for publishing, sharing, and interlinking different types of knowledge representation using Semantic Web technologies (Heath and Bizer, 2022). One key aspect of LD is the use of URIs (Uniform Resource Identifiers) as identifiers for resources (Heath and Bizer, 2022). URIs provide unique and globally resolvable identifiers for resources, ensuring their unambiguous identification and enabling linking and integration of data across different sources (Heath and Bizer, 2022). This graph-based representation facilitates the linking of resources and the expression of complex relationships between them. To query and retrieve data from LD sources, the SPARQL query language is commonly used. SPARQL enables powerful and flexible querying capabilities, allowing users to express complex patterns and retrieve specific information from linked datasets (Hartig, Bizer, and Freytag, 2009).

2.4.3 Knowledge Representation in Education

The adoption of knowledge representation principles and technologies has shown promise in addressing the challenges of interoperability in educational contexts (Pereira et al., 2017). Knowledge representation, which builds upon the foundations of the Semantic Web, provides a set of well-established principles and standards for sharing data on the Web (Shadbolt, Berners-Lee, and Hall, 2006). In educational domain, knowledge representation plays a crucial role in capturing and organizing data, metadata, and relationships between resources (Katagall et al., 2015).

By adopting Semantic Web and Linked Data principles and technologies, educational communities can benefit from enhanced data sharing, integration, and reuse (Dietze et al., 2013). LD promotes the publication of structured data on the Web, enabling researchers, teachers, and learners to access and combine diverse sources of information seamlessly (Greenhow, Robelia, and Hughes, 2009). While challenges related to heterogeneity and the need for standardization persist, the LD approach provides a promising methodology for realizing a framework for achieving interoperability in educational contexts. In recent years, the focus of research in technology-enhanced learning (TEL) has primarily revolved around

promoting interoperability and reusability of learning resources and data. Consequently, this has resulted in a fragmented landscape characterized by competing metadata schemas. These schemas range from general-purpose ones like Dublin Core (Board, 2018) to educational-specific schemas such as IEEE Learning Object Metadata (LOM) (RISK, 2002). Educational resource repository providers leverage these technologies to facilitate interoperability. Moreover, the incorporation of social data and attention metadata, which captures perceptions of learning resources, has become increasingly valuable for tailoring learning experiences to meet specific user needs and requirements. Despite the vast amount of educational content and data shared openly on the web, the integration process remains costly due to the isolation of different learning resources from each other and their reliance on diverse implementation standards (De Santiago and Raabe, 2010).

Despite the abundance of educational data already accessible on the Web through diverse schemas and interface mechanisms, the primary challenge lies in adopting Linked Data (LD) principles and vocabularies, while harnessing the existing non-LD compliant data available on the Web (Dietze et al., 2013). The key research challenges are addressed through the adoption of LD principles in science and education (Dietze et al., 2013). These challenges are encapsulated as infrastructural interoperability, semantic and syntactic interoperability, and formalizing domain knowledge.

Interoperability challenges in educational contexts are multifaceted and addressed through Linked Data (LD) principles. These challenges encompass infrastructural, semantic, and syntactic dimensions. Infrastructural interoperability confronts the heterogeneity of storage and interface methods in educational data (Yu et al., 2011). LD employs RDF, SPARQL, and standardized HTTP endpoints to facilitate data exposure, integration, and sharing. Semantic and syntactic interoperability tackle variations in schemas, vocabularies, and representation languages. LD adopts RDF(S) as a unified data representation language and promotes consistent dereferenceable URIs for broad data and schema element reuse. Features like `owl:sameAs` statements aid in linking heterogeneous datasets, supporting educational processes' citing and extension needs. Moreover, LD's concept of Linked Open Data emphasizes data's public and open nature, aligning with the transparency goals of the Open Educational Resources (OER) approach in education. Formalizing domain knowledge remains challenging, but LD has generated domain-specific vocabularies and facilitates alignment through standards like RDFS and SKOS, eliminating the necessity for a single domain model (Dietze et al., 2013).

Linked Data and ontologies have a strong connection in the educational domain. Ontologies provide a formal and structured representation of knowledge within a specific domain, capturing concepts, relationships, and properties (Guarino, 1995). They serve as a foundation for organizing and categorizing information in a meaningful way. In the educational domain, ontologies play a crucial role in facilitating the representation and organization of pedagogical resources, learning activities, teacher and learner profiles, and other relevant information (Aroyo et al., 2006). Linked Data leverages ontologies by using them as the basis

for creating linked educational data (Keßler, d'Aquin, and Dietze, 2013). Ontologies define the concepts, relationships, and properties within a specific educational domain, enabling data to be linked and related in a meaningful way. This linkage facilitates more accurate and contextually relevant recommendations, personalized learning experiences, and improved educational analytics.

Ontologies have been used in different fields with various purposes. Particularly, ontologies are integrated with recommender systems to connect users with their preferences according to the recommended item, or to represent the selected knowledge (Martinez-Cruz et al., 2015). Most researchers tend to use ontologies to model the relation of learners with their preferences for learning resources. On the other hand, the modeling of the relationship between teachers and learning resources is mostly neglected as researchers prefer to couple teachers with pedagogical planning and design.

Some remarkable educational ontologies are not used with resource recommender systems but for data representation in an e-learning framework with a task-oriented approach. An example for this type of ontologies was introduced by (Bhattacharya, Tiwari, and Harding, 2012) which represents the learning objects and learning methodology in an e-learning system. The introduced ontology provides the representation of context of learners but with the consideration of limited factors such as lesson, location, and environment. Other ontologies represent the resources in the context of an e-learning system such as the ontology introduced by (Mouromtsev et al., 2013) to represent learning resources for a distance-education system. The system represents and structures three main educational process concepts: curriculum, exam, and library. Educational resources are transferred to distance learners encouraging the mobility of learners. Another ontology was offered by (Dermeval et al., 2019) to facilitate the connection between e-learning systems and the innovative gamification activities which help the learners to be more motivated to engage in the learning process. A teaching model was proposed by (Chang et al., 2020) to represent teaching rules for an intelligent tutoring system (ITS). The system uses data mining techniques to extract rules from educational sessions and map it to the system ontology.

As previously stated, most researchers favor the ontology models of learners instead of teachers. One of these examples is the learner model which was introduced in (Clemente, Ramírez, and De Antonio, 2011) to represent an adult learner's characteristics. The ontology is integrated into a system for distance learning. The Adaptive-Courseware Tutor learning system uses the knowledge information about a learner to construct a learner representation using both Bayesian network and ontology (Grubišić, Stankov, and Peraić, 2013). The system is adapting the learning contents according to the collected information about the learner's knowledge. The learner context is one of the important aspects to consider in a distance e-learning system. Learning styles are an important aspect to be considered in a learner's ontology. The ontology representation of (Labib, Canós, and Penadés, 2017) explains the different learner models and the corresponding learning styles. The research proposes that the usage of this ontology provides the teacher with the main characteristics of learners. The ontology that was proposed by (Akharraz, El Mezouary, and Mahani, 2018) is concerned with describing the learner context according to multiple descriptive points such as computing necessity, location, time, physics, user activity, user

profile, and social contexts. A domain-independent ontology is used to model the learner profile and the corresponding preferences (Grivokostopoulou et al., 2019). The ontology is used in an intelligent tutoring system for higher education learners. The ontology, introduced by (Ilkou et al., 2021), is designed for representing online learning resources in the context of personalized learning systems, particularly focusing on educational and career-oriented aspects. Within the Automatic Domain Ontology Learning (ADOL) framework, (Chen and Gu, 2021) introduces an ontology that aims to represent the systematic knowledge architecture of different subjects for educational purposes. The ontology, designed by (Dascalu et al., 2022), reflects the career landscape in Romania and is intended for use in the context of career guidance and educational recommendations.

It is obvious that the teacher modeling is neglected from the research interests, and it is replaced with an e-learning system as the tutor. However, there is a need to represent the teacher context into an ontology to benefit from their expertise and to provide them with helpful resources.

2.4.4 Summary

Ontology and Linked Data principles emerge as pivotal tools to represent the knowledge of the multiple contexts of teachers. Ontologies provide a formal and structured representation of knowledge, which is crucial in capturing and organizing the diverse information related to teachers' demographics, teaching styles, and contextual dimensions within the educational domain. By adopting Linked Data principles, educational data can be interconnected using standardized formats and unique identifiers, facilitating the integration of diverse educational resources and information.

Table 2.7 provides a summary of surveyed knowledge representation models in the educational domain. These models encompass various aspects such as resources, learners, and learning processes. The representation models primarily fall into two categories: Linked Data (LD) and Ontology. LD models target resources and metadata, while ontologies focus on learners, learning processes, and resources. The context in which these models are applied varies, with some considering physical, social, and activity-related contexts, while others are context-agnostic.

The concluded observation highlights an important gap in current educational research. While there has been considerable focus on learner-centered models and e-learning systems, the role of teachers and the representation of their contexts have often been overlooked. An ontology that captures the teacher context is a valuable step in addressing this gap. Overall, ontologies offer several key advantages in the educational landscape. They provide a formal specification of deductive educational knowledge in terms concepts, properties, and relationships, ensuring precise representation of educational information.

2.5 Feature Selection

Feature selection plays a crucial role in extracting meaningful information from vast amounts of data in the data analysis and machine learning approaches (Cai

Table 2.6: Summary of surveyed knowledge representation models.

Articles	Model Type	Target	Focus	Context
(RISK, 2002)	LD	Resource	LOM	×
(Clemente, Ramírez, and De Antonio, 2011)	Ontology	Learner	Adult learner	×
(Bhattacharya, Tiwari, and Harding, 2012)	Ontology	Learner	Learning process	Physical
(Grubišić, Stankov, and Peraić, 2013)	Ontology	Learner	Adaptive-Courseware tutoring	×
(Mouromtsev et al., 2013)	Ontology	Resource	Learning resources	×
(Labib, Canós, and Penadés, 2017)	Ontology	Learner	Learning styles	×
(Board, 2018)	LD	Resource	Learning Resource Metadata	×
(Akharraz, El Mezouary, and Mahani, 2018)	Ontology	Learner	Learner context	Physical Social Activity
(Dermeval et al., 2019)	Ontology	Learner	Learning gamified resources	×
(Grivokostopoulou et al., 2019)	Ontology	Learner	Higher education learner	×
(Chang et al., 2020)	Ontology	Learner	Teaching rules (recommendations for teachers)	Physical
(Ilkou et al., 2021)	Ontology	Resource	Learning resource (recommendations for learner)	×
(Chen and Gu, 2021)	Ontology	Course	Knowledge architecture (educational process)	×
(Dascalu et al., 2022)	Ontology	learner	Career landscape	Physical Social

et al., 2018). By identifying and selecting relevant features, we can uncover patterns, make accurate predictions, and gain valuable insights. This process is particularly important in educational contexts, where understanding the factors that contribute to effective teaching and learning is of utmost significance (Hanushek and Rivkin, 2010).

Educational knowledge encompasses various aspects, including teacher and learner profiling, institutional aspects, and environmental factors (Dudek, 2000). Within this rich dataset, selecting the most informative features becomes essential to gain a deeper understanding of the educational landscape and improve educational outcomes. However, the challenge lies in identifying which features are truly relevant and impactful amidst the complexity and diversity of educational contexts.

By linking feature selection with the available knowledge in education, we can identify the key factors that contribute to effective representation of the educational context from the teacher's perspective (Ramaswami and Bhaskaran, 2009). Accordingly, the feature selection allows us to prioritize relevant features based on their theoretical foundations and empirical evidence (Liu et al., 2009). Moreover, it enables us to consider context-specific factors, such as teacher and student demographics, and environmental conditions, that influence the educational process.

2.5.1 Definitions

In the realm of feature selection, it's crucial to grasp fundamental definitions that underpin this process (Cai et al., 2018). A feature represents a measurable characteristic of an object or phenomenon and is a key component of datasets used in machine learning and data analysis (Cai et al., 2018). On the other hand, feature selection involves the careful curation of a subset of relevant features from the original set, aiming to retain those with strong associations with the target variable while discarding redundant or irrelevant ones (Cai et al., 2018). These relevant features are those with significant influence on the target variable, enhancing the accuracy and performance of machine learning models, whereas irrelevant features contribute little or no value and may introduce noise and complexity (Cai et al., 2018). Assessing feature importance quantifies the relevance of each feature in influencing a model's predictive performance or the variability of the target variable, providing valuable insights for feature selection and model optimization (Cai et al., 2018).

Grasping these definitions lays a solid groundwork for comprehending the principles and techniques of feature selection, highlighting its pivotal role in enhancing the effectiveness of machine learning and information systems.

2.5.2 Types of Features

In feature selection, it's essential to categorize features based on their nature and representation, as different types of features require distinct treatment (Cai et al., 2018). These commonly encountered feature types include numerical, categorical, and textual features. Numerical features, representing numeric values, can be either continuous or discrete and are typically quantitative, encompassing data like age, income, or test scores (Khalid, Khalil, and Nasreen, 2014). Categorical features, on the other hand, deal with discrete values that belong to specific categories or groups, and they can be further classified into nominal (no inherent order) and ordinal (with defined order) categories. Examples of these include gender (nominal) and education level (ordinal) (Cai et al., 2018).

Lastly, textual features involve unstructured text data, necessitating preprocessing techniques like tokenization and vectorization for machine learning suitability (Khalid, Khalil, and Nasreen, 2014). These textual features are particularly relevant in natural language processing, sentiment analysis, or document classification applications.

Applying these feature categorizations to the educational context, various elements within the multiple contexts of teachers can be considered potential features. Numerical features could encompass age and years of experience, while categorical features may include teaching styles, field of science, educational level, and institution type. Additionally, textual features could represent pedagogical resources through attributes such as titles and descriptions. These diverse feature types contribute to a comprehensive understanding of the factors influencing effective teaching and learning outcomes. Importantly, the selection of specific features should align with preserving the unique characteristics of each teacher's context, ensuring that feature selection methods cater to the intricacies of the educational environment (Ramaswami and Bhaskaran, 2009).

2.5.3 Feature Selection Approaches

Feature selection methods encompass a wide spectrum, from simpler filter-based techniques that utilize statistical measures to assess feature relevance to more intricate wrapper-based approaches that leverage machine learning algorithms to select features based on their predictive capabilities (Cai et al., 2018; Kuhn and Johnson, 2019). These methods can be broadly categorized into supervised, semi-supervised, and unsupervised techniques (Ang et al., 2015). Supervised approaches utilize target variables to select pertinent features from labeled datasets, with the crucial distinction lying in the proportion of labeled data employed in the algorithm. Conversely, unsupervised feature selection methods do not rely on labeled datasets and can be categorized into filter, wrapper, and hybrid techniques (Dong and Liu, 2018). However, it's important to note that the classification of feature selection algorithms into supervised, semi-supervised, and unsupervised methods is not universally recognized. Instead, these methods are typically categorized based on their underlying approach or methodology (Kuhn and Johnson, 2019).

Common types of feature selection algorithms include filtering approaches, which evaluate feature relevance based on statistical measures or heuristics; wrapper approaches, which select features by assessing the performance of machine learning models using different feature subsets; embedded/intrinsic approaches, which incorporate feature selection within the model-building process itself, often through regularization techniques; and dimensionality reduction approaches, which aim to reduce the dimensionality of the feature space while preserving relevant information. While these methods are not inherently categorized as supervised or unsupervised, they can be adapted to work in various contexts, depending on the availability of labeled data and the specific goals of the feature selection process (Kuhn and Johnson, 2019).

In the supervised setting of feature selection, three primary approaches are classified as filter, wrapper, and embedded/intrinsic methods. Filter approaches

assess feature relevance by examining the relationship between individual features and the target variable, utilizing statistical measures like correlation or mutual information to gauge feature significance (Ang et al., 2015; Kuhn, Johnson, et al., 2013). Wrapper approaches, on the other hand, are inherently supervised, involving the evaluation of machine learning algorithms' performance on feature subsets using labeled data as a criterion for feature selection, exemplified by Recursive Feature Elimination (RFE) (Ang et al., 2015; Kuhn, Johnson, et al., 2013). Embedded or intrinsic methods, which can also operate in a supervised context, incorporate feature selection within the model training process, employing regularization techniques like L1 regularization (Lasso) or L2 regularization (Ridge) to promote automatic feature selection during training, often requiring labeled data for techniques such as decision trees (Ang et al., 2015; Kuhn, Johnson, et al., 2013).

In the unsupervised setting of feature selection, three main types of approaches are identified: filter, wrapper, and hybrid methods. Filter approaches can be adapted for unsupervised settings, where feature relevance is assessed based on their relationships with each other rather than with a specific target variable, utilizing measures like variance or mutual information between features (Ang et al., 2015; Kuhn, Johnson, et al., 2013). Wrapper approaches can also be adapted for unsupervised use, where the elimination phase ignores the outcome but leverages unlabeled data for additional insights (Ang et al., 2015; Kuhn, Johnson, et al., 2013). Hybrid approaches involve a combination of multiple feature selection methods, particularly useful when dealing with datasets containing various types of features (Ang et al., 2015). In the context of representing the multiple contexts of teachers using ontologies, unsupervised feature selection, especially filter-based approaches, presents advantages. They do not require labeled data, which is often challenging to obtain in educational contexts due to privacy and annotation concerns. Unsupervised methods enable the exploration of data's inherent structure, particularly useful for capturing hidden relationships between features and contexts. These approaches can effectively handle the heterogeneity of features present in the multiple teacher contexts, including numerical, categorical, and textual, making them suitable for this thesis. Filter-based methods are computationally efficient and scalable, aligning well with large-scale educational systems. Additionally, hybrid strategies can be employed to maintain a consistent level of accuracy across various feature types.

2.5.4 Unsupervised Filter Approaches

Different feature selection algorithms has been classified as unsupervised filter approaches. These algorithms handle different categorizes of features such as chi-squared test, analysis of variance (ANOVA), R-regression and F-regression, mutual information, and variance-based threshold (Li et al., 2017).

2.5.4.1 Chi-squared

The Chi-squared test is a statistical test used to examine if a category feature and a categorical target have a significant relationship. The Chi-squared statistic is computed by the test and is defined as the sum of the squared differences between the observed and predicted frequencies of the feature and target classes.

The p-value is calculated using the Chi-squared statistic, which reflects the likelihood that the observed correlation between the feature and the target is due to chance. As the p-value decreases, the more likely the characteristic is associated with the target as shown in Eq.2.1, where f_o and f_{ex} are the observed and expected frequencies of features.

$$Chi^2 = \sum((f_o - f_{ex})^2 / f_{ex}) \quad (2.1)$$

2.5.4.2 R-regression

R-regression is a feature selection approach that uses the Pearson correlation coefficient r to calculate the linear connection between two continuous variables as shown in Eq.2.2. The correlation coefficient between the feature and the target (teacher) is calculated using R-regression, and the absolute value of the coefficient is used to indicate the strength of the association. The closer the absolute value of coefficient is to one, the greater the linkage between the feature and the target.

$$r = \frac{\sum (X_i - X_{mean}) * (Y_i - Y_{mean})}{\sqrt{\sum (X_i - X_{mean})^2} * \sqrt{\sum (Y_i - Y_{mean})^2}} \quad (2.2)$$

2.5.4.3 F-regression

F-regression is a feature selection method based on the univariate linear regression, which is used to determine whether there is a significant linear relationship between a continuous feature and a continuous target. The F-statistic represents the rate of $variance_{explained}$ to $variance_{unexplained}$ in F-regression. The $F - statistic$ is used to determine the p-value, which indicates the probability that the observed relationship between the feature and target is due to chance as shown in Eq.2.3, where N_f and N_o are the total number of features and observations respectively.

$$F - statistic = (variance_{explained} / N_f) / (variance_{unexplained} / (N_o - N_f - 1)) \quad (2.3)$$

2.5.4.4 Analysis of Variance (ANOVA)

ANOVA is a statistic-based method that is used to evaluate the significant difference between group feature structures. In the context of feature selection, ANOVA may be used to assess whether there is a significant variation in the means of the target variable across different levels of a categorical feature. When the goal variable is continuous, ANOVA is commonly utilised. The p-value is calculated using the F-statistic, which reflects the likelihood that the observed difference in means is attributable to chance as shown in Eq.2.4, where $variance_{explained}$ and $variance_{unexplained}$ are the explained and unexplained variances respectively, and N_g and N_o are the total number of groups and observations respectively.

$$F - statistics = (variance_{explained} / N_g - 1) / (variance_{unexplained} / N_o - N_g) \quad (2.4)$$

2.5.4.5 Mutual Information

Mutual information is a statistical measure that quantifies the amount of information shared between two random variables (Duboue, 2020). In the context of feature selection, mutual information assesses the dependency between a feature and the target variable without considering any learning process. It's particularly useful for identifying features that have a strong relationship with the target variable.

The formula for mutual information (I) between a feature (X) and a target variable (Y) is given by:

$$I(X; Y) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \left(\frac{p(x, y)}{p(x) \cdot p(y)} \right) \quad (2.5)$$

Where: $p(x, y)$ is the joint probability distribution of the feature X and the target Y . $p(x)$ and $p(y)$ are the marginal probability distributions of the feature X and the target Y , respectively.

2.5.4.6 Variance-based Threshold

Variance-based thresholding is a method for reducing the dimensionality of a dataset by removing features with low variance (Duboue, 2020). The intuition is that features with low variance may not carry significant information for distinguishing between different samples, and therefore, they can be safely discarded without impacting the overall performance of a machine learning model.

The variance of a feature X_i is calculated as follows:

$$\text{var}(X_i) = \frac{1}{N} \sum_{j=1}^N (X_{ij} - \bar{X}_i)^2 \quad (2.6)$$

Where: N is the number of samples in the data. X_{ij} is the value of feature X_i for the j th sample. \bar{X}_i is the mean value of feature X_i across all samples.

The variance-based threshold (thr) can be calculated based on the mean importance of information gain ($mean_I$):

$$thr = mean_I \cdot (1 - mean_I) \quad (2.7)$$

Features with variance below the threshold $\text{var}(X_i) < thr$ are considered low-variance features and can be removed from the data.

2.5.5 Ontology-based Feature Selection

Ontologies are formal representations of knowledge in a specific domain, typically organized as a hierarchy of concepts and their relationships (Brewster and O'Hara, 2004). They are commonly used in artificial intelligence, natural language processing, and other areas where understanding and reasoning about complex domains is required. Ontologies provide a powerful means of leveraging domain knowledge to improve machine learning performance (Fung and Bodenreider, 2019). By encoding the relationships between domain concepts, ontologies can guide the selection of relevant features from high-dimensional input data. This approach is particularly useful in situations where traditional feature

selection methods become less effective, such as in domains with a large number of features in the model (T-BOX) and limited instances (A-BOX) in a knowledge base (T-BOX + A-BOX). By exploiting the structure and semantics of the domain ontology, ontology-based feature selection can effectively filter out irrelevant features and retain only those that are most informative for the learning task at hand. This can result in higher classification accuracy, less overfitting, and shorter training periods.

Ontology-based feature selection is a technique that exploits the structure and semantics of the domain ontology to guide the selection of relevant features from high-dimensional input data (Sikelis, Tsekouras, and Kotis, 2021). By mapping input data features to ontology concepts, it becomes possible to identify and select the most informative features based on their semantic relationships within the domain. This approach not only improves the accuracy and efficiency of feature selection but also provides a more interpretable and explainable model, as the selected features can be directly traced back to their corresponding ontology concepts. Moreover, ontology-based feature selection can be used to discover previously unknown relationships between features and concepts, providing new insights into the underlying structure of the data. Overall, ontology-based feature selection offers a promising avenue for enhancing the performance of machine learning algorithms on complex, real-world datasets, by utilizing prior knowledge in the form of domain-specific ontologies. However, the success of this approach depends on the ability to effectively map input data features to ontology concepts, as well as the availability and quality of the domain ontology. A domain ontology is a formal representation of the concepts, entities, and relationships within a specific domain of knowledge. It defines a common vocabulary and structure for the domain, enabling more effective communication and knowledge sharing among stakeholders.

2.5.6 Feature Selection in Education

Variance-based feature selection is a filtering-based unsupervised technique that excludes features with variance lower than a given threshold among all data. Furthermore, it counts the large variance as an indicator of the presence of relevant information. A search query combining the terms "education," "variance-based," "feature," and "selection" is used to find relevant studies on variance-based feature selection. Acquiring appropriate findings in the educational area is a challenging task. Therefore, the majority of the discussed research is related to medical, textual, imaging, networking, or manufacturing applications. As a result, regardless of domain of application, the most relevant results to our case are assessed as shown in Fig.2.4.

The approach by (Lakshmi Padmaja and Vishnuvardhan, 2019) uses random forest (RF) technique for feature reduction and the random subset feature selection (RSFS) technique for feature extraction for cancer-related datasets. This technique takes into account feature variation without defining a threshold based on it, which raises the computational complexity and its cost. Another approach by (Roberts, Catchpoole, and Kennedy, 2018) offers a novel feature selection strategy for colon and lung cancer classification. The differential variation of

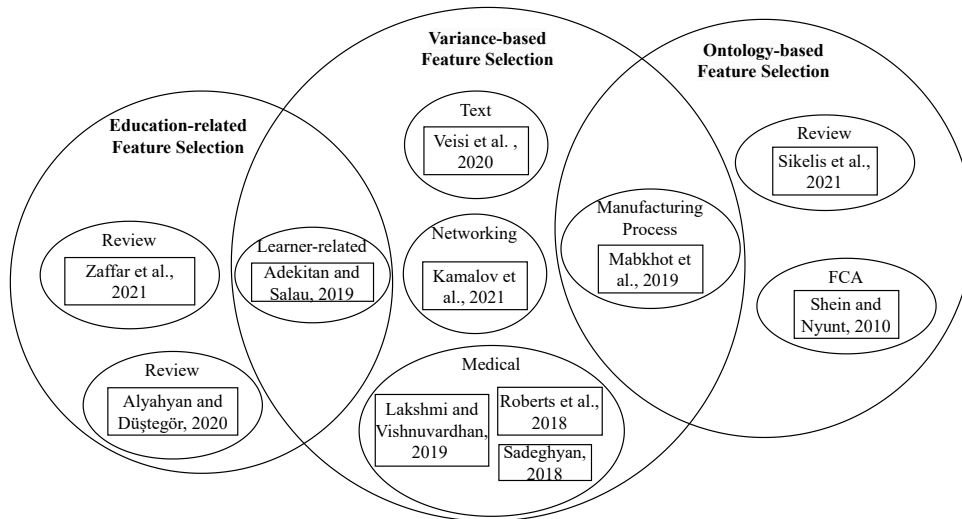


Figure 2.4: An overview of connections between the selected related articles to our work.

classification subgroups is used to choose features that is calculated from the ratio between cancerous and non-cancerous samples. As a result, they chose not to employ the across-features variance as the selection criterion.

The approach, introduced by (Sadeghyan, 2018), integrates sensitivity analysis (SA) with extended Fourier amplitude sensitivity test (EFAST) for variance-based feature selection in medical and biological datasets. The variance of input features is evaluated using the contribution of each feature, where a fixed threshold is chosen once for each dataset without taking into account any other changeable parameters. The approach, by (Kamalov et al., 2021) and enhanced in (Kamalov, 2021), employs an orthogonal variance-based decomposition for feature selection to find network traffic characteristics for intrusion detection systems (IDS). To differentiate DDoS assaults, they employ uniform distribution with variation of traffic properties. This method merely analyses the variance of all characteristics without providing a detailed explanation of how the threshold is chosen.

The approach by (Veisi, Aflaki, and Parsafard, 2020) presents a keyword extraction technique from Farsi and English text materials. Keywords are chosen based on token weighting and the accompanying variance. This study presents a hybrid variance-based feature selection strategy using Term Frequency Inverse Document Frequency (TF-IDF). The variation of terms serves as a selection criterion, but no intelligent identification of a more specific threshold is available. The approach proposed by (Mabkhot, Al-Samhan, and Hidri, 2019), offers a variance-based manufacturing process feature selection for a decision support system. The presented technique uses ontology to model the manufacturing process and adds a set of reasoning principles to decrease the computational cost of algorithm execution for each new insertion. The model only takes into account 9 features, which is considered relatively low number of features and is a disadvantage of the generalization of this approach in other domains.

Throughout our search, three essential reviews are chosen, one of which covers ontology-based feature selection in various domains and the other two of which represent major contributions to feature approaches in educational data processing and mining.

The first review by (Sikelis, Tsekouras, and Kotis, 2021), presents an overview of notable ontology-based feature selection approaches in selected application domains. The researchers conclude that ontologies may successfully discover dominating features in varied knowledge domains and that they can be incorporated into current feature selection and classification algorithms. The majority of the relevant papers addressed in this review article are Web-related and concern ontology-based feature selection for text document categorization. The paper discusses the critical problem of scalability which assesses the influence of increasing the quantity of the training data on an algorithm's computational performance in terms of accuracy and memory. The fundamentals of feature selection were created prior to the advent of Big Data. As a result, most feature selection techniques are inefficient when scaling high-dimensional data since their efficiency seems to decrease rapidly.

One notable reviewed research is the approach introduced by (Shein and Nyunt, 2010). This review uses formal concept analysis to create an ontology in OWL (FCA). The algorithmic framework aims to create semantic structures that are formal abstractions of language concepts, as well as to find conceptual patterns among data. The ultimate result is an ontological framework capable of effectively analyzing complicated text structures and revealing data relationships. Such approach lacks the utilization of prior feature selection before applying FCA which may result in high computational overhead in FCA and any following classification step.

The second review by (Alyahyan and Düştegör, 2020) addresses contemporary ways to predicting academic achievement at the higher education level. The study highlights new approaches for achieving such a goal at several phases of prediction, such as collection of data, data preprocessing, feature selection, and data mining for various student datasets. This study shows that, regardless of the relevance of using either feature selection techniques (filtering or wrapper approaches), it is preferable to employ embedding methods with data mining tools to decrease computational cost which results in losing relatively important information. One of the reviewed articles by (Adekitan and Salau, 2019) introduces a variance-based feature selection approach to predict cumulative GPA for learner using learner-related features only without any mention of teacher-related ones. The last review presented by (Zaffar et al., 2021) gives a state of the art of developments in educational data mining (EDM), however it focuses on classification and feature selection techniques. This study provides a guidance for researchers who want to develop an EDM with all of its components. The feature selection section reveals that wrapper techniques are the most popular, accounting for 78.5% of all examined approaches. The wrapper approaches, however, are not the fastest when dealing with massive and real-time datasets, according to the researchers. All selected techniques target data of learners and offer valuable predictions to assist both learners and teachers.

According to the reviewed research, the importance of using feature selection can be summarized into three scenarios: decreasing computational time by

selecting the most important features, combining feature selection with feature extraction that significantly improves its performance, and finally, the effect of feature selection on classification accuracy. Several approaches that use two feature selection strategies at the same time follow the two-stage feature selection. This method improves the efficiency and quality of the chosen features. None of these studies explain the significance of precise threshold selection or provide a detailed explanation of threshold selection and the amount of flexibility with other applications.

2.5.7 Summary

Feature selection is a critical step in machine learning and information systems, especially in education, where it aids in understanding the factors influencing effective teacher representation. Different types of features, including numerical, categorical, and textual, require specific selection approaches. Numerical features can be assessed using variance-based thresholds, while categorical features benefit from rank-based algorithms like mutual information. Textual features often demand techniques like TF-IDF or word embeddings. Various methods for feature selection exist, such as filter, wrapper, embedded/intrinsic, or dimensionality reduction approaches. Ontology-based feature selection stands out, using domain-specific ontologies to improve accuracy, interpretability, and insights.

In the context of education, feature selection has been applied in domains like learner prediction and educational data mining, often using wrapper approaches in combination with data mining tools. Choosing the right feature selection method depends on specific goals, available data, and computational constraints. Adopting an ontology-based hybrid feature selection approach, combining mutual importance calculation and variance-based thresholding, for the complex scenarios of multiple teachers is grounded in several compelling reasons. Educational contexts are multifaceted, encompassing teacher and student profiles, institutional dynamics, and environmental influences. The hybrid approach ensures comprehensive coverage of these facets, capturing the essence of multiple teacher contexts. Ontology-based methodologies add a layer of semantic understanding, enhancing accuracy and relevance, and leveraging ontology hierarchies for insights into feature importance.

Table 2.7 illustrates the various feature selection approaches discussed in this section. This table categorizes these approaches based on their classification, such as supervised/unsupervised and filter/wrapper/intrinsic/hybrid methods.

Existing research emphasizes the importance of feature selection in scenarios such as reducing computational time, improving classification accuracy, and enhancing performance through the combination of feature selection with feature extraction. Two-stage feature selection approaches that use multiple strategies concurrently are particularly notable for their efficiency and quality. However, precise threshold selection and its impact on various applications remain under-explored in these studies, highlighting a potential area for further research.

Table 2.7: Summary of surveyed feature selection approaches.

Articles	FS Approach	FS Type	Application
(Lakshmi Padmaja and Vishnuvardhan, 2019)	Variance-based	Unsupervised, Filter	Medical (Cancer-related data)
(Roberts, Catchpoole, and Kennedy, 2018)	Variance-based	Unsupervised, Filter	Medical (Colon and lung cancer data)
(Sadeghyan, 2018)	Variance-based	Unsupervised, Filter	Medical (Biological data)
(Kamalov et al., 2021)	Variance-based	Unsupervised, Filter	Networking (Intrusion detection data)
(Veisi, Aflaki, and Parsafard, 2020)	Variance-based + TF-IDF	Unsupervised, Hybrid	Text (Farsi and English text data)
(Mabkhot, Al-Samhan, and Hidri, 2019)	Ontology-based	Unsupervised, Hybrid	Manufacturing process
(Adekitan and Salau, 2019)	Variance-based	Supervised, Filter	Educational (Predicting cumulative GPA)
Review (Alyahyan and Düştegör, 2020)	Various approaches	Unsupervised, Filter or Wrapper	Educational (Learner data)
Review (Zaffar et al., 2021)	Various approaches	Various types	Educational (Learner data)
(Sikelis, Tsekouras, and Kotis, 2021)	Multiple approaches + Ontology-based	Supervised	Various domains

2.6 Recommender Systems

Recommender systems are sophisticated information filtering systems that employ computational algorithms to predict and suggest items or content to users based on their preferences, interests, and historical behavior (Gupta and Pandey, 2019). These systems have gained significant attention and importance due to the explosive growth of digital content and the need to assist users in navigating vast amounts of available information (Ricci, Rokach, and Shapira, 2015).

Recommender systems rely on various techniques and algorithms to generate personalized recommendations (Ricci, Rokach, and Shapira, 2015). They aim to address the information overload problem by filtering and presenting relevant items to users, thereby enhancing user satisfaction, engagement, and decision-making processes (Ricci, Rokach, and Shapira, 2015). These systems are widely

employed in diverse domains such as e-commerce, entertainment, social media, and education.

2.6.1 Types of Two-dimensional Recommender Systems

Two-dimensional (2D) recommender systems encompass several distinct types based on their underlying algorithms and methodologies, such as collaborative filtering, content-based filtering, knowledge-based, and hybrid approaches (Aggarwal et al., 2016). Collaborative filtering approaches leverage user preferences and behaviors from a user community to make recommendations (Schafer et al., 2007). These methods assume that users with similar past preferences or behaviors will have similar future preferences. Collaborative filtering includes memory-based and model-based approaches (Schafer et al., 2007). Memory-based methods use similarity metrics to identify users or items with similar preferences, while model-based approaches employ statistical and machine learning models for pattern recognition.

In contrast, content-based filtering approaches recommend items by analyzing the intrinsic characteristics and features of items (Lops, De Gemmis, and Semeraro, 2011). User profiles are constructed based on preferences for specific item features, and items with similar attributes are recommended. Content-based filtering often involves techniques like natural language processing, information retrieval, and machine learning to extract relevant features from items. Knowledge-based approaches use explicit knowledge representation and reasoning methods, incorporating domain-specific knowledge and user preferences into a knowledge base or ontology (Burke, 2000). Hybrid approaches combine multiple recommendation techniques, such as collaborative filtering and content-based filtering, to provide diverse and personalized recommendations (Burke, 2002).

Hybridization techniques have emerged as powerful solutions to enhance the accuracy and effectiveness of two-dimensional (2D) recommender systems, which aim to provide personalized recommendations by considering multiple dimensions of data (Burke, 2002). In the context of this thesis, which focuses on pedagogical resource recommendations for teachers, hybridization plays a pivotal role in integrating different recommendation approaches to optimize the recommendation process (Tarus, Niu, and Mustafa, 2018). Several hybridization strategies are employed, each with distinct characteristics and advantages.

Weighted hybridization combines recommendation scores from various approaches by assigning specific weights to each approach (Burke, 2002). In teacher resource recommendations, it often fuses content-based and collaborative filtering outputs, with weights determined by factors like performance or domain expertise. Switching hybridization dynamically selects the best-performing approach based on predefined rules or algorithms, adapting to the available data (Burke, 2002). Mixed hybridization merges outputs of different approaches, yielding a unified recommendation. Feature combination hybridization combines features extracted from contributing recommender systems, creating an augmented feature space. Feature augmentation hybridization expands the feature space by incorporating additional dimensions from various data sources. Cascade hybridization involves sequential application of different methods, with one method's output serving as input for another. Meta-level hybridization employs

higher-level algorithms to combine multiple recommendation methods based on data characteristics and historical performance.

Lastly, context-aware approaches take into account contextual factors, including physical, social, interactivity, and sentimental contexts, to deliver more personalized recommendations (Adomavicius and Tuzhilin, 2010). These systems aim to adapt their recommendations based on the specific circumstances in which users interact with the system.

2.6.2 Context-Aware Recommender Systems

Context-aware recommender systems (CARS) have gained significant attention for their ability to deliver highly personalized recommendations by integrating contextual information (Adomavicius and Tuzhilin, 2010). Context in these systems refers to relevant information related to the user, items, or the recommendation process itself (Adomavicius and Tuzhilin, 2010). This context can encompass factors such as time, location, device, social connections, and environmental conditions, among others (Kulkarni and Rodd, 2020). By considering these contextual factors, these systems can adapt their recommendations to specific situations, enhancing their utility and user satisfaction (Kulkarni and Rodd, 2020).

CARSs employ various machine learning and data mining techniques to model and interpret contextual information, including collaborative filtering with context, content-based filtering with context, and hybrid approaches that incorporate contextual features into recommendation algorithms (Villegas et al., 2018). These systems dynamically adjust their recommendations as the context evolves, ensuring that the suggestions remain pertinent and valuable to the user's current situation.

CARSs utilize different contextual filtering approaches to enhance the level of personalization (Adomavicius and Tuzhilin, 2010). These approaches include contextual prefiltering, contextual postfiltering, and contextual modeling. Contextual prefiltering involves filtering resources based on the user's context before applying the main recommendation algorithm, streamlining the recommendation process. Contextual postfiltering refines recommendations based on the user's context, ensuring they align with specific needs. Contextual modeling integrates contextual features directly into the recommendation algorithm, creating a joint model that considers both user preferences and contextual information. While contextual modeling provides highly accurate recommendations, it comes with higher implementation complexity and computational demands. The choice of approach depends on factors like data sparsity, filtering effectiveness, computational overhead, and the relevance of recommendations to each teacher's context (Adomavicius and Tuzhilin, 2010).

2.6.3 Recommender Systems in Education

Recommender systems have also found applications in the field of education, where they are commonly referred to as educational recommender systems (Adomavicius and Tuzhilin, 2005b). These systems aim to assist learners in finding relevant learning materials, courses, or educational resources based on their individual needs and preferences (Adomavicius and Tuzhilin, 2005b). Educational

recommender systems play a crucial role in addressing the challenges of information overload and personalization in education (Tarus, Niu, and Mustafa, 2018). By analyzing learner data, including past performance, interests, learning style, and goals, these systems can generate personalized recommendations to optimize the learning experience. Educational recommender systems are utilized in various educational settings, including schools, universities, and online learning platforms (Tarus, Niu, and Mustafa, 2018).

Educational recommender systems can recommend suitable courses and learning paths based on a learner's interests, prior knowledge, and career aspirations. They can suggest learning resources, such as textbooks, articles, videos, or interactive materials, that align with the learner's current topics of study (Drachler, Hummel, and Koper, 2008). Collaborative learning recommender systems can facilitate group formation by recommending potential study partners or project collaborators based on shared interests and complementary skills. Adaptive learning recommender systems can personalize the learning experience by recommending content and activities that cater to the learner's individual strengths, weaknesses, and learning style. By leveraging recommender systems in education, learners can discover and engage with educational materials that are more aligned with their preferences, goals, and learning needs. These systems can enhance the learning experience, increase learner motivation and engagement, and ultimately contribute to better learning outcomes.

The review papers that were published during the past ten years (Dutt, Ismail, and Herawan, 2017; Vaidhehi and Suchithra, 2018; George and Lal, 2019), discussed the educational recommendation systems trends from 2010 until 2018 with only 3 papers, 1 paper, and 1 paper respectively that take the teacher as the targeted user and none of these RSs took into consideration all the mentioned factors. We surveyed educational resources recommender systems that provide the teacher with either educational resources or training courses during the last ten years. After an excessive search, only 11 out of 34 ERSs met the search criteria as shown in Table 2.8, while the omitted ones either do not provide the precedent two types of recommendations or target learners as well as teachers. In order to conclude the essential criteria for constructing an efficient RS, four critical points should be discussed; user personalized experience, RS model, data sources, and items measurability as shown in Figure 2.5 (Manouselis et al., 2012). Every teacher gets resources recommendations according to the teacher profile and activities creating a realistic experience for the teacher (Sergis and Sampson, 2015b). From the surveyed papers, 4 ERSs only provide a personalized experience for teachers (Sergis, Zervas, and Sampson, 2014; Zhao et al., 2014; Rodriguez et al., 2015; Sabourin, Kosturko, and McQuiggan, 2015). As our problem has multiple aspects, it is not sufficient to use one RS strategy to address all these aspects. A hybridization of two strategies is a necessity to achieve acceptable results (Burke, 2002) which is followed by 3 ERSs from our list (Sergis and Sampson, 2015b; Bozo, Alarcón, and Iribarra, 2010; Pursel et al., 2016) but none of them considered the geographical dimension of the problem. All 11 ERSs use only one source of educational resources datasets which limits the availability and variety of the resources. The variety of data sources creates a reliable experience for users (Manouselis, Vuorikari, and Van Assche, 2010). The last point to be discussed is the measurability of educational resources which is important to

deliver the right resources to the targeted teacher. This important criterion had been taken into consideration by only one ERSs (Rodriguez et al., 2015).

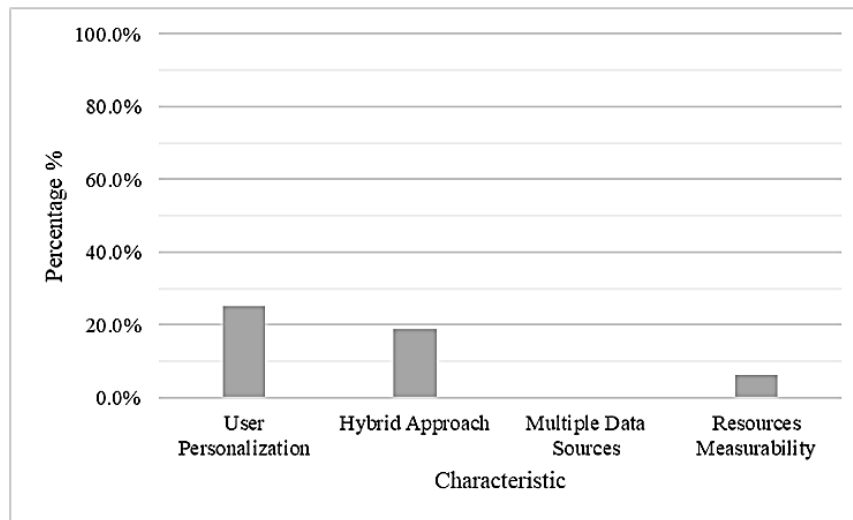


Figure 2.5: Percentage of recommender systems fulfilling the required characteristics.

The article by (De Medio et al., 2020) introduces MoodleRec, a hybrid recommender system integrated as a plug-in within the Moodle Learning Management System. MoodleRec streamlines the search process by searching supported standard-compliant Learning Object Repositories and providing a ranked list of Learning Objects based on keyword-based queries. Recommendations are made using two levels of strategies: the initial ranking is based on relevance to the query and the quality as indicated by the repository, followed by the use of social-generated features to showcase how these Learning Objects have been utilized in other courses. One of the latest approaches published by (Ali et al., 2022) emphasizes the importance of providing quality resources during the training phase and recognizes the lack of online assistance from service providers as a significant issue. The proposed architecture introduces semantic recommendations generated by virtual agents, taking into account user requirements and preferences, to assist users, including teachers, learners, and administrative staff, in finding appropriate courses.

Two latest reviews by (Urdaneta-Ponte, Mendez-Zorrilla, and Oleagordia-Ruiz, 2021) and (Ko et al., 2022) are considered among the important research efforts regarding recommender systems and more specifically educational recommender systems. The review by (Urdaneta-Ponte, Mendez-Zorrilla, and Oleagordia-Ruiz, 2021) discusses the research trends in education services, particularly in the context of e-learning. The review has witnessed significant growth and interest since 2010. E-learning has evolved to offer personalized learning experiences with the proliferation of smart devices and the emergence of conferencing apps facilitating real-time communication with teachers. This evolution has led to a substantial increase in studies related to learning recommendation systems. Notably, the COVID-19 pandemic in 2019 further accelerated the adoption of e-learning, as schools turned to conferencing apps and Learning Management

⁷<http://www.eun.org/>

Table 2.8: Summary of reviewed ERRSs

Article	Year	Item	Data Model	Approach
(Manouselis, Vuorikari, and Van Assche, 2010)	2010	Evaluation of current resources	European Schoolnet portal data model ⁷	Collaborative-based
(Bozo, Alarcón, and Iribarra, 2010)	2010	Learning Objects (LOs)	IEEE learning object metadata (IEEE LOM) (RISK, 2002)	Hybrid
(Zhao et al., 2014)	2014	Teaching Resources	Knowledge tree	Knowledge-based
(Sergis, Zervas, and Sampson, 2014)	2014	LOs	IEEE LOM	Competence-based (hybrid)
(Rodríguez et al., 2015)	2015	Educational Resources	Ontology	Knowledge-based
(Sabourin, Kosturko, and McQuiggan, 2015)	2015	Educational Resources (SAS)	Descriptive Models (SAS Enterprise Miner)	Hybrid
(Sergis and Sampson, 2015b)	2015	Learning Objects Repositories (LORs)	---	Competence-based (hybrid)
(Pursel et al., 2016)	2016	Open Educational Resources (OER)	---	Design-based (hybrid)
(Villalba et al., 2017)	2017	LOs	IEEE LOM	Collaborative filtering
(Ali et al., 2017)	2017	Educational Resources	MEMORAe SoIS (Abel, 2022)	Hybrid
(Iniesto and Rodrigo, 2019)	2019	Modifications to MOOC courses	Massive Open Online Courses (MOOCs) (Porter, 2015)	Collaborative filtering
(De Medio et al., 2020)	2020	Learning Object Repositories (LORs)	---	Hybrid
(Ali et al., 2022)	2022	Courses for teachers and learners	---	Hybrid

Systems (LMS) to provide remote learning. This reflects the growing fascination with alternative educational approaches and the potential of e-learning to complement offline education. The review observed that 50% of educational RSs use a web platform in which 80% of all approaches used a 2D recommender system without any hybridization. Another notable observation from this review is the

limited number of RSs that target teachers, only 2%. The RS introduced by (Sergis and Sampson, 2015a) is one of the teacher-centered RSs that offers teaching practices recommendations for teachers. The approach used memory-based collaborative filtering approach that uses adaptive neighbor selection to recommend learning objects (LOs). The other teacher-centered RS is the approach published by (Dang, 2018) which recommends learning resources to non-formal learners including teachers but without any dedicated treatment of data for users who act as teachers. Overall, this review highlights the research gaps in the educational recommender systems such as the limited usage of hybrid RSs that can elevate the performance of these RSs, and also, the consideration of the social information of users which can lead to a more complete profiling of users.

The review by (Ko et al., 2022) discusses the research trends in education services, particularly in the context of e-learning. The review has witnessed significant growth and interest since 2010. E-learning has evolved to offer personalized learning experiences with the proliferation of smart devices and the emergence of conferencing apps facilitating real-time communication with teachers. This evolution has led to a substantial increase in studies related to learning recommender systems. Notably, the COVID-19 pandemic in 2019 further accelerated the adoption of e-learning, as schools turned to conferencing apps and Learning Management Systems (LMS) to provide remote learning. This reflects the growing fascination with alternative educational approaches and the potential of e-learning to complement offline education. According to this review, researchers favor e-learning recommendations for learners rather than teachers. Additionally, it is noticeable that 60% of the surveyed approaches use 2D RSs without any hybridization. The review observed the increase of market value of e-learning services with 12.1% of market interest during the surveyed period, however, this increase cannot cope with the amount of demand. This observation leads to the need to include further research attention towards the educational recommender systems that target the teacher or the instructor.

2.6.4 Peer Recommendations in Education

The initial understanding of the term "peer recommendations", is directed towards the peer-to-peer recommendation which is not the purpose of this research. In this paper's context, peer recommendation is defined as introducing a peer (person) as the recommended item to a user (person). This definition is not adapted by any of the reviewed approaches, but many approaches benefit from the power of the peer-based recommendations. The term "peer learning" is defined as the process of collaboration between two or more equals or matched companions to acquire knowledge and skills in a certain context in which they provide the required support and help to each other Boud, Cohen, and Sampson, 1999; Topping, 2005. From the previous definition, it can be deduced that most of the research input in this aspect, is directed towards the different types of learners. Therefore, this subsection discusses the latest research in this aspect and connects them to our research.

The concept of Recommendation in Personalised Peer Learning Environments (RiPPLE) was introduced by Khosravi, 2017 to describe the information systems, including recommender systems, that provide the required personalized learning support for the targeted users. In another work related to the previous

concept, Khosravi, Cooper, and Kitto, 2017 introduced a new question sharing approach based on a collaborative filtering algorithm. In this approach, learners express their knowledge gaps by posting questions through a forum. Potts et al., 2018 introduced a course-level peer recommender system to provide peer-to-peer collaboration between learners to fill the knowledge gaps of these learners. The algorithm recommends questions for learners to explore, review or answer, based on their skills. Labarthe et al., 2016 conducted an experiment on a project management course that was held over a MOOC system. Their recommender provided peer recommendations within a group of learners of the same interest and based on the learners' overall performance. During later research related to the previous experiment Bouchet et al., 2017, learners were asked if they want to chat with other peers of the same interest. The research proved through this experiment that the learners tend to use this peer recommendation to express their emotion rather than the learning support.

The systematic review conducted by Deschênes, 2020, concludes that there is a need to invest more research effort towards peer recommender systems in the learning context. Moreover, we could not find any teacher peer recommender systems in the educational context. Therefore, this paper introduces a new research input towards a topic that lacks attention, and it introduces a needed help for teachers.

2.6.5 Sentimental Context Importance in Recommender Systems

The physical and mental health of teachers are strongly related to their working efficacy (McIntyre et al., 2017; Lachowska et al., 2018; Oberle et al., 2020). The resultant mood swings, during the current times, are responsible for teacher efficacy and performance (Frenzel, 2014). In addition, it is directly related to the teacher's burnout which directly affects the learners (Mérida-López and Extremera, 2017; Heutte et al., 2016). Educational/training resources can act as a mitigator to COVID-19 effects on teachers such as burnout and stress (Lizana et al., 2021). These educational resources are categorized into internal resources, such as classroom management and instructional resources, and external resources, such as supporting educational resources. This approach is said to be efficient if teachers are provided with personalized strategies and resources to acquire new skills and facilitate the classroom management (Lizana et al., 2021).

During the current COVID-19 era, digital technologies are essential support for teaching and learning processes in various fields and different educational levels (Perifanou, Economides, and Tzafilkou, 2021; Mukhtar et al., 2020). Educational resources recommender systems (ERRS) aid this purpose by providing personalized educational resources recommendations for each teacher. Traditionally, ERRS targets learners by recommending learning objects or materials in addition to performance evaluation (Zhang, Lu, and Zhang, 2021). The importance of the context-aware recommenders arises recently to consider the social and environmental conditions (Nilashi et al., 2020). (De Meo et al., 2017) introduced a social networking context-aware recommender system approach by combining trust relationships and skills evaluation to reform the online classrooms. On the other hand, (Klašnjaja-Milićević, Vesin, and Ivanović, 2018) combines social tagging and sequential patterns to provide context-aware recommendations in an e-learning environment. (Moore, Zhao, and Pham, 2019) monitors student's

progress in order to provide personalized performance evaluation without using face-to-face tutorial sessions and then recommends learning materials. As for the teacher centered ERRS, (Cobos et al., 2013) introduced a recommender approach to teachers by detecting the pedagogical patterns from the online recorded information based on the class context. The systematic reviews conducted by (Zhang, Lu, and Zhang, 2021) and (Imran et al., 2021) prove that research efforts are directed towards learners' context with nearly neglecting to teacher context which is one of the motives for this research.

2.6.6 Summary

Recommender systems have gained prominence in the field of education. These systems are designed to assist learners in discovering relevant learning materials, courses, or educational resources tailored to their individual needs and preferences, addressing the challenges of information overload and personalization in education. They achieve this by analyzing learner data, including past performance, interests, learning style, and goals, to generate personalized recommendations and optimize the learning experience. Educational recommender systems find applications in various educational settings, including schools, universities, and online learning platforms. However, it's worth noting that research in this domain has predominantly focused on recommendations for learners, with limited attention given to teacher-centered recommender systems (Adomavicius and Tuzhilin, 2005b; Tarus, Niu, and Mustafa, 2018).

Teacher-centered educational recommender systems are crucial in helping teachers access suitable courses, learning resources, and teaching practices based on their unique profiles and requirements. These systems can offer personalized teaching strategies and resources to enhance classroom management and instructional effectiveness, ultimately benefiting both teachers and students (Sergis and Sampson, 2015a; Sergis, Zervas, and Sampson, 2014; Zhao et al., 2014; Rodriguez et al., 2015; Sabourin, Kosturko, and McQuiggan, 2015; Pursel et al., 2016). However, there are gaps in the existing research, including the limited use of hybrid recommender system strategies that could enhance system performance and the need for more comprehensive user profiling, particularly in the teacher context (Manouselis et al., 2012).

Furthermore, the COVID-19 pandemic has underscored the importance of personalized educational resources for teachers to combat issues like burnout and stress. These resources can support teachers in acquiring new skills and effectively managing their classrooms (Lizana et al., 2021). However, the research in this area has mainly focused on students and learners, neglecting the specific needs of teachers (Perifanou, Economides, and Tzafilkou, 2021; Mukhtar et al., 2020).

Context-aware recommender systems have emerged as a crucial aspect of educational recommender systems, considering social and environmental conditions. These systems adapt recommendations based on trust relationships, skills evaluation, and classroom context (De Meo et al., 2017). Social tagging and sequential patterns are also employed to provide context-aware recommendations in e-learning environments (Klašnja-Milićević, Vesin, and Ivanović, 2018). Monitoring student progress and providing personalized performance evaluations

without face-to-face tutorials have become feasible with context-aware recommendations (Moore, Zhao, and Pham, 2019). However, the majority of context-aware recommender system research still focuses on learners rather than teachers, emphasizing the need for more teacher-centered approaches (Zhang, Lu, and Zhang, 2021; Imran et al., 2021).

In summary, while educational recommender systems have made significant strides in assisting learners, there is a clear gap in research and development for teacher-centered recommender systems. These systems have the potential to enhance teaching practices, classroom management, and overall teacher well-being. Context-aware recommendations, which consider social and physical contexts, also have the potential to improve the relevance and effectiveness of educational recommendations for both teachers and learners.

Part III
Contribution

Chapter 3

Contribution Overview

The context-aware pedagogical resources recommender system is a complex task to achieve, therefore, it is divided into several approaches to address our research questions. Through this chapter, an overview of the contribution of this thesis is discussed and a roadmap of the related upcoming chapters is paved to present its different approaches. Section 3.1 introduces an overall view of our contribution and the connectivity between its components, while Section 3.2 summarizes each component separately. Section 3.1 provides an overview of the chapter.

3.1 Approach Overview

In the ever-evolving landscape of education, the pursuit of personalized and contextually informed teaching methodologies has become a paramount endeavor. To address this, our contribution embarks on a comprehensive exploration of a novel Context-Aware Pedagogical Resource Recommender System (CAPRRS), as presented in Figure 3.1. By integrating advanced knowledge engineering, ontological modeling, semantic reasoning, precise feature selection methodology and cutting-edge recommendation techniques, this section delves into the intricate fabric of tailoring pedagogical resource recommendations to the unique contexts of teachers.

In response to pressing research questions (Chapter 1), our contribution elucidates the intricacies of addressing and integrating teacher contexts guided by the state of the art (Chapter 2). The first research question, "How to consider the multiple contexts of teachers? And how to integrate data from these different contexts?", is met with a comprehensive approach, entailing ontological modeling and semantic reasoning. The Teacher Sentimental and Collaboration Context Ontology (TSCCO) (Chapter 4) stands as a testament to this integration, providing a structured framework that harmoniously encapsulates diverse teacher contexts. This involves encoding various dimensions of teacher-related information, including living environments, working environment, and sentimental state, into a coherent ontological framework. Semantic reasoning mechanisms further enable the integration of disparate contextual data, ensuring a harmonious synthesis of information from diverse sources (Chapter 5). Through this approach, the intricate tapestry of teacher contexts is woven together, facilitating a holistic understanding that underpins the subsequent stages of the pedagogical resource recommendation process.

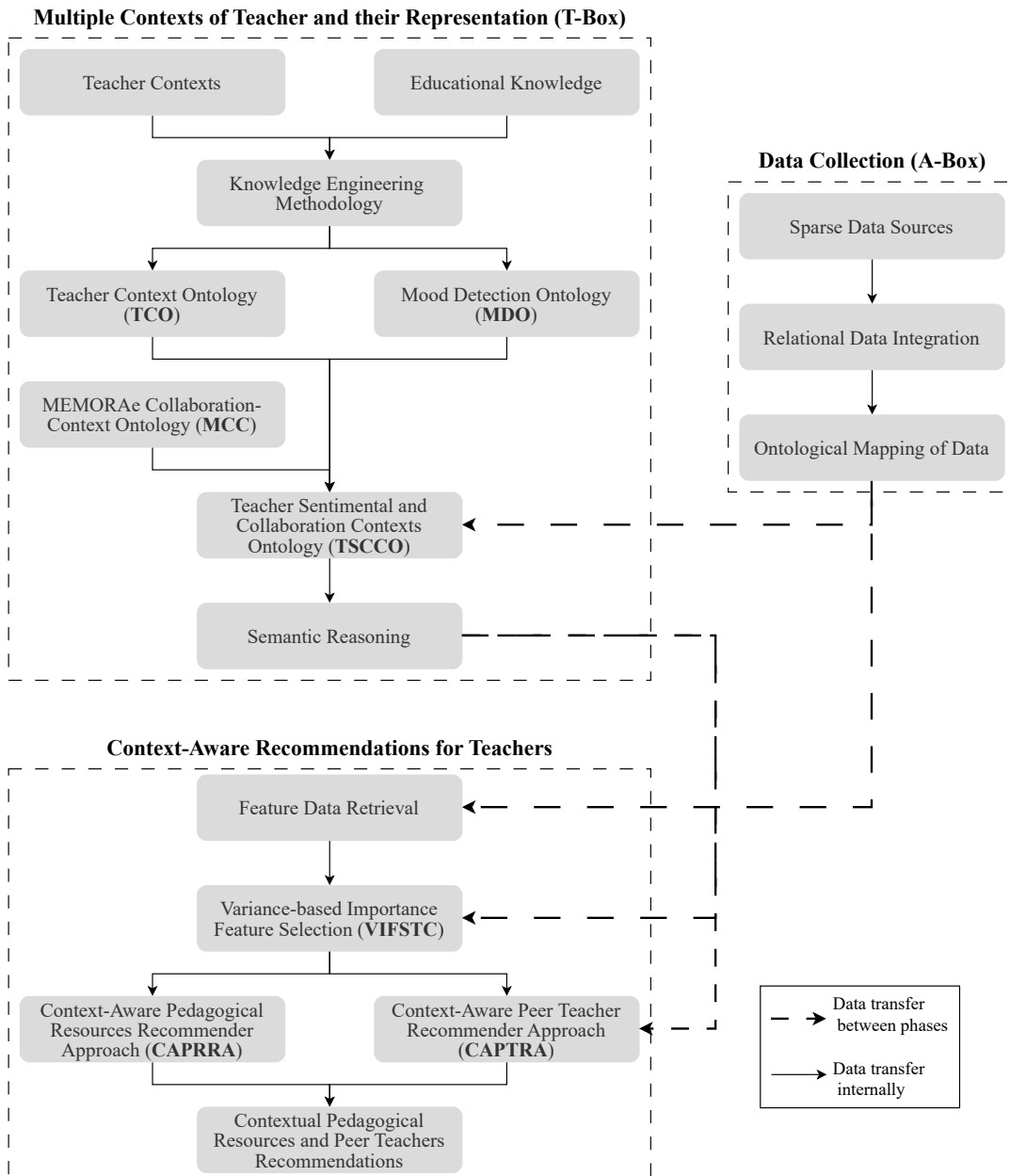


Figure 3.1: An overview of the contributed approach of this thesis.

In pursuit of the second research question, "How to take into account the multiple contexts of teachers to recommend suitable pedagogical resources using the most descriptive features?", our contribution unveils an intricate approach that draws upon the fusion of various recommender approaches (Chapter 6). The amalgamation of collaborative and content-based filtering techniques, underpinned by the rich tapestry of teacher contexts, forges the foundation of the Context-Aware Pedagogical Resource Recommender System (CAPRRS). Through the innovative Variance-based Importance Feature Selection for Teacher Contexts (VIFSTC) methodology, relevant features aligned with teacher contexts are systematically identified and weighed. Collaborative filtering leverages historical

interactions, while content-based filtering harnesses resource metadata, resulting in a comprehensive yet contextually nuanced recommendation approach.

3.2 Approach Components

Our contribution falls within three categories: knowledge representation, data collection, and recommender systems. Consequently, each component of our approach addresses one of these categories as illustrated by Figure 3.1. Each of these components is explained through the following three chapters to this one. In Chapter 4, Multiple Contexts of Teacher and their Representation is introduced by Teacher Sentimental and Collaboration Contexts Ontology (TSCCO). TSCCO is designed to comprehensively represent the diverse contexts impacting teachers. The integration of three constituent ontologies—Teacher Context Ontology (TCO), Mood-Detection Ontology (MDO), and MEMORAE Collaboration Context Ontology (MCC)—is explored, highlighting their synergy. This chapter acts as the terminology or the schema-level knowledge of the ontology (T-Box). Chapter 5 focuses on Data Collection, detailing the extraction and harmonization of data from various sources, including environment-related and institutional datasets, to provide a comprehensive foundation for informed recommendations and their ontological-based mapping to TSCCO. This chapter represents the assertion-level knowledge of TSCCO (A-Box) as it explores data relevant to each context of teacher and their mapping to TSCCO. In Chapter 6, the approach of Context-Aware Recommendations for Teachers is presented, addressing the intricacies of pedagogical resource recommendation through a hybridization approach, considering both explicit and implicit feedback. The chapter emphasizes the importance of integrating teacher contexts and examines different contextual filtering approaches, culminating in the exploration of feature selection techniques to optimize the context-aware recommender system. These chapters collectively offer an integrated approach to enhancing pedagogical resource recommendations within the context of diverse teacher environments.

3.2.1 Multiple Contexts of Teacher and their Representation (T-Box)

The knowledge representation of the multiple contexts of teacher follows the detailed knowledge engineering methodology introduced by (Noy, McGuinness, et al., 2001). Through this methodology, various rules are respected, and multiple steps are followed to construct the Teacher Sentimental and Collaboration Contexts Ontology (TSCCO). TSCCO introduces a novel approach to comprehensively represent and integrate the diverse contexts in which teachers operate. TSCCO acts as the T-Box terminology of the collected knowledge within the educational domain, as illustrated in Figure 3.2. The corresponding chapter (Chapter 4) delves into the construction and utilization of TSCCO, which is designed to

encapsulate various facets of a teacher’s environment, emotional state, and collaborative interactions. Employing the Web Ontology Language (OWL) and developed within the Protégé framework, TSCCO forms the backbone for context-aware pedagogical resource recommendation. TSCCO encompasses three distinct ontologies: the Teacher Context Ontology (TCO), the Mood-Detection Ontology (MDO), and the MEMORAE Collaboration Context Ontology (MCC). These ontologies synergistically capture the intricate interplay between teacher contexts, emotional states, and collaboration dynamics. Through the integration of these ontologies, TSCCO establishes a robust foundation for context-aware recommendations, paving the way for personalized and informed pedagogical resource suggestions.

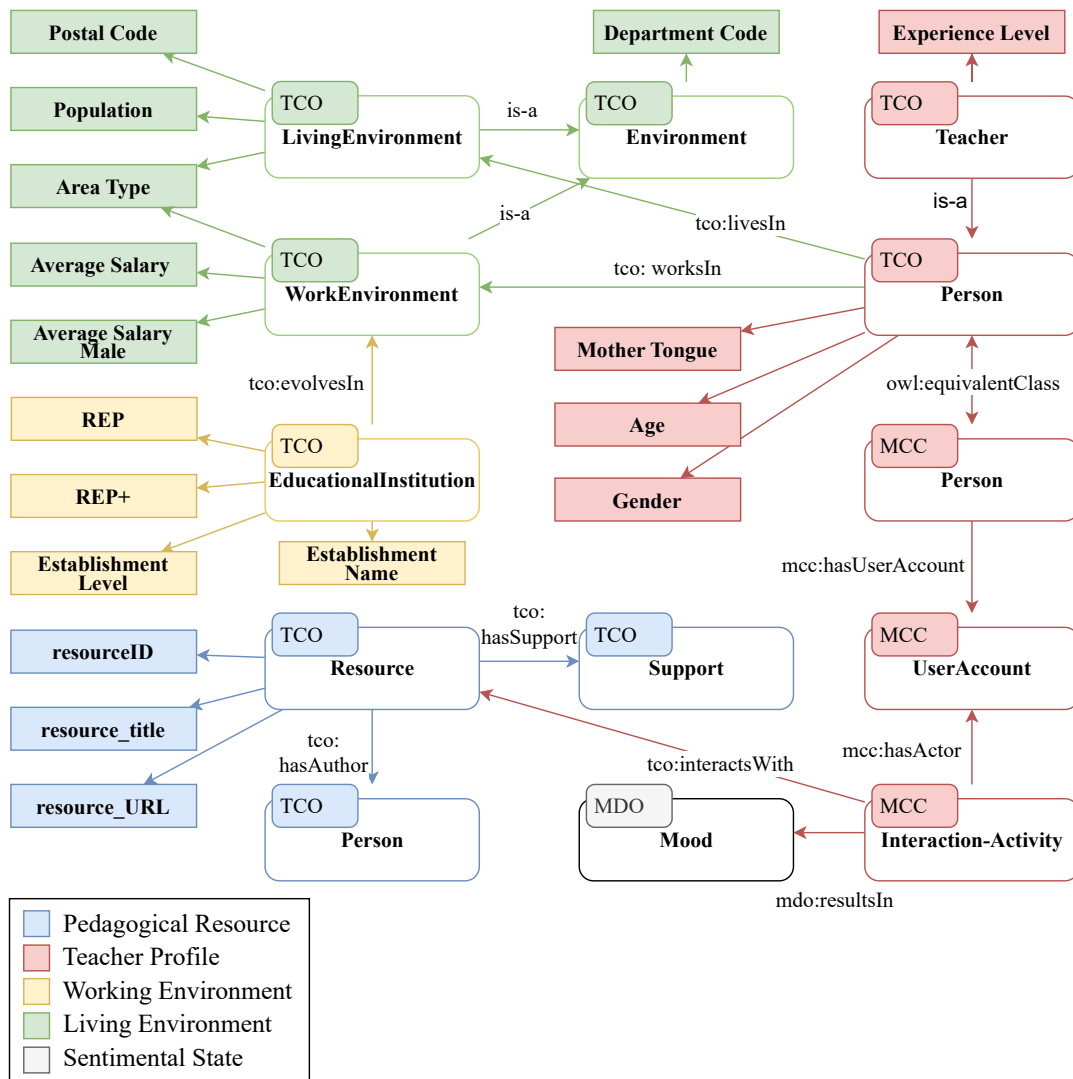


Figure 3.2: A partial T-Box snippet of TSCCO highlighting multiple contexts of teacher.

Additionally, Chapter 4 provides a detailed exploration of the Teacher Sentimental and Collaboration Contexts Ontology (TSCCO), focusing on the integration and synergy of its three constituent ontologies: TCO, MDO, and MCC. The

TCO elucidates the multifaceted contexts within which teachers operate, encompassing factors such as professional experience, teaching style, living conditions, and working environment. The MDO introduces a groundbreaking perspective by linking mood detection systems to sentimental states, enabling a deeper understanding of a teacher's emotional disposition. The MCC, functioning as a management ontology within the MEMORAe platform, facilitates collaborative knowledge sharing among educators. This section provides insights into the development and integration of these ontologies, unveiling the synergistic relationship that underpins TSCCO's holistic approach to contextual representation.

3.2.2 Data Collection (A-Box)

The coexistence of teachers within a myriad of contexts necessitates a comprehensive approach to data collection and integration. Chapter 5 delves into the intricacies of sourcing and harmonizing data from various contexts, constructing a foundation for informed recommendations. The data collection phase unfolds the mechanisms employed to extract data from different sources, including environment-related and institutional datasets. These datasets encompass living and working environments, teacher profiles, resource interactions, and sentiment analysis. The data collection approach elucidates the integration process, mapping relational databases into virtual RDF graphs, and establishing SPARQL endpoints for data access. This data collection phase forms the A-Box instances of TSCCO, as illustrated in Figure 3.3. By weaving together data from disparate contexts, this chapter lays the groundwork for context-aware pedagogical resource recommendations.

3.2.3 Context-Aware Recommendations for Teachers

The context-aware recommender approach of this thesis introduces a pivotal framework for context-aware recommendations tailored to the unique contexts of teachers, as shown in Figure 3.1. The approach commences by elucidating the intricate nature of pedagogical resource recommendation, encompassing both explicit and implicit feedback mechanisms. It leverages a hybridization two-dimensional approach that combines content-based and collaborative filtering techniques. This approach addresses the challenges of personalized resource suggestions. The consideration of teacher contexts becomes paramount, leading to the conceptualization of context-aware recommendations. Three distinct types of contextual filtering - prefiltering, postfiltering, and contextual modeling - are explored in the context of recommendation, offering insights into the integration of contextual features in the recommendation process. This exploration results in a combination of these different contextual filtering approach called context-aware pedagogical resources recommender approach (CAPRRA). The corresponding chapter culminates by delving into a variance-based importance feature selection (VIFSTC) methodology that optimizes data used by the context-aware recommender system. Through this holistic exploration, the chapter establishes a comprehensive framework for enhancing pedagogical resource recommendations through the lens of teacher contexts.

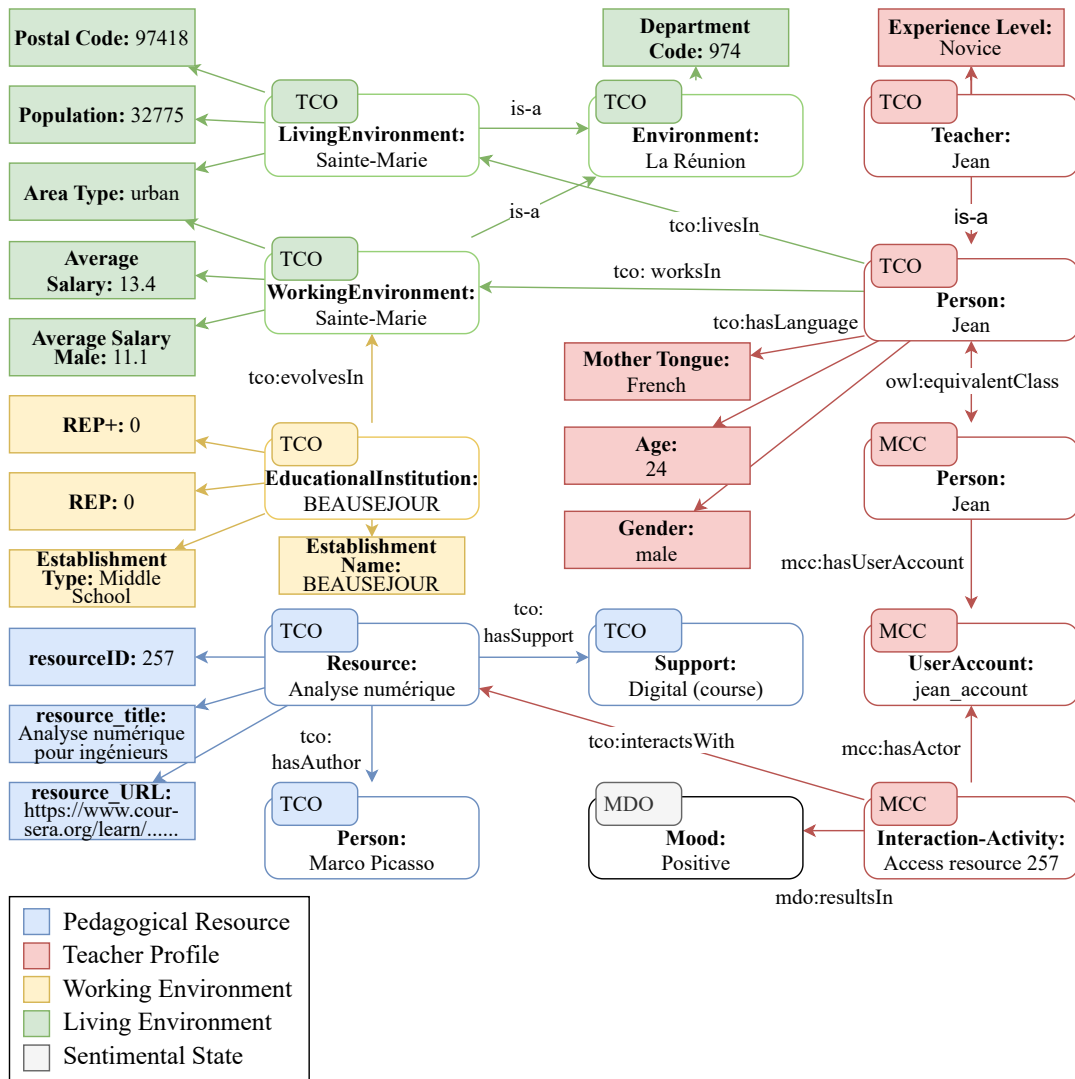


Figure 3.3: A partial A-Box snippet of mapping collected data into TSSCO highlighting multiple contexts of teacher.

Chapter 4

Teacher Contexts and Their Representation

4.1 Knowledge Engineering Methodology

Ontologies have transitioned from being limited to Artificial Intelligence laboratories to being widely used on the Web. Our motivation for developing an ontology includes sharing a common understanding of the multiple contexts of a teacher and their structures among different systems, enabling knowledge reuse, making domain assumptions explicit, separating domain knowledge from operational knowledge, and analyzing domain knowledge. By having a shared underlying ontology, different recommender systems and websites can extract and aggregate information, allowing agents to answer user queries or provide data to other applications.

As previously stated, an ontology is a structured representation of knowledge that includes concepts, properties, and their relationships within a domain, while a knowledge base is built by populating the ontology with individual instances and specific values for properties. Hence, throughout the development of our ontology, we follow the knowledge engineering guidelines provided by (Noy, McGuinness, et al., 2001). According to the guide (Noy, McGuinness, et al., 2001), there is no one "correct" approach to develop an ontology but some fundamental rules need to be respected, as follow:

- A domain may be modeled in more than one correct way depending on the application.
- An iterative process is necessary to develop an ontology.
- Concepts and relationships should be close to the domain of interest.

First, we start by defining our domain of interest and the scope of the ontology. Four questions need to be answered for the purpose of defining the domain and scope. The first question is "What is the domain of the ontology and its coverage?". According to our first research question (Section 1.1.2), the domain is the representation of the multiple contexts of a teacher. However, the second question provides a specification for the domain "What is the use of the ontology?". We refer to the second research question in an effort to answer this question which indicates that the ontology will be used to offer context-aware recommendations to teachers. This answer leads to the third question "What are the questions that the ontology will answer?". According to the guidelines,

a set of competency questions (Grüninger and Fox, 1995) are sketched to define the scope of the ontology and its knowledge base. The following list of competency questions illustrates possible required knowledge in view of the answers provided to the first two questions. These competency questions are answered accordingly through the sections of this chapter.

- Which contexts should be considered when describing a teacher?
- Is this teacher an expert?
- What field of science is this teacher specialized in?
- In which educational level does this teacher work?
- What is the suitable teaching style for the targeted learners?
- Is the working environment of a teacher suitable for this living environment?
- What are the teacher characteristics that affect his capability to perform his job in a certain working environment?
- Does the work environment affect the sentimental state of a teacher?
- Is this teacher similar to another? And why?
- What are the required characteristics to recommend personalized pedagogical resources to a teacher?

The final question pertains to the identification of potential ontology users and maintainers. The users are envisioned as researchers employing it in conjunction with recommender systems, whereas the maintainers are the research team tasked with its ongoing development. Consequently, the requirement for a mapping methodology to diverse languages is deemed unnecessary.

Subsequently, the guide on ontology development outlines a series of procedural stages as follows:

- Step 1: Contemplate the possibility of reusing existing ontologies.
- Step 2: Enumerate essential terms relevant to the ontology.
- Step 3: Establish the classes and their hierarchical organization.
- Step 4: Specify the properties of classes.
- Step 5: Define the attributes or facets of the properties.
- Step 6: Generate instances of the ontology.

Regarding the first two steps and as per the findings from the literature review, no ontologies have been identified that specifically address the educational context from the perspective of teachers. However, we can leverage existing ontologies to expand the concepts within our ontology. For instance, The Modern Science Ontology (modsci)¹ offers a comprehensive representation that can replace

¹https://saidfathalla.github.io/Science-knowledge-graph-ontologies/doc/ModSci_doc/index-en.html

the field of science associated with teachers in our ontology. Moreover, the Semantic Sensor Network Ontology (SSN)² facilitates the representation of various sensors and their corresponding observations, thereby extending the representation of sentimental state detection in our ontology. Upon examining different teacher contexts, it is evident that certain vocabularies are present in other existing ontologies. For instance, the country attribute can be aptly represented using the DBpedia ontology³. Finally, the primary objective of our ontology is to be utilized in conjunction with a recommender system, a purpose that aligns efficiently with the MEMORAE Collaboration-Context Ontology (MCC) (Abel, 2022).

As mentioned earlier, our intended utilization of the ontology involves its integration with a recommender system, necessitating the consideration of diverse contextual factors (Agarwal, Chen, and Long, 2011). These factors can be categorized into four primary groups: teacher profile, living environment, working environment, and teacher feedback (Han and Yin, 2016), and depicted in Figure 4.1. These factors serve as the foundational underpinnings for the development of our ontology.

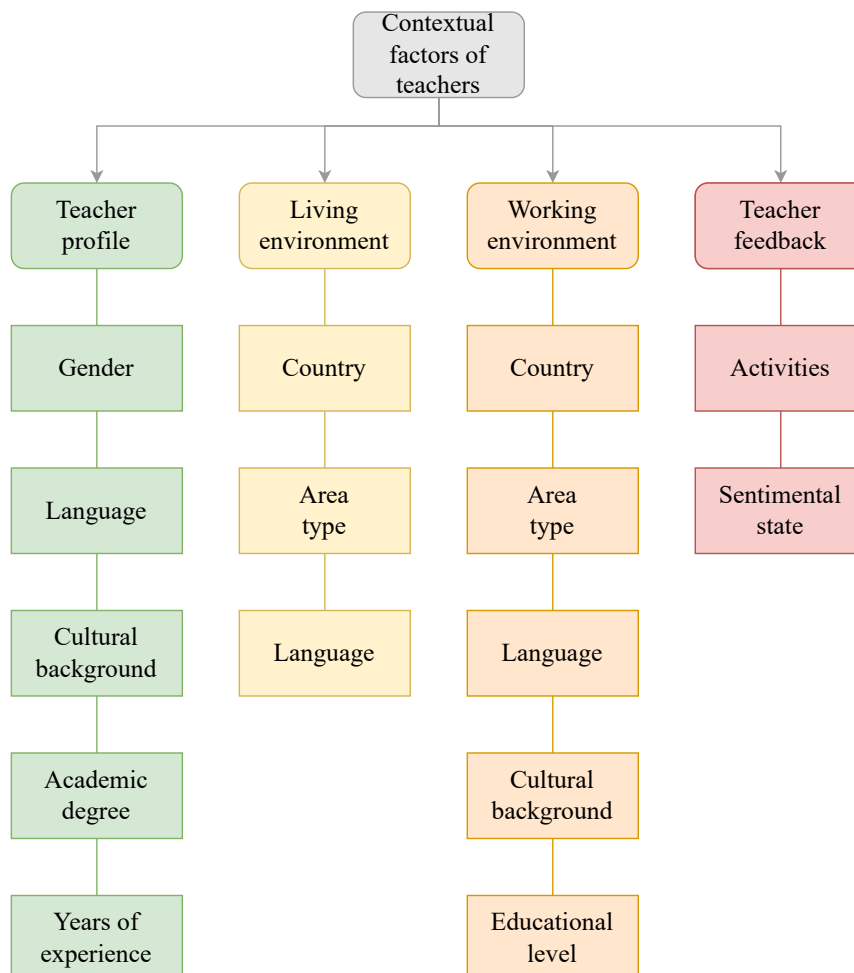


Figure 4.1: Contextual factors of teachers.

²<https://www.w3.org/TR/vocab-ssn/>

³<https://dbpedia.org/ontology/>

In pursuit of a comprehensive top-down perspective on domain ontologies, the present approach adheres to the same principles concerning the concepts within this ontology. This chapter provides an illustration of these concepts, aligning them with the steps outlined in the guidelines for ontology development.

The subsequent sections of this chapter are structured as follows: Section 4.2 offers an introduction to the Teacher Sentimental and Collaboration Contexts Ontology (TSCCO) and its integration with other ontologies. Section 4.3 provides an overview of the distinct components comprising the Teacher Context Ontology (TCO), while Section 4.4 presents a detailed exposition of the sentimental state monitoring approach and its Mood Detection Ontology (MDO). Section 4.5 delves into the practical application of the MEMORAe Collaboration Context ontology (MCC) in conjunction with TSCCO. Lastly, Section 4.7 elucidates the role and function of the SWRL reasoning rules within TSCCO.

4.2 Teacher Sentimental and Collaboration Contexts Ontology (TSCCO)

Teacher Sentimental and Collaboration Contexts Ontology (TSCCO) expounded in this chapter provide an enhanced comprehension of the different contexts of a teacher, encompassing the contextual factors that impact a profile of a teacher and his sentimental state. The ontologies are encoded using the Web Ontology Language (OWL) and have been developed using the Protégé framework, as depicted in Figure 4.2.

TSCCO centers its attention on the exposition of three ontologies: the Teacher Context Ontology (TCO), the Mood-Detection Ontology (MDO), and the MEMORAe Collaboration Context ontology (MCC), as shown in Figure 4.3. TCO is dedicated to offering a comprehensive depiction of the context surrounding a teacher working in an educational institution and teaching learners enrolled at a particular educational level. The primary objective of this ontology is to provide a comprehensive portrayal of the diverse environmental and contextual factors that shape a teacher's profile, encompassing aspects such as their professional experience, teaching style, and living and working environments.

Conversely, the development of MDO centers on representing intricate relationships between mood detection systems and sentimental state representation in terms of mood. Advancements in mood detection technologies have enabled the identification of bio signals indicative of diverse mood states, and MDO endeavors to design an ontology that precisely interprets these mood detection systems, thereby determining a person's mood level.

MCC is purposefully developed to capture the collaboration context among various users within an organization. Functioning as a management ontology within the MEMORAe platform, MCC facilitates knowledge sharing among collaborators by means of resources and annotations. This section thoroughly examines the design of TSCCO and explores the integration between the three different ontologies.



Figure 4.2: Screenshot of Protégé view of TSCCO upper level concepts.

4.2.1 Integration of Three Ontologies

The Teacher Context Ontology (TCO) delineates the coexistence of teachers in various contexts, encompassing both their work and living environments. However, these contexts alone may not suffice to comprehensively describe a teacher's current situation, requiring the inclusion of their affective or non-affective state within the contextual description. By integrating mood with a teacher's context, novel opportunities arise for enhanced comprehension and evaluation of their current circumstances. In this regard, an integration approach is introduced between TCO and MDO, as illustrated in Figure 4.4.

In TCO, a person can be a teacher or a learner that interacts with resources. These activities and interactions, which are part of the integration between TCO and MEMORAe (MCC) (Nashed, Lahoud, and Abel, 2021), results in the mood changes of a person. As a result of this integration, the teacher full context is represented by describing both work and living conditions/environment from TCO, in addition to its associated mood.

MCC ontology provides a semantic representation for an account of a user MCC:UserAccount which extends the profiling of a teacher in TCO. With MCC, it becomes easier to share and compare profiles across different contexts and applications. For example, gender representation with ontological concepts can help to ensure that the professional development programs are inclusive and equitable for all teachers. By analyzing the specific gender factors that impact teaching effectiveness, it becomes possible to develop programs that support all teachers, regardless of gender, to thrive in their roles. Moreover, MCC associates an

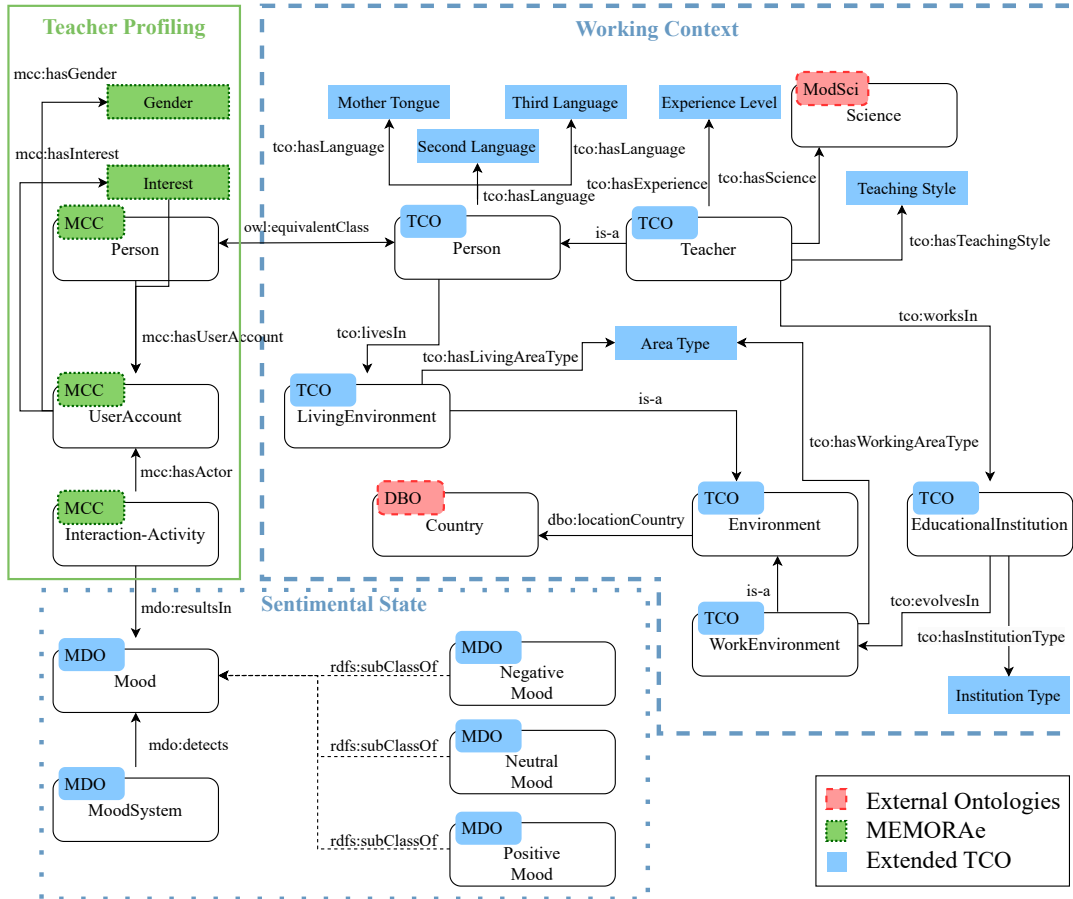


Figure 4.3: A partial T-Box representation of Teacher Sentimental and Collaboration Contexts Ontology (TSSCO)

interaction or an activity of teacher in an information system `MCC:Interaction-Activity` with each teacher profile `MCC:UserAccount` which facilitates the linkage between the mood of a teacher from `MDO` and the teacher from `TCO` as shown in Figure 4.5. Through the following sections, the present thesis proceeds to individually expound upon two developed ontologies, namely `TCO` and `MDO`, alongside the `MCC` ontology.

4.3 Teacher Context Ontology (TCO)

`TCO`, as shown in Figure 4.5, provides a context description of teacher who works in an educational institution and teaches learners enrolled in a specific educational level. The ontology introduces the concept of intersecting contextual information for a teacher profile by describing multiple environmental presence of a teacher (Li, Abel, and Negre, 2019). Also, `TCO` demonstrates teacher interactions with the resources to enrich the courses and consequently the educational process. The ontology was encoded using the Web Ontology Language (OWL)⁴

⁴<https://www.w3.org/TR/owl-ref/>

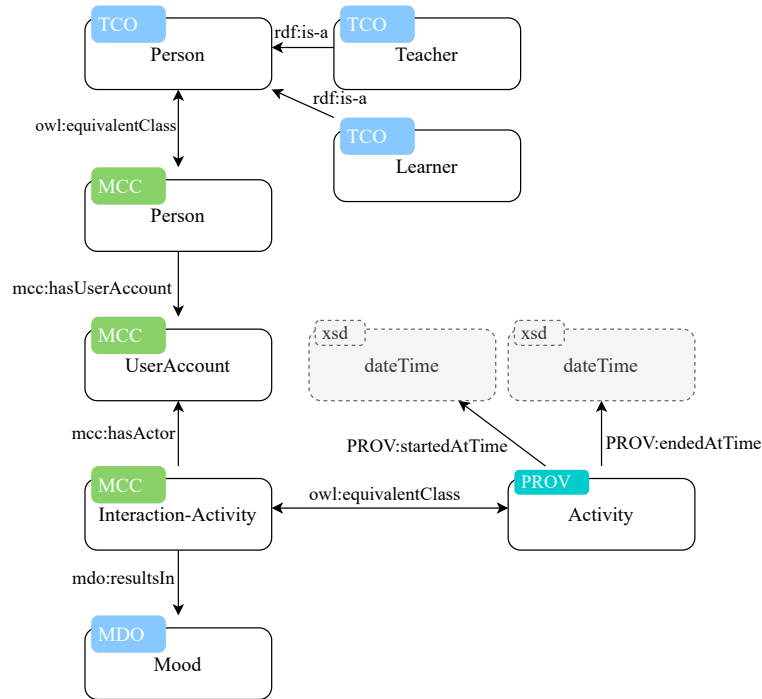


Figure 4.4: T-Box of TCO and MDO integration.

and was developed and reasoned by the Protégé framework⁵. The next subsections explain the main concepts of TCO and their role in a teacher context.

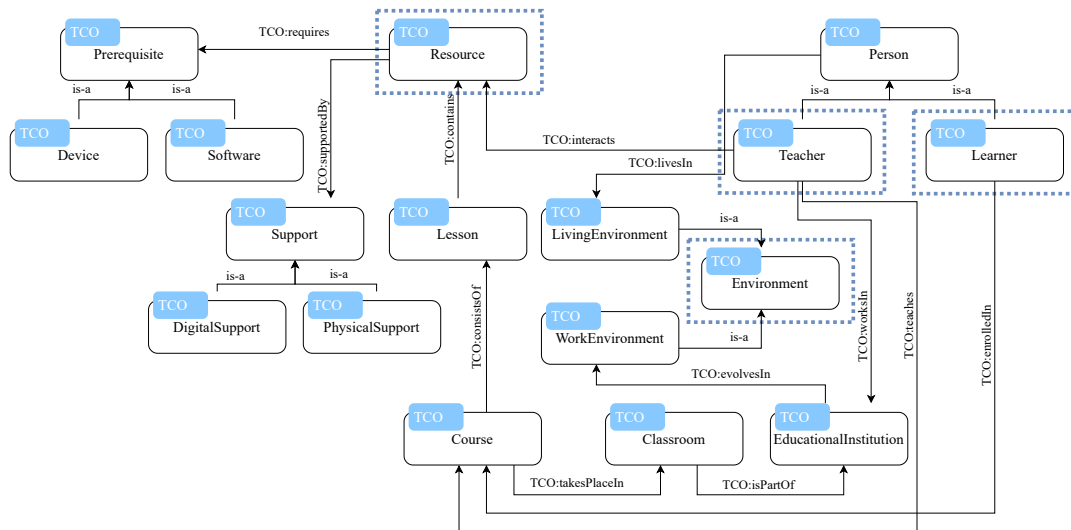


Figure 4.5: A T-Box overview of TCO highlighting the main concepts.

4.3.1 Teacher

A teacher participates in different contexts beside the educational context. When the teacher’s living environment is combined with her/his teaching profile, the

⁵<http://protege.stanford.edu>

context description is realistic and reliable. Thus, these contexts must be considered in the TCO. A teacher profile contains the basic information that is available in a user account such as name, age, gender, contact information, education, etc. In addition to this, other important information describing teacher's work experience and teaching style must be included as shown in Figure 4.6.

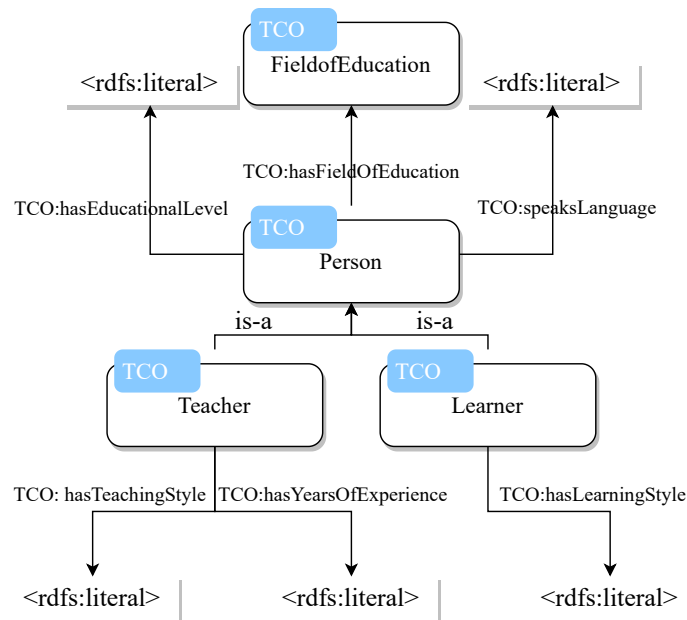


Figure 4.6: A T-Box example of teacher and learner representation in TCO.

In TCO, the experience of a teacher is measured according to two characteristics: years of experience during teaching career, and age of this teacher. According to these two characteristics, teachers are classified into three level: novice (beginner), intermediate, and expert levels (Rice, 2003; Carton and Fruchart, 2014). If a teacher has less than five years of teaching career experience for the same educational level and ages less than or equal 25 years old, a teacher is classified as a novice. For more than five years of teaching experience and less than 10 years, and an age within the range of 26 and 36 years, a teacher is classified as an intermediate teacher. A teacher is considered as an expert, if he has an age more than 35 years old and more than 10 years of experience. On the other hand, teachers who obtain a higher education degree, are promoted to the higher experience level despite the number of years. Other characteristics can be debatably included in this classification such as teaching style and the nature of targeted learners, but they are opted out due to high complexity concerns.

In order to consider more personalized information about the teacher, we must explore the teaching styles and its diversity. Teaching styles are divided into five main types according to the instructors' preferences in their courses: authority, demonstrator, facilitator, delegator, and hybrid styles (Opdenakker and Van Damme, 2006). The authority style teachers tend to deliver their knowledge on the form of one-way long lectures, while the demonstrator style ones prefer the coaching style using activities and practical demonstrations. Teachers using facilitator style always try to provide resources and self-learning activities to learners. Delegators do not prefer teacher-centered methods rather they

proceed toward group style activities such as laboratory and feedback activities where learners are dependent on each other. The final type of teaching styles is the hybrid or blended style where teachers blend their personal experience and interests of different styles with learners.

The teacher concept has multiple relations with other concepts as shown in Figure 4.5. In order to represent the workplace of a teacher, the relation TCO:worksIn is created to illustrate such connection. Normally, teacher teaches a certain educational course and hence the relation TCO:teaches. For a better illustration of this course, a teacher TCO:interacts with available resources.

4.3.2 Learner

In order to achieve an efficient description of the teacher context, the learners are present in TCO. Learner's profiling is defined also by the basic information that is included in a user's account. Learners in TCO are characterized by the educational level and the learning styles as shown in Figure 4.6. The combination of this knowledge about the target learners relates to the teacher context description. Accordingly, educational levels are classified according to the 2011 international standard classification of education (ISCED) (Schneider, 2013). ISCED delineates a hierarchical arrangement encompassing a total of seven levels within the educational spectrum. These levels are further consolidated into four overarching categories: early childhood education (elementary/primary), lower secondary education (preparatory), upper secondary education (secondary), and tertiary education (undergraduate and postgraduate university levels).

Learners achieve better understanding if the teacher's style of teaching matches their learning style (Peacock, 2001; Awla, 2014; Dinçol et al., 2011). The learning styles are represented according to one of these three terms: cognitive, personality, and sensory (Peacock, 2001). The cognitive type borders the logical type of learners where they use logic and reasoning to understand a learning problem. While according to personality, learners tend to study either in group or individually. The learners preferring learning in group interpersonally, are called the social learners, while the others preferring interpersonal self-studying, are called the solitary learners. The sensory type is the last category of learners which contains four types: visual, aural, verbal, and physical learners. A visual style learner favors the spatial representations to understand a given subject using images or videos. On the other hand, aural learners choose to search for auditory and musical resources of learning such as music or sounds in general. Verbal learners can easily recognize linguistics in the form of speeches and written explanations. The final type of the sensory category is the physical style where learners prefer kinesthetic activities using their bodies to get familiar with new topics.

The learner concept intersects with the teacher through the relation TCO:enrolledIn which connects the learner with the course in which the teacher TCO:teaches as shown in Figure 4.5. Accordingly, learner and teacher contexts' representations are misleading without providing an accurate description of the contextual information which exists outside the educational process. In order to realize these connections, the relations are created to connect the environment of teacher, learner, and educational institution concepts together.

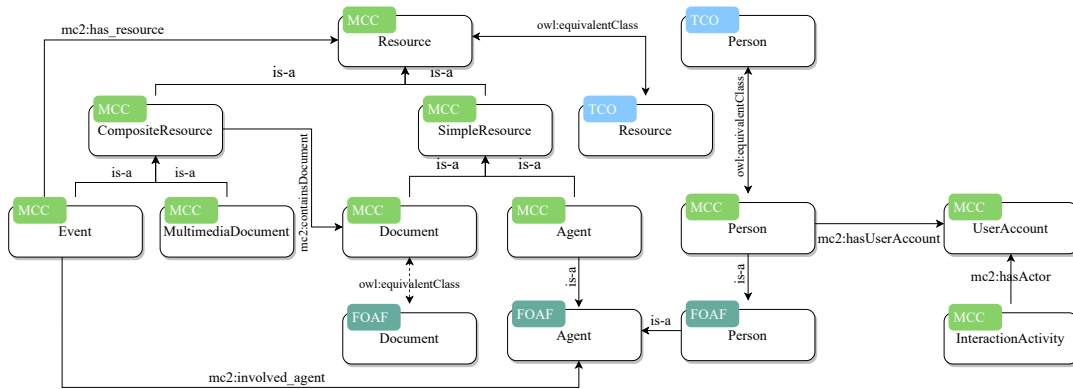


Figure 4.7: A T-Box example of resource representation in TCO.

4.3.3 Environment

A user profile contains information from the user current context along with other contextual information from other contexts where the user acts (Li, Abel, and Negre, 2019). Therefore, TCO's environment is classified into living and working environment as shown in Figure 4.5. The living environment describes the surroundings and conditions in which the person had been raised, while the working environment describes the similar concept for the educational institution. Country, language, and nature of the region are the main data required to describe an environment. As for the working environment, the infrastructure and available technologies are included to extend the description of the working environment.

The role of the environment concept representation in TCO is raised to have a full description of all teacher contexts; living environment, working environment, and the learners living environment. Identification of the three intersecting contexts facilitates the condition matching for the different environments where the teacher interacts with the surroundings.

4.3.4 Resource

An educational course consists of multiple lessons, each with a different goal for learners. Thus, teacher includes pedagogical resources of different forms to enrich the educational content and motivate learners (Khan, Yusoff, and Khan, 2014). These pedagogical resources are found in different forms and are used under certain conditions and requirements (Benayache, 2005). TCO demonstrates the construction of a resource, its support and prerequisites as shown in Figure 4.5. Then, it is extended by the MCC representation of MCC:Resource as shown in Figure 4.7. Every resource is TCO:supportedBy a certain support digitally or physically. A digital support takes the form of a multimedia file, an executable file, a text, or URL. While the physical support is a tape, CD/DVD, or a paper. A resource also TCO:requires a form of prerequisite to use it probably. This prerequisite is met by using a device, a software, or both. The importance of including the support and the prerequisite is raised due to the changeable conditions in the various environments where the educational process takes place. Resources must be well-represented for the future integration of TCO with a recommender system.

4.4 Mood-Detection Ontology (MDO)

Recent advances in brain imaging technology have enabled the detection of brain signals indicative of various mood states. Therefore, mood detection systems have been widely developed during recent years with a persisting need to accurately interpret these brain signals to determine a person's emotional state. To achieve this, it is necessary to design a mood detection ontology that represents the complex relationships between brain signals and emotions (Al-Nafjan et al., 2017). In this section, the mood detection mechanism is discussed and the development of MDO is presented.

4.4.1 Mood Detection Model

The mood-monitoring platform (Moodflow@doubleYou) is followed as a reference model (Andres et al., 2021; Frangeto et al., 2021), as shown in Figure 4.8. Moodflow@doubleYou platform allows the collection of user data and the detection of his mood which facilitates the integration with TCO to offer better representation of the collected mood data (Nashed et al., 2021).

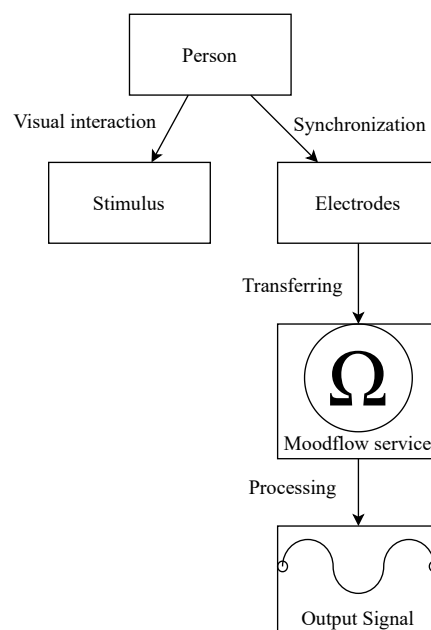


Figure 4.8: An overview diagram of Moodflow@doubleYou platform.

The detected mood by this platform is categorized into three levels: positive, neutral, and negative Figure 4.9 (Corson, 2002). The positive mood is detected by positive value readings, and it results from the personal desire to accomplish a goal or the searching for fulfillment. When a person is expecting an action to happen or doing a legitimate task, the platform outputs a neutral reading to indicate the neutral mood. However, the negative mood is an indication of boredom, repulsion, or reaching high satisfaction level.

Emotional commitment is the person's ability to maintain the positive mood for a certain period (Corson, 2002). Average absolute deviation AAD is used to detect the emotional commitment of a certain person during a specific activity

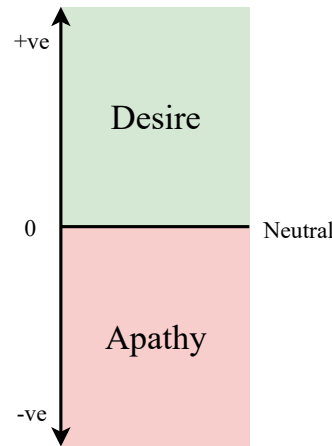


Figure 4.9: Mood levels obtained by Moodflow@doubleYou.

Equation(4.1). According to Equation(4.2), if the AAD is nearly equal to zero, the person is said to be emotionally committed for the current task t . If the person keeps the emotional commitment over multiple activities, it is said that the person is in flow (Gnoth et al., 2000). The number of activities is decided according to the total number of recorded activities for each person.

$$AAD_t = \frac{1}{n} \sum_{i=i_0}^n |x_i - m(X)|; x_i \geq 0 \quad (4.1)$$

$$EC = \begin{cases} 0 & AAD_t \neq 0 \\ 1 & AAD_t \simeq 0 \end{cases} \quad (4.2)$$

4.4.2 Ontological Representation of Mood Concepts

Mood is one of the sentimental state forms that provides a recommender system with real-time feedback of user. However, mood can be detected using different systems and approaches which differs in its methodology and nature. Consequently, this thesis introduces mood detection ontology (MDO) that represents the three levels of mood in addition to emotional commitment, and flow Figure4.10. This ontology was implemented with Web Ontology Language (OWL) ⁶ in Protégé framework ⁷. The main classes of the ontology are $MDO:Mood$, $MDO:EmotionalCommitment$, $MDO:Flow$, $MDO:MoodSystem$, and $MDO:Process$. The ontology is centred around two concepts: mood representation and mood detection system. In the mood representation, mood with its types, emotional commitment and flow are connected through relational properties. However, in mood detection system representation, the classes extend its meaning from W3C-SSN⁸. One of the remarkable concepts in this concept is the $MDO:process$ class, which is divided into three sub-processes (sub-classes): $MDO:measuring$: the measuring process of the readings of electrodes, $MDO:transferring$: transferring the readings to the system, and $MDO:analysis$: analysing the transferred measurement.

⁶<https://www.w3.org/TR/owl-ref/>

⁷<https://protege.stanford.edu/>

⁸<https://www.w3.org/TR/vocab-ssn/>

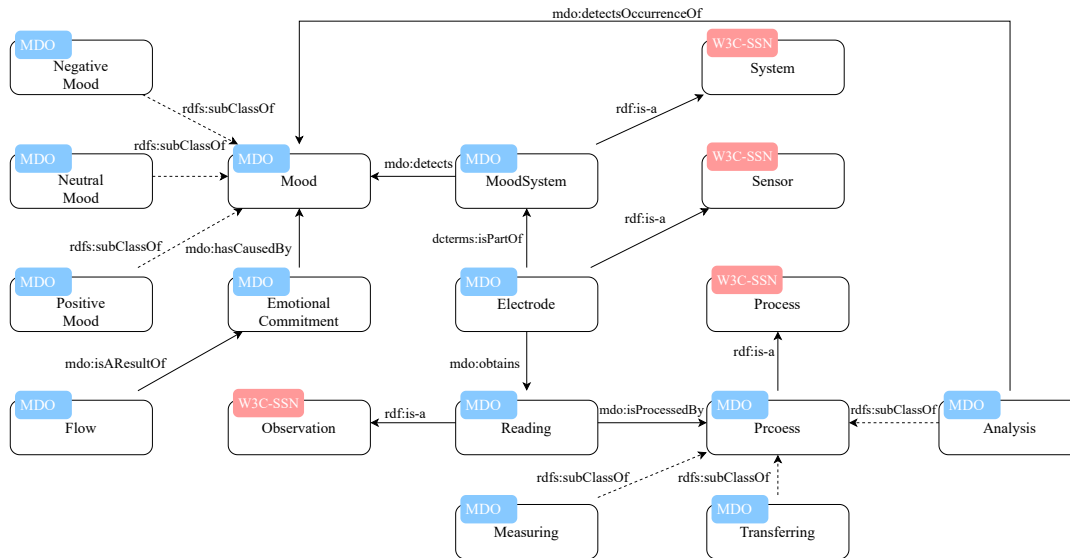


Figure 4.10: A T-Box example of mood detection ontology (MDO).

4.5 MEMORAE Collaboration-Context Ontology (MCC)

MEMORAE Collaboration-Context Ontology (MCC) is an ontology developed to support the creation of collaborative experience sharing environments and based on the concepts of MC2 and SoIS ontologies (Abel, 2022). MCC has the capability to define and encompass contextual information, factors, and dimensions relevant to collaboration. MCC adeptly describes collaborative attempts of a user along with the contextual settings in which these collaborations occurs. The ontology provides a representation of the context in which collaboration takes place, including the roles of the participants, the tasks they perform, and the artifacts they use. The MCC ontology is designed to be used in conjunction with other ontologies, such as TCO and MDO ontologies. By using these ontologies together, it is possible to create a rich and detailed representation of a learning environment for teachers as well as learners, which can be used to support a variety of learning activities. One of the main benefits of the MCC ontology is that it provides a standardized way of representing collaboration context information. This makes it easier for developers to create collaborative learning environments that are interoperable and can work together seamlessly. The ontology is also useful for researchers who are interested in studying collaboration and its impact on learning. Overall, the MEMORAE Collaboration-Context Ontology (MCC) is an important tool for the development of collaborative learning environments and has the potential to improve the effectiveness of collaborative learning. More details of this ontology concerning data integration is discussed in Section 5.5.1 of Chapter 5.

4.6 Complete Contexts of Teacher

The complete representation, as shown in Figure 4.11, can be divided into three main sections: teacher profiling, living/social environment, working environment, and sentimental state (Nashed, Lahoud, and Abel, 2021).

4.6.1 Teacher Profiling

The profile of a teacher involves the main information of a teacher such as name, age, experience, gender, and competences. These concepts identify the key features of a teacher and avoid overlapping and redundancy of features with the other integrated ontologies. Also, it becomes possible to match teachers with learners based on their individual learning needs and preferences. TCO represents experience and competences elements as illustrated in the following list while other elements are addressed by MCC.

- i *Experience level* are used to detect the level of expertise of a teacher: novice, intermediate, or expert.
- ii *Competences* are divided into science, teaching style and spoken languages by a teacher. Teaching style is represented through 5 types: authority, demonstrator, facilitator, delegator, and hybrid style (Opdenakker and Van Damme, 2006). Languages are divided into three levels: mother tongue, second language, and third language.

The list of selected elements from MCC for this approach is as follows:

- i *Gender* of a teacher can be a female or a male.

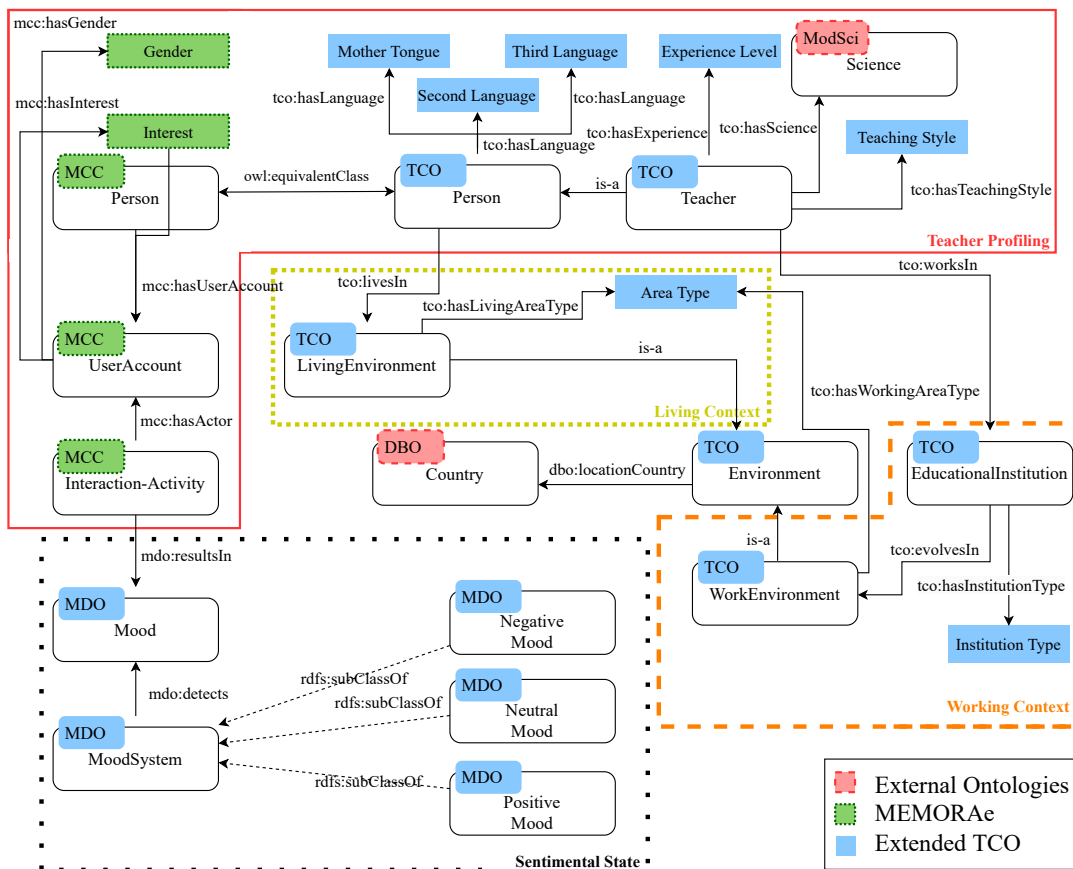


Figure 4.11: Different contexts partial T-Box representation of Teacher Sentimental and Collaboration Contexts Ontology (TSSCO)

- ii *Interests* are list of keywords which are provided by teacher to represent the points of interest for this teacher.

4.6.2 Living Context

The living context-related concepts can help with feature selection by allowing the identification of the most relevant features for a particular analysis. For example, if an analysis is interested in examining the impact of socio-economic status on teacher effectiveness, the ontology can be used to identify relevant features such as country and area type which forms a direct indication of accessibility to resources. The living context of a teacher in TCO, is represented by personal address and type of the living area as follows:

- i *Country* is obtained from the personal address provided by teacher.
- ii *Area Type* is detected using the same address and is classified into three main types: rural, urban, or suburban areas.

4.6.3 Working Context

The working context can include factors such as the location of educational institution and level of education which directly indicates the required technology and support services, for example. The working context of a teacher is accentuated by address and affiliation which forms the connection between the geographical location and educational level of this work environment, or the educational institution as shown in the following list.

- i *Institution Level* indicates the level of education of the educational institution where the teacher works. There are four levels of educational institutions: primary school (elementary school), preparatory school (middle school), secondary school (high school), and university.
- ii *Area Type* is detected using the same address and is classified into three main types: rural, urban, or suburban areas.

4.6.4 Sentimental State

The sentimental state of a teacher is described by the mood-detection ontology (MDO) through three main concepts: mood level, emotional commitment, and flow occurrence (Nashed et al., 2021). Through this proposition, the mood level is used without considering the other two concepts for the purpose of simplification. This concept is described as following.

1. *Mood level* measures mood of teacher according to readings of used electrodes and computational system. This method of measurement introduces three level of mood: positive, neutral, and negative moods.

Afterwards, SWRL reasoning rules are applied to check for possible similarities between teachers.

4.7 Semantic Reasoning

Semantic web rule language (SWRL) offers a practical way to define casual relations and has been used by multiple knowledge management systems (Horrocks et al., 2004). Therefore, the ontology representation provides a structured way to represent the context of a teacher, which can help to reduce the noise and redundancy in the data. By using SWRL reasoning to find similarities between teachers' contexts, it becomes possible to identify patterns and relationships that may not be immediately obvious. This can lead to more accurate and reliable results. In this approach, SWRL rules find similarity between teachers in terms of features and properties. These SWRL rules are implicated between antecedents and consequent in the form *antecedent* \rightarrow *consequent*. Conjunctive atoms are used to express these antecedents and consequents. Conjunctive antecedents are represented as the following atoms: $a_1 \wedge a_2 \wedge \dots \wedge a_i \wedge \dots \wedge a_n$, where $1 \leq i \leq n$. In our case, atoms are classified into three forms: class $C(?x)$, property $P(?x, ?y)$, and "sameAs" relation $sameAs(?x, ?y)$. Class atom $C(?x)$ holds if x is considered an instance of class C , while property atom $P(?x, ?y)$ holds if x and y have a mutual relation P . SWRL rules are used in different phases of our approach: preprocessing of data, resource recommendation and peer recommendation. Some rules are used in an exchangeable manner while others are used for a certain phase only. Through the following paragraphs, we explain the function of each rule and its placement within our approach.

4.7.1 Preprocessing SWRL Rules

The application of the first three SWRL reasoning rules identifies the degree of experience of each teacher. This identification can aid in tailored professional development and training. It is possible to discover areas where a teacher may want further help or training by studying the individual traits and aspects that contribute to their competence. These first three SWRL rules in Table 4.1 define the levels of expertise for a new teacher: `tco:noviceTeacher`, `tco:intermediateTeacher`, and `tco:expertTeacher`. These rules use the age of a teacher and the years of experience to classify each new teacher into one of the predefined expertise levels. For example, the first rule means that if a teacher is younger than 25 years old and has less than 5 years of experience, this teacher can be classified as a `tco:noviceTeacher`. The other expertise levels are identified as follows: `tco:intermediateTeacher` for teachers with an age that is older than 25 years and younger than 35 years and experience of more than 5 years and less than 10 years, `tco:expertTeacher` for teacher with an age that is older than 35 years and experience of more than 10 years.

The fourth SWRL rule, in Table 4.1, is applied to any new teacher entry against all saved teachers. This rule checks the similarity of the new teacher with other teachers in terms of experience, teaching style, spoken languages, living environment information, working environment information, and field of science in which a teacher is specialized. The rule starts by matching two teachers, `?t` and `?ts`, who have the same years of experience `?ex` and teach with the same style or using the `MixedStyle` teaching style. Additionally, both teachers must have at least one language in common. Next, the rule checks that both teachers live

in environments that have the same type and country. The type and country information is retrieved from the environment `?env` and country `?coun` that each teacher lives in. The rule also checks that both teachers work in educational institutions that have the same education level and work environments that have the same type. The education level and work environment type information is retrieved from the educational institution `?inst` and work environment `?workenv` that each teacher works in. Finally, the rule checks that both teachers specialize in the same field of science, based on the science `?sci` that they each have expertise in. If all of these conditions are met, the rule concludes that the two teachers are the same person (using the `sameAs` built-in function). This means that the system will treat `?t` and `?ts` as identical variables and return the same value for both of them in any subsequent queries. The utilization of such rule prevents the computational overhead accompanied by executing the totality of this approach for each new teacher.

4.7.2 Resource and Peer Recommendation SWRL Rules

The next set of rules provided in Table 4.1 consists of a series of extraction rules that are used by both types of recommendation to select the matching teachers from a list. The fifth rule is used to find teachers with higher experience than a selected teacher. The rule takes as input two teachers, their respective years of experience, and returns the teacher with the highest number of years of experience. The sixth rule is used to find teachers who have a better mood level. The rule takes as input the mood levels of both parties and returns the second teacher if he obtained more positive mood level. The seventh rule is used to find teachers who have similar teaching style or use a mixed teaching style. The rule takes as input two teachers and their teaching styles and returns the second teacher if he uses similar teaching style or a mixed one. The eighth rule is used to find teachers who are proficient in a similar language. The rule takes as input two teachers and their respective languages and returns the second teacher if he has a matching language. The ninth rule is used to find teachers who live in a given type of living environment. The rule takes as input two teachers and their respective living environments and returns the second teacher if he lives in the given type of living environment. The tenth rule is used to find teachers who live in a given country. The rule takes as input two teachers and their respective living environments and returns the second teacher if he lives in this given country.

The eleventh rule can be interpreted as follows: If there is a teacher who works in an educational institution and the educational institution evolves in a working environment, and this working environment is a type of environment that has the same type as the environment of another teacher who works in another educational institution, then select the second teacher. A set of conditions and actions are defined in this rule. The conditions include the `tco:teacher` and `tco:worksIn` predicates to identify a teacher who works in an educational institution. The `tco:EducationalInstitution` predicate is used to specify that the institution is an educational one. The `tco:evolvesIn` predicate is used to relate the educational institution to a working environment, which is also defined as a type of environment using the `tco:is-a` predicate. The last part of the condition compares the environment of the first teacher to the environment of another teacher, using the `tco:hasType` predicate. If the two environments

have the same type, as determined by the `swrlb:equal` function, then the rule selects the first teacher using the `sqwrl:select` function. Overall, this rule aims to identify teachers who work in similar environments, as defined by their respective educational institutions' working environments, and are therefore potentially interested in collaborating or sharing ideas.

Finally, the twelfth rule is used to find teachers who have a given level of education. The rule takes as input two teachers and their respective educational institutions and returns the second teacher if he has this given level of education. The final rule is used to find teachers who specialize in a given field of science. The rule takes as input two teachers and their respective fields of science and returns the second teacher if he specializes in this given field of science. Overall, these rules can be used to query a knowledge base about teachers and their characteristics. By reasoning about this knowledge, it is possible to identify teachers who possess certain attributes or have certain skills. This can be useful for a variety of recommendation purposes, such as identifying potential mentors/peers for new teachers or matching context of certain teacher with another who both have particular skills or experience.

Table 4.1: List of SWRL rules used in our approach

#	Antecedent	Consequence
1	tco:teacher(?t)^tco:hasYearsOfExperience(?t,?ex)^swrlb:lessThan(?ex,5)^tco:hasAge(?t,?age)^swrlb:lessThan(?age,25)	sameAs(tco:noviceTeacher,?t)
2	tco:teacher(?t)^tco:hasYearsOfExperience(?t,?ex)^swrlb:lessThan(?ex,10)^swrlb:greaterThan(?ex,5)^swrlb:greaterThanOrEqual(?age,25)^swrlb:lessThan(?age,35)	sameAs(tco:intermediateTeacher,?t)
3	tco:teacher(?t)^tco:hasYearsOfExperience(?t,?ex)^swrlb:greaterThan(?ex,10)^swrlb:greaterThanOrEqual(?age,35)	sameAs(tco:expertTeacher,?t)
4	tco:teacher(?t)^tco:teacher(?ts)^tco:hasYearsOfExperience(?t,?ex)^tco:hasYearsOfExperience(?ts,?exs)^swrlb:greaterThanOrEqual(?exs,?ex)^tco:hasTeachingStyle(?t,?tch)^tco:hasTeachingStyle(?ts,?tchs)^swrlb:stringEqualIgnoreCase(?tch,?tchs)^swrlb:stringEqualIgnoreCase(?tch,"MixedStyle")^tco:hasLanguage(?t,?lan)^tco:hasLanguage(?ts,?lans)^swrlb:contains(?lan,?lans)^tco:livesIn(?t,?livenv)^tco:LivingEnvironment(?livenv)^tco:is-a(?livenv,?env)^tco:environment(?env)^tco:hasType(?env,?envtype)^tco:hasCountry(?env,?coun)^owl:country(?coun)^tco:livesIn(?ts,?livenvs)^tco:LivingEnvironment(?livenvs)^tco:is-a(?livenvs,?envs)^tco:environment(?envs)^tco:hasType(?envs,?envtypes)^tco:hasCountry(?envs,?couns)^owl:country(?couns)^swrlb:equal(?envtype,?envtypes)^swrlb:equal(?coun,?couns)^tco:worksIn(?t,?inst)^tco:EducationalInstitution(?inst)^tco:evolvesIn(?inst,?workenv)^tco:WorkingEnvironment(?workenv)^tco:is-a(?workenv,?wenv)^tco:environment(?wenv)^tco:hasType(?wenv,?wenvtype)^tco:hasEducationLevel(?inst,?edulvl)^dcterms:EducationLevel(?edulvl)^tco:worksIn(?ts,?insts)^tco:EducationalInstitution(?insts)^tco:evolvesIn(?insts,?workenvs)^tco:WorkingEnvironment(?workenvs)^tco:is-a(?workenvs,?wenvs)^tco:environment(?wenvs)^tco:hasType(?wenvs,?wenvtypes)^tco:hasEducationLevel(?insts,?edulvls)^dcterms:EducationLevel(?edulvls)^swrlb:equal(?envtype,?envtypes)^swrlb:equal(?edulvl,?edulvls)^tco:hasScience(?t,?sci)^modsci:Science(?sci)^tco:hasScience(?ts,?scis)^modsci:Science(?scis)^swrlb:equal(?sci,?scis)	sameAs(?t,?ts)
5	tco:teacher(?t)^tco:hasYearsOfExperience(?t,?ex)^tco:teacher(?ts)^tco:hasYearsOfExperience(?ts,?exs)^swrlb:greaterThan(?exs,?ex)	sqwrl:select(?ts)
6	mcc:InteractionActivity(?e)^mcc:hasActor(?e,?acc)^mdo:resultsIn(?e,?m)^mdo:mood(?m)^mdo:hasValue(?m,?v)^mcc:hasActor(?e,?accs)^mdo:resultsIn(?e,?ms)^mdo:mood(?ms)^mdo:hasValue(?ms,?vs)^swrlb:greaterThan(?vs,?v)	sqwrl:select(?ts)
7	tco:teacher(?t)^tco:hasTeachingStyle(?t,?tch)^tco:teacher(?ts)^tco:hasTeachingStyle(?ts,?tchs)^swrlb:stringEqualIgnoreCase(?tch,?tchs)^swrlb:stringEqualIgnoreCase(?tch,"MixedStyle")	sqwrl:select(?ts)
8	tco:teacher(?t)^tco:hasLanguage(?t,?lan)^tco:teacher(?ts)^tco:hasLanguage(?ts,?lans)^swrlb:contains(?lan,?lans)	sqwrl:select(?ts)
9	tco:teacher(?t)^tco:livesIn(?t,?livenv)^tco:LivingEnvironment(?livenv)^tco:is-a(?livenv,?env)^tco:environment(?env)^tco:hasType(?env,?envtype)^tco:teacher(?ts)^tco:livesIn(?ts,?livenvs)^tco:LivingEnvironment(?livenvs)^tco:is-a(?livenvs,?envs)^tco:environment(?envs)^tco:hasType(?envs,?envtypes)^swrlb:equal(?envtype,?envtypes)	sqwrl:select(?ts)
10	tco:teacher(?t)^tco:livesIn(?t,?livenv)^tco:LivingEnvironment(?livenv)^tco:is-a(?livenv,?env)^tco:environment(?env)^tco:hasCountry(?env,?coun)^owl:country(?coun)^tco:teacher(?ts)^tco:livesIn(?ts,?livenvs)^tco:LivingEnvironment(?livenvs)^tco:is-a(?livenvs,?envs)^tco:environment(?envs)^tco:hasCountry(?envs,?couns)^owl:country(?couns)^swrlb:equal(?coun,?couns)	sqwrl:select(?ts)
11	tco:teacher(?t)^tco:worksIn(?t,?inst)^tco:EducationalInstitution(?inst)^tco:evolvesIn(?inst,?workenv)^tco:WorkingEnvironment(?workenv)^tco:is-a(?workenv,?env)^tco:environment(?env)^tco:hasType(?env,?envtype)^tco:teacher(?ts)^tco:worksIn(?ts,?insts)^tco:EducationalInstitution(?insts)^tco:evolvesIn(?insts,?workenvs)^tco:WorkingEnvironment(?workenvs)^tco:is-a(?workenvs,?envs)^tco:environment(?envs)^tco:hasType(?envs,?envtypes)^swrlb:equal(?envtype,?envtypes)	sqwrl:select(?ts)
12	tco:teacher(?t)^tco:worksIn(?t,?inst)^tco:EducationalInstitution(?inst)^tco:hasEducationLevel(?inst,?edulvl)^dcterms:EducationLevel(?edulvl)^tco:teacher(?ts)^tco:worksIn(?ts,?insts)^tco:EducationalInstitution(?insts)^tco:hasEducationLevel(?insts,?edulvls)^dcterms:EducationLevel(?edulvls)^swrlb:equal(?edulvl,?edulvls)	sqwrl:select(?ts)
13	tco:teacher(?t)^tco:hasScience(?t,?sci)^modsci:Science(?sci)^tco:teacher(?ts)^tco:hasScience(?ts,?scis)^modsci:Science(?scis)^swrlb:equal(?sci,?scis)	sqwrl:select(?ts)

Chapter 5

Data Collection

5.1 Data Sources Layer

The coexistence of a teacher in multiple contexts forces the data extraction from different sources and their integration into one form of data format (Varelas, 2012) as shown in Figure 5.1. A person can be either a teacher or a learner and is represented by the integration of user-provided information, organizational information, and user-recorded activities. These analytics are collected during the real-time usage of teacher to an educational system such as a recommender system (Romero and Ventura, 2020). The multiple contexts of a teacher are illustrated through environment-related and institutions datasets that are publicly published by governmental and research entities. This data integration approach highlights the necessity of resources to aid teachers and enhance their performance. Accordingly, data resources are collected from multiple datasets that target different types of resources in the educational field.

5.2 Questionnaire - Teacher Profile

We recognize the significant role that the contexts of teachers play in shaping their minds and their performance. As such, understanding their experiences and the factors that influence their well-being is crucial. With this in mind, we have developed a comprehensive questionnaire designed to gather a wealth of information about teachers and their work environment. This questionnaire is adapted to suits both the French and Egyptian contexts as this thesis takes place within a collaboration between two universities in both countries. By delving into various aspects of their personal and professional lives, this questionnaire aims to shed light on the diverse backgrounds, living conditions, educational institutions, professional development opportunities, student advancement, and sentimental well-being of teachers. The insights gleaned from this research will not only contribute to a deeper understanding of the real-life context of teachers, but also help identify areas for improvement and implement measures that can enhance their overall teaching experience. This work presents the questionnaire as a tool for capturing valuable data and aims to provide a foundation for future research and evidence-based practices in the field of education. Through the participation of teachers and the collection of comprehensive data, we aspire to address the gaps in current knowledge and understanding of the complexities faced by teachers.

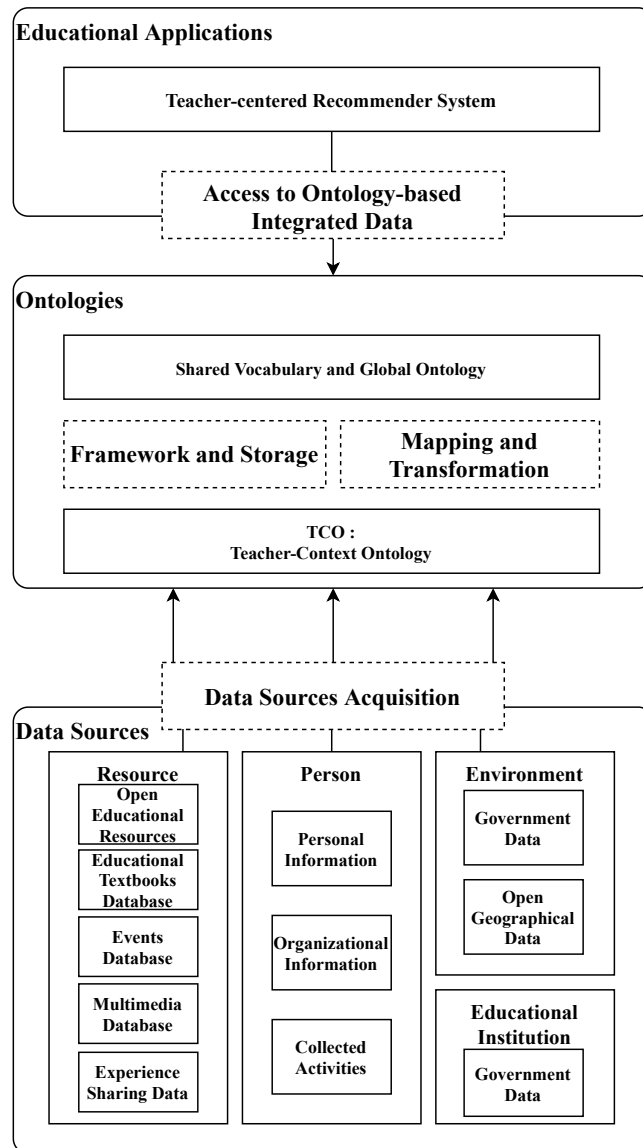


Figure 5.1: An overview of the proposed approach's layers.

5.2.1 Questionnaire Methods

Questionnaires are a method of data collection in research and there are several methods to conduct such questionnaires such as self-administrated, interviewer-administrated, online, phone, paper-based, and mixed-mode (White et al., 2005; Gomm, 2008). The self-administrated questionnaires are completed independently by respondents, with no direct interaction with the researcher. This method of questionnaire can be distributed in person, via mail, or shared over the internet. They provide respondents privacy and allow them to finish the questionnaire in their convenient method and time. One of the popular methods of self-administered questionnaires is online questionnaires which are administered via the internet, using a special platform or a tool. This method of questionnaire is convenient, cost-effective, and offers easier data collection and analysis. Another method is paper-based questionnaires which are printed and distributed manually. Respondents fill out the questionnaire on paper, which is later collected

and processed by the researcher. Interviewer-administered questionnaires are conducted through an interviewer who reads the questions to the respondent and records the correspondent answers in this manner. The interviewer might explain any ambiguities and probe for further details. This method is associated with cost overhead and expenses but allows for more in-depth data collection with complex and sensitive topics. Interviewer-administered questionnaires can be carried over a phone call (phone questionnaire). The interviewer reads the questions and records the answer of the respondent. It is widely used in market research and social studies. Mixed-mode questionnaires involve using multiple methods of questionnaire administration. For example, a researcher may combine self-administered questionnaires with in-person interviews or use a combination of online and paper-based questionnaires. This approach allows for reaching a broader range of respondents and accommodating their preferences or accessibility limitations.

Each questionnaire method has its own advantages and considerations, such as cost, response rate, data quality, and suitability for the research objectives and target population. The appropriate method should be carefully selected based on the nature of the study and the specific requirements of data collection. In our case, a mixed-mode questionnaire method is selected to combine the online self-administered and paper-based methods. In this manner, a larger distribution sample is achieved to address different types of respondents who cannot access the questionnaire via the internet. However, the distribution of questionnaires requires participants recruitment using probabilistic and non-probabilistic modes. The questionnaire is said to be probabilistic if the respondents are targeted through a group of interest. On the other hand, the respondents of non-probabilistic questionnaires are considered from an unknown sample. The later mode cannot be controlled to target a certain group which leads to the selection of the probabilistic mode. The online questionnaire is classified as a probabilistic mode that is controlled through sending the questionnaire to a targeted group of respondents or through specialized mailing lists. For the traditional methods, paper-based questionnaire recruits respondents by providing them with face-to-face handing of the questionnaire (Lehdonvirta et al., 2021). In contrast during non-probabilistic mode, the sample of respondents cannot be controlled in online questionnaires. In our case, probabilistic non-strict mode is selected to choose the targeted sample of respondents through online questionnaire and paper-based one.

5.2.2 Elements Identification of Questionnaire

The multiple contexts of a teacher has been discussed through different researches to determine the optimal definition of these contexts (Ben-Peretz, Mendelson, and Kron, 2003; Lindblom-Ylänne et al., 2006). Researchers differ a unified model for such problem, however various aspects have been agreed on. These aspects encompass four key dimensions:

- i Profile: The personal information and preferences are the main aspects that shape this dimension.
- ii Social/Living environment: This dimension pertains to the surrounding environment of the living place of a teacher and its impact.

- iii Educational institution/Working environment: This dimension relates to the workplace and cognitive workload of a teacher. This dimension is further divided into three sub-dimensions, namely workplace, professional development, and the targeted students. Some of the features in these sub-dimensions are shared with the sentimental dimension,
- iv Sentimental impact: This dimension denotes the sentimental state of a teacher in terms of level of satisfaction of surrounding conditions which forms an explainable understanding of the current sentimental state of this teacher.

Research has highlighted that social and work environments act as crucial factors that identify the major dimensions needed to define contexts of a teacher (Ben-Peretz, Mendelson, and Kron, 2003). However, the importance of sentimental state has arisen during the past few years due to the pandemic effect on the educational work environment (Peimani and Kamalipour, 2021). Therefore, this questionnaire considers the three main factors that describe the context of a teacher.

In our questionnaire, we incorporate other factors over sections such as "personal and professional development", and "students' advancement". This addition is done to determine whether the working environment facilitates skill development through training and scientific events, and if teachers have the necessary resources from colleagues and superiors to perform their job effectively, as in section "personal and professional development". Additionally, students are addressed in the section titled "students' advancement", which aims to assess the workload of teachers.

5.2.3 Teachers Questionnaire Trends

In order to properly measure job conditions for teachers, three well-known models are used to find the best representation of all factors previously mentioned. The three questionnaires are Teacher as Social Context Questionnaire (TASCQ) (Belmont et al., 1988; Belmont et al., 1992), TIMSS and PIRLS Context Questionnaire Framework (Mullis et al., 2020; Mullis and Martin, 2019), and OECD's Teaching and Learning International Survey (TALIS) (Ainley and Carstens, 2018).

Teacher as Social Context Questionnaire (TASCQ) (Belmont et al., 1988; Belmont et al., 1992) is a well-known research approach used to evaluate the social environment established in the educational institution by teachers (Belmont et al., 1988; Belmont et al., 1992). The questionnaire is designed to collect quantitative and qualitative information about the teacher's behavior, instructional practices, and institution environment. TASCQ includes multiple dimensions or measures that assess several components of the teacher's social environment, such as teacher support, institution organization, good sentimental state, and student participation. Multiple items or statements are used to measure these aspects, to which teachers reply based on their observations and experiences. The utilization of TASCQ can provide useful insights on the mutual affection between teachers and students in terms of social and emotional experiences. The questionnaire provides a scientific method for evaluating the quality of teacher-student interactions, student engagement, and the overall institution environment.

Trends in International Mathematics and Science Study (TIMSS)¹ and Progress in International Reading Literacy Study (PIRLS)² (Mullis et al., 2020; Mullis and Martin, 2019) are two context questionnaire frameworks that are released every 4 years and 5 years respectively. They allow researchers to gather contextual information about students, teachers, institutions, and educational systems. These questionnaires are carried out on a global scale to examine educational progress and the factors that influence it in mathematics, science, and reading literacy. Both frameworks comprise questions for a variety of participants, such as students, teachers, school principals, and parents. The teacher questionnaire, in particular, is designed to gather information about the background of the teacher, teaching practices, educational resources, and institution environment. It collects data on teacher qualifications, teaching experience, professional development, teaching strategies, assessment practices, and other elements. The relationship between teacher characteristics, teaching practices, and student outcome can be analyzed through administering the TIMSS and PIRLS frameworks to teacher which provides valuable insights into the factors that contribute to effective teaching and learning.

Teaching and Learning International Survey (TALIS)³ (Ainley and Carstens, 2018) is carried out by the organization for economic co-operation and development (OECD) with its latest published version of 2018. Its primary goal is to collect data and insights regarding the working environment of teachers, their practices, and professional development opportunities in participating countries. A teacher questionnaire is included in the TALIS survey, which covers several areas of the teaching profession. It gathers data on teacher demographics, education, training, workload, job satisfaction, teacher cooperation, teaching practices, and support from institution heads and the larger institutional community. This questionnaire intends to offer insightful information and a global understanding of the teaching profession and its challenges.

5.2.4 Questionnaire Adaptation

Field studies are subject to different regulations and laws according to the country in which these studies take place. In the context of this questionnaire, the French and Egyptian laws have to be respected and followed. For the French context, researchers are allowed to conduct questionnaires unless they create a computerized data file with personal information directly related to a certain citizen. If the later is the case, researchers are obligated to report it this data file to France's national data protection agency (CNIL)⁴ according to the new legal framework of General Data Protection Regulation (GDPR) (Voigt and Bussche, 2017). Moreover, it is not allowed to ask directly and indirectly about the race of ethnic of a person which is regulated by Chapter II, Section 2 of no. 2004-801⁵.

¹<https://timssandpirls.bc.edu/timss-landing.html>

²<https://timssandpirls.bc.edu/pirls-landing.html>

³<https://www.oecd.org/education/talis/index.htm>

⁴<https://www.cnil.fr/en/media>

⁵Article 8: "Loi n° 2004-801 du 6 août 2004 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel et modifiant la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés."

As for the Egyptian context, any activity regarding collection of data and questionnaires, online or interviewed based ones, are subject to prior approval from the Central Agency for Public Mobilization And Statistics (CAPMAS)⁶ according to the tenth article of the Republican Decree no. 1964-2915. In order to get the required approval, a request must be applied to CAPMAS with a set of obligatory documents including:

- An official letter from the institutional research organization which is responsible for the applicant's research project. This letter should directly address the director of security department of CAPMAS and includes an extensive detailed explanation of the nature of this questionnaire and the sample of respondents that will be included in the questionnaire (number and regions), and a list of participants.
- Photocopies of national identification for all data collection personals.
- A translated copy of the questionnaire, if the questionnaire is not originally in Arabic.

However, this process can be time consuming and even if you get the approval, conducting the questionnaire can be a challenging task due to the cultural differences between the different regions in Egypt. Challenges are not limited to cultural differences but also financial overhead. Therefore, this task is said to be time and money consuming despite its effectiveness. But first, an adaptation of the questionnaire to assure its appropriateness to both contexts, French and Egyptian. In order to adapt the three previously mentioned questionnaires to our context, main sections of questions are identified to cover the multiple contextual dimensions of teachers. Six sections are selected as follows:

- i Teacher profile: It meets the required information for the profile dimension of teacher and it exists in the three well-known questionnaires.
- ii Living environment: This section shares the same features of the dimension of social/living environment. The content of this section cannot be found directly in the other questionnaires as these questionnaires are targeting a predefined location and audience.
- iii Educational institution / Working environment: This section belongs to the dimension with the same name but other important features can be found in the next section to identify the development of teachers. It exists in the other questionnaires, but it is either a direct section or in the form of distributed questions over multiple sections.
- iv Personal and professional development: This section is complementary section for the working environment dimension. We can find this section in the other questionnaires as a standalone section.
- v Students' advancement: In order to form an understanding of the sentimental state of teachers, the advancement of students are included which impacts the sentimental state of teachers in the context of institution and work environment. It exists in the three questionnaires within the "teaching in class" sections.

⁶<https://www.capmas.gov.eg/>

vi Sentimental impact: This section consists of direct questions about the satisfaction of teacher about multiple factors which indicate the sentimental impact of the surroundings over the state of teachers. The factors address the sentimental strategies by the teacher such as social support, self-control, acceptance, and personal evolution (Carton and Fruchart, 2014). This section satisfies the dimension of sentimental state of teachers; however, it is neglected in the other questionnaires.

The protocol of adapting questions to both French and Egyptian contexts is divided into three phases (Epstein, Santo, and Guillemin, 2015):

1. Selecting questions that are suitable to the predefined semantic model of teachers, TSCCO.
2. Summarizing questions to consume a time interval of 7 to 10 minutes for each respondent.
3. adding, neglecting, or adapting questions to suit both contexts.

Four researchers realize this protocol across different phases: 1 professor of computer science from France, 1 associate professor of computer science from France but lives in Egypt, 1 associate professor of humanities from Egypt, and a PhD student in computer science from Egypt but lives in France.

The first phase includes the selection of suitable questions according to TSCCO semantic model and was performed by the PhD student only. It results in the selection of 57 questions across all three questionnaires and other questions regarding the sentimental state. This large number of questions needs to be carefully refined to achieve an accurate description of context and removing questions that do not directly related to TSCCO representation. The filtering of questions results in 47 questions only at the end of this phase.

The second phase of this process was conducted through multiple meetings between the associate professor who lives in Egypt, the associate professor of humanities and the PhD student. The main purpose of these meetings is to adapt the questions and their types to respect the time interval of 7-10 minutes to answer the questionnaire. The associate professor of humanities was the main contributor of this phase by summarizing multiple questions into one and eliminating others. At the end of this phase, the questionnaire reached 27 questions without neglecting important factors. The selection of questions was done through a voting process with an upper hand for the associate professor of humanities regarding the validity of some questions. With the help of the associate professor of humanities, we perform a standardization of types of questions by minimizing the free text input in the responses for easier analysis of collected responses. Additionally, various questions that share the same type, are concatenated into one question such as concatenating radio list and scale-based questions into array questions. Such approaches offer a comprehensive questionnaire for the respondents.

The third phase is divided into two stages: adapting the selected questions to the Egyptian context and similarly to the French context. The stage, involves the Egyptian part, was handled by the same group of the last phase, and the group discussed the appropriateness of each question to match the Egyptian culture.

This process results in some modifications to the prologue of questions to accomplish an absolute understanding of questions from the respondents in English language. The second stage was performed by the professor from France, the associate professor who lives in Egypt, and the PhD student. The group discussed the perspective of French context and adapted the French version of the questionnaire. The resulting questionnaire is represented through two languages: English and French. Both these versions are the exact same except for the language and the paraphrasing of some questions to be convenient to the respective context. The next section discusses the resulting questionnaire and its relation to the other mentioned questionnaires.

5.2.5 Questionnaire Realization

The purpose of this section is to compare our questionnaire teacher's context questionnaire, with the other three questionnaires: TALIS 2018, (TIMSS 2019, PIRLS 2021), and TASC. This comparison is essential to explain the origin of each question and to gain insights into the similarities and differences in the structure of these questionnaires from the perspective of our questionnaire. Moreover, the relevance of questions is explained according to the respective concept of the TSCCO ontology. As previously mentioned, the three-phase protocol results in a questionnaire of 27 questions distributed across 6 sections related to the different concepts of TSCCO ontology.

In order to achieve a clear understanding of the questionnaires, we need to perform a comparative analysis of all questionnaires. Table 5.1 and Table 5.2 provide an extensive comparison of the questionnaires in terms of questions, their sections and their types. Table 5.1 highlights the similarities and differences across the sections of the questionnaires, enabling a thorough examination of their design and content. On the other hand, Table 5.2 offers a comparison of the questionnaire with respect to different types of questions. Using these two tables, similarities as well as differences are discussed equally to identify the strengths and limitations of our questionnaire in relation to our aim.

Despite the diversity in the targeted respondents and research objectives, several common sections are evident across the questionnaires. For instance, all questionnaires include demographic questions such as gender and age group, which allow for an analysis of the relationship between these factors and context of teachers. Additionally, there is a shared emphasis on work environment, professional development, and the impact of the learning environment on student achievement. These shared areas of interest provide opportunities for comparative analysis and the potential for cross-referencing data across studies. Furthermore, a notable similarity across the questionnaires is the utilization of multiple-choice list and scale-based array questions to gather quantitative data. This similarity facilitates data comparison and statistical analysis, promoting consistency in measurement and increasing the reliability of findings. The use of standardized question types allows for greater integrity between different data collection and processing of data and enhances the generalizability of research findings.

While there is notable dependency of our questionnaire on the others, the questionnaires also exhibit differences in terms of sections, question types, and areas of focus. Teacher's context questionnaire is restricted to the concepts of

TSSCO ontology, such as Teacher, Living environment, Educational Institution, Working environment, Learner, Classroom, and Sentimental state. TALIS 2018 places a significant emphasis on teachers' current work and professional development activities and their impact on instructional practices. This questionnaire includes specific questions related to the types and duration of professional development programs attended by teachers as well as work conditions and teaching practices. In contrast, the teacher versions of TIMSS 2019 and PIRLS 2021 focus solely on teaching generally and teaching a certain class specifically in the fields of mathematics and science, and literature, respectively. These questionnaires include subject-specific questions assessing student knowledge and skills in these domains from the perspective of teachers. The inclusion of subject-specific questions allows for a deeper exploration of the relationship between teacher context and student achievement in specific subject areas but that is not the case with our questionnaire.

Another noteworthy difference is found in TASSCO, which explores the sentimental impact of teaching. This questionnaire includes questions related to teachers' emotional experiences, job satisfaction, but it neglects the overall well-being of teacher. This emphasis on emotional factors provides a unique perspective on the teaching profession, highlighting the potential impact of teachers' emotional state on instructional practices and student outcomes. By considering the emotional aspects of teaching, TASSCO adds a valuable dimension to the study of teacher context and complements the more traditional focus of the other questionnaires. Therefore, the section "Students' Advancement" of our questionnaire is mainly based on the questions of this questionnaire due to its precise coverage of most sentimental aspects related to teaching practices. The main differences between our questionnaire and the other ones, can be summarized as follows:

1. Our questionnaire collects various questions related to the profile of teacher into one section, while other questionnaires distribute these questions across different sections.
2. Our questionnaire concatenates similar scale-based or radio list questions into collective questions as in the sections "Teacher Profile", "Professional Development", "Students' Advancement" and "Sentimental Impact". On the other hand, the other questionnaires follow a separation approach in such questions except for TASSCO questionnaire.
3. Teaching and learning styles are included in our questionnaire due to their importance and effective impact on the performance of teachers (Peacock, 2001; Awla, 2014; Dinçol et al., 2011). No mention of such styles is found in the other questionnaires.
4. The sentimental impact is considered in our questionnaire due to the recent effect of such aspect in the job of teachers (Dinçol et al., 2011). Other questionnaires neglect such important aspect.

The comparison of these questionnaires holds great relevance to this research on teacher context. By examining the similarities and differences, relevant constructs and measurement scales can be easily identified to reach the best practice of integrating this questionnaire to the TSSCO ontology. The inclusion of

questions related to the availability of teaching and learning spaces in multiple questionnaires, for instance, underscores the importance of this aspect in understanding the working environment's impact on teacher performance and context. Furthermore, the variations in focus across the questionnaires allow for a more comprehensive exploration of the multifaceted nature of teacher context with respect to the concepts of TSCCO. Additionally, the differences in question types and areas of focus provide opportunities for conducting nuanced analyses and addressing specific research questions. For instance, the subject-specific questions in TIMSS 2019 and PIRLS 2021 enable a detailed examination of the relationship between teacher context and student achievement in specific subject areas, allowing for a deeper understanding of teaching practices and curriculum effectiveness. Similarly, the inclusion of emotional factors in TASCOCO allows for an exploration of the impact of teachers' emotional well-being on teaching quality and student engagement. These variations in focus and question types towards students are not relevant to this research work and hence, they are neglected in favor of the context of teacher.

The comparative analysis of the Teacher's Context Questionnaire with TALIS 2018, TIMSS 2019, PIRLS 2021, and TASCOCO provides valuable insights into the similarities and differences in the design, content, and focus areas of these questionnaires. Understanding these similarities and differences enhances the scientific rigor and validity of the adaptation of these questionnaires to the context of this research. Furthermore, the comparison contributes to a comprehensive understanding of teacher context by exploring diverse dimensions from the different perspectives of these questionnaires. This knowledge enables us to make informed decisions regarding the selection and utilization of questions that best aligns with the research aims, thereby enhancing the validity and reliability of the integration of the collected data to TSCCO ontological model.

Table 5.1: The sections of questions across all questionnaires

Question	Section				
	Teacher Context Questionnaire	TALIS 2018	TIMSS 2019	PIRLS 2021	TASCO
What is your gender?	Teacher Profile	General questions	About you	About you	×
What is your age group?	Teacher Profile	General questions	About you	About you	×
What is the highest level of formal education that you have completed?	Teacher Profile	General questions	About you	About you	×
What are your language preferences?	Teacher Profile	×	×	×	×
What is your country of origin?	Teacher Profile	×	×	×	×
Living location	Living environment	×	×	×	×
How do you feel about the following factors?	Living environment	×	×	×	×
What education level(s) do you teach?	Educational institution/Working environment	General questions	Questionnaire description	Questionnaire description	Questionnaire description
Institution name	Educational institution/Working environment	General questions	Questionnaire description	Questionnaire description	Questionnaire description
Institution location	Educational institution/Working environment	General questions	Questionnaire description	Questionnaire description	Questionnaire description
What type of transportation do you use to move between your living and working places?	Educational institution/Working environment	×	×	×	×
How long do you usually take to move between your living and working places?	Educational institution/Working environment	×	×	×	×

Continuation of Table 5.1

Question	Teacher Context Questionnaire	TALIS 2018	TIMSS 2019	PIRLS 2021	TASCO
Where do you categorize your educational institution's system?	Educational institution/Working environment	General questions	Questionnaire description	Questionnaire description	Questionnaire description
In your institution, which of the following spaces are available for teaching and learning?	Educational institution/Working environment	Teaching	About teaching a class	About teaching a class	×
Which equipment is available in (any) classrooms/halls?	Educational institution/Working environment	Teaching	About teaching a class	About teaching a class	×
How long have you been working as a teacher?	Personal and professional development	Current work	About you	About you	×
Into what category does your area of expertise fall?	Personal and professional development	Current work	About you	About you	×
In a typical working week, estimate the number of hours you spend on the following: Teaching, Planning, Administrative	Personal and professional development	Current work	Professional development + About teaching mathematics and science	About teaching literature	×
During the last 12 months, did you participate in any of the following activities?	Personal and professional development	Professional development	Professional development	About you	×
Was the teaching of your courses part of your academic training?	Personal and professional development	Background and qualifications	About you	About you	×

Continuation of Table 5.1

Question	Teacher Context Questionnaire	TALIS 2018	TIMSS 2019	PIRLS 2021	TASCO
How would you characterize your teaching style? Authority, Demonstrator, Facilitator, Delegator, Hybrid	Personal and professional development	×	×	×	×
On average throughout the year, how many students are in your courses?	Students' advancement	Teaching in class	Teaching in a class	Teaching in a class	×
In your opinion, what are the preferred students' learning styles which suits your teaching style? Visual, Aural, Verbal, Physical, Logical, Social, Solitary	Students' advancement	×	×	×	×
How strongly do you agree or disagree with the following statements about your courses?	Students' advancement	Teaching in class	×	×	All questionnaire
Please select the most appropriate answer options	Sentimental impact	×	×	×	×
Please select Yes or No for each of the following questions.	Sentimental impact	×	×	×	×

Table 5.2: The types of questions across all questionnaires

Question	Type				
	Teacher Context Questionnaire	TALIS 2018	TIMSS 2019	PIRLS 2021	TASCO
What is your gender?	Gender list	Gender list	Gender list	Gender list	×
What is your age group?	Age group (Radio list)	Numerical text input	Age group list	Age group list	×
What is the highest level of formal education that you have completed?	4 educational levels (Radio list)	UNISCO ISCCD 2011 levels list	UNISCO ISCCD 2011 levels list	UNISCO ISCCD 2011 levels list	×
What are your language preferences?	5 languages array (5-point choice array)	×	×	×	×
What is your country of origin?	Countries list (Dropdown list)	×	×	×	×
Living location	Longitude and latitude (Map selector)	×	×	×	×
How do you feel about the following factors?	living factors array (5-point choice array)	×	×	×	×
What education level(s) do you teach?	Multiple choice list	Multiple choice questions	Directed question-naire	Directed question-naire	Directed question-naire
Institution name	Short free text input	Directed question-naire	Directed question-naire	Directed question-naire	Directed question-naire
Institution location	Longitude and latitude (Map selector)	Directed question-naire	Directed question-naire	Directed question-naire	Directed question-naire
What type of transportation do you use to move between your living and working places?	Means of transport (Multiple choice list)	×	×	×	×

Continuation of Table 5.2

Question	Teacher Context Questionnaire	TALIS 2018	TIMSS 2019	PIRLS 2021	TASCO
How long do you usually take to move between your living and working places?	Time duration in minutes/hours (Radio list)	×	×	×	×
Where do you categorize your educational institution's system?	Summarized 4 types (Radio list)	Directed questionnaire	Directed questionnaire	Directed questionnaire	Directed questionnaire
In your institution, which of the following spaces are available for teaching and learning?	List of spaces (Multiple choice list)	Availability (Radio List)	Mathematics and science spaces (Multiple choice list)	Library spaces (Multiple choice list)	×
Which equipment is available in (any) classrooms/halls?	List of equipment (Multiple choice list)	Availability (Radio List)	Mathematics and science equipment (Multiple choice list)	Library equipment (Multiple choice list)	×
How long have you been working as a teacher?	Time duration in years (Radio list)	Numerical text input	Numerical text input	Numerical text input	×
Into what category does your area of expertise fall?	Field of science of ModSci ontology (Dropdown list)	List of categories (Radio list)	List of categories related to mathematics and science (Radio list)	List of categories related to literature (Radio list)	×
In a typical working week, estimate the number of hours you spend on the following: Teaching, Planning, Administrative	Time duration in hours (Numerical list array)	Time duration in hours (Numerical text input)	Time duration in hours (Numerical text input)	Time duration in hours (Numerical text input)	×
During the last 12 months, did you participate in any of the following activities?	List of professional activities (Yes/No/Uncertain array)	List of professional activities (Yes/No array)	Number of hours (Radio list)	Categories of activities (Yes/No array)	×

Continuation of Table 5.2

Question	Teacher Context Questionnaire	TALIS 2018	TIMSS 2019	PIRLS 2021	TASCO
Was the teaching of your courses part of your academic training?	Yes/No list	Yes/No list	Different fields of training (Radio list)	Different fields of training (Radio list)	×
How would you characterize your teaching style? Authority, Demonstrator, Facilitator, Delegator, Hybrid	5 teaching styles (Dropdown list)	×	×	×	×
On average throughout the year, how many students are in your courses?	Numerical groups (Radio list)	Numerical text input	Numerical text input	Numerical text input	×
In your opinion, what are the preferred students' learning styles which suits your teaching style? Visual, Aural, Verbal, Physical, Logical, Social, Solitary	7 learning styles (5-point choice array)	×	×	×	×
How strongly do you agree or disagree with the following statements about your courses?	6 statements about behavior of learners (5-point choice array)	Multiple statements (Radio list)	×	×	Multiple statements about the behavior of learners and teachers (Yes/No array)
Please select the most appropriate answer options	10 statements about teacher satisfaction (5-point choice array)	×	×	×	×
Please select Yes or No for each of the following questions.	9 statements about current sentimental state of teacher (Yes/No/Uncertain array)	×	×	×	×

5.3 French Teacher Context Dataset (Multiple Contexts of Teacher)

The French Teacher Context Dataset (FTCD) is a comprehensive dataset that describes the different contexts in which a teacher coexists in France. The french context is one of the two contexts that are discussed before and is selected due to the availability of the open data. The dataset is divided into two parts: the working context and the living context. The working context dataset provides geolocalisation of educational establishments and their classification in prioritized zones such as réseaux d'éducation prioritaire (REP, REP+) (Stéfanou, 2017) in France. The living context dataset, on the other hand, covers multiple aspects such as establishments, salaries, and population. In this section, we will provide an in-depth description of the dataset and its sources. Data was collected from different sources with respect to each context. All sources are classified as governmental data sources with open access licenses or collected from governmental sources with public domain dedication license. All data is in comma-separated values file format (CSV) which facilitates the reusability of such data in other research-related experiments.

5.3.1 Working Context Data

First, working context of a teacher is described in our case as geolocalization of educational establishments and their classification in a prioritized zone such as réseaux d'éducation prioritaire (REP, REP+) in France. We choose governmental data sources only for this context. These sources are the French ministry of national education and youth (ministère de l'éducation nationale et de la jeunesse)⁷, the French national office of educational and professional information (l'office national d'information sur les enseignements et les professions - ONISEP)⁸, and the French academic and ministerial directory of the educational system (le répertoire académique et ministériel sur les établissements du système éducatif - RAMSESE)⁹. Two datasets describe the geolocalization of educational establishments. These educational establishments are categorized as first and second degrees establishments (écoles, collèges, lycées), and higher educational establishments (universités, écoles supérieures, ...) in both public and private sectors. The prioritized educational networks (REP, REP+) require special educational methodologies. Therefore, it is considered a better approach to point out these establishments in our dataset. Three more datasets are used to indicate establishments that are included in these networks. The latest updates are performed to these datasets as recent as November 2022. The following list includes all data sources for the working context of a teacher. Table 5.3 shows the list of sources that are used to construct this data. On the other hand, Table 5.4 illustrates the different data columns in the same dataset.

The integration of these datasets is realized over a two-step process. First, the datasets concerning geolocalization of educational establishments are merged

⁷<https://www.education.gouv.fr/annuaire/>

⁸<https://www.onisep.fr/>

⁹<https://www.data.gouv.fr/fr/datasets/annuaire-de-leducation/>

Table 5.3: Sources of working context datasets

Type of data	Dataset name	Format
Geolocalization of educational establishments	Adresse et géolocalisation des établissements d'enseignement du premier et second degrés (source RAMSESE)	CSV
	Établissements d'enseignements supérieurs (source ONISEP)	CSV
Prioritized educational establishments (REP, REP+)	Écoles éducation prioritaire (source Education Nationale)	CSV
	Collèges éducation prioritaire (source Education Nationale)	CSV
	Académies éducation prioritaire (source Education Nationale)	CSV

into the first resultant dataset (Result 1). Then, the same merging process is performed over the datasets concerning the prioritized educational establishments (REP, REP+) to result in the second resultant dataset (Result 2). Finally, the two resultant datasets are concatenated, and similar columns are unified.

Merging of "Adresse et géolocalisation des établissements d'enseignement du premier et second degrés" dataset (Dataset 1.1) and "Établissements d'enseignements supérieurs" dataset (Dataset 1.2) results in the creation of the first part of a new dataset, as shown in Figure 5.2. This new dataset contains information about educational institutions at various levels of education, ranging from primary and secondary schools to institutions of higher education. The new partial dataset includes several columns that provide valuable information about each institution. The unique identifier for each institution is represented by the "Code" column, which serves as the primary key to merge the two datasets. The official name of each institution is included in the "Nom" column, while the complete address, including the street name and number, is provided in the "Adresse" column. The postal code and name of the town or city where the institution is located are represented by the "Code postal" and "Commune" columns, respectively. Additional information about the location of each institution is available in Dataset 1.1, which includes columns for latitude and longitude coordinates, as well as the name of the department and region where the institution is located. This information can be used to map the location of each institution. On the other hand, the XY coordinates from Dataset 1.2 are converted to the respective latitude and longitude using the WGS84 coordinate system. Finally, the "Type" column, which is only available in Dataset 1.2, provides more specific information about the level of education provided by each institution. This column distinguishes between primary and secondary schools and institutions of higher education, allowing for more detailed analysis of the educational landscape. Overall, the merging of these two datasets concerning geolocalization of educational establishments provides a comprehensive view of educational institutions at various levels of education and can be used to support research, planning, and analysis in the education sector.

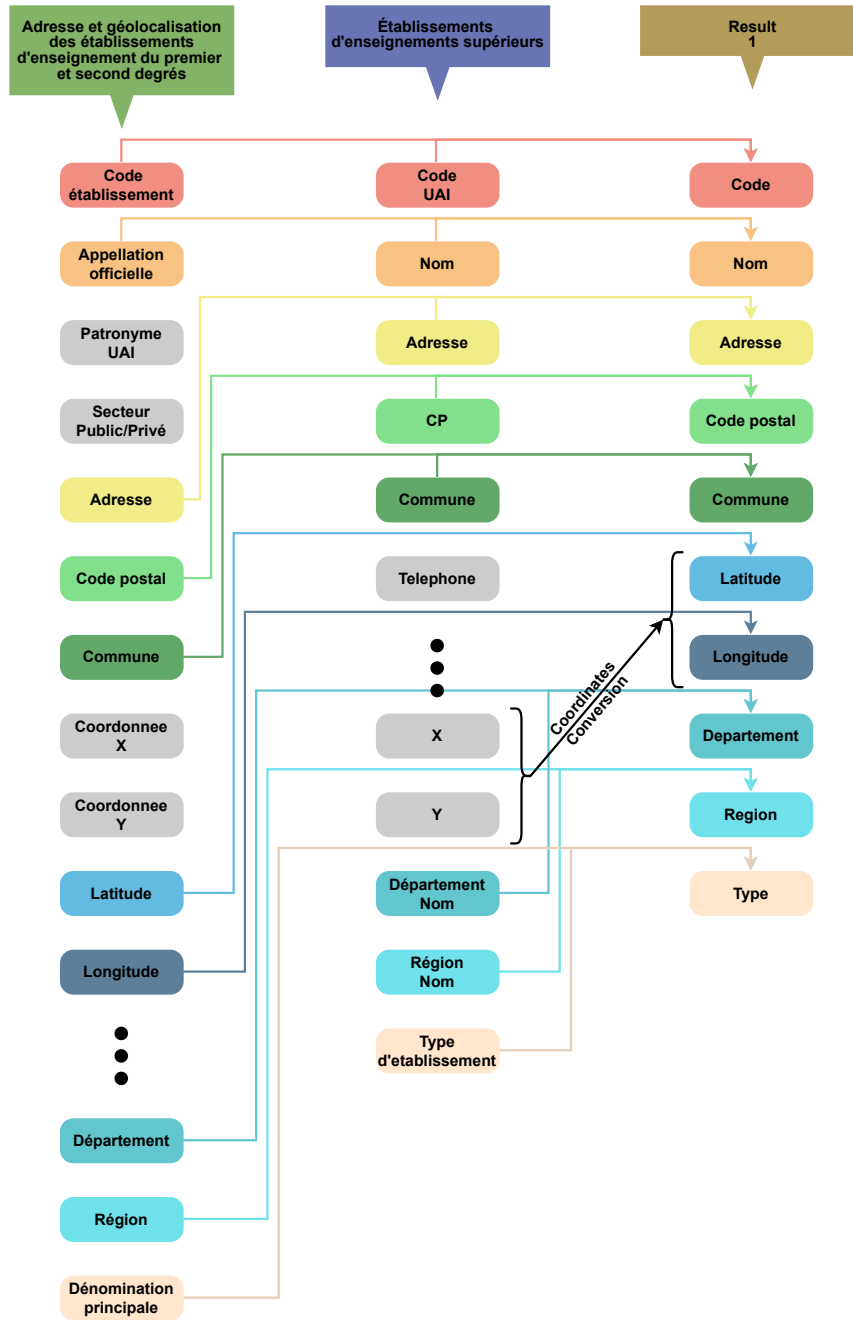


Figure 5.2: Data selection from datasets concerning geolocalization of educational establishments.

The merge of three datasets, "Écoles éducation prioritaire" (Dataset 1.3), "Collèges éducation prioritaire" (Dataset 1.4), and "Académies éducation prioritaire" (Dataset 1.5), into one consolidated dataset presents a rich source of information on prioritized educational institutions and their corresponding details, as shown in Figure 5.3. The resultant dataset provides a comprehensive overview of prioritized educational institutions and sensible ones. Additionally, it includes repetitive data as the previous resultant dataset such as the institution's name, address, postal code, type, and the institution's location, including the latitude and longitude coordinates. Dataset 1.3 represents the primary or elementary schools and includes data related to the academic year, school number, and type

of establishment. It also provides details on the location of the institution, including the coordinates and the commune and arrondissement code and name. Furthermore, the dataset includes information on the institution's specialized nature and its rural or urban location. This dataset also includes information on the number of students and classes, with a breakdown by grade and the presence of pre-elementary education. Dataset 1.4 represents the college or preparatory schools and provides data on the academic year, the academy name, and the institution's affiliation with the education priority system. This dataset also includes information about the institution's location, including its sector code and name, as well as its academy code and region code. Dataset 1.5 includes data about secondary or high schools and provides information on the academic year, institution number, and type of establishment. The dataset includes information on the institution's location, including the coordinates, the commune and arrondissement code and name, the department code and name, and the academy code and name.

The consolidated dataset resulting from the merge of these three datasets provides a complete view of the educational institutions in the priority education system, with a range of information that can be analyzed for research or used for planning purposes. It includes data that can help identify different types of prioritized institutions. All institutions included in these datasets are classified as REP while institutions marked as sensible establishments are considered as part of the REP+ network. Taking that into account, two new columns distinguish the difference between these institutions. At the end, this resultant dataset is concatenated with the first resultant one to form an extended final version that describes the educational institutions across different levels, as shown in Figure 5.4 and Table 5.4.

5.3.2 Living Context Data

Multiple aspects should be considered to describe the living context of a teacher. According to these aspects, data can be categorized as establishments, salaries, and population. The dataset (French employment, salaries, population per town) by Etienne Le Quéré, is collected from several governmental datasets of the French national institute of statistics and economical studies (Institut National de la Statistique et des Etudes Economiques - INSEE)¹⁰. The data sources are implemented from various studies performed by INSEE. First dataset discusses the database of firms in France, and it provides an indicator that classifies the firms into nine categories according to the size of employment for each town. Second introduces an overall indicator of salaries per town as well as an indicator for each gender (female, male) and for each level of employment (executive, middle manager, employee, worker). Third dataset (Population) represents a demographical description of all French communes. Last dataset contains the geographical information for all communes in France.

In order to find a well-represented living context of teachers, five datasets from French governmental open source data, are integrated into the final resultant dataset. The resultant dataset, as shown in Figure 5.5 and Table 5.6, is a comprehensive compilation that brings together information from various sources, allowing for a deeper understanding of regions, departments, municipalities, and

¹⁰<https://www.insee.fr/>

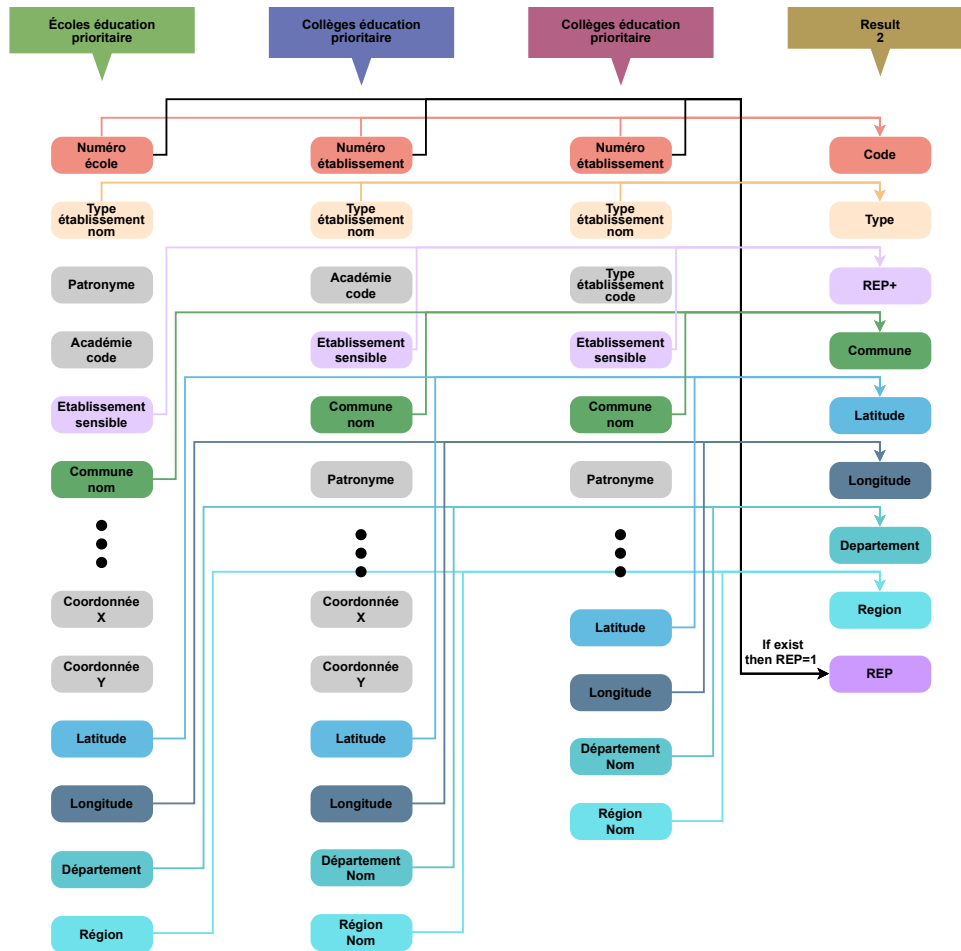


Figure 5.3: Data selection from datasets concerning Prioritized educational establishments (REP, REP+).

their corresponding demographics, economic activity, and average salaries. By combining these datasets, a more comprehensive overview of the regions and their characteristics is obtained.

The dataset includes several key fields that provide valuable insights into different aspects of each municipality. The "Code" field serves as a unique identifier for each entry, ensuring data integrity and facilitating further analysis. The "Nom Region" field specifies the name of the region to which the municipality belongs, while the "Lieu Region" field provides information about the main capital/city of the region. The "Numero Departement" field contains a numerical code assigned to each department, enabling easy identification and categorization of municipalities based on their administrative division. The "Nom Commune" field represents the name of the municipality itself, and the "Code Postal" field includes the postal code associated with each municipality, aiding in postal services and geographical reference. Geographic coordinates are essential for spatial analysis, and therefore, the merged dataset includes the "Latitude" and "Longitude" fields. These fields provide the precise geographic location of each municipality, facilitating mapping and spatial visualization of the data.

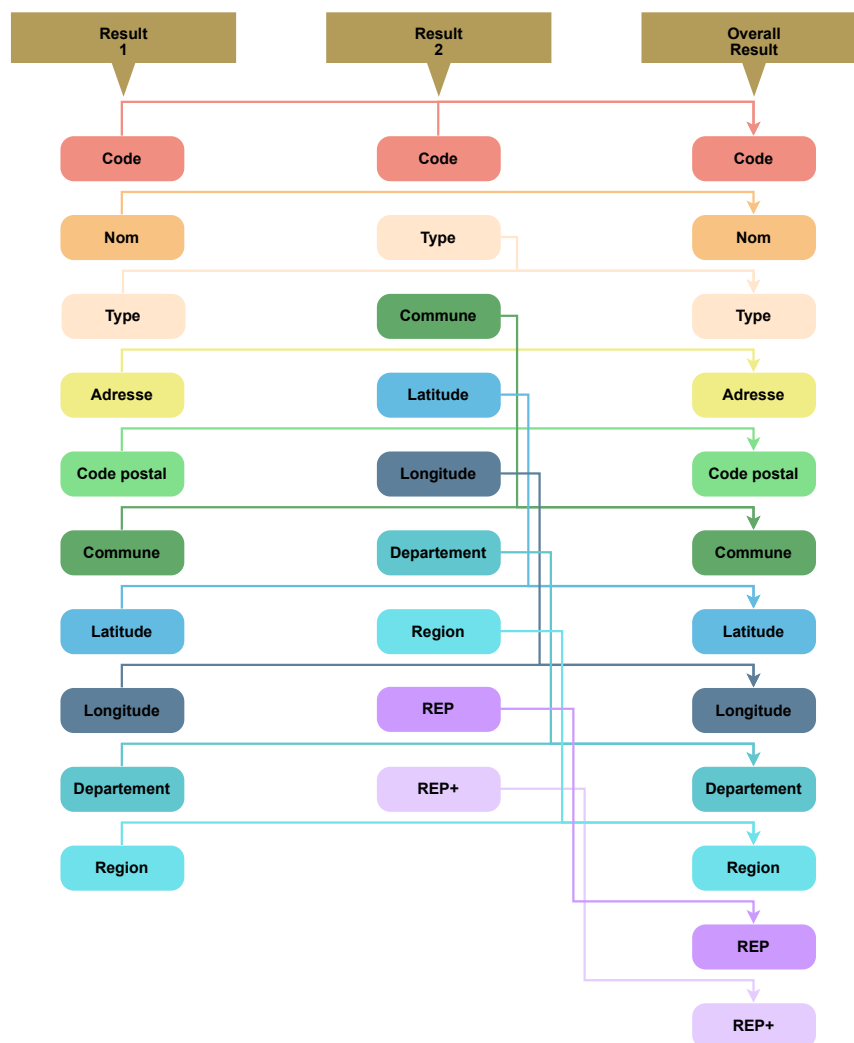


Figure 5.4: Overall data selection from datasets concerning working context of a teacher.

Economic indicators are also incorporated into the merged dataset. The "Entreprise" field represents the total number of active establishments as of December 31, 2020, within each municipality. This information provides insights into the level of economic activity and business presence in different areas. Salary information is another crucial aspect of the merged dataset. The "Salaire Moyen" field provides the average net hourly salary for the municipality as a whole, giving an overview of the economic well-being of the workforce. Additionally, the dataset includes specific average salary values for different demographic groups, such as women and men, as indicated by the "Salaire Moyen Femme" and "Salaire Moyen Homme" fields. This information enables the analysis of gender-based wage disparities within municipalities. Demographic data is an essential component of the merged dataset. The "Population" field contains the total population of each municipality in 2019. But we choose to neglect additional demographic breakdowns, such as age groups and gender, allowing for a detailed understanding of the population composition within each municipality. This important information can be easily included in the future to extend the research scope and the examined features for the demographical data of teachers.

Table 5.4: Columns of working context datasets

Column	Datatype	Description
Index	Integer	a unique index for each data row
Code	String	educational establishment code given by the French ministry of national education and youth
Nom	String	educational establishment name
Adresse	String	address of educational establishment
Code Postal	String	postal code of each educational establishment
Commune	String	name of the town where the educational establishment is located
Latitude	Float	latitude coordination of the educational establishment (WGS84 coordinate system)
Longitude	Float	longitude coordination of the educational establishment(WGS84 coordinate system)
Type	String	type of educational establishment (école, college, lycée, université, ...)
Departement	String	name of the department where the educational establishment is located
Region	String	name of the region where the educational establishment is located
REP	Boolean	1 -> if the educational establishment belongs to a prioritized education network, 0 -> if not
REP+	Boolean	1 -> if the educational establishment belongs to a high prioritized education network, 0 -> if not

The merged dataset provides a comprehensive resource for analyzing and comparing economic and demographic factors across regions and municipalities. This dataset acts as an explicit representation of the living and working environments of teachers.

5.4 Pedagogical Resources Dataset

The integration of diverse educational datasets plays a crucial role in enriching the different resources data acquisition. In this context, we aim to collect and integrate data from several datasets related to online courses, textbooks, and pedagogical resources from various sources. The integration process involves merging and organizing data from different platforms to create a comprehensive and unified database. The combined dataset has a wide range of pedagogical resources, including MOOC courses, textbooks, and other resources, all accessible through our ontological model.

The resource data retrieved for integration are divided into two types: repository data retrieval and Application Programming Interface (API) data retrieval.

Table 5.5: Sources of living context datasets

Type of data	Dataset name	Format
French employment, salaries, population per town (source Etienne Le Quéré)	Nombre d'établissements et de postes salariés par secteur d'activité détaillé et tranche d'effectifs salariés détaillés (source INSEE)	CSV
	Salaire net horaire moyen (source INSEE)	CSV
	Population en 2019 (source INSEE)	CSV
	Informations géographique (source INSEE and Code Officiel Géographique (COG))	CSV
Geolocalization of communes	Base officielle des codes postaux (Source: La Poste)	CSV

The repository data refers to existing relational datasets that previously collected and static data. This type includes data from Coursera¹¹, edX¹², Udemy¹³, Udemy IT¹⁴, Goodreads¹⁵, and Wikibooks¹⁶. Each dataset contains unique information about courses, books, authors, ratings, durations, languages, and other relevant attributes. API data retrieval indicates the process of collecting data through API server requests which ensures the retrieved data is up to date. This process is applied to different sources that offer free or open-source APIs such as IEEE¹⁷, Springer¹⁸, and Scratch¹⁹. By merging these data sources, we can create a more extensive collection of pedagogical resources, offering teachers a diverse and comprehensive selection to choose from.

To accomplish this integration, we employ various techniques and methods tailored to each dataset's structure and schema. The data is collected, cleaned, and standardized to ensure consistency and accuracy across all records. One of the challenges encountered during the integration process is the entity resolution problem, particularly when dealing with multiple records referring to the same course or book from different platforms. To address this challenge, similarity-based techniques are employed to identify and combine similar records. Weighted similarity measures, such as cosine similarity, are utilized to compare attributes like course titles, authors, lengths, and book titles, among others. By applying a predefined threshold, matching records are identified and merged, eliminating duplicates, and ensuring a unified representation of courses and books.

¹¹<https://www.coursera.com/>

¹²<https://www.edx.org/>

¹³<https://www.udemy.com/>

¹⁴<https://www.udemy.com/courses/it-and-software/>

¹⁵<https://www.goodreads.com/>

¹⁶<https://www.wikibooks.org/>

¹⁷<https://developer.ieee.org/>

¹⁸<https://dev.springernature.com/>

¹⁹https://en.scratch-wiki.info/wiki/Scratch_API

Table 5.6: Columns of living context datasets

Column	Datatype	Description
Index	Integer	a unique index for each data row
Code	Integer	regional code attached to a town
Nom Re- gion	String	given name to the region attached to a town
Lieu Region	String	given name to the administrative center
Numero De- partement	String	number of the department of the town
Nom Com- mune	String	given name of the town
Code Postal	String	postal code of the town
Latitude	Float	latitude coordination of the town (WGS84 coordinate system)
Longitude	Float	longitude coordination of the town (WGS84 coordinate system)
Entreprise	Integer	total number of firms in the town (if available)
Salaire Moyen	Float	average salaries in the town (if available)
Salaire Moyen Femme	Float	average salaries of female workers in the town (if available)
Salaire Moyen Homme	Float	average salaries of male workers in the town (if available)
Population	Integer	total number of people living in the town (if available)

The resulting integrated dataset is mapped onto the TCO (Triple, Class, and Object) representation using D2RQ mapping language. This mapping facilitates data querying, retrieval, and manipulation, enabling seamless integration with educational platforms and applications.

Overall, the integration of these datasets aims to create a comprehensive and cohesive pedagogical resource database. Teachers can benefit from bringing together a diverse range of online courses, textbooks, and other resources. This integration allows new possibilities for personalized learning, enhanced discovery of relevant resources, and improved educational experiences for teachers.

5.4.1 Repository Data Retrieval

The integration of datasets from Coursera, edX, Udemy, Udemy IT, Goodreads, and Wikibooks plays a crucial role in providing a comprehensive view of educational courses and books. These datasets were collected using repository data

retrieval methods, ensuring a standardized and systematic approach to data collection followed by the authors of these datasets, as shown in Table 5.7.

The Coursera dataset²⁰ comprises information about courses offered on the platform, including the course index, name, course link, university name, course type, university logo, time required for completion, course language, course subtitles, and course skills. The edX dataset²¹ includes details such as the course index, course name, course URL, course photo, course level, effort required, course length, course language, course subtitle, and course price. Similarly, the Udemy dataset²² provides data on course ID, course title, course URL, course payment status, course level, number of subscribers, content duration, number of lectures, number of reviews, and course price. To enhance the dataset's breadth, the Udemy IT dataset²³ offers additional information on course ratings, number of subscribers, number of reviews, wishlisted status, average rating, and recent average rating.

The Goodreads dataset²⁴ provides valuable insights into books, including unique book identifiers, titles, ISBNs, ISBN13, authors (with multiple authors delimited by "-"), average ratings, number of pages, language codes, total ratings count, and text reviews count. Lastly, the Wikibooks dataset²⁵ contributes book-related information, such as the book index, title, URL, abstract, body content, body HTML content, and language.

To integrate these datasets effectively, a data integration process is followed. This process involves data cleaning, and entity resolution. First, data cleaning eliminates duplicates, irrelevant records, and inconsistencies. Next, the entity resolution problem arises because of the existence of different records from the different platforms referring to the same course or book. Therefore, a similarity-based technique is required to combine similar records. For example, given two courses c_1 and c_2 from two different datasets with the following common properties:

$$(\text{course_title}, \text{course_url}, \text{course_level}, \text{course_length}) \quad (5.1)$$

²⁰<https://www.kaggle.com/datasets/santoshapatil31/coursera-courses-all5164-courses-in-all-languages>, version 2.0

²¹<https://www.kaggle.com/datasets/santoshapatil31/edx-all-courses-3082-courses>, version 1.0

²²<https://www.kaggle.com/datasets/andrewmvd/udemy-courses>, version 3.0

²³<https://www.kaggle.com/datasets/jilkothari/it-software-courses-udemy-22k-courses>, version 1.0

²⁴<https://www.kaggle.com/datasets/jealousleopard/goodreadsbooks>, version 1.0

²⁵<https://www.kaggle.com/datasets/dhruvildave/wikibooks-dataset>, version 1.0

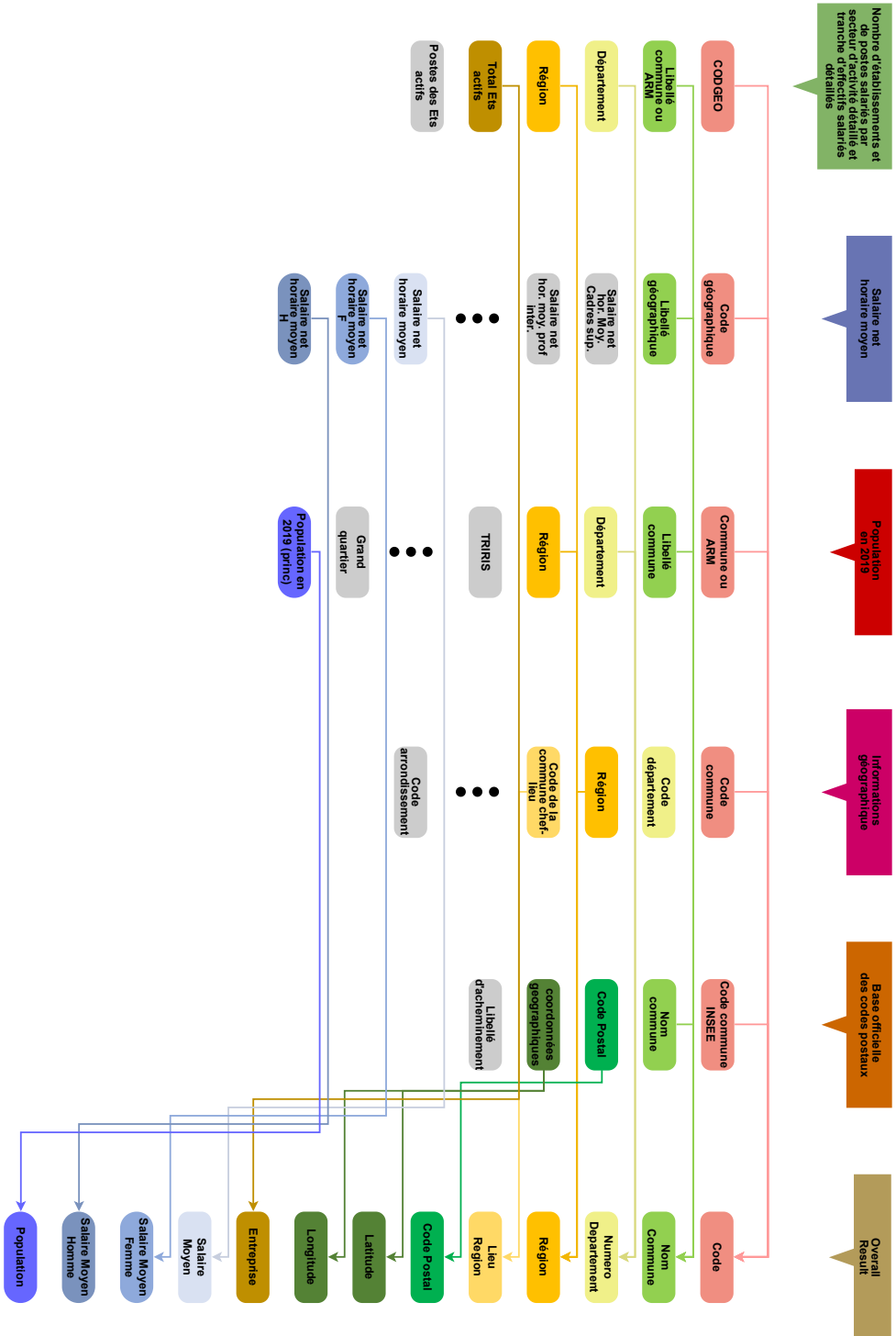


Figure 5.5: Overall data selection from datasets concerning living context of a teacher.

Table 5.7: Columns of resources repository datasets with description as provided by authors

Courseera	edX	Udemy	Udemy IT	Goodreads	Wikibooks
Index (Index of the course)	Index (Index Number)	course_id (Course ID)	id (Shows the ID of a course)	bookID (A unique Identification number for each book)	Index (Unique book index)
Course_name (Name of the course; Name of the courses is not unique to each course)	Course_name (Name of the course)	course_title (Course Title)	title (Title of the course.)	title (The name under which the book was published)	Title (Book title)
Course_link (Courses URL)	Course_url (Course link)	url (Course URL)	url (URL of course.)	isbn (Another unique number to identify the book, the International Standard Book Number.)	url (Book URL)
University_name (Name of the University or Institute Teaching the course)	Course_photo (Each course on Edx has a photo or cover photo for the course and this is the url of that photo location)	is_paid (Whether the course is free or paid)	is_paid (Returns true if course is paid and false otherwise.)	authors (Names of the authors of the book. Multiple authors are delimited with -.)	Abstract (Book abstract)
Course_type (Type of Course: Guided Project-Short course with a simple hand's on project COURSE: Long course with)	Course_level (Course level like Introductory,Advanced, Intermediate)	level (Course difficulty)	rating (Shows course rating)	average_rating (The average rating of the book received in total.)	Body_text

Continuation of Table 5.7

Coursera	edX	Udemy	Udemy IT	Goodreads	Wikibooks
University_logo (University Logo of the institute)	course_effort (Effort required to com- plete the course like No. of hours to put in a week to com- plete the course)	num_subscribers (Number of sub- scribers)	num_subscribers (Number of sub- scribers)	isbn13 (A 13-digit ISBN to identify the book, instead of the standard 11-digit ISBN.)	Body_html
Time required (Ap- proximate Time re- quired to complete the course)	course_length (Number weeks required to com- plete the course)	content_duration (Duration of all course materials)	num_reviews (Number of reviews a course received.)	num_pages (Num- ber of pages the book contains.)	Language
Course_language (Main course lan- guage and language of the Course main Audio or Course main Language)	course_language (Main Course Lan- guage)	num_lectures (Num- ber of lectures)	is_wishlisted (Shows if course is wishlisted.)	language_code (Helps understand what is the primary language of the book. For instance, eng is standard for English.)	
Course_subtitle (All Available course Subtitles)	course_subtitle (Course Transcript or Subtitle)	num_reviews (Num- ber of reviews)	avg_rating (Average rating of a course)	ratings_count (To- tal number of rat- ings the book re- ceived.)	
Course_skills (All the skills the course will teach once com- pleted)	course_price (Price of the course)	price (Course Price)	avg_rating_recent (Recent Average rating of a course.)	text_reviews_count (Total number of written text reviews the book received.)	

The two courses are declared a match if

$$\begin{aligned} &w_1.f_1(\text{course_title}) \\ &+ w_2.f_2(\text{course_level}) \\ &+ w_3.f_3(\text{course_length}) \geq \tau \end{aligned} \quad (5.2)$$

where f_i is the cosine similarity function of two characters' vectors in Equation 5.3 (Li and Han, 2013), w_i is the weight for each property p_i , and a τ is the threshold.

$$\text{Cosine}(\vec{v}, \vec{w}) = \frac{\vec{v} \bullet \vec{w}}{|\vec{v}| |\vec{w}|} = \frac{\sum_{i=1}^N v_i \times w_i}{\sqrt{\sum_{i=1}^N v_i^2} \sqrt{\sum_{i=1}^N w_i^2}} \quad (5.3)$$

Similarly for the sources of books, the entity resolution problem is solved using the similarity-based technique using the form of properties as follows:

$$(\text{book_title}, \text{book_author}, \text{book_language}) \quad (5.4)$$

and is followed by the cosine similarity function of two characters' vectors in Equation 5.3. This step is followed by the data fusion concatenation.

Afterwards, each dataset is projected to the unified schema as follows:

$$\begin{aligned} &<\text{resource_title}, \text{resource_url}, \text{resource_type}, \\ &\quad \text{resource_external_system}, \text{resource_description}> \end{aligned} \quad (5.5)$$

where the column `resource_type` can take one of two values referred to the source dataset: `course`, and `book`. The column `resource_external_system` refers to the source corresponding to each resource: Coursera, edX, Udemy, Goodreads, and Wikibooks. The last column `resource_description` contains a collection of other columns that are not listed in the unified schema with the following form: `<column_name>:<column_value>`.

At the end, we apply the concatenation function `CONCAT` to perform a vertical concatenation that involves stacking the unified datasets with the same structure on top of one another. The concatenate function for vertical concatenation appends the rows of one dataset below the rows of another one. Once the integration is complete, the integrated dataset can be accessed through the unified schema.

5.4.2 Application Programming Interface (API) Data Retrieval

The dynamic approach of data collection is an integral component that enables the extraction, discovery, and collection of new resources and trends in specific fields (Lomborg and Bechmann, 2014). This approach plays a crucial role in aggregating data from various sources and unifying them into a coherent data model. Our objective is to gather data from different APIs and integrate it with the repository data according to the unified data model.

APIs act as intermediaries between applications, allowing them to communicate and exchange data (Ong et al., 2015). Open APIs or free APIs offer a free

to use services to third-party developers to access proprietary data from different applications (Qiu, 2017). In our case, we selected the three APIs: IEEE, Springer, and Scratch. IEEE (Institute of Electrical and Electronics Engineers) is indeed the world's largest technical professional organization. With a global presence and a membership of professionals from various fields, IEEE is committed to advancing technology for the betterment of humanity. The IEEE Open Access initiative provides researchers with the opportunity to review and access full-text articles that have been designated as Open Access. Open Access articles are scholarly publications that are freely available to the public, allowing for increased dissemination of knowledge and research findings. Springer is a global publisher with about 3,000 journals and 13,000 books in various domains, including science, technology, and medicine. In the academic and professional society domain, it has approximately 500 journals. The Springer Nature Open Access API is a valuable resource for developers seeking access to freely available content for noncommercial use. This API grants access to a vast collection of metadata and full-text content from over 649,000 online documents. Scratch is a block-based visual programming language and educational platform aimed at young learners and beginners. Due to limitations in the Scratch API, data extraction involved identifying the 30 most followed users and retrieving the projects created by them.

The integration process, as described in Algorithm 1, involves merging and combining data obtained from multiple APIs. The algorithm takes two inputs: *API_path*, which represents the API paths for data retrieval, and *thr_similarity*, which denotes the similarity threshold. Equation 5.6 is provided to calculate the threshold *thr_similarity*. The equation uses the average (*avg*) and standard deviation (σ) of the similarity matrix (*similarity_matrix*), along with a coefficient α for adjusting the threshold. This equation is based on the Grouping Genetic Algorithm (GGA) clustering approach (Mutingi and Mbohwa, 2017). The desired output is the final integrated dataset, represented as *integrated_data*. At the beginning, it converts the obtained data from each API into a dataframe format *df*. Then, the description column, *df_description*, is composed from the existing columns of *df* but excluded from the unified schema 5.5. Afterwards, if there are any empty cells, they are filled using the Not a Number (NaN) value. The resulting dataframe *df_description* is fitted and transformed using the Term Frequency - Inverse Document Frequency (TFIDF) vectorizer (Christian, Agus, and Suhartono, 2016) to obtain the TF-IDF matrix, *tfidf_matrix*. The cosine similarity matrix, *similarity_matrix*, is computed based on the *tfidf_matrix* to get the similarity between each data entry. If the similarity value *sim* is greater than or equal to the threshold *thr_similarity*, the corresponding entry is dropped from the dataframe. After processing all the APIs, the algorithm concatenates all the dataframes of *dfs* to obtain the final integrated dataset, *integrated_data*. This integrated dataset is concatenated to the previous dataset from the repository data retrieval.

$$thr_similarity = avg(similarity_matrix) + \alpha \cdot \sigma(similarity_matrix) \quad (5.6)$$

Algorithm 1 Data integration of API data retrieval**Procedure** APIDataIntegration()**Input:** *API_path*: API paths for data retrieval*thr_similarity*: Similarity threshold**Output:** *integrated_data*: Final integrated dataset**begin**initialize(*dfs*)**for all** *path* in *API_path* **do**initialize(*df*)*data_API* \leftarrow GetData(*path*)*df* \leftarrow SaveAsDataFrame(*data_API*)*df_description* \leftarrow ComposeDescriptionColumn(*df*, *columns*)*complete_df* \leftarrow FillMissedColumns(*df_description*, *columns*)*tfidf* \leftarrow TfidfVectorizer(*stop_words*='English')*tfidf_matrix* \leftarrow FitTransform(*tfidf*, *complete_df*)*similarity_matrix* \leftarrow CosineSimilarity(*tfidf_matrix*)**for all** *sim,index* in *similarity_matrix* **do****if** *sim* \geq *thr_similarity* **then**DropEntry(*index*, *complete_df*)**end if****end for***dfs* \leftarrow Append(*complete_df*)**end for***integrated_data* \leftarrow CONCAT(*dfs*)**end**

5.5 Activity and Feedback

Tracking the behavior of teachers as well as collecting their feedback is the focus of this section. We explore the vital role of monitoring interactions and sentimental variability to gain valuable insights that can drive improvements and enhancements. By understanding teacher behavior and sentiment, the recommender system can optimize future recommendations to tailor the experience of each teacher. Two types of behavior activity tracking through MEMORAe platform (Abel, 2022), and sentimental feedback through MoodFlow@doubleYou (Andres et al., 2021). Activity tracking delves into the mechanisms and processes involved in capturing, recording, and categorizing teacher activities within the system. It examines various types of interactions, such as accessing, creating, adding, modifying, deleting, sharing, and navigating through MEMORAe platform. Sentimental feedback focuses on the collection and analysis of teacher mood within the recommendation process. By employing sentiment analysis techniques, the system can evaluate the overall mood expressed by teachers through the MoodFlow@doubleYou approach. This analysis allows recommender systems to understand teacher satisfaction levels regarding the recommended resources.

5.5.1 Activity Tracking

MCC ontology (Li, 2021) integrates the activity concepts from MC2 and SOIS ontologies (Deparis, 2013; Atrash, 2015; Saleh, 2018). According to these ontologies, activities are classified into three categories: `mc2:ProceduralActivity`, `mc2:InteractionActivity`, and `sois:NavigationActivity`. Interaction activities of users on resources, represented by `mc2:InteractionActivity` and later `mcc:InteractionActivity`, are reorganized and updated through actions such as annotating and discussing resources in MCC ontology. Interaction activities are identified by six types: creating, deleting, modifying, accessing, adding, and sharing (Wang, 2016). Additionally, MCC ontology enables the recording of activities outside the Collaborative Working Environment (CWE), represented by `mc2:ProceduralActivity`. Navigational activities, represented by `sois:NavigationActivity`, involve browsing activity of the user within an organization.

In TSCCO ontology, the interaction activities are used to extend the profile of a teacher to track the different types of interactions with resources. The concept `mcc:InteractionActivity` is associated with `mcc:UserAccount` which leads to the profile of teacher, as shown in 5.6. In our data collection phase, two categories of activities are targeted: the interaction activities, and navigation activities. Each activity in MEMORAE platform is represented with different properties such as actor, date, resource name and type, organization, location, and type of action, as shown in Table 5.8. This data is inserted to and retrieved from the online linked data using SPARQL queries, such as the given example in Figure 5.7, which retrieves all activities happen within a certain organization. Figure 5.8 shows an example of an activity representation within an organization called PRE103 that represents an introductory course to computer programming. Through this example, we can notice that the activity is related to three other concepts: `mcc:User` the actor of this activity, `mcc:Resource` the resource which this activity affected, and `mcc:SharingSpace` the sharing group where this activity took place.

5.5.2 Sentimental Feedback

Interaction activity is extended with the sentimental state of the user through mood detection using MoodFlow@doubleYou approach (Nashed et al., 2021). As explained in Section 4.4.1, sentimental state is detected through the representation of mood using the three-level approach (Corson, 2002). Recorded data in this case takes the form of time-series signal recorded from special electrodes and it indicates the level of mood with respect to the calibrated neutral for each person.

Figure 5.9 shows an example of the recorded signal which is then processed through the MoodFlow@doubleYou system to detect the different levels of mood. The calculated average reading is compared against the neutral value predefined during calibration, as shown in Figure 5.10. For example, during "Activity 1", the user experiences a swing in his mood which results in an average recorded value which approximately equals to the neutral level, therefore, the user is said to be in neutral mood level during "Activity 1". Similarly, the user is said to be in a positive mood during "Activity 3" due to the higher average of recorded values.

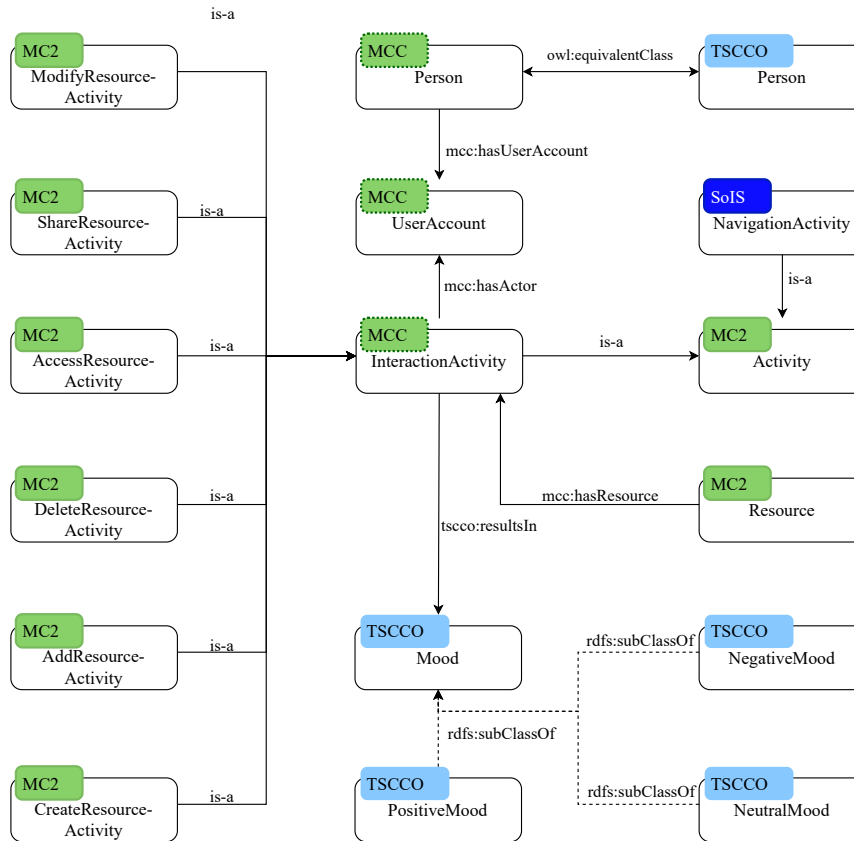


Figure 5.6: Activity and mood representation from TSCCO. (T-Box component)

In order to integrate the mood with the recorded activity, Figure 5.11 shows A-Box assertion component of the "add resource" activity `APIConcept61963d02137-c6` which was performed by the account of Teacher 1. This activity is indexed to a new resource `APIConcept61963d0145d09` and results in `TSCCO:NeutralMood` for the teacher.

5.6 Data Ontological Mapping

TSCCO is developed to facilitate the representation of various educational contexts from a teacher's perspective. TSCCO is designed to be integrated into an educational resources recommender system (Nashed, Lahoud, and Abel, 2021). Figure 5.12 provides a partial representation of selected concepts from TSCCO that are relevant to the integrated datasets. Specifically, the representation focuses on the environment concept, which distinguishes between the living and working environments. The living environment refers to the geographical location of the places of residence of teachers, while the working environment encompasses factors such as educational institution, location, educational level, and type of prioritized area. Within the working environment, teachers interact with resources of different types.

To establish a connection between the integrated datasets and TSCCO, two

Table 5.8: Properties of Activity concept in MCC ontology

Property	Description	Values
element	A-Box assertion component of recorded activity	URI
actor	Corresponding user to activity	URI
date	Date stamp of activity	date
Resource-Name	Title of resource targeted by activity	string
Resource-Type	Type of resource targeted by activity	string "Weblink", "Document", "Vote", "Annotation", "Satisfaction", "Concept"
Resource-System	Original source of resource	string "MEMORAe CWE", "Upload File", "Google Search", "Create Annotation", "HAL", "Youtube Search", "Google Drive", "Google Calendar"
organization	Organization name which a user belongs to	string
location_type	Type of location where activity took place	string "Digital Location", "Geographical Location"
action_type	Type of interaction activity depending on the illustrated in Figure 5.6	string "Access", "Add", "Create", "Update", "Share", "Delete"
SharingSpace	A group where resource exists	URI

fundamental steps are involved in the implementation. First, the relational database is mapped into virtual Resource Description Framework (RDF) graphs. Second, a SPARQL endpoint is set up for accessing the relational data. In this mapping process, the living context table from the French teacher context dataset is mapped to the TSCCO:Environment concept in the RDF graph, while the working context table from the same dataset is mapped to the TSCCO:WorkingEnvironment and TSCCO:EducationalInstitution concepts. The integrated resources dataset is directly mapped to the TSCCO:Resource concept. Finally, data from the questionnaire is mapped to the TSCCO:Teacher concept.

To accomplish the mapping, the Database to RDF Query (D2RQ) Platform, which provides a declarative language for describing the link between relational models and ontologies, is utilized (Bizer and Cyganiak, 2006). It provides a way to

Figure 5.7: Example of SPARQL query to retrieve all activities within an organization

```

prefix tmcc: <http://www.hds.utc.fr/mcc/tbox#>

SELECT ?element (group_concat(distinct ?Actor;separator=',') as ?Actor) (
  group_concat(distinct ?Date;separator=',') as ?Date) (group_concat(
  distinct ?Location;separator=',') as ?Location) (group_concat(
  distinct ?Index;separator=',') as ?Index) (group_concat(distinct ?
  SharingSpace;separator=',') as ?SharingSpace) ?ResourceID ?
  ResourceName ?ResourceSystem ?ResourceType ?Organization ?LocationType
  ?ActionType
FROM <http://www.hds.utc.fr/pre103/abox#>
{
    OPTIONAL {?element tmcc:has_resource_id ?ResourceID}
    OPTIONAL {?element tmcc:has_resourceSystem ?ResourceSystem}
    OPTIONAL {?element tmcc:organization ?Organization}
    OPTIONAL {?element tmcc:has_index_concepts ?Index}
    OPTIONAL {?element tmcc:has_locationID ?Location}
    OPTIONAL {?element tmcc:has_space ?SharingSpace}
    ?element a tmcc:ResHistory;
    tmcc:has_resourceName ?ResourceName;
    tmcc:has_resourceType ?ResourceType;
    tmcc:has_action ?ActionType;
    tmcc:has_locationType ?LocationType;
    tmcc:Date ?Date;
    tmcc:has_actor ?Actor.
}
group by ?element ?ResourceID ?ResourceName ?ResourceSystem ?
  ResourceType ?Organization ?LocationType ?ActionType

```

expose the contents of a relational database as RDF, allowing users to query and interact with the data using Semantic Web technologies. D2RQ offers a declarative language for specifying the mapping between the relational database schema and the RDF model. This mapping defines how tables, columns, and relationships in the database correspond to classes, properties, and relationships in the RDF graph. The platform includes a map-generation tool that automatically generates a default mapping file based on the structure of the database. This mapping file can then be customized to align with the ontology or vocabulary being used. It specifies the rules for converting the relational data into RDF triples and defines the properties and relationships of the resulting RDF graph. Once the mapping file is created, the D2RQ Server component of the platform can be used to publish the relational database as a linked data endpoint. The server acts as a bridge between the relational database and the RDF representation, allowing users to query the data using SPARQL queries. By utilizing D2RQ, relational databases can be integrated into the semantic ecosystem, enabling interoperability with other RDF-based resources and applications. It provides a means to leverage the benefits of RDF, such as the ability to represent and reason over data using ontologies, linked data principles, and semantic technologies.

The automatic map-generation tool creates a customizable mapping file through the generate-mapping command. For example, the resources data are mapped

Figure 5.8: Example of retrieved activity data

```

1 [
2   {
3     "element": "http://www.hds.utc.fr/pre103/abox#reshistoryAPIConcept61
4       963d02137c6",
5     "Actor": "http://www.hds.utc.fr/mcc/abox#testAPIConcept61952c62ba011",
6     "Date": "11-18-2021",
7     "Location": "",
8     "Index": "http://www.hds.utc.fr/pre103/tbox#TP2",
9     "SharingSpace": "http://www.hds.utc.fr/pre103/abox#spaceAPIConcept61
10      95304802065",
11     "ResourceID": "http://www.hds.utc.fr/pre103/abox#documentAPIConcept6
12      1963d0145d09",
13     "ResourceName": "TP1 Enonce",
14     "ResourceSystem": "Upload File",
15     "ResourceType": "Document",
16     "Organization": "apre103",
17     "LocationType": "Digital Location",
18     "ActionType": "Add"
19   }
20 ]

```

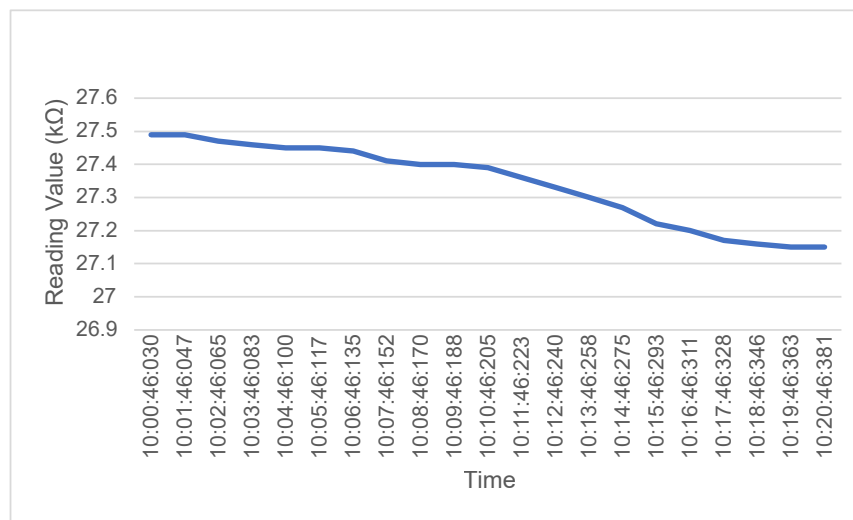


Figure 5.9: An example of time-series signal recording for sentimental state monitoring by MoodFlow@doubleYou

based on the TSCCO:Resource representation, as shown in Figure 5.13. The generated mapping file is then customized to align with the concepts defined in TSCCO. Figure 5.14 illustrates an example of the content in the customized mapping file from D2RQ. The file specifies the mapping rules for the TSCCO:Resource concept, including properties such as label, ID, title, external source, type, and URL.

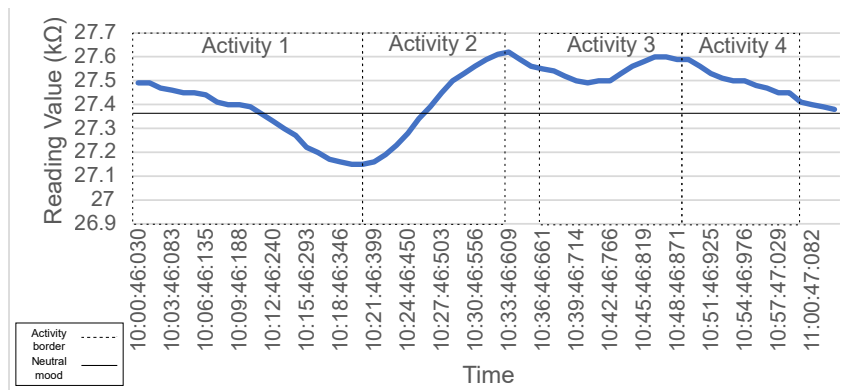


Figure 5.10: An example of time-series signal with respect to different activities and showing the neutral mood level of the person

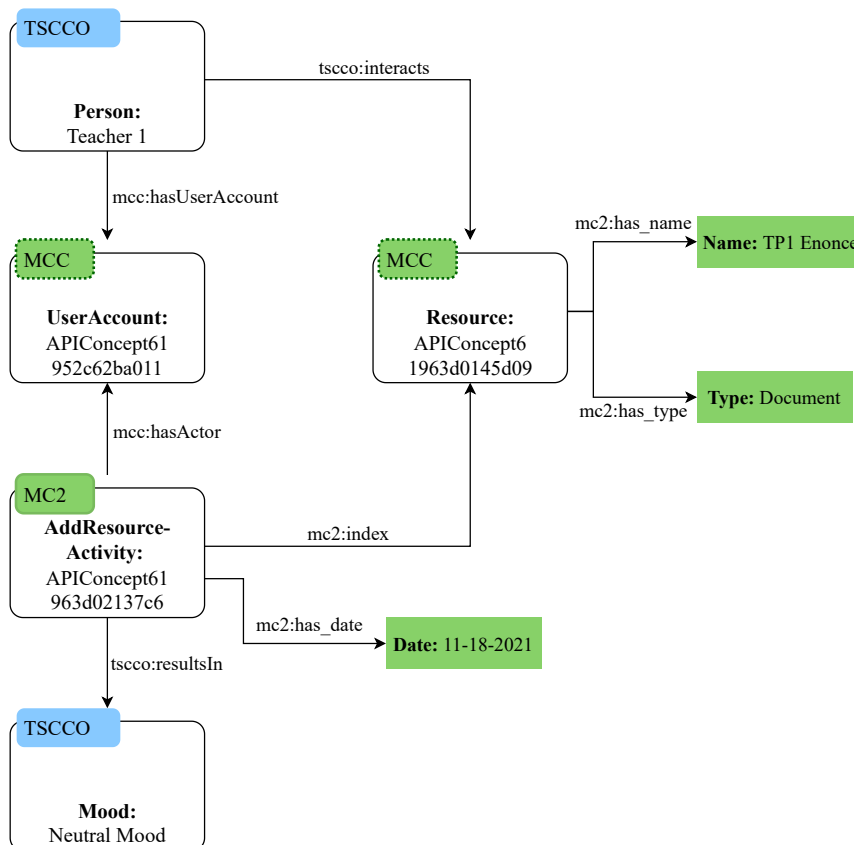


Figure 5.11: An example of activity and mood assertion to TSCCO (A-Box assertion component)

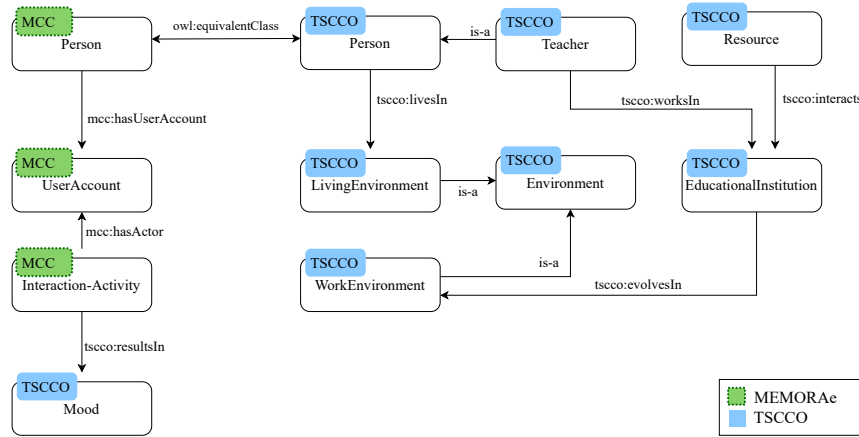


Figure 5.12: Selected concepts representation from TSCCO. (T-Box component)

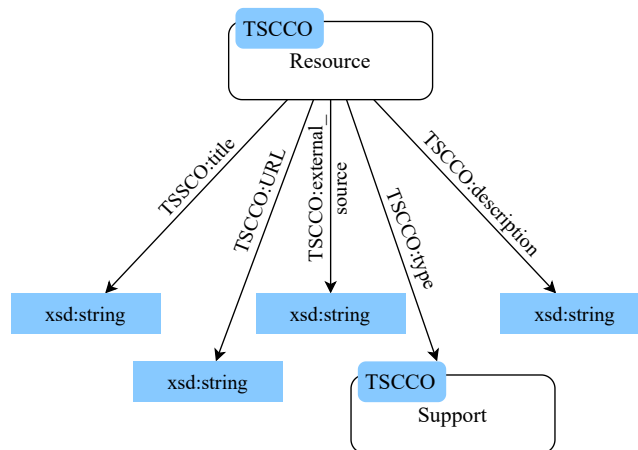


Figure 5.13: Resource concept representation in TSCCO. (T-Box component)

Figure 5.14: Example of D2RQ customized mapping file for TSCC0:Resource

```
# Table Resource
map:Resource a d2rq:ClassMap;
  d2rq:dataStorage map:database;
  d2rq:uriPattern "Resource/@@Resource.id@";
  d2rq:class tscco:Resource;
  d2rq:classDefinitionLabel "Resource";
.
map:Resource_label a d2rq:PropertyBridge;
  d2rq:belongsToClassMap map:Resource;
  d2rq:property rdfs:label;
  d2rq:pattern "Resource_@@Resource.id@";
.
map:Resource_id a d2rq:PropertyBridge;
  d2rq:belongsToClassMap map:Resource;
  d2rq:property tscco:ID;
  d2rq:column "Resource.id";
.
map:Resource_title a d2rq:PropertyBridge;
  d2rq:belongsToClassMap map:Resource;
  d2rq:property tscco:title;
  d2rq:column "Resource.title";
.
map:Resource_description a d2rq:PropertyBridge;
  d2rq:belongsToClassMap map:Resource;
  d2rq:property tscco:description;
  d2rq:column "Resource.description";
.
map:Resource_type a d2rq:PropertyBridge;
  d2rq:belongsToClassMap map:Resource;
  d2rq:property tscco:type;
  d2rq:column "Resource.type";
.
map:Resource_URL a d2rq:PropertyBridge;
  d2rq:belongsToClassMap map:Resource;
  d2rq:property tscco:url;
  d2rq:column "Resource.url";
.
map:Resource_Source a d2rq:PropertyBridge;
  d2rq:belongsToClassMap map:Resource;
  d2rq:property tscco:external_source;
  d2rq:column "Resource.external_source";
.
```


Chapter 6

Context-Aware Recommendations for Teachers

A pedagogical resources recommender system for teachers offers tailored suggestions of pertinent pedagogical materials for each individual teacher. The level of personalization achieved by this system is determined based on the available feedback history of a given teacher u_t . This feedback can be categorized into two forms: explicit and implicit. Explicit feedback encompasses ratings provided by teachers, including likes, dislikes, scaled ratings, and written reviews. This type of feedback offers comprehensive insights into the specific reactions of teachers but collecting it can be challenging. In contrast, implicit feedback pertains to the historical interactions of a teacher, such as resource access. This kind of feedback is more abundant but can be less precise and more prone to noise compared to explicit feedback. The system generates a predictive model of ratings, or a rating matrix R_{UI} , which represents the anticipated ratings of a teacher u_t for a specific resource i_r , drawing insights from their feedback history.

Various recommendation techniques, as elaborated upon in Section 2.6, can be employed by this recommender system. Given the intricacy of our challenge, a hybridization procedure becomes necessary. The metadata of resources holds a significance equivalent to that of the historical interactions between teachers and resources to furnish personalized and exceptionally precise resource recommendations. Employing a content-based approach is appropriate for processing metadata, enabling the recommendation of akin resources to those previously favored by the teacher. In parallel, a collaborative-filtering approach utilizes teachers' feedback history to identify user-based recommendations, drawing parallels with analogous teachers. Consequently, the hybridization approach combines the merits of content-based and collaborative filtering approaches, thereby amplifying the performance capabilities of the recommender system.

The meticulous choice of an appropriate hybridization technique holds great importance in enhancing the accuracy of recommendation for teachers. Consequently, it becomes imperative to opt for a method, drawn from the several possibilities expounded upon in Section 2.6.1, that seamlessly integrates the recommendations derived from the content-based recommender approach with those emanating from collaborative filtering. Among the hybridization possibilities, the adoption of a weighted approach emerges as the most suitable course of action due to the inherent linearity characterizing our problem. This linearity is a result of our dataset lacking item-level sensitive information and the fact that the volume of recommended items does not attain excessive proportions. The weighted

hybridization technique involves assimilating the outputs derived from each recommender approach by employing static weightings. Typically, these weightings are determined based on predicted ratings from each approach and are balanced by a scaling factor of 50%.

However, none of the previous approaches consider the multiple contexts of teachers, encompassing contextual factors such as mood, living location, work location, etc. The hybridization of the two approaches affects the personalization based on teacher preferences and requisites, which are inherently context dependent. Accordingly, context-aware recommendations incorporate the environmental context in the recommendation process. Traditional recommender systems function with two-dimensional datasets that solely focus on teacher and resources. In contrast, context-aware recommender systems adopt multi-dimensional datasets encompassing the multiple contexts of teachers with the two-dimensional dataset.

Contextual factors are integrated at various phases of the recommendation process depending on the amount of available data and frequency of new data insertion. Contextual prefiltering, postfiltering and contextual modeling are the three types of contextual filtering, each applied during different phases of the recommendation process, as discussed in Section 2.6.2. The contextual factors need to be abundant and specific to provide high accuracy recommendations. Unfortunately, this elevated level of context specificity causes the problem of finding similar contexts to that of the current teacher. This problem can be solved through selecting the most suitable factors using an appropriate feature selection algorithm.

Feature selection algorithm for our context-aware recommender system involves considering nature of the dataset and complexity of the selected algorithm to ensure the effectiveness and relevance of the recommendations as illustrated in Section 2.5.7. A more balanced selection of features needs to be achieved by considering both the categorical and numerical features. Mutual information measures the mutual dependence between categorical features which identifies features that are informative for recommendation. MI is suitable for capturing complex relationships between different features within the multiple contexts of teachers. Conversely, variance-based feature selection focuses on identifying numerical features with high variance, as they tend to carry more information and might be more informative for the recommendation process. Combining variance-based feature selection with mutual information obtains an enhanced selection of features that considers both types of features in our dataset.

In the subsequent sections of this chapter, we present the recommender approach, as shown in Figure 6.1. It starts with the feature selection approach (Section 6.1). Afterwards, the overall recommender system approach is presented in Section 6.2. Then, the components of the 2D recommender system are illustrated by the hybridized approach (Section 6.2.1) including content-based filtering approach (Section 6.2.1.2), collaborative-filtering approach (Section 6.2.1.1). At the end, the different contextual filtering approaches are discussed in Section 6.2.2.

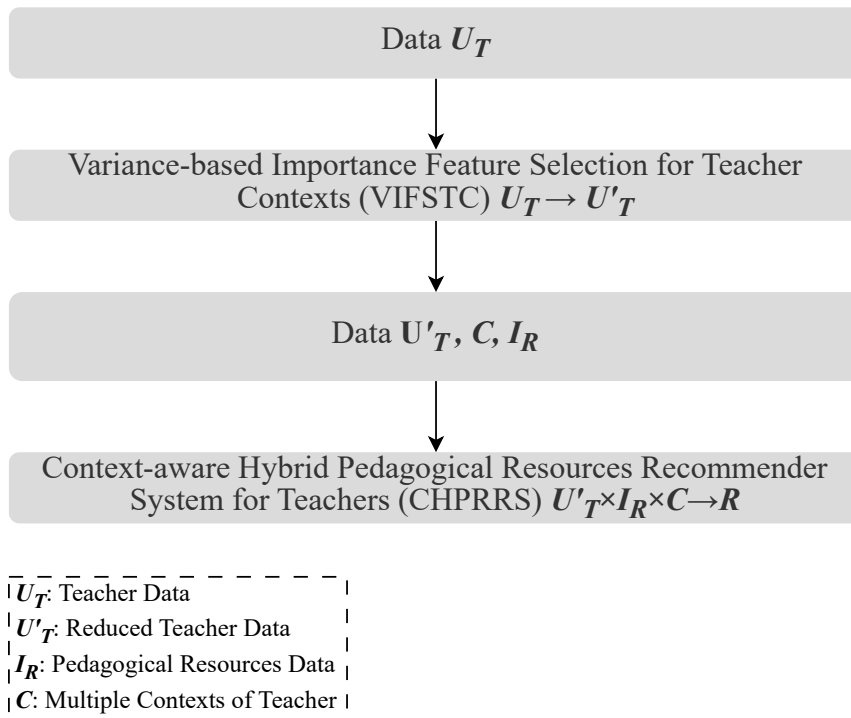


Figure 6.1: An overview of the recommender approach.

6.1 Variance-based Importance Feature Selection for Teacher Contexts (VIFSTC)

Feature selection is a fundamental pre-processing step that optimizes the pedagogical resources context-aware recommender system by reducing dimensionality, enhancing interpretability, mitigating overfitting, improving computational efficiency, elevating recommendation quality, accommodating mixed data types, and filtering out noise. The feature selection approach, outlined in Figure 6.2, is designed to systematically enhance the efficacy of feature selection within the pedagogical resources context-aware recommender system. The process commences with the extraction of ontological feature data, which serves as the foundation for subsequent analyses. These features, encapsulating pertinent features of teachers, undergo a transformation from their original feature-value representation to a vectorized format optimized for seamless integration with feature selection algorithms.

To facilitate this transformation, binary one-hot encoding is chosen, a method well-suited for converting categorical feature values into Boolean indicators. Each distinct potential value of a categorical feature is encoded as a binary feature, aligning with prior research by (Seger, 2018). This encoding approach ensures that the data is amenable to subsequent feature selection procedures.

The core of the feature selection methodology encompasses a hybrid fusion of two filtering unsupervised algorithms. The initial algorithm, mutual information, computes the collective significance of all features, drawing inspiration from the work of (Kraskov, Stögbauer, and Grassberger, 2004; Ross, 2014). MI serves as a yardstick for evaluating the relative importance of each feature within the

recommendation context.

The second constituent of this hybrid approach involves variance-based feature selection. By analyzing the variance exhibited by features, this algorithm pinpoints those with substantial variance exceeding a predetermined threshold, guided by the average importance computed through mutual information. This technique (Guyon and Elisseeff, 2003), specializes in selecting numerical features while concurrently incorporating categorical features.

The rationale behind amalgamating these disparate methods stems from their distinct strengths and weaknesses. Mutual information excels in discerning the significance of categorical or binary features but might falter in handling multi-valued features (Guyon and Elisseeff, 2003). Conversely, variance-based feature selection shines in identifying numeric variations but may inadvertently overlook crucial categorical features (Zebari et al., 2020). Through their harmonious integration, a harmonized, balanced selection of features emerges, thereby accounting for both categorical and numerical features.

Furthermore, this hybrid approach transcends the constraints posed by each individual method, mitigating their limitations. The combined methodology surmounts the challenges posed by multi-valued features and enhances the visibility of vital categorical features.

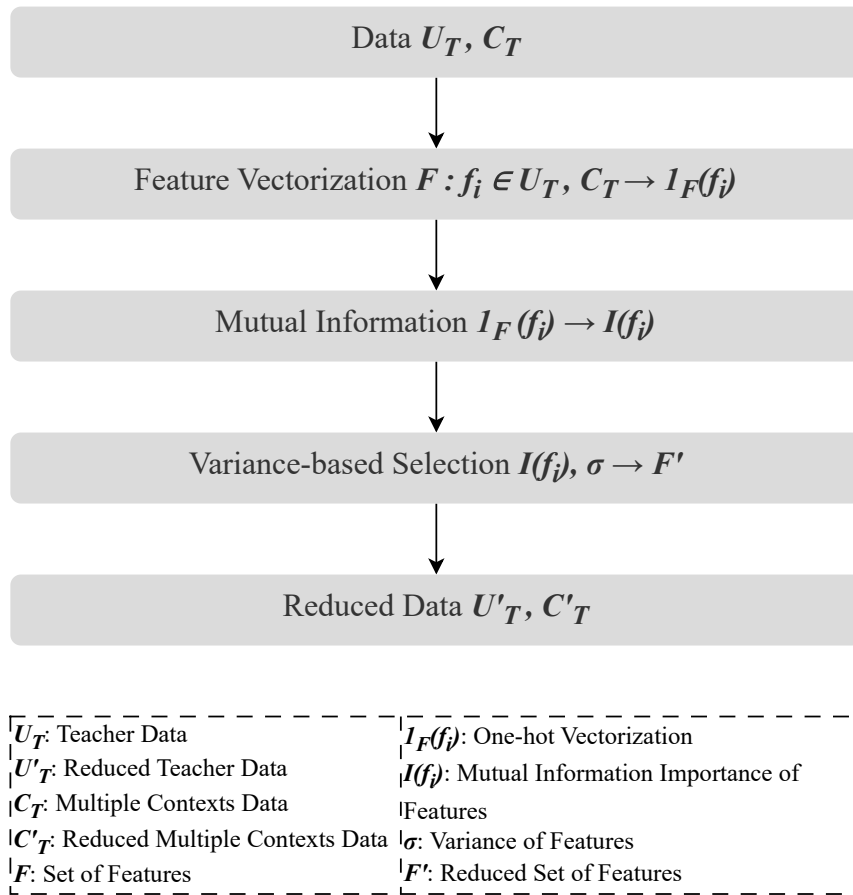


Figure 6.2: An overview of the Variance-based Importance Feature Selection for Teacher Contexts (VIFSTC) approach.

6.1.1 Features Data Retrieval and Vectorization

Data retrieval is a fundamental step aimed at extracting relevant information from a data source, in order to inform subsequent analyses and processing. In the context of the pedagogical resources context-aware recommender system, data retrieval serves the purpose of accessing the intricate features and preferences associated with multiple contexts of teachers.

The objective of data retrieval is to gather comprehensive instances of a teacher's profile, encompassing various features such as experience levels, spoken languages and teaching style. This rich pool of information is essential for tailoring personalized recommendations, as it provides insights into the nuanced factors guiding a teacher's selection of resources.

Vectorization, on the other hand, is pivotal for translating this heterogeneous and diverse set of teacher features into a uniform, structured format suitable for computational analysis. Vectorization transforms textual, categorical, and numerical features into a cohesive numerical representation, facilitating the application of mathematical and statistical techniques.

The aim of vectorization is twofold: homogenization and algorithm compatibility. Vectorization ensures that all extracted features, irrespective of their original data type, assume a standardized numerical form. This homogenization enables seamless integration of features into subsequent data processing and modeling steps. Most machine learning and statistical algorithms operate on numerical data. By vectorizing diverse features, the data becomes amenable to a wide array of computational techniques, allowing for efficient feature selection and other analyses.

Overall, data retrieval furnishes the requisite information foundation, while vectorization transforms this information into a harmonious numerical representation, thus enabling the subsequent analytical processes that drive the pedagogical resources context-aware recommender system for teachers.

6.1.1.1 Features Data Retrieval

The retrieval of teacher data from TSCCO ontological representation is achieved through the utilization of the OpenLink Virtuoso Universal Server¹, an advanced platform renowned for its fusion of relational database management functionalities with those dedicated to handling linked data and knowledge graphs. This integration capacitates the efficient management and querying of Resource Description Framework (RDF) data via the SPARQL querying language², a pivotal tool for ontological knowledge graphs manipulation.

The ensuing knowledge derived from SPARQL queries is subsequently distilled into the concise yet expressive JavaScript Object Notation (JSON) format³. This transformation accords the knowledge portability and structure required for seamless integration with the feature selection algorithm. Each piece of retrieved data culminates as a distinct feature for selection, each holding import in characterizing the preferences and features of individual teachers within the recommender system.

¹<https://virtuoso.openlinksw.com/>

²<https://www.w3.org/TR/rdf-sparql-query/>

³<https://www.json.org/json-en.html>

A categorical breakdown of the features slated for retrieval is elucidated in Table 6.1. This table delineates diverse facets of a teacher’s profile, encompassing features such as experience level, spoken languages, and other features. Each categorical domain encapsulates a set of features, with the number of features therein serving as a reflection of the diversity and granularity of information extracted. For example, within the "Experience Level" category, the availability of three distinct values—novice, intermediate, and expert—manifests as three individual features within the framework.

The strategic amalgamation of OpenLink Virtuoso Universal Server and SPARQL, underpinned by the transformative prowess of JSON, orchestrates a streamlined and comprehensive data retrieval mechanism. This concerted synergy furnishes the requisite foundation for the ensuing feature selection algorithm. Moreover, this adept integration augments the potential for the development of sophisticated recommender systems and analogous algorithms, elucidated in subsequent chapters, capitalizing on the robust capabilities afforded by knowledge graphs and semantic data representation.

Table 6.1: List of features

Features	Related Concept	Range	# of corresponding features
Experience level	Teacher profile	{novice, intermediate, expert}	3
Gender	Teacher profile	{male, female, other}	3
Mother tongue	Teacher profile	{Arabic, English, French, Italian, Spanish}	5: only 5 languages exist in the used dataset
Second language	Teacher profile	No language, Arabic, English, French, Italian, Spanish	6: only 5 languages exist in the used dataset
Third language	Teacher profile	No language, Arabic, English, French, Italian, Spanish	6: only 5 languages exist in the used dataset
Science	Teacher profile	{<list_of_science_fields>}	6: only 6 fields of science exist in the used dataset
Teaching style	Teacher profile	authority, demonstrator, facilitator, delegator, hybrid	5
Country	Living environment	{<list_of_countries>}	5: only 5 countries exist in the used dataset
Living area type	Living environment	urban, suburban, rural	3
Work area type	Working environment	urban, suburban, rural	3
Institution type	Working environment	primary, preparatory, secondary, university	4
Mood	Sentimental state	negative, neutral, positive	3

6.1.1.2 Features Data Vectorization

Data vectorization is the method of transforming a list of features into vectors. If the list of features, or dictionary of features, is a set of strings, the vectorization takes a binary one-hot encoding, as shown in Equation 6.1. For each string value, a feature is created according to legal combination '1' or otherwise '0'. Firstly,

it uses the list of features for each teacher to regard the occurrence of a string-valued feature. For each feature, the corresponding values are encoded into the following pair $\langle feature_name \rangle = \langle feature_value \rangle$. If the string value of this feature exists for a teacher, the value of '1' is placed in the corresponding cell in the vectorized matrix. Figure 6.3 shows the resulting matrix for 20 teachers from the used dataset. Each row of the matrix represents the features of teachers while columns represent the existence of the corresponding feature with the context of a teacher. For example, the feature "lang1=Spanish" is associated with teachers #0, #3, #5, and #16 only. These vectorized features are transmitted into the feature selection approach.

$$1_F(f) := \begin{cases} 1 & : f \in F, \\ 0 & : f \notin F \end{cases} \quad (6.1)$$

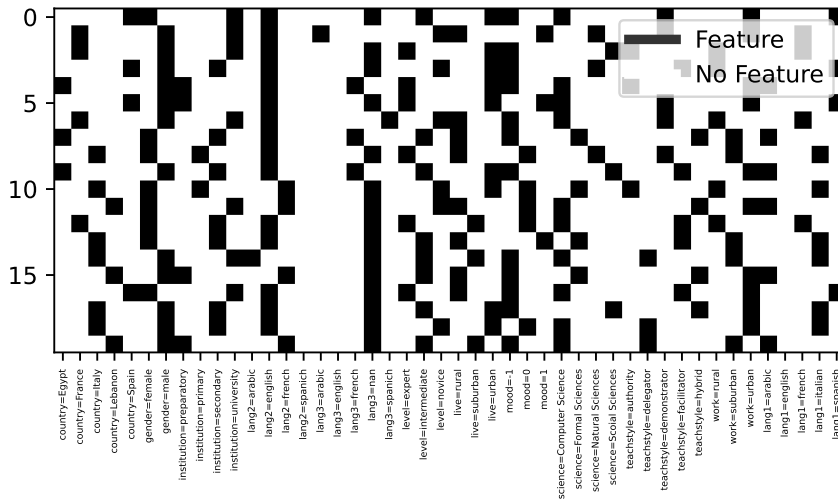


Figure 6.3: An example of vectorized list of features.

6.1.2 Mutual Information

The approach starts with mutual information, a prominent measure of statistical dependence. Mutual Information (MI) is a robust assessment method that transcends mere correlations, encompassing complex relationships within the dataset (Duboue, 2020). It stands as a fundamental building block for feature selection, where its core lies in capturing the intrinsic information shared between two distinct features.

The fundamental underpinning of MI is rooted in the k -nearest neighbors distance estimator, which embodies the essence of proximity-based feature interrelation. As eloquently depicted in Equation 6.2, MI quantifies the pertinence of a given feature f_i by elucidating its information contribution across categorical divisions (Duboue, 2020). In essence, MI gauges the extent to which a feature accentuates discernment among various categories, thereby unveiling the crux of feature importance.

$$I(f_i, f_j) = \int_{f_i} \int_{f_j} p(f_i, f_j) \log \frac{p(f_i, f_j)}{p(f_i)p(f_j)} dx dy, \quad (6.2)$$

where f_i, f_j are two features,

$p(f_i, f_j)$ is the joint probability density function of the two features,

$p(f_i)p(f_j)$ are the marginal density function.

Moreover, MI expounds its prowess in the realm of exploring the entwined dynamics of feature pairs. A non-negative value, yielded through rigorous computation, signifies the intensity of mutual dependency binding two features. This symbiotic relationship encapsulates intricate patterns that potentially underlie a teacher's preferences and resource interaction. Evidently, such insights underscore the paramount significance of this metric in effecting the judicious selection of pivotal features.

For instance, an incisive examination of the correlation coefficients portrayed in Figure 6.4 unveils profound interplays. Notably, the hybrid teaching style feature evinces dependence on the intermediate expertise level of teachers. This revelation, depicted in Figure 6.5, elucidates a relatively diminished computed importance, a tangible demonstration of MI's prowess in discerning nuanced feature dynamics.

To arrive at a comprehensive perspective, the culmination of these individual importances amalgamates into the average importance $mean_I$ of all features. This pivotal metric lays the groundwork for the subsequent endeavor: the determination of the threshold for variance-based feature selection. This synthesis signifies the convergence of mutual information and variance, underscoring a holistic approach that seamlessly integrates statistical interdependence with discerning variability.

In summary, the intended mutual information approach navigates the intricate landscape of feature relevance, deftly unraveling the interwoven fabric of dependencies within the pedagogical context. By harnessing the power of mutual information, this endeavor augments the efficacy of feature selection, ultimately fostering the precision and discernment requisite for the pedagogical resources context-aware recommender system for teachers.

6.1.3 Variance-based Feature Selection

Variance-based threshold feature selection emerges as a potent dimensionality reduction strategy, strategically tailored to filter a dataset (Duboue, 2020). The substructure of this technique rests upon the premise that features characterized by low variance, denoting limited variability across instances, contribute minimally to the prediction task. Elegantly encapsulated by Equation 6.3, the rationale behind this approach is that features f_1, \dots, f_i, f_n with low variance σ across different instances in contexts of teachers are not significantly contribute to the accuracy of prediction tasks.

$$\sigma_i = \sqrt{\frac{\sum (f_i - \bar{f}_i)^2}{n}} \quad (6.3)$$

Akin to a hybrid process, the variance-based threshold approach delineates a delineation whereby features harboring variance below a predefined threshold confront elimination from teacher data. It puts forward that the inadequate information carried by low-variance features compromises their potential to furnish

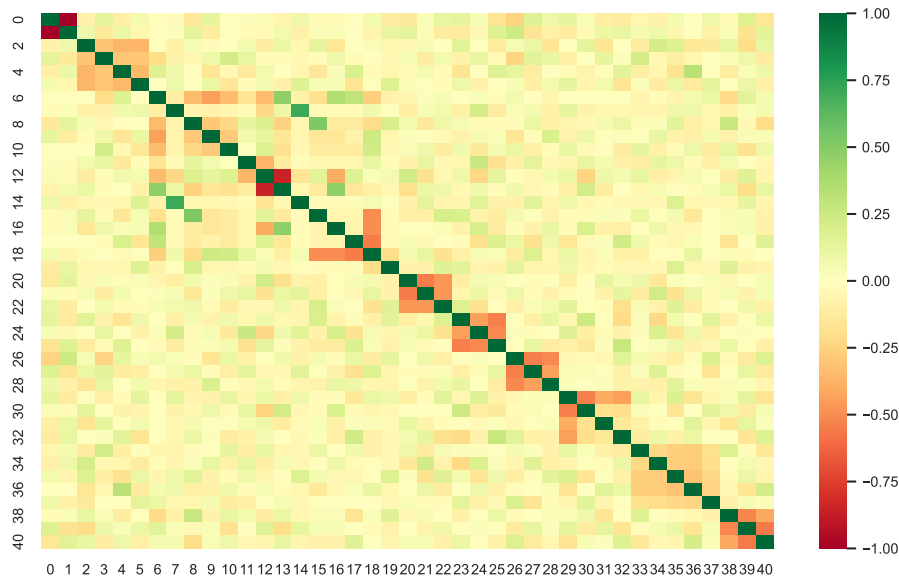


Figure 6.4: Correlation between features.

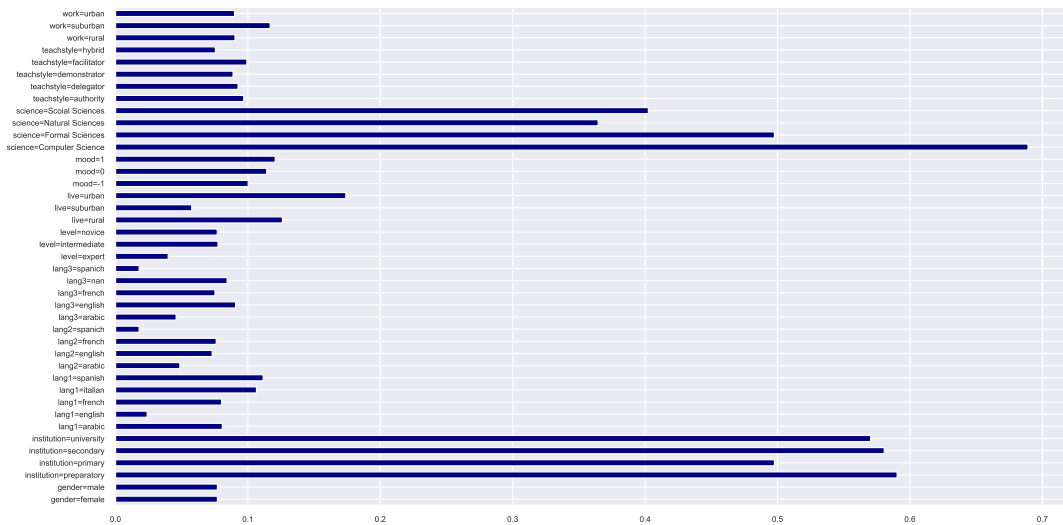


Figure 6.5: Mutual information importance of features.

meaningful insights for predictive endeavors. The established threshold, often set as a fraction of the total variance intrinsic to the dataset, wields the authority to demarcate the threshold of significance.

Through VIFSTC approach, the threshold thr is like a quality control mechanism, as illustrated by Equation 6.4. It ensures that only features exhibiting a certain level of variability, deemed significant by the combination of mutual information and variance, are retained. This way, the dataset becomes more streamlined and focused on the most informative attributes.

$$thr = mean_I(1 - mean_I) \quad (6.4)$$

After the thresholding process, the dataset is left with a reduced set of features that possess the desired level of variability and uniqueness between teachers. These features are then used in subsequent steps of the recommender system. By narrowing down the feature set in this manner, the subsequent analysis becomes more efficient and effective, enabling the system to make more accurate recommendations for teachers based on their preferences and interactions while reducing the computational complexity.

In summary, variance-based threshold feature selection acts as a data refining phase, concentrating on the features that truly matter in the pedagogical resources context-aware recommender system. It ensures that the recommendations are driven by the most informative features while disregarding those that contribute less to the overall prediction quality.

6.2 Context-Aware Pedagogical Resources and Peer Teachers Recommender System

The task of providing personalized recommendations in education is a challenging and complicated process, particularly when considering the multiple contexts of teachers. In order to address this challenge, the recommender approach introduces a methodology that aims to provide two types of recommendations: pedagogical resources and peer teachers recommendations. These two types address the potential solutions elicited through the problem statement in Section 1.1.1 to offer help to teachers in the form of aiding pedagogical resources and collaboration with experienced teachers. This approach utilizes the integration of TSCCO ontology to represent multiple contexts of teachers such as living environment, working environment, and sentimental state within the collaboration management ontology, MCC. The approach first detects the mood of teacher using MoodFlow@doubleYou and tracks their activity using the MEMORAE platform. Then, the contextual features are extracted and selected carefully, as previously illustrated in Section 6.1, keeping the mood-related features due to its importance as discussed in Section 2.6.5. The contextual features are used at different phases of the context-aware recommendation process, but the approach starts with a hybrid 2D recommender approach. The 2D recommender approach captures the teacher-to-teacher similarity easily captured by user-based collaborative filtering recommender approach as well as the pedagogical resources suitability for a teacher excelled by content-based filtering recommender approach. Accordingly, the countability of both types of similarities combines the advantages of both approaches.

Therefore, this section starts with the representation of the hybrid 2D pedagogical resources recommender approach in Section 6.2.1 which is composed of user-based collaborative filtering (Section 6.2.1.1) and content-based filtering (Section 6.2.1.2) recommender approaches. Then, we discuss the consideration of multiple contexts during different phases of recommendation process in Section 6.2.2. Finally, we present how the peer teacher recommendations are obtained in Section 6.2.3.

6.2.1 Hybrid 2D Pedagogical Resource Recommender Approach

The two-dimensional hybrid recommender (2DHR) offers personalized recommendations for pedagogical resources most relevant to each teacher using predictive ratings matrix R_{UI} of a teacher t on each item in I . As previously discussed, there are two types of rating or feedback, but we use each type of rating with different recommender approaches to capture the different relationships between a teacher and items or between a teacher and other peers. Therefore, two recommender approaches are hybridized together as illustrated in Figure 6.6: collaborative filtering (2DCFR) and content-based filtering (2DCBR). Collaborative filtering approach offers high accuracy recommendations using the teacher-resource interactions history. But it requires this interactions history to operate which causes the cold-start problem (Lika, Kolomvatsos, and Hadjiefthymiades, 2014). A well-known solution for the cold-start problem is using the content-based approach that operates using the resources meta-data without the use of the interactions history (Lika, Kolomvatsos, and Hadjiefthymiades, 2014).

Due to the linearity of our problem, weighted hybridization is used to combine recommendations from each approach. The hybridization merges complementary information relying on pedagogical resource features and explicit ratings in content-based filtering approach, and explicit user-item interactions as well as implicit feedback in collaborative filtering approach. The weighting hybridization splits the predictive rating of recommendations from each approach by a static weighting factor of 50%. Through the rest of this section, the two recommender approaches are presented along with their algorithms.

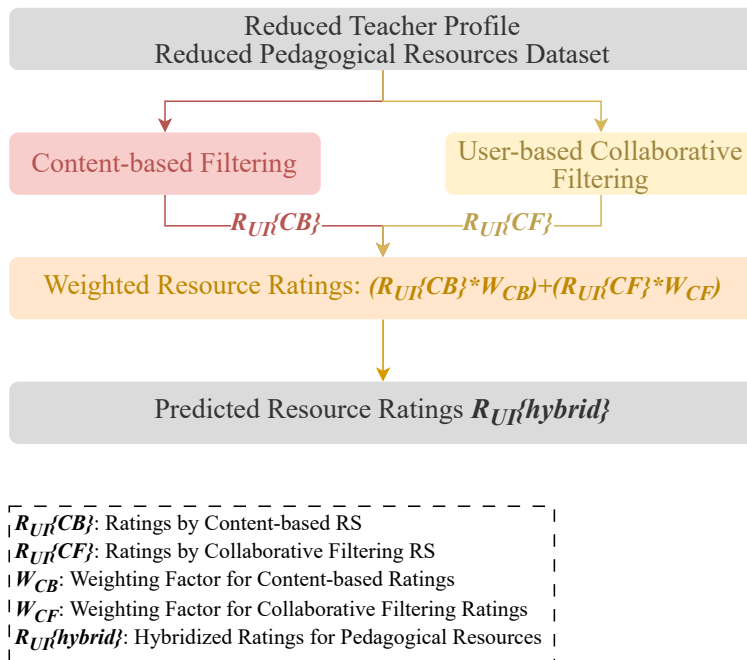


Figure 6.6: Overview of the 2D hybrid recommender approach.

6.2.1.1 User-based Collaborative Filtering Recommender Approach

User-based collaborative filtering recommender approach (2DCFR) uses all available feedback from teachers like explicit scaled ratings of pedagogical resources as well as the implicit access from the interaction history (Schafer et al., 2007). It searches for unseen pedagogical resources relative to a teacher that are highly rated by similar teachers which leverages collective interaction history of all teachers. The 2DCFR approach, as summarized in Figure 6.7, starts with retrieving the interactions history H_t of teacher t . Then, the approach searches for teachers with similar interactions history $U'_T(t)$. Parallely, another search process takes place to find unseen pedagogical resources $I'_R(t)$ by teacher t . Pearson correlation coefficient is selected to compute similarity between teachers over other classic similarity approaches as it excels in user-based CF recommender approaches (Koochi and Kiani, 2016). Pearson similarity is computed between two teachers t_1 and t_2 as proved in (Koochi and Kiani, 2016) as follows:

$$sim(t_1, t_2) = \frac{\sum_{i \in I_R} (r_{ui}(t_1) - \overline{r_u(t_1)})(r_{ui}(t_2) - \overline{r_u(t_2)})}{\sqrt{\sum_{i \in I_R} (r_{ui}(t_1) - \overline{r_u(t_1)})^2} \sqrt{\sum_{i \in I_R} (r_{ui}(t_2) - \overline{r_u(t_2)})^2}} \quad (6.5)$$

where $I_R = \{i_1, \dots, i_m\}$ are the set of pedagogical resources,

$r_{ui}(t_i)$ is the rating provided by teacher t_i on pedagogical resource i ,
 $\overline{r_u(t_i)}$ is the average rating for teacher t_i .

Afterwards, the results from each search process are merged to be used for ratings prediction. K-nearest neighbor (K-NN) algorithm predicts the ratings of unseen pedagogical resources according to the highest k Pearson similarities of teachers. K-NN offers a simple and effective method to leverage teacher similarity for personalized recommendations (Park et al., 2015). Normalized rating prediction of pedagogical resource i_{R_i} are computed for teacher t_1 using the k teachers with highest Pearson similarity (Koochi and Kiani, 2016) as follows:

$$\hat{r}_{ui}(t_1) = \mu_{t_1} + \frac{\sum_{t_2 \in N_k(t_1)} sim(t_1, t_2) \cdot (r_{ui}(t_2) - \mu_{N_k(t_1)})}{\sum_{t_2 \in N_k(t_1)} |sim(t_1, t_2)|} \quad (6.6)$$

where $sim(t_1, t_2)$ is the Pearson similarity between teachers t_1 and t_2 ,

$r_{ui}(t_2)$ is the rating provided by teacher t_2 on pedagogical resource i ,

μ_{t_1} is the mean rating for teacher t_1 ,

$\mu_{N_k(t_1)}$ is the mean rating for top- k neighbor teachers to teacher t_1 .

Finally, a list of pedagogical resource recommendations is offered to teacher t_1 . These recommendations are to be combined with the content-based ones.

6.2.1.2 Content-based Filtering Recommender Approach

Content-based Filtering recommender approach (2DCBR) uses explicit ratings and meta-data of pedagogical resources to recommend similar ones to teacher t based on the highly rated resources (Pazzani and Billsus, 2007). As previously explained in Chapter 5-Section 5.4, a pedagogical resource is described with a

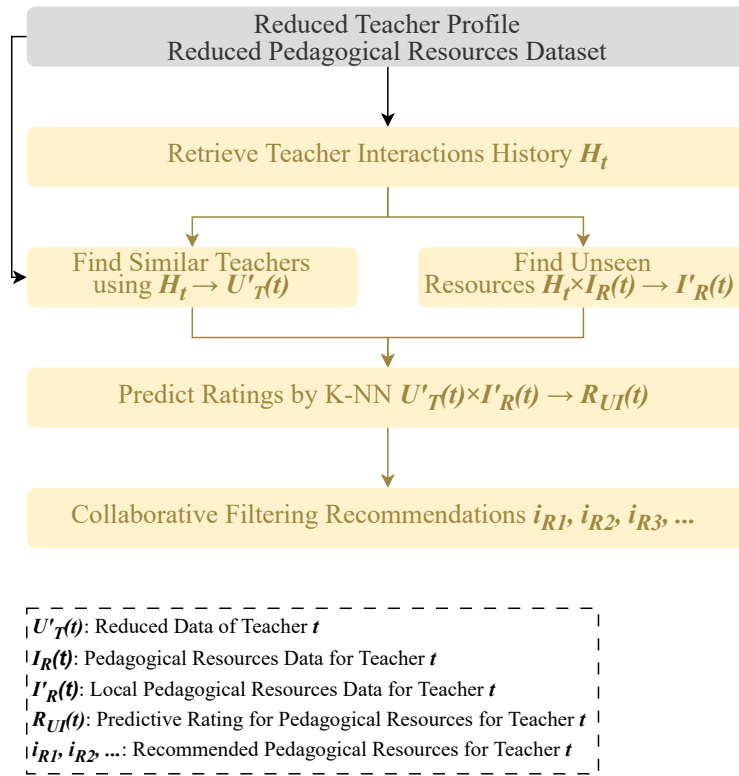


Figure 6.7: Overview of user-based collaborative filtering recommender approach (2DCFR).

unified form due to the diversity of the data sources and is represented as follows: `<resource_title, resource_url, resource_type, resource_external_system, resource_description>`. These features lack descriptive meta-data to easily distinguish one resource from another. Therefore, the 2DCBR approach, as illustrated in Figure 6.8 starts with a preprocessing phase. This phase is required at first to facilitate the processing of text-based data such as `resource_title` and `resource_description` which contain more informative features. Therefore, the 2DCBR follows a systematic preprocessing algorithm, as in Algorithm 2 that starts with the concatenation of all text-based data. The algorithm tokenizes the text into individual words which is followed by unification of the case of characters. Afterwards, through lemmatization, each word is reduced to its root form. Additionally, stop words are removed to eliminate common and non-informative words with a length less than 3 characters. Finally, Term Frequency-Inverse Document Frequency (TF-IDF) captures the importance of each word in the text-based data relative to each pedagogical resource (Pazzani and Billsus, 2007). It emphasizes unique words that carry high levels of significance by transforming these data into a numerical representation to be easily processed by Singular Value Decomposition (SVD). SVD reduces the dimensionality through extracting latent features in addition to the reduction of the high dimensional space created by TF-IDF features (Pazzani and Billsus, 2007).

The approach continues with predicting ratings for each pedagogical resource from the preprocessed features $I''_R(t)$. Additional step of TF-IDF vectorization is performed to combine the teacher profile features with the ones of resources.

Algorithm 2 Text-based data preprocessing algorithm of 2DCBR**Input:** $I'_R(t)$: local pedagogical resources meta-data for teacher t **Output:** $I''_R(t)$: preprocessed local pedagogical resources meta-data for teacher t **begin**

- 1: $I'_R(t).resource_title_description = I'_R(t).resource_title.concat(I'_R(t).resource_description);$
- 2: $tokenize = I'_R(t).resource_title_description.word_tokenize();$
- 3: $lower_case = tokenize.casefold();$
- 4: $lemmatizer = lower_case.lemmatizer();$
- 5: $filtered_stopwords = lemmatize.apply(stopwords_dict);$
- 6: $filtered_words = filtered_stopwords.apply(filter,len < 3);$
- 7: $I'_R(t).preprocessed = filtered_words;$
- 8: $tfidf_vectorizer = TfidfVectorizer();$
- 9: $tfidf = tfidf_vectorizer.fit_transform(I'_R(t).preprocessed);$
- 10: $svd = SVD();$
- 11: $I''_R(t) = svd.fit_transform(tfidf);$
- 12: **return** $I''_R(t);$

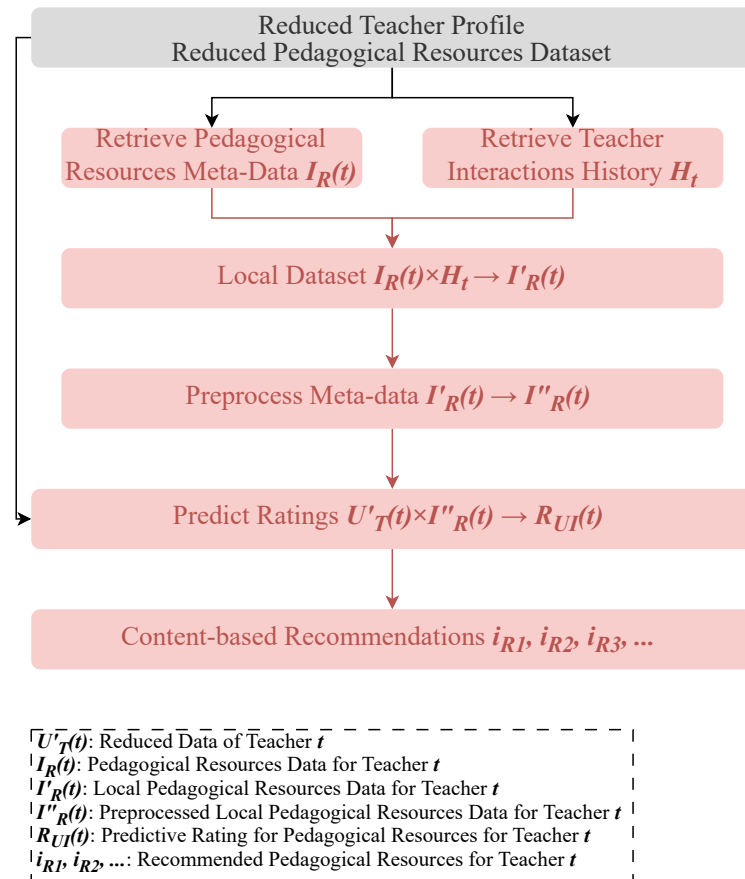
end

Figure 6.8: Overview of content-based filtering recommender approach (2DCBR).

Linear least squares regression and L2 regularization (Ridge regression) are employed to predict ratings for pedagogical resources (Paradarami, Bastian, and Wightman, 2017; Xu et al., 2019). The linear least squares method offers a straightforward methodology to assume linear relationship between pedagogical resources features and their corresponding ratings which is the case after features vectorization (Xu et al., 2019). L2 regularization is used to prevent overfitting in the regression model by adding a penalty attribute to the objective function to eliminate large coefficients (Paradarami, Bastian, and Wightman, 2017). At the end, the rating predictions $R_{UI}(t)$ are used to offer content-based recommendations that are mixed with the collaborative filtering ones using weighted hybridization.

6.2.2 Context-Aware Recommender Approaches

The 2D recommender approach considers teacher-pedagogical resource interactions history or content of resources only. However, the multiple contexts of teachers encapsulate important information related to living context, working context, or sentimental state. These multiple contexts must be considered to enhance the recommender performance to achieve high accuracy. Preferences of a teacher are content-dependent and greatly affect the rating predictions (Krause, Smailagic, and Siewiorek, 2005).

Through the approach of this thesis, three types of contextual filtering are integrated to the hybrid 2D recommender approach at different phases of the recommendation process. Figure 6.9 illustrates the different phases of contextual filtering: prefiltering, postfiltering, contextual modeling, and mix-filtering approach named context-aware pedagogical resources recommender approach (CAPRRA). The contextual prefiltering, as explained in Algorithm 3, refines the set of related teachers at the beginning of the recommendation process by selecting a reduced set of similar teachers taking into account the multiple contexts of the current teacher. This contextual filtering approach decreases the computational complexity of the recommendation process by reducing the original set of similar teachers of the 2D recommender approach (Adomavicius and Tuzhilin, 2010). Unfortunately, the sparsity nature of data across the multiple contexts can result in less accurate recommendations raising the question of whether contextual prefiltering is better or 2D recommender is sufficient (Negro, 2021).

On the contrary, contextual postfiltering, as explained in Algorithm 4, approach refines the 2D recommendations with respect to the multiple contexts of the current teacher (Adomavicius and Tuzhilin, 2010). As the 2D recommender is independent of the actual contexts, the rating predictions remain unchanged and hence, no significant improvement of accuracy (Negro, 2021). These two contextual filtering approaches have low implementation complexity and easily integrated with any 2D recommender, but there is uncertainty of their effectiveness over the 2D recommenders (Negro, 2021).

The last type of contextual filtering is contextual modeling which constructs a contextual model to be applied on 2D recommendation processes as well as a mixture of prefiltering and postfiltering phases to ensure the accurateness of recommendations (Adomavicius and Tuzhilin, 2010; Negro, 2021; Campos et al., 2013). This approach processes a multi-dimensional matrix (multipartite graph)

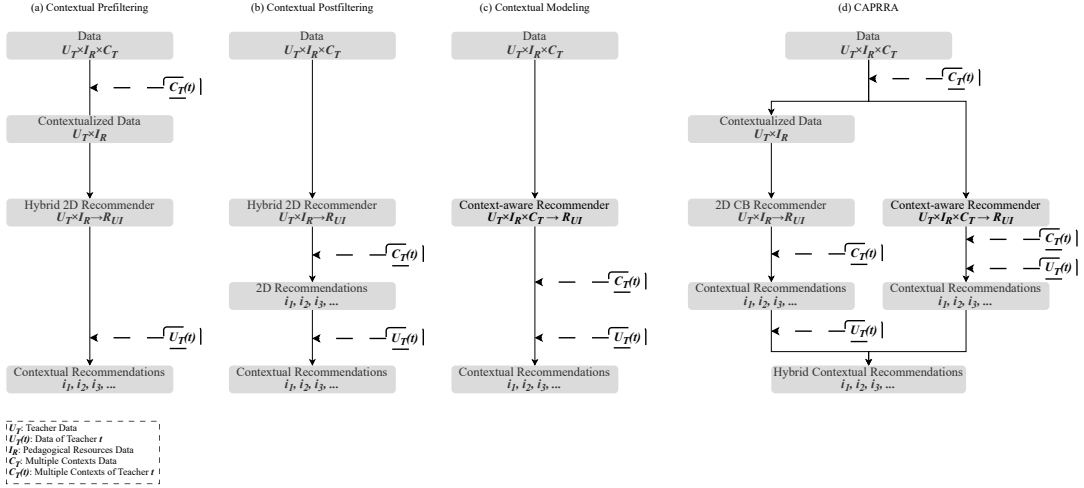


Figure 6.9: Overview of different context-aware filtering recommender approach.

Algorithm 3 Contextual prefiltering recommender algorithm

Input: t : teacher t

$C_T(t)$: multiple contexts of teacher t

H_t : history of current teacher T

k : number of pedagogical resources to be recommended to teacher t

Output: context_recommendations : list of contextual recommendations

begin

1: prefiltering_model = ContextPreFiltering($C_T(t)$);

2: data_filtered = prefiltering_model.apply(H_t);

3: CB_context_recommendations = ContentBasedRS(t , data_filtered, k);

4: CF_context_recommendations = CollaborativeFilteringRS(t , data_filtered, k);

5: context_recommendations = weightedHyrid(CB_context_recommendations,
CF_context_recommendations, 0.5);

6: **return** context_recommendations;

end

of different features or factors in teacher \times resource \times context during the recommendation process (Negro, 2021). This multipartite graph, as illustrated in Figure 6.10, models features from the multiple contexts of teachers in addition to features of the pedagogical resources. This representation of the features/factors is integrated with the 2D collaborative-filtering approach to enhance the teacher-based recommendation process. This combination prevents data sparsity problem and speeds up the contextual filtering process. The integration of such approach with the content-based approach suffers from high complexity and results in high computational power (Negro, 2021).

Therefore, in this thesis, context-aware pedagogical resources recommender approach (CAPRRA), as explained in Algorithm 5, is followed to avoid high computational complexity and power. CAPRRA introduces a hybrid combination of the different contextual filtering types with the presented 2D hybrid recommender. CAPRRA is divided into two parallel recommendation processes: contextual mix-filtering with the 2D content-based recommender, and contextual modeling with

Algorithm 4 Contextual postfiltering recommender algorithm

Input: t : teacher t

$C_T(t)$: multiple contexts of teacher t

H_t : history of current teacher T

k : number of pedagogical resources to be recommended to teacher t

Output: context_recommendations : list of contextual recommendations

begin

- 1: CB_2D_recommendations = ContentBasedRS(t, H_t, k);
- 2: CF_2D_recommendations = CollaborativeFilteringRS(t, H_t, k);
- 3: hybrid_2D_recommendations = weightedHyrid(CB_2D_recommendations, CF_2D_recommendations, 0.5);
- 4: postfiltering_model = ContextPostFiltering($C_T(t)$);
- 5: context_recommendations = postfiltering_model.apply(hybrid_2D_recommendations);
- 6: **return** context_recommendations;

end

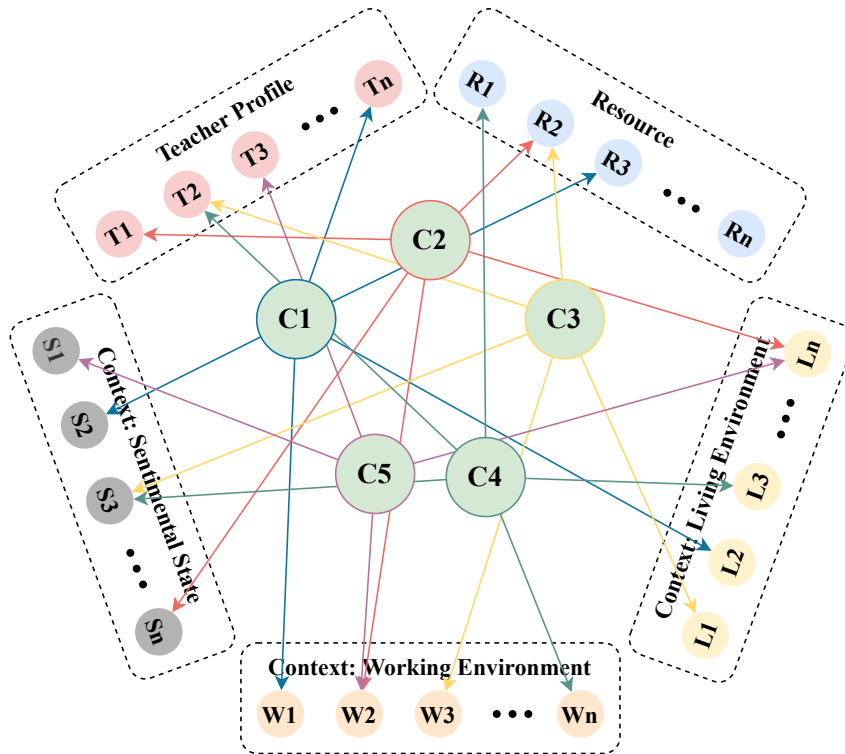


Figure 6.10: Representation of the multi-dimensional matrix (multipartite graph) of the multiple contexts of a teacher.

the 2D collaborative-filtering recommender. Contextual mixfiltering is a combination of contextual prefiltering and postfiltering with the content-based recommender. This combination relies essentially on the contextual prefiltering that is responsible for the obtained recommendations while the contextual postfiltering is an enforcing phase to prevent anomalies in the recommended pedagogical resources.

On the other hand, the contextual modeling phase creates multi-dimensional matrices or the contextual model to be used during the factorization machines

Algorithm 5 CAPRRA recommender algorithm**Input:** t : teacher t $C_T(t)$: multiple contexts of teacher t H_t : history of current teacher T k : number of pedagogical resources to be recommended to teacher t **Output:** hybrid_context_recommendations : list of contextual recommendations**begin**1: prefiltering_model = ContextPreFiltering($C_T(t)$);2: data_filtered = prefiltering_model.apply(H_t);3: preCB_context_recommendations = ContentBasedRS(t , data_filtered, k);4: postfiltering_model = ContextPostFiltering($C_T(t)$);

5: postCB_context_recommendations = postfiltering_model.apply(preCB_context_recommendations);

6: CF_context_recommendations = ContextAwareCollaborativeFilteringRS(t , $C_T(t)$, H_t , k);

7: hybrid_context_recommendations = weightedHyrid(postCB_context_recommendations, CF_context_recommendations, 0.5);

8: **return** hybrid_context_recommendations;**end**

(FMs) phase (Rendle, 2010). Instead of the Pearson similarity and K-NN, the new combination uses FMs for contextual modeling and filtering of the multiple contexts of teachers. FMs are particularly suitable for handling interactions between various contextual features and resource features, making them an effective choice for capturing complex relationships and providing accurate recommendations in a context-aware approach (Pasricha and McAuley, 2018). FMs are designed to capture interactions between features using factorized parameters for our regression task of ratings prediction as follows (Rendle, 2010):

$$\hat{y}(x_t) = w_0 + \sum_{i=1}^n w_i x_i + \sum_{i=1}^n \sum_{j=i+1}^n \langle v_u, v_j \rangle x_i x_j \quad (6.7)$$

where $\hat{y}(x_t)$ is the predicted rating for the features vector x for teacher t ,

w_0 is the bias term,

w_i are the weights for individual features x_i ,

v_i are the latent factor vectors associated with features x_i ,

w_i are the weights for individual features x_i ,

n is the dimensionality of features space.

At the end, FMs Regressor predicts ratings of each unseen pedagogical resource according to the computed rating of the teacher features vector. These contextual recommendations are combined with the other contextual recommendations from the mixfiltering approach using the same weighted hybridization previously applied in the 2D recommender. This weighted hybridization splits the importance weight equally between the recommendations from both context-aware recommender phases.

Algorithm 6 An overview of the CAPTRA algorithm

Input: C_{resr}^T : current context of a teacher T and concerning the current pedagogical resource $resr$

H^T : history of current teacher T

Output: sortedList : sorted list of all matching teachers

begin

1: **if** (getMood(T) == 'negativeMood' && offerHelp == true) **then**

2: $W_{global} = \text{ContextualFeaturesWeighting}(C_{resr}^T, H^T)$;

3: matchingList = contextMatching(W_{global});

4: sortedList = SWRLSort(matchingList);

5: **end if**

6: **return** sortedList;

end

6.2.3 Context-Aware Peer Teacher Recommender Approach

The contextual matching between peer teachers, is a complicated process that can be changeable depending on the context of a teacher. Multiple contexts of a teacher are described by a set of features. These features are dependent on the situation of this person which prevents the unification of the selected ones. Moreover, predictions of context-aware peer recommendations, can be imprecise if all contextual features were weighted equally. The personalized process of weighing and extraction of contextual features improves the quality of recommendations offered to each teacher. Therefore, the context-aware peer teacher recommender approach (CAPTRA) introduces an algorithm to weigh and extract contextual features, in addition to a sorting algorithm using SWRL reasoning rules using the ontological representation of TSCCO (Nashed, Lahoud, and Abel, 2021; Nashed et al., 2021).

First, a teacher T is asked if she/he needs help when a negative mood is detected during her/his interaction with a pedagogical resource $resr$. A negative mood can be defined as the sentimental commitment level below the personal average neutral commitment level that is calibrated at first teacher interaction. If the teacher accepts the help or she/he is not sure, the algorithm starts to find a peer recommendation for this teacher. The algorithm is divided into three steps: extraction/weighting of features for T with respect to $resr$. It returns a list of contextual features pairs composed of f associated with a weight w . Afterwards, it generates a matching peer teacher list by computing the similarity between the teacher T and other peer teachers. At the end, it sorts the matching list according to the pre-defined SWRL reasoning rules using scoring and ranking for the peer teachers, as shown in Algorithm 6 which is explained through the following sections.

CAPTRA extracts the contextual features and weights it in correspondence to a specific teacher T who interacts with a specific pedagogical resource $resr$ which introduces precise description of this teacher's context C_{resr}^T and high personalization level of peer recommendations. Algorithm 7 starts with the *factorsExtraction* algorithm that selects the factors and computes the initial weights W_{init} of each factor $f_i \in F$. The initial weights are optimized by an optimization algorithm based on the particle swarm optimization (PSO) (Zheng, Burke, and Mobasher,

Algorithm 7 ContextualFeaturesWeighting algorithm

Input: C_{resr}^T : current context of a teacher T and concerning the current resource $resr$
 H^T : history of current teacher T
 FF : fitness function

Output: $W_{global} = \{ \langle f_1, w_1 \rangle, \dots, \langle f_{n'}, w_{n'} \rangle \}$: set of final weights w_i for each of the selected n' context factor f_i from the n initial context factors where $n' < n$.

begin

- 1: $W_{init} = \text{ContextualFeatureExtraction}(C_{resr}^T, H^T)$;
- 2: $W_{global} = \text{weightsOptimization}(W_{init}, FF)$;
- 3: **return** W_{global}

end

2012; Zheng, Burke, and Mobasher, 2013b). The root mean squared error (RMSE) is used as the fitness function for the optimization algorithm for error minimization.

The *factorsExtraction* function, as shown in Algorithm 8, extracts a set of contextual features n' from all contextual features n under the condition $n' < n$ for a teacher's context C_{resr}^T . A weight w_i is calculated for each contextual feature f_i with respect to all teachers T_s that interacts with a pedagogical resource $resr$. The algorithm excludes teachers and features from the definitive list according to a given similarity threshold and a weight threshold, respectively. The similarity threshold thr_1 is used to select teachers with similarity score greater than a pre-computed similarity score for a teacher T , while the weight threshold thr_2 is responsible for the extraction of features with a calculated initial weight greater than the average of all weights of the contextual features. The similarity function between two teachers *calculateSimilarity* computes the semantic similarity to ensure an effective representation of the actual context of both teachers. It is computed by calculating the distance between similar teachers using TSCCO contextual instances. This semantic similarity algorithm is repeated to get the matching list of peer teachers *matchingList* in Algorithm 6.

In order to enforce the sentimental state of the peer recommendation, a coefficient $coef_{+ve}$ is multiplied by the calculated similarity for similar teachers T_s to prioritize the mood feature (Zheng, Burke, and Mobasher, 2013a). The usage of such coefficient is reflected on the computed initial weight w_i corresponding to the mood feature f_i . On the other hand, if there is few or no results for teachers with a positive mood, the algorithm adapts normally with the new situation, extracting the features and assigning their initial weights.

The initial weight of each feature is determined by averaging the mean semantic similarity scores of all selected teachers Sim^{resr} and the variance of their ratings for the current pedagogical resource $resr$ as follows:

$$Sim_{var}^{resr} = \frac{\sum (r_i^{resr} - \overline{r^{resr}})^2}{nb_T} \quad (6.8)$$

$$w_{kinit} = \frac{Sim^{resr} + Sim_{var}^{resr}}{2} \quad (6.9)$$

where r_i^{resr} is the rating for selected teacher i for the pedagogical resource $resr$,

Algorithm 8 ContextualFeatureExtraction algorithm

Input: $C_{resr}^T = (T, F, rating, resr)$: current context of a teacher T and concerning the current resource $resr$ with rating r , represented by context factors F .
 H^T : history of current teacher T .
 thr_1 : given similarity threshold.
 thr_2 : given weight threshold.
 $coef_{+ve}$: positive mood contextual factor enforcing coefficient.

Output: $W_{init} = \{ \langle f_1, w_{1_{init}} \rangle, \dots, \langle f_{n'}, w_{n'_{init}} \rangle \}$: set of initial weights $w_{i_{init}}$ for each of the n' context factor f_i .

begin

- 1: **for all** $f_k \in F$ **do**
- 2: $nb_T = 0$;
- 3: $Sim_{total} = 0$;
- 4: **for all** $\langle resr_j, c_j^{T_s} \rangle \in H^T$ **do**
- 5: **if** $(resr_j == resr)$ **then**
- 6: $similar(T_s, T) = \text{calculateSimilarity}(C_{resr}^T, C^{T_s})$;
- 7: **if** $(similar(T_s, T) \geq thr_1)$ **then**
- 8: $nb_T ++$;
- 9: $rating_{T_s} = \text{getRating}(T_s, resr)$;
- 10: $similarTeachersList.add(rating_{T_s})$;
- 11: **if** $(f_k == 'mood' \&\& (\text{getMood}(T_s) == 'positiveMood' || \text{getMood}(T_s) == 'neutralMood'))$ **then**
- 12: $similar(T_s, T) *= coef_{+ve}$;
- 13: **end if**
- 14: $Sim_{total} += similar(T_s, T)$;
- 15: **end if**
- 16: **end if**
- 17: **end for**
- 18: $sim^{resr} = \frac{Sim_{total}}{nb_T}$;
- 19: $sim_{var}^{resr} = \text{getVariance}(similarTeachersList)$;
- 20: $w_{k_{init}} = \text{average}(sim^{resr}, sim_{var}^{resr})$;
- 21: **if** $(w_{k_{init}} \geq thr_2)$ **then**
- 22: $W_{init} = \text{addFactor}(\langle f_k, w_{k_{init}} \rangle)$;
- 23: **end if**
- 24: **end for**
- 25: **return** W_{init} ;

end

$\overline{r^{resr}}$ is the average rating for this resource, nb_T is the number of selected teachers.

SWRL rules are used to sort the peer list of each teacher that is generated from *contextMatching* function, as shown in Algorithm 9. For each teacher T_i in the generated list *matchingList*, the rules, as illustrated in Table 4.1, are applied to reason the A-Box instance of this teacher. A total score is calculated by counting the number of satisfied rules except for rule 5 with double counts, as it is responsible for the sentimental state detection. Then, the list is sorted according to a descending order of the obtained scores. At the end, the teacher is asked for the preferred level of expertise for the recommended peer.

Some of the rules in Table 4.1 are used in the recommendation approach. The fifth rule detects if the compared teacher's years of experience is more than the current one to determine the knowledge and beneficial level for this match. The sixth rule checks if the recommended teacher has a better sentimental state (mood) than the current teacher. This rule enforces the positive influence of the sentimental state of one person on another person or an organization (Corson, 2002). The seventh rule matches the teaching styles of both teachers as it is

Algorithm 9 SWRSort algorithm**Input:** T_0 : current teacher $MatchingTeacherList$: list of all matching teachers to current teacher T_0 $R_{SWRL} = \{r_1, \dots, r_8\}$: set of SWRL rules thr_{swrl} : given SWRL rules score threshold**Output:** $sortedList$: sorted list of matching teachers to current teacher T_0 with respect to the obtained SWRL rules score**begin**

```

1: MatchingTeacherList = { $\phi$ };
2: for all  $T_i \in MatchingTeacherList$  do
3:    $score_{T_i} = 0$ ;
4:   for all  $r_j \in R_{SWRL}$  do
5:      $result = applySWRL(r_j, T_i, T_0)$ ;
6:     if  $T_i \in result$  then
7:        $score_{T_i} ++$ ;
8:     end if
9:   end for
10:  if  $score_{T_i} \geq thr_{swrl}$  then
11:     $MatchingTeacherList[T_i, score] = score_{T_i}$ ;
12:  else
13:     $MatchingTeacherList.delete(T_i)$ ;
14:  end if
15: end for
16:  $sortedList = MatchingTeacherList.sortByScore()$ ;
17: return  $sortedList$ ;

```

end

said the difference in styles has its impact on the collaboration between two peer teachers (Pratt et al., 2017; Pratt, 2014). The eighth rule checks the spoken languages by the two teachers which facilitates the collaboration and enhances the communication between them.

The living and working environments have a great impact on the personality of a teacher as well as the working methodologies (Dorman, Kennedy, and Young, 2015). Therefore, the ninth rule checks if there is a matching between the type of the two teachers' living environment, such as rural, urban, sub-urban, etc., while the tenth rule matches the country where each teacher lives in. The eleventh and twelfth rules compare the working environment in terms of the environment type and the educational institution level. The final rule checks if the two teachers belong to the same field of science. For each teacher fulfilling a rule, his score is increment by one. At the end, the list is sorted according to the final score. These rules work as an enforcer for the list order but are not used to exclude any of the recommended teachers.

Part IV

Experimentation and Results

Chapter 7

Experimentation

In this chapter, we present the experimental configuration to obtain results from a series of experiments conducted to evaluate the effectiveness of our approach throughout the different components of dataset, feature selection, and recommender systems.

Our own dataset is used throughout the entire experimentation process which arises the need to ensure the completeness of the dataset as well. In order to validate the answer provided in Chapters 4 and 5 of our first research question (Section 1.1.2), we use the ontological representation of the multiple contexts of teachers throughout all experiments.

An evaluation process is required for the feature selection approach as well as the multiple contexts choice which is essential as the recommendation performance is dependent on it. The illustrated experiments allow thorough analysis of the Variance-based Importance Feature Selection for Teacher Contexts (VIFSTC) that represented in Chapter 6 in terms of representability of the multiple contexts of teacher, and effectiveness and impact on the identifying of teachers.

In addition, we illustrate the required experiments to validate the answers that offered in Chapter 6 to the second research question (Section 1.1.2). Accordingly, we present the evaluation criteria for the context-aware pedagogical resources recommender approach (CAPRRA) against different setups such as pre-filtering and post-filtering approaches.

This chapter is organized as follows: Section 7.1 explores the used data and their retrieval. Additionally, the experimental setup of the mood detection is presented to evaluate the precision of the selected methodology. Section 7.2 introduces the feature selection experimental setup and the evaluation methodology. Section 7.3 shows the experimental setup to test the recommender approach and evaluate its performance for CAPRRA approach and the different contextual filtering models.

7.1 Experimental Data

In these experiments, data is categorized into two main types: teacher-related and resource-related, as shown in Figure 7.1. The resource-related data (Chapter 5, Section 5.4) are collected from different sources and covers several categories of pedagogical resources. On the other hand, the teacher-related data are handled to represent the multiple contexts along with teacher profiling and feedback tracking. The contextual data are divided into two types: environmental and educational institutions data. The environmental data covers the whole

French territories and provides summarized information about the population, area type, and financial conditions for all French provinces and cities. The environmental data include 34,142 out of 36,681 metropolises that are distributed over the French communes. These data are integrated with the 466,357 educational institutions from French prioritized education network.

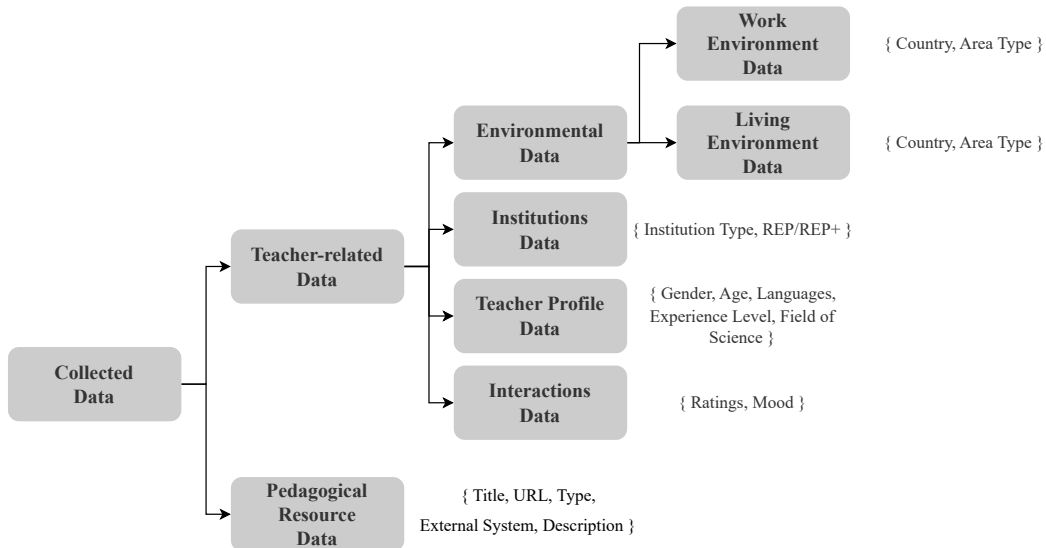


Figure 7.1: The hierarchy of the used data.

The experimentation data of teachers consist of data from the generative interactions of 100 teachers with different contexts to represent different scenarios. Other data for teacher profiling includes demographics, interests, and teaching styles data. Teacher feedback is collected using activity tracking from MEMORAe platform (Chapter 5, Section 5.5.1) and mood detection from Mood-Flow@doubleYou approach (Chapter 5, Section 5.5.2). The experimental teacher data are manipulated according to the purpose of each experiment as follows:

- i All features of all contexts and sentimental state
- ii All features except sentimental state ones
- iii All features except living context ones
- iv All features except working context ones.

SPARQL queries are used to retrieve teacher-related data and pedagogical resource data from mapped ontological data into TSCCO ontologies (Chapter 4). The first SPARQL query example, as shown in Figure 7.2, represents the teacher-related data from the MEMORAe user account `tmcc:login` related to a teacher `ttscco:Teacher`. This query retrieves all data associated with the teacher which are filtered according to each of the four variations of the experimental teacher data. The second SPARQL query example, as shown in Figure 7.3, retrieves all data associated with a pedagogical resource `ttscco:Resource` that exists within an organization `PRE103`. The organization is changeable according to the list of organizations wherein a teacher collaborates with other peers. The retrieved pedagogical resources data are used without any modifications.

Figure 7.2: Example of SPARQL query to retrieve teacher-related data

```

prefix tmcc: <http://www.hds.utc.fr/mcc/tbox#>
prefix ttscco: <http://www.hds.utc.fr/tsscco/tbox#>

SELECT * FROM <http://www.hds.utc.fr/mcc/abox#>
WHERE {
  ?teacher rdf:type ttscco:Teacher;
  ?teacherAccount rdf:type tmcc:User;
  tmcc:login ?login .
  FILTER(?login = "teacher_account")
}

```

Figure 7.3: Example of SPARQL query to retrieve pedagogical resource data within the organization "PRE103"

```

prefix tmcc: <http://www.hds.utc.fr/mcc/tbox#>
prefix ttscco: <http://www.hds.utc.fr/tsscco/tbox#>

SELECT * FROM <http://www.hds.utc.fr/pre103/abox#>
WHERE {
  ?resource rdf:type ttscco:Resource;
  tmcc:ResourceID ?resID .
  FILTER(?resID = "resource_ID")
}

```

In order to validate the effectiveness of mood monitoring, we asked two male teachers to undergo a sequence of search processes while Moodflow@doubleYou platform monitors their mood. Teacher A is a novice teacher aged under 30 years old with less than 5 years of experience while Teacher B is an intermediate teacher over the age of 30 with 5-10 years of experience. Both teachers are in the field of computer science. The two teachers are staying in two separate rooms and are trying to search for resources to illustrate some concepts in their courses. The two teachers are asked to use the same four keywords and they get the same list of results for easier comparison. At the end, each teacher stated which search result has come to his satisfaction. Through the results of this experiment in the next chapter (Chapter 8, Section 8.1), we explore the possible usefulness of sentimental state with contexts of a teacher, and we illustrate its integration with other contexts.

7.2 Experimental Setup for VIFSTC Feature Selection

A dedicated experimental setup is configured to evaluate the Variance-based Importance Feature Selection for Teacher Contexts (VIFSTC) where the experiment model is illustrated in Figure 7.4 along with the experimental procedure in Figure 7.5, is designed to compare the VIFSTC approach against other algorithms. Initially, data of teachers is retrieved from the TSSCO ontology using SPARQL

queries as previously stated. Then, the SWRL reasoning rules are applied to all teachers to determine their level of expertise and eliminate the repetition among them. Next, the data is divided into 75% training data and 25% testing data which allows experiments to be conducted to compare the proposed architecture to other filtering-based feature selection methods using the same retrieved data, such as normal variance-based threshold, normal mutual gain information, chi2, r-regression, f-regression and ANOVA-based one.

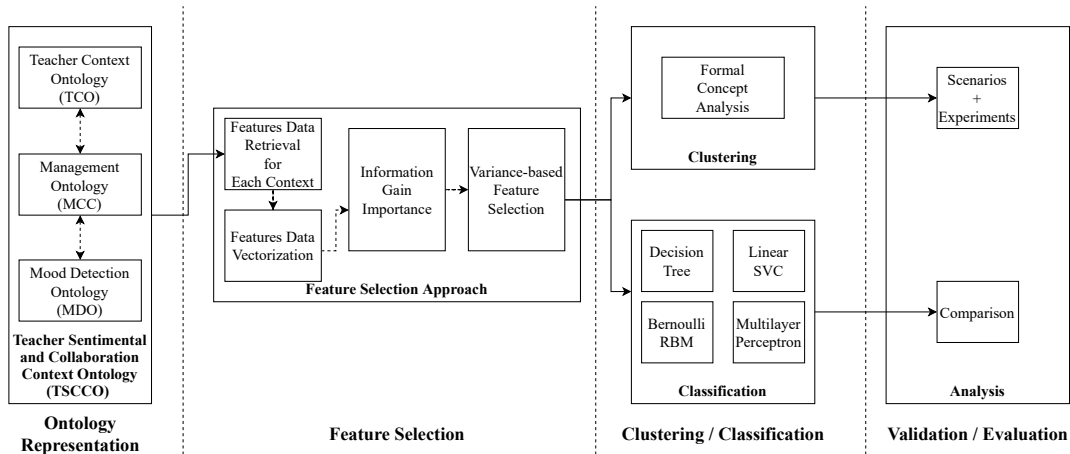


Figure 7.4: Experiment model for VIFSTC approach.

For comparison purposes, data of teachers is pre-labelled and used with multiple classification methods to classify teachers to the predefined classes. This data contains the details of 100 teachers that cover the mentioned categories of features in Table 6.1. For each teacher, there are 46 corresponding features that represent the multiple contexts of teacher. After the initial execution of this approach, it is repeated with each insertion of a new teacher. Afterwards, the selected features are stored and are used for any further processes. For privacy concerns, names and contact information of each teacher are eliminated from the retrieved data and are not considered in the VIFSTC approach. Finally, the training data is used to train the classifiers: Decision Tree Classifier, Linear Support Vector Classification (SVC), Bernoulli Restricted Boltzmann Machine (RBM), and Multilayer Perceptron (MLP). The described experiment is performed for each variation of data and the obtained results are discussed in the next chapter.

7.3 Experimental Setup for Recommender System

Another experimental setup is configured to evaluate the context-aware pedagogical resources recommender approach (CAPRRA) which is compared against other approaches. In this thesis, three types of contextual filtering are integrated to the hybrid 2D recommender approach at different phases of the recommendation process (Chapter 6, Section 6.2.2). Three different phases of contextual filtering are investigated: prefiltering, postfiltering, and mix-filtering approach named CAPRRA. Initially, reduced data of teachers are retrieved from the VIFSTC feature selection phase as previously stated. Then, the reduced data are used to conduct two types of experiments: scenario, and accuracy evaluation. A scenario

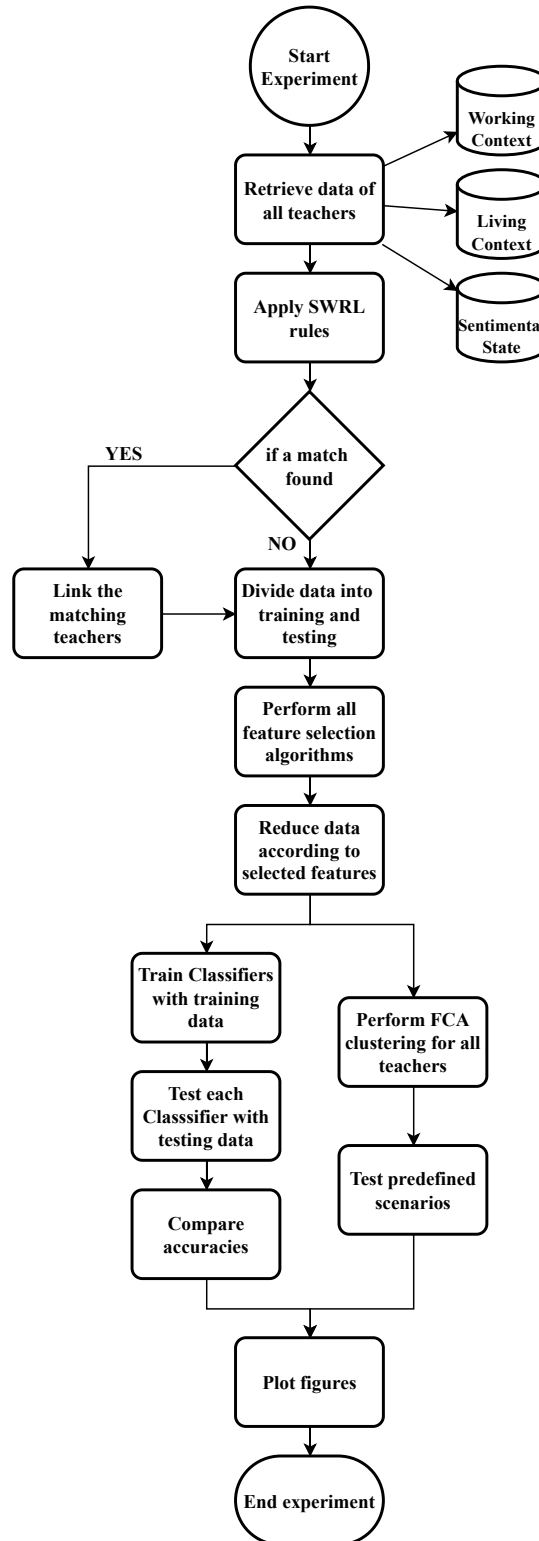


Figure 7.5: Flowchart of the experimental procedure for VIFSTC approach.

is designed to evaluate the peer teacher recommendations where teachers collaborate with others according to the set of reasoning rules as stated in Algorithm 6. Scenarios provide a predictable and controlled environment for experimentation.

This control ensures that variations in peer recommendations can be attributed to the specific changes made within the scenario, rather than external factors. On the other hand, evaluation metrics should be carefully selected to precisely evaluate the performance of the pedagogical resource recommendations.

The interactions history of teachers is divided into two subsets: 80% of interactions for recommendation process and 20% of interactions for evaluation process. F1 score, root mean squared error (RMSE), and mean absolute error (MAE) are among the commonly used evaluation metrics for recommender systems (Lü et al., 2012). However, RMSE and MAE measure the accuracy of predicted ratings against the actual ratings which do not align with the objective of the pedagogical resource recommendations (Lü et al., 2012). The success of a pedagogical resource recommender system depends on suggesting resources that match preferences and contexts of teachers. RMSE and MAE lack the ability to capture the nuances of teacher preferences and contexts.

F1 score serves as a more suitable evaluation metric for recommender systems as it considers the balance between precision and recall (Lü et al., 2012). The F1 score has the ability to balance the trade-off between precision and recall which harmonizes the significance of both false positives and false negatives, aligning perfectly with the core objective of a recommender system (Lü et al., 2012). Therefore, the choice of F1 score is well-suited for CAPRRA due to its incorporation of precision and recall, as both aspects are crucial in the educational context. Precision ensures that the pedagogical resources recommended are highly relevant, reducing the risk of overwhelming teachers with irrelevant content. Recall, on the other hand, guarantees that a substantial portion of relevant resources is not overlooked. F1 score can be computed as follows:

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7.1)$$

Where:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad \text{and}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}.$$

Throughout this evaluation process, the terms True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) are widely used in the precision and recall calculations. True positives (TP) are pedagogical resource recommendations where the recommender system correctly recommends relevant pedagogical resources to teachers. False positives (FP) are the cases where the recommender system incorrectly recommends resources that are not relevant to the multiple contexts of teachers. True negatives (TN) are the pedagogical resources that are not recommended as they are not relevant to a teacher. False negatives (FN) are the cases where the system fails to recommend relevant pedagogical resources that teachers would find beneficial where they coexist in the 20% subset of the interactions history.

The experimentation of the different recommender approaches is divided into two types: evaluation with different number of recommendations, evaluation with the four variations of the data according to contexts as previously mentioned. The next chapter discusses the results of these experiments in addition to the discussion of such results.

Chapter 8

Results

This chapter focuses on the experimental results and findings of the experimentations regarding the different phases of the pedagogical resources recommender system for teachers. Section 8.1 discusses the behavior and mood of teachers during the experiments. Mood experiments are evaluated according to the time spent and resources selected by different teachers for different keywords where the observed mood and behavior patterns are discussed.

Section 8.2 discusses the selection of appropriate features for representing and distinguishing teachers. It describes the experimentation process and the impact of different contexts on the accuracy of the selection. The effectiveness of the VIFSTC feature selection approach is evaluated and compared to other methods.

The last Section 8.3 discusses the experimentation related to the pedagogical resources recommender system for teachers. The evaluation of the context-aware pedagogical resources recommender approach (CAPRRA) is discussed by comparing its results to three different approaches 2D recommender, prefiltering recommender, and postfiltering recommender. Using F1 score as an evaluation metric for the pedagogical resource recommender system is a suitable choice. F1 score balances the trade-off between precision and recall, which is crucial in an educational context. Additionally, the evaluation is performed according to different experiments with different number of recommendations, and different variations of the data. Lastly, we discuss the scenario evaluation to evaluate the performance of peer recommendations.

8.1 Experiment with Mood Detection

Teachers A and B took 56 and 54 minutes respectively to finish the experiment which are approximately the same amount of time, as shown in Table 8.1 and Table 8.2. However, we observe that Teacher B spent a balanced amount of time searching for each keyword. On the contrary, Teacher A spent 30 minutes for the first keyword and afterwards, he was in a hurry to finish the subsequent keywords. The choices of each teacher are used as a reflection of his experience. Additionally, it is shown that Teacher A preferred to select only one resource while Teacher B selected multiple resources in three search operations.

The detected mood explains the search and choice behavior of both teachers. Worthwhile, a person is exposed to a negative or lower-positive mood if that person accomplished a task or reached a goal. For example, the time spent by Teacher A to search for keyword 1, is not necessary because he is already

Table 8.1: Experiment summary for Teacher A

Keyword (#)	Time (minutes)	Selected resource (index)	Accumulative mood (level)
1	30	1	Neutral
2	14	5	Positive
3	7	4	Neutral
4	5	3	Negative

Table 8.2: Experiment summary for Teacher B

Keyword (#)	Time (minutes)	Selected resource (index)	Accumulative mood (level)
1	20	2,3	Neutral
2	12	3	Neutral
3	12	1,2	Positive
4	10	2,4,5,6	Positive

satisfied with the first search result after 3 minutes only as shown in Fig. 8.1. The wasted time is an indication of the inexperienced behavior of Teacher A. Also, this chosen result is the general explanation of multiple concepts and it is not inclusive of keyword 1. On the other hand, Teacher B took his time to explore the list of resources and stopped the search process after he reached more than one satisfying result for keyword 1 as shown in Fig. 8.2. Therefore, more experienced teachers tend to efficiently evaluate the provided resource recommendations which saves time and effort, while less experienced ones struggle to act similarly by consuming more time to reach one acceptable result.

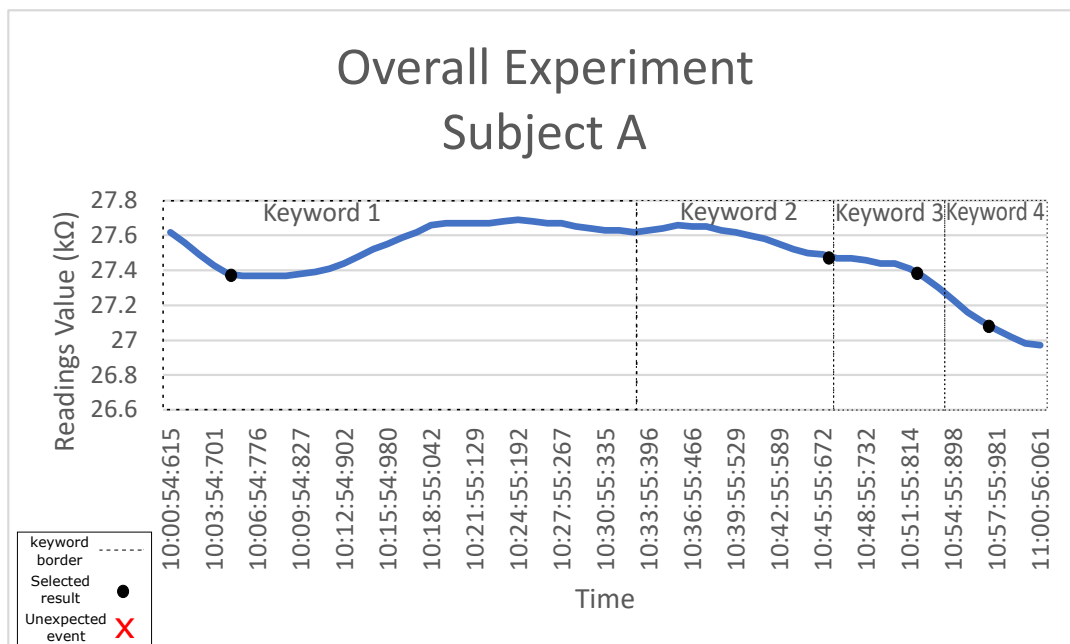


Figure 8.1: Moodflow@doubleYou output for Teacher A.

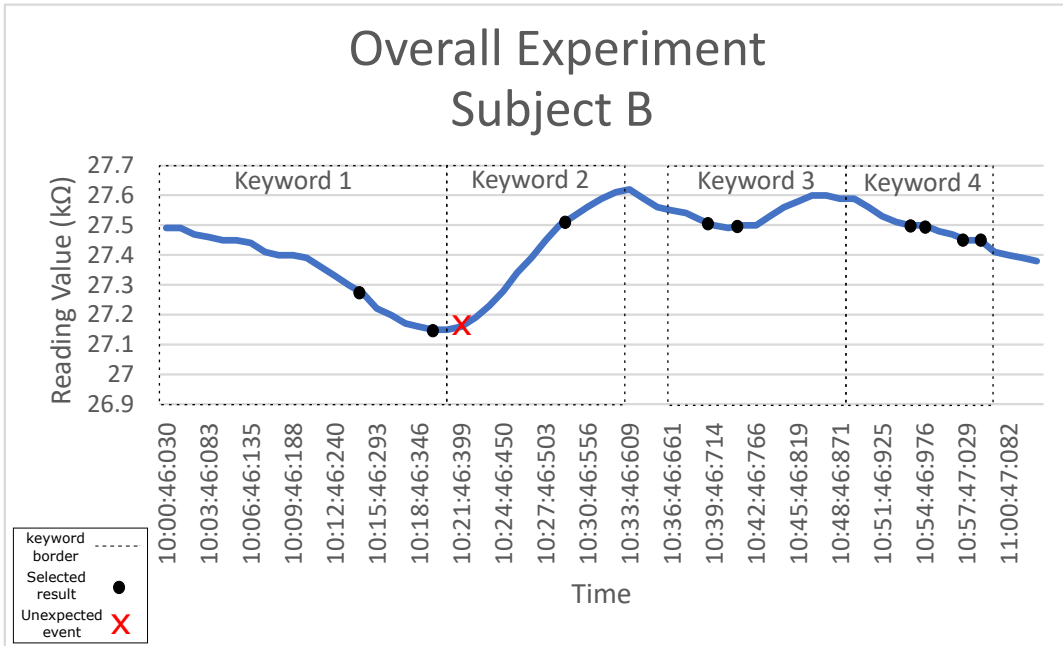


Figure 8.2: Moodflow@doubleYou output for Teacher B.

The level of experience helps each subject to achieve a state of emotional commitment. Teacher A achieved an average absolute deviation AAD of 0.024 for the four keywords while Teacher B achieved an AAD of 0.0208. Accordingly, we can say that Teacher B was more committed to the assigned task than Teacher A. The mood is affected by unexpected events, creating unexplained negative mood in the detected measurement. We can detect an undeniable observation for Teacher B at the beginning of the search process of keyword 2. He stated that this event affected his mood negatively which provides a better representation of the actual context of Teacher B. Through mood detection, we can state that the mood of a teacher is directly proportional to his level of experience. Throughout this experiment, we conclude that mood integration while providing the context of a teacher, allows an expressive description of his actual feedback. The obtained observations can explain behavior and recommendation choices but knowing the mood of each teacher summarizes this process.

Hence, the integration of the sentimental state of teachers, as shown in Figure 8.3 and Figure 8.4, offers an opportunity to provide recommendations according to realistic feedback of teachers. For example, a young teacher, who is searching for a specific educational resource, will be able to find simpler resources but if his mood is in the negative region, the provided results can be more personalized and mind-teasing to enlighten his mood.

8.2 Experiment with VIFSTC Feature Selection

The main objective of VIFSTC approach is to select the appropriate features to represent and distinguish one teacher from another which facilitates the recommendation process. In order to achieve this objective, Section 8.2.1 illustrates the experimentation with different contexts while Section 8.2.2 discusses the obtained results.

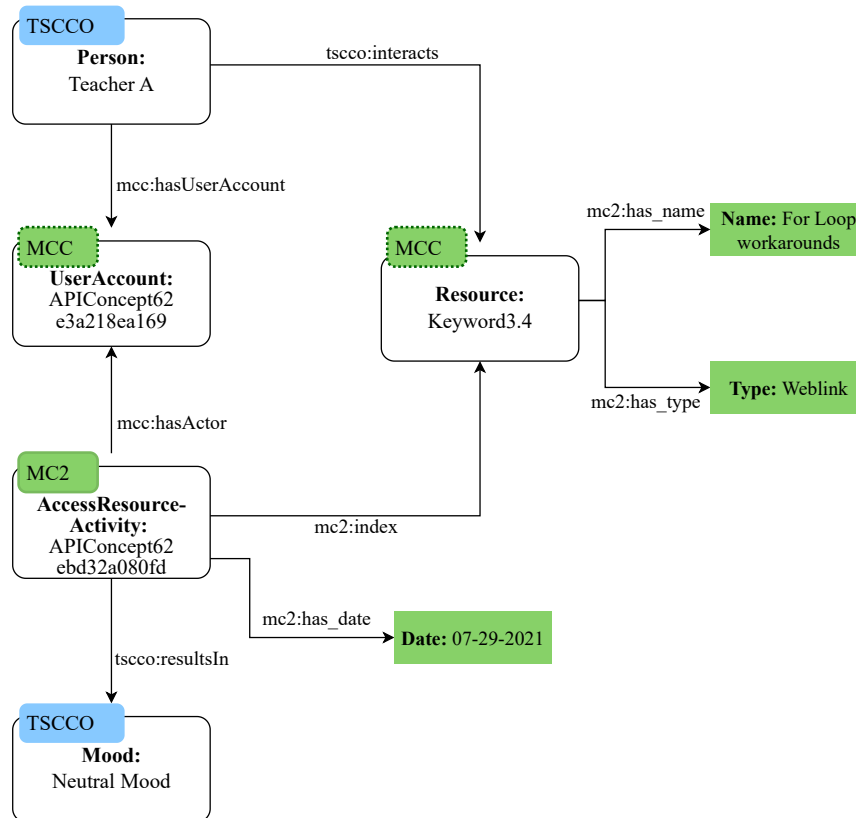


Figure 8.3: A-Box example of integrating the mood of Teacher A with TSCCO.

8.2.1 Experiments with different contexts

At the beginning, the correlation coefficient is computed for all features, as shown in Figure 8.5. The illustrated heatmap shows that there are dependencies between features which should be eliminated by the feature selection algorithms. Afterwards, the information gain importance is calculated for all features, as shown in Figure 8.6. For example, the feature number 12 "lang1=English", represents English language as the mother tongue of a teacher, is dependent on several other features, in addition to the obtained low importance using the information gain method. Therefore, this feature is eliminated by all feature selection methods except for the variance-based threshold method according to the used scoring algorithm, as shown in Figure 8.7. After the above data analysis, the VIFSTC approach is applied to the dataset as well as the other methods.

The dataset is reduced according to the selected features, as shown in Figure 8.7, corresponding to each method, then, the output data is classified using various classification methods. Figure 8.8 represents the scored accuracy of each feature selection method with the various classification methods. The results show that the VIFSTC approach outcores the other approaches across two classification methods: decision tree classifier and Bernoulli RBM with accuracies 80.76% and 15.38% respectively. The obtained accuracies with linear SVC and multilayer perceptron classifiers are positioned at second best with 65.38% and 53.65% respectively. When the contextual features related to sentimental

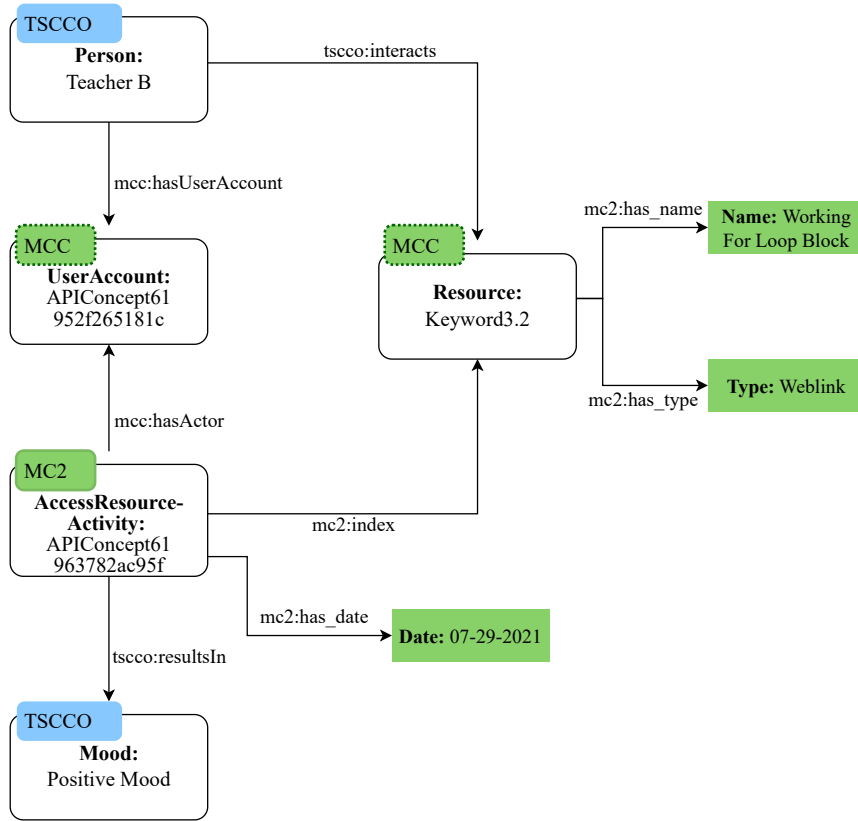


Figure 8.4: A-Box example of integrating the mood of Teacher B with TSCCO.

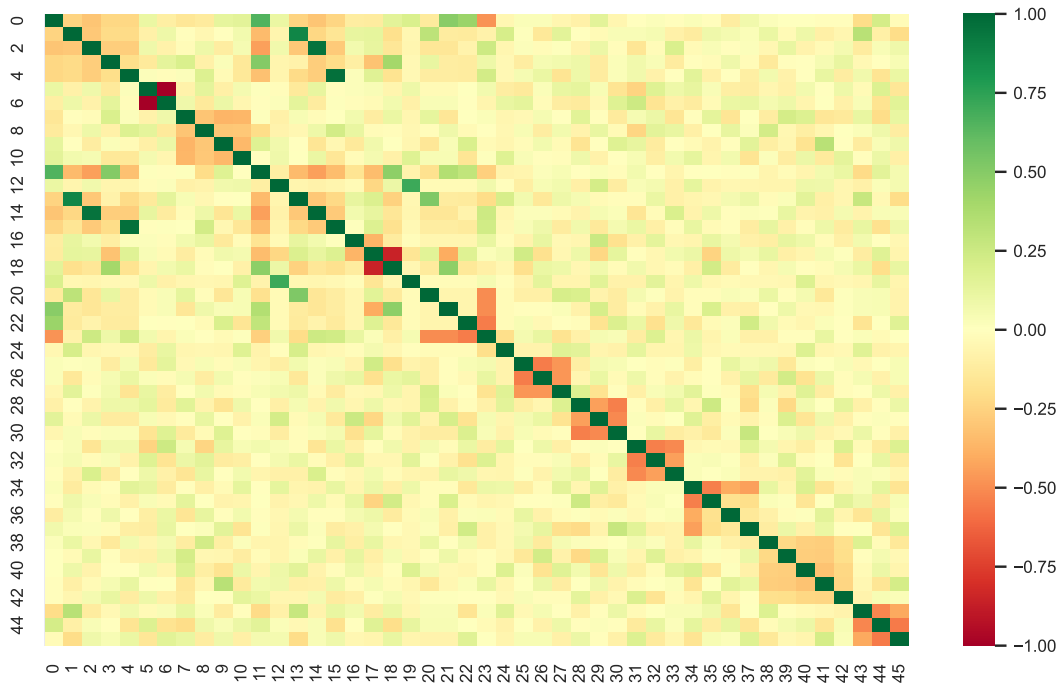


Figure 8.5: Correlation coefficient of all features.

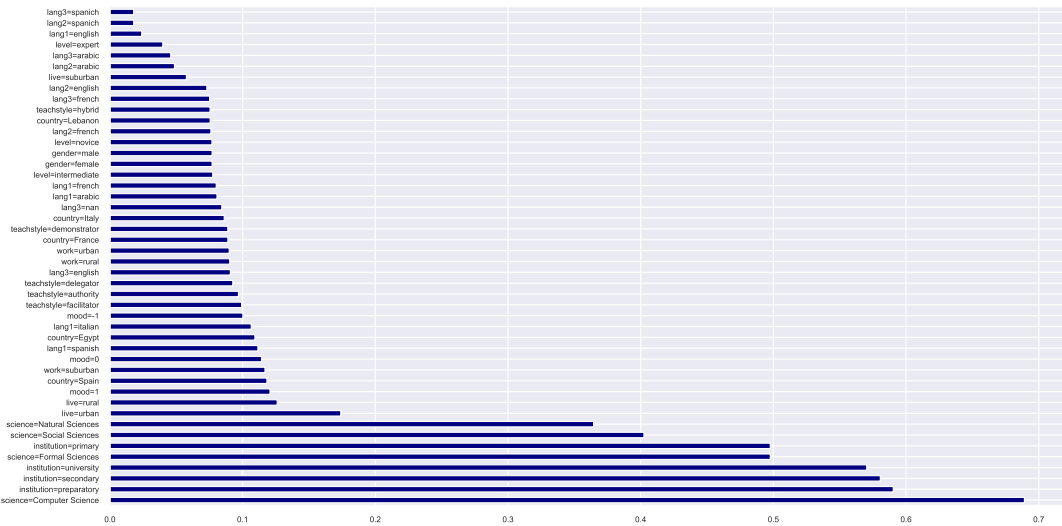


Figure 8.6: Information gain importance for all features.

state are omitted, the accuracies of the four classifiers are affected by an average decrement of 1.9% for the VIFSTC approach. The other FS approaches are greatly affected by the absence of sentimental state with an average decrement of 3.52% across all FS approaches, as shown in Figure 8.9. Variance-based threshold and ANOVA are the most affected with an average decrement of 5.77% and 8.65% respectively. The absence of features related to the living context, greatly affects the performance of the VIFSTC approach with an average decrement of 3.84% while other FS approaches are not equally affected with an average decrement of 0.96% as shown in Figure 8.10. The absence of features concerning working context has the most noticeable impact on the performance of all classifiers with an average decrement of 34.5% across all approaches as shown in Figure 8.11. The average overall performance of the VIFSTC approach outperformed the other approaches with approximately 1.5% as shown in Figure 8.12. The verdicts of these results are discussed in the following discussion subsection.

8.2.2 Discussion of Feature Selection Experiments

The discussion of the proposed approach is illustrated using three points: importance of ontological representation, effectiveness of feature selection, and functionality of FCA clustering. The ontological representation before preprocessing of data helps ensuring the consistency of data and its standardization. Also, it can semantically enrich the quality of data by finding relations between different concepts. For example, in our knowledge base, the ontological representation comprises concepts such as "teacher", "living environment", "working environment", and "mood", as well as their relationships (e.g., "a teacher, who lives in a certain living environment, teaches in a certain working environment and has a certain level of mood"). By developing an ontological representation, it becomes simpler to find and analyze patterns in data, as well as take personalized decisions regarding the preprocessing and analyzing data meaningfully. Moreover, ontology provides an acceptable level of automation by transformation and integration of data during the preprocessing phase. In the context of

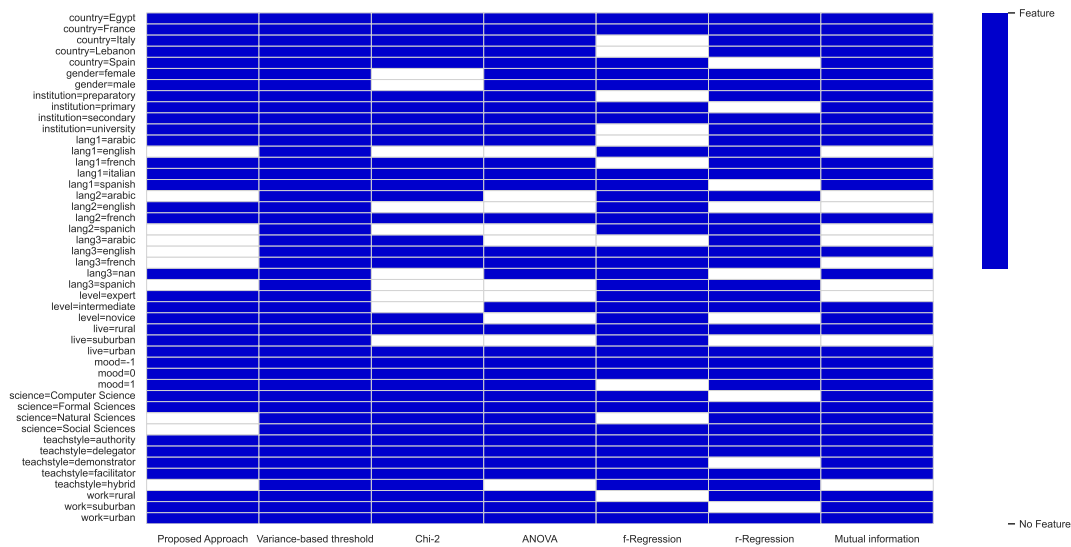


Figure 8.7: The selected features across all methods.

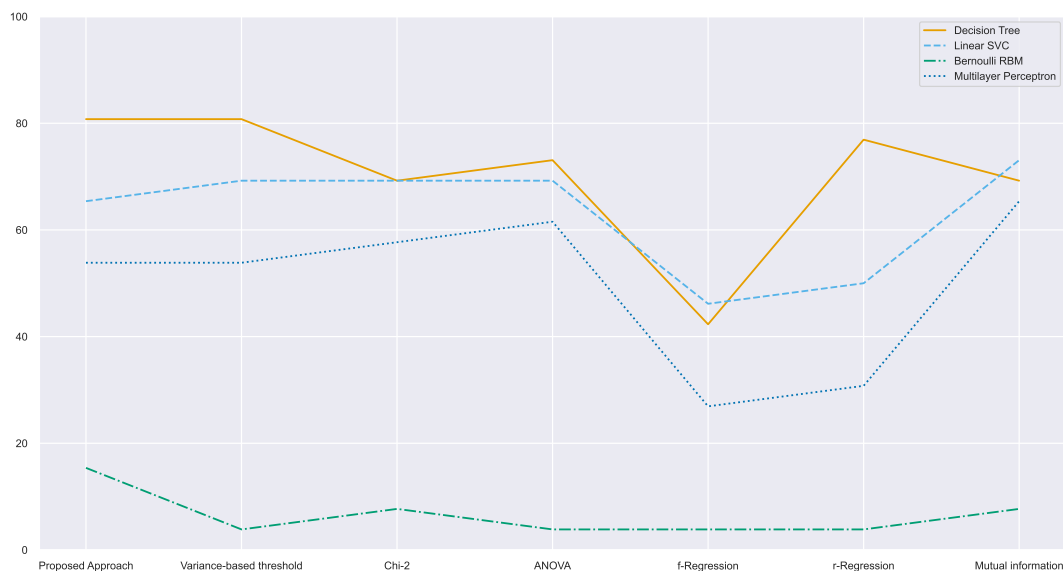


Figure 8.8: Classification accuracy of all contexts across all feature selection methods.

teacher, ontology is used to transform data from different forms and integrate it into a unified ontological representation. SWRL rules are used to formalize some concepts such as the level of experience of a teacher as introduced in Table 4.1. Additionally, it is used to reduce the computational overhead associated with the following phases by preventing repetition of the insertion process for each new teacher. The last SWRL rule in Table 4.1 finds if the context of a new insertion already exists in the dataset. Afterwards, if this teacher exists, the new teacher is linked with the existing one which stops the execution of the remaining steps.

The effectiveness of the proposed feature selection approach is evaluated by

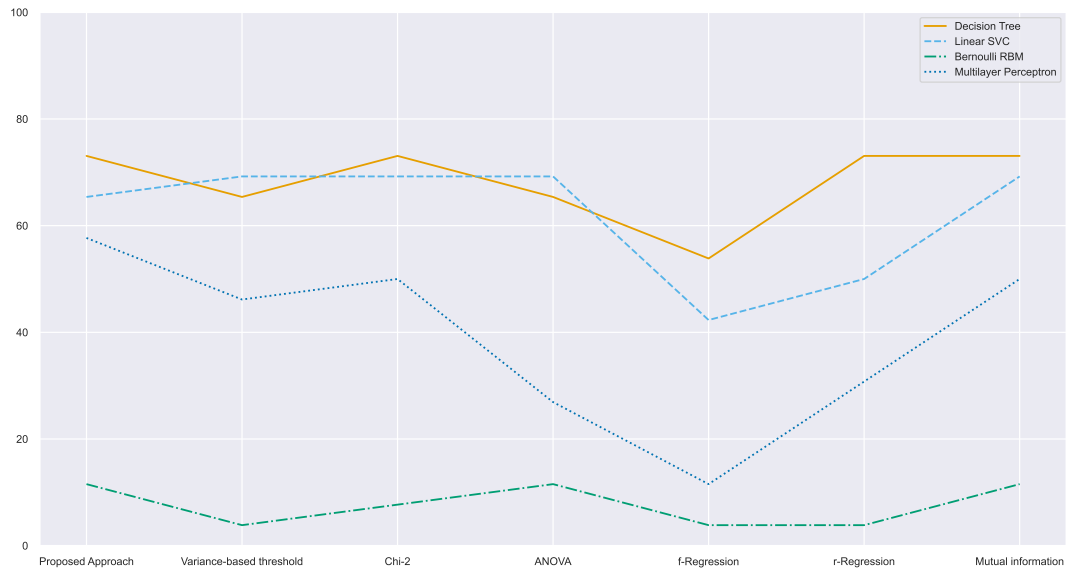


Figure 8.9: Classification accuracy without sentimental context across all feature selection methods.

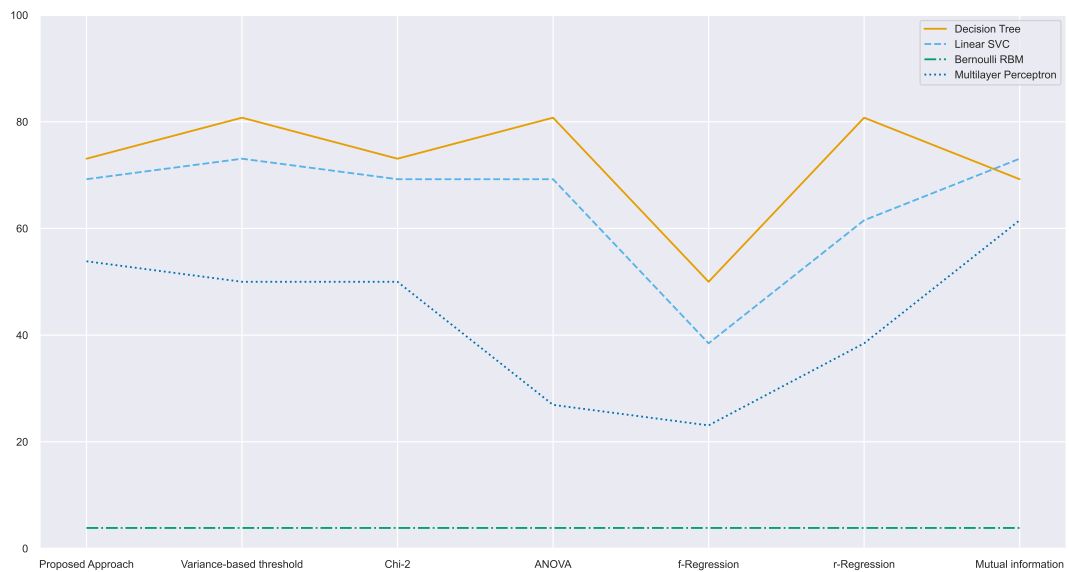


Figure 8.10: Classification accuracy without living context across all feature selection methods.

two types of evaluation: targeted scenarios and accuracy of classification. Evaluation by experience assesses these scenarios by checking consistency of the selected features with their importance in real life. Accordingly, FCA clustering evaluates the consistency of these scenarios by finding proper intentions and extensions between selected features and teachers respectively as shown in Figure 8.13. During this evaluation phase, some observations are made that need to be highlighted concerning the importance of some features during the selection process. The sentimental state of a teacher is an essential criterion to be considered due to the recent developments in this modern era (Nashed,

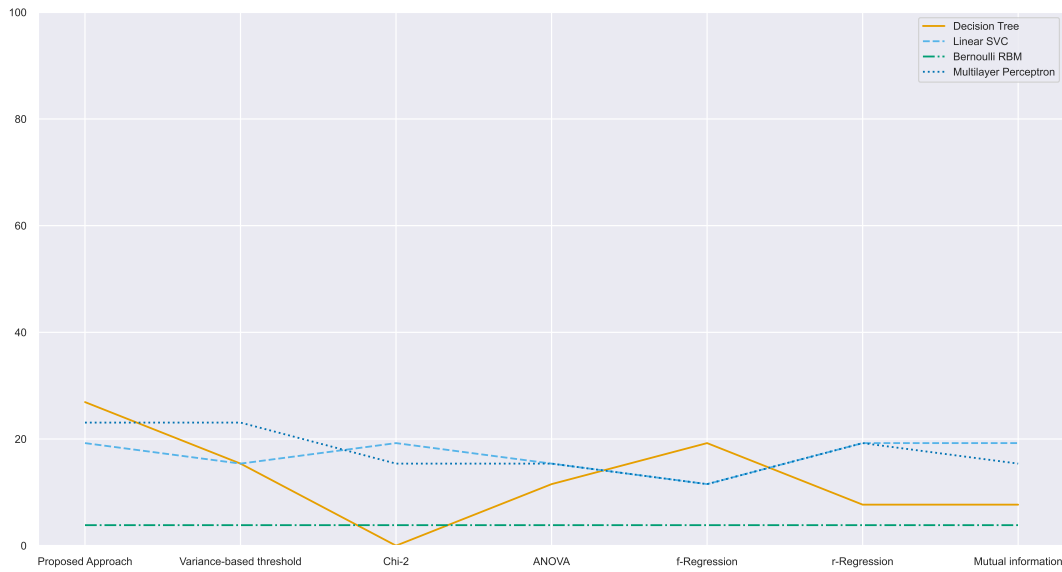


Figure 8.11: Classification accuracy without working context across all feature selection methods.

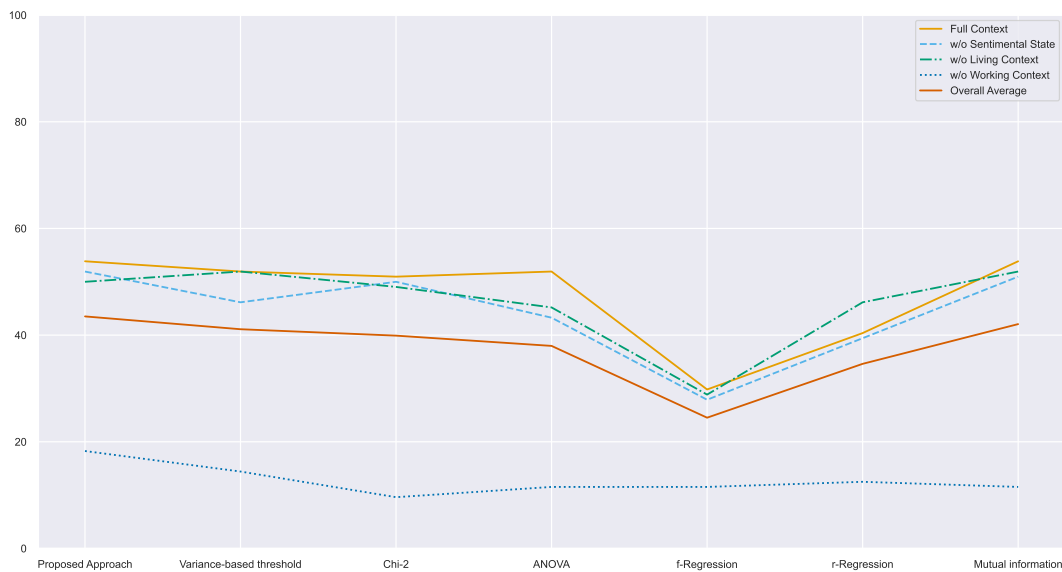


Figure 8.12: Average classification accuracy across all feature selection methods.

Lahoud, and Abel, 2022). Therefore, features related to the mood of a teacher should be prioritized over less-important features. The proposed approach succeeds in selecting mood-related features while other approaches fail to achieve that such as r-Regression, as shown in Figure 8.7. The elimination of the hybrid teaching style is another example of the effectiveness of the proposed approach, as it is not a unique style that is used to distinguish one teacher from another. Only mutual information and ANOVA eliminate this feature from all the experimented approaches along with the proposed one. The features of working context are equally important as shown by the obtained classification accuracies in

Figure 8.11. However, most of the compared FS algorithms fail to select all these features while our approach succeeds in this task. Therefore, the evaluation by experience shows that our approach successfully selects essential features to describe contexts of a teachers.

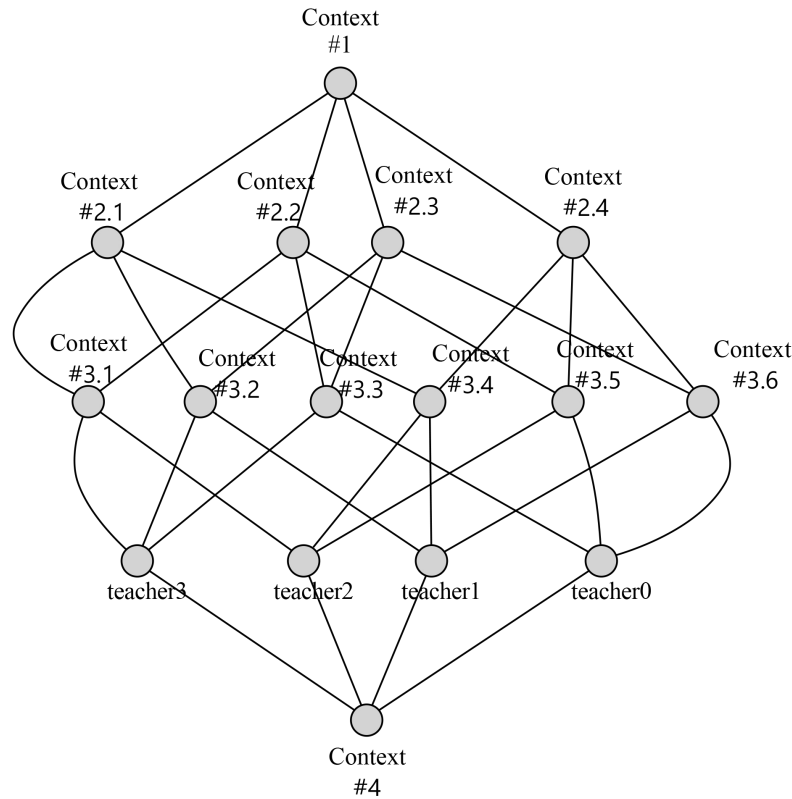


Figure 8.13: FCA generated concepts lattice for the contexts of 4 teachers.

The usage of classification methods acts as a universal metric for evaluation and testing of the proposed method against other feature selection methods. As previously illustrated in Section 8.2.1, the effect of each context can be measured by obtaining the average difference of accuracy across all classifiers for full dataset and the other three variants. These average differences are 1.37%, 3.3%, and 34.75% for the dataset without living context, sentimental state, and working context respectively. We can notice that the performances of all classifiers are highly impacted by the absence of the working context. This observation highlights the fact that features of the working context cannot be eliminated during feature selection. All compared FS approaches fail to guard features of working context while the proposed approach selects all of them, as shown in Figure 8.7.

At the end, FCA clustering is performed to construct an ontology-like lattice structure which is used to interpret the results of FS approaches. Moreover, the generation of such lattice structures enhances the performance of recommender systems (RSs) and in this case, teacher-centered ones. FCA profiles a teacher based on available characteristics which can be used to personalize recommendations for these teachers. One of the benefits of FCA is finding incorporating

contextual information by identifying common features during a certain context such as negative mood or specific working area which can help context-aware recommender systems (CARS) to increase the level of personalization for offered recommendation. In general, FCA clustering is useful for RSs and especially CARS which needs more experiments to measure its effect on the level of personalization for teachers.

8.3 Experiments with Recommender System

In this section, we delve into the results of the experimental phase of evaluating the performance of the context-aware pedagogical resources recommender approach (CAPRRA). Through a series of experiments, we assess the effectiveness of CAPRRA in providing relevant pedagogical resource recommendations to teachers based on varying conditions. Specifically, we investigate CAPRRA performance under different numbers of recommendations and explore how different variations of contextual factors influence its recommendations. Additionally, we present experiments conducted within peer scenarios to showcase the impact of semantic rules on the recommendation process. The section is structured into sub-sections, each focusing on a specific aspect of experimentation and evaluation.

8.3.1 Experiments with Different Number of Recommendations

In the pursuit of evaluating the recommendations provided by the CAPRRA approach, we turn our attention to the investigation of different numbers of recommendations. The number of recommendations a recommender system generates (value of K) can significantly impact the user experience. While a higher number of recommendations K might ensure coverage of diverse resources, it might also risk overwhelming the teacher with choices. Conversely, a lower number of recommendations may provide a more focused selection but could potentially miss out on valuable recommendations.

To explore this trade-off, we conducted a series of experiments where we systematically varied the number of recommendations (K) generated by the CAPRRA approach along with the compared approaches. For each value of K , the system generated recommendations for a group of teachers, and we evaluated the performance using the F1 score. The final F1 score is calculated as the average across all F1 scores of all teachers.

The outcomes of these experiments, presented in Figure 8.14, exhibit trends that provide valuable insights into the behavior of the CAPRRA approach with respect to different numbers of recommendations. As K increases, the precision tends to decrease while the recall generally increases. This phenomenon is anticipated as the approach aims to cover a broader spectrum of resources, which might introduce some false positives (resources that are recommended but not relevant). However, by doing so, the approach enhances its ability to capture more relevant pedagogical resources, thereby increasing recall.

Interestingly, the point at which the F1 score is maximized ($K = 10$) can be identified as an optimal trade-off between precision and recall. This observation guides the determination of the ideal number of recommendations that should be

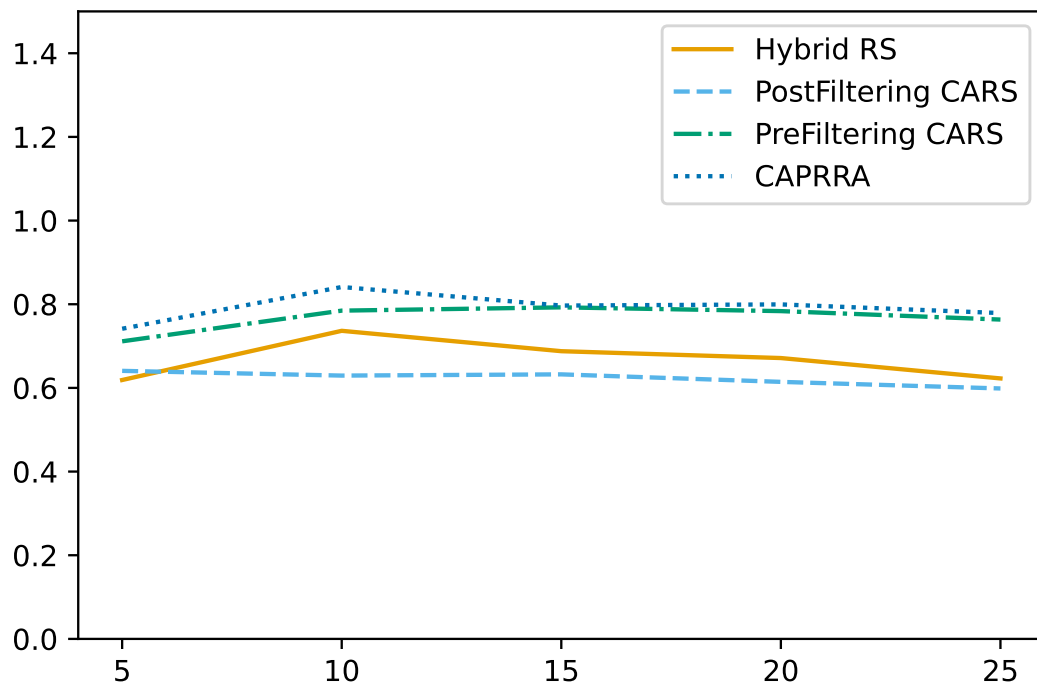


Figure 8.14: F1 scores obtained by the compared recommender approaches with respect to the different number of recommendations K .

presented to teachers to strike a balance between the quantity of recommendations and their relevance.

The Hybrid approach consistently performs competitively across different K values. It demonstrates decent F1 scores, showing its ability to provide relevant recommendations while incorporating multiple strategies. Interestingly, the postfiltering approach performs well at lower values of K (e.g., $K = 5$) with an F1 score of 0.6407. However, its performance drops as K increases, indicating that the postfiltering technique might not effectively handle larger numbers of recommendations. The prefiltering approach consistently performs well, especially at higher K values. This suggests that filtering contextual features in advance based on teacher preferences and contexts can lead to accurate recommendations when more options are provided. The CAPRRA approach, which is the context-aware pedagogical resources recommender system proposed in this thesis, shows competitive performance across various K values. It achieves the highest F1 score in all cases, especially at $K = 10$, demonstrating its ability to leverage contextual information to enhance recommendations.

In conclusion, the choice of the number of recommendations K has a notable impact on the performance of the different approaches. While the prefiltering and CAPRRA approaches consistently demonstrate competitive results, the performance of the postfiltering approach seems to degrade as the number of recommendations increases. The results also suggest that context-aware approaches like CAPRRA can effectively leverage contextual information to enhance the quality of recommendations. These observations highlight the importance of considering the specific number of recommendations K of the pedagogical resource domain when designing and evaluating recommendation systems.

8.3.2 Experiments with Different Context Variations

In this section, we delve into the impact of different contextual variations (as illustrated in Chapter 7) on the performance of the CAPRRA system. Understanding how changes in various contextual features influence the recommendations is crucial for tailoring the approach to accommodate diverse teacher preferences.

To explore this aspect, we conducted experiments using different variations of contextual features. Each variation represented a specific configuration of contexts: living environment, working environment, and sentimental state (mood). Each approach generates recommendations by neglecting one context at a time, and we evaluated the provided recommendations using the F1 score.

The first scenario neglects the contextual features related to the living environment from the recommendation process, its results are illustrated in Figure 8.15. The 2D hybrid approach maintains relatively consistent performance across different K values, indicating a moderate dependency on the living environment context. While the postfiltering approach demonstrates a decrease in F1 scores across different K values, suggesting that living environment context might play a role in refining postfiltering recommendations. However, the prefiltering approach shows a consistent performance with higher F1 scores compared to the other approaches, indicating that it is less affected by the absence of living environment context. Finally, CAPRRA maintains competitive performance across different K values, showcasing its ability to leverage other available contextual factors effectively even when one context is missing.

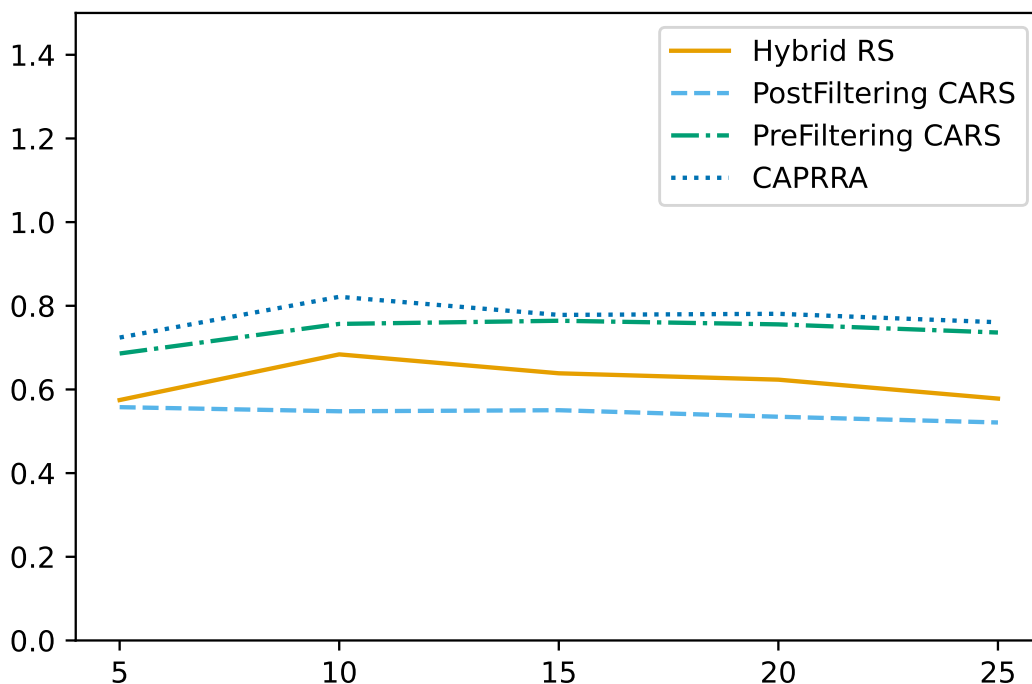


Figure 8.15: F1 scores obtained by the compared recommender approaches while neglecting the contextual features related to the living environment.

The second scenario involves excluding the contextual features related to

working environment, its results are visually represented Figure 8.16. The pre-filtering and CAPRRA approaches have relatively the highest and consistent F1 scores across different K values, indicating that they are less influenced by the absence of working environment context. Additionally, the postfiltering approach demonstrates a decline in F1 scores, suggesting that the working environment context might contribute to refining postfiltering recommendations. Finally, the 2D hybrid approach displays the moderate F1 scores, showing a moderate effect of the absence of the contextual features of working environment.

The elimination of the contextual features related to working environment has a profound impact on the F1 scores obtained from all compared approaches. The working environment encompasses critical features such as the type, level of educational institution, field of science, and specialized education in the institution. These features strongly influence teachers' preferences and behavior within the educational setting. Removing the working environment diminishes the recommendations process ability to tailor recommendations to individual teachers, leading to less relevant and applicable suggestions. The absence of this context also disregards the varying pedagogical resource types, and teaching styles, which are crucial for providing effective and personalized recommendations. Consequently, the resulting F1 scores exhibit a significant drop, underlining the pivotal role of the working environment context in enhancing the accuracy and usefulness of pedagogical resource recommendations. In contrast, by acknowledging and utilizing the working environment context, the recommender approaches can offer more relevant, valuable, and tailored suggestions, enhancing the overall effectiveness of the pedagogical resource recommendations provided to teachers.

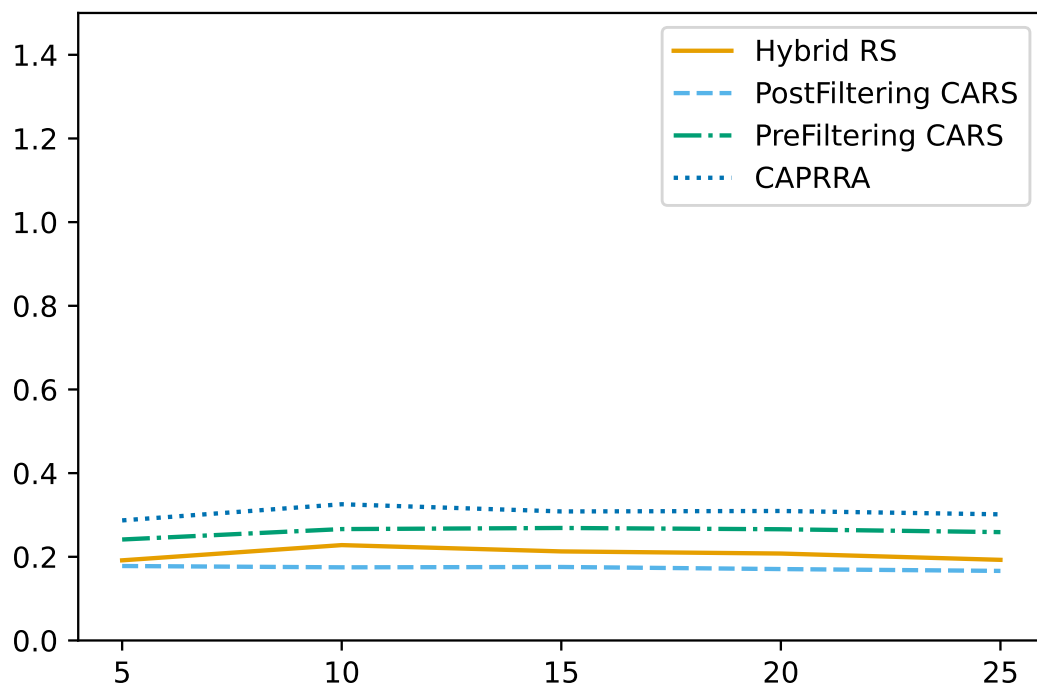


Figure 8.16: FF1 scores obtained by the compared recommender approaches while neglecting the contextual features related to the working environment.

In the last scenario, the contextual features related to sentimental state (mood) are eliminated which results in the F1 scores illustrated in Figure 8.17. The pre-filtering and CAPRRA approaches maintain the highest competitive performance across different K values, implying that they can provide meaningful recommendations even without considering mood. On the other hand, the postfiltering approach demonstrates relatively consistent performance, with a slight drop in F1 scores. This suggests that mood could play a role in enhancing postfiltering recommendations. The 2D hybrid approach showcases moderate F1 scores, indicating the effect of the absence of sentimental state on its obtained F1 score.

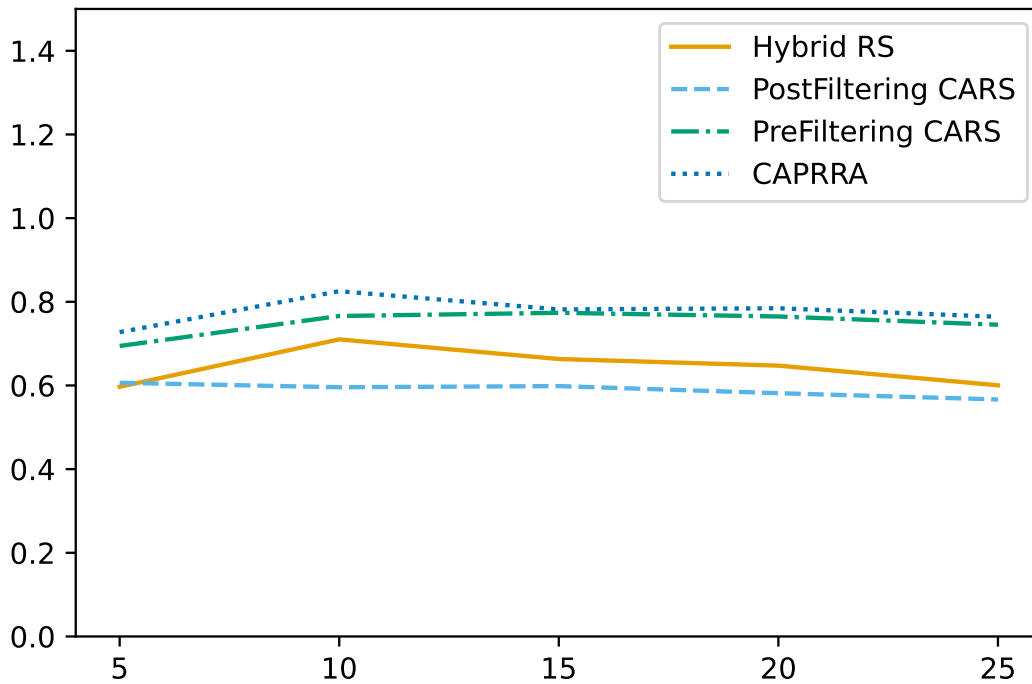


Figure 8.17: F1 scores obtained by the compared recommender approaches while neglecting the contextual features related to the sentimental state.

The average performance of the different recommendation approaches across all experiments, as shown in Figure 8.18, highlights the impact of varying contexts on their effectiveness. Among the approaches, CAPRRA consistently achieves the highest average F1 scores, indicating its robustness in adapting to different context variations. CAPRRA's ability to integrate and leverage multiple contextual features, including living environments, working environments, and sentimental states, contributes to its superior recommendation quality.

The prefiltering approach also exhibits strong performance across the experiments, showcasing its capacity to maintain high F1 scores even when certain context features are removed. This suggests that its initial data filtering process, which focuses on retaining only relevant pedagogical resources, is effective in generating valuable recommendations regardless of contextual variations.

On the other hand, the 2D hybrid approach, which combines both collaborative and content-based filtering, displays moderate average F1 scores. This suggests that while the 2D hybrid approach benefits from combining two recommendation techniques, it might not be as sensitive to context variations as

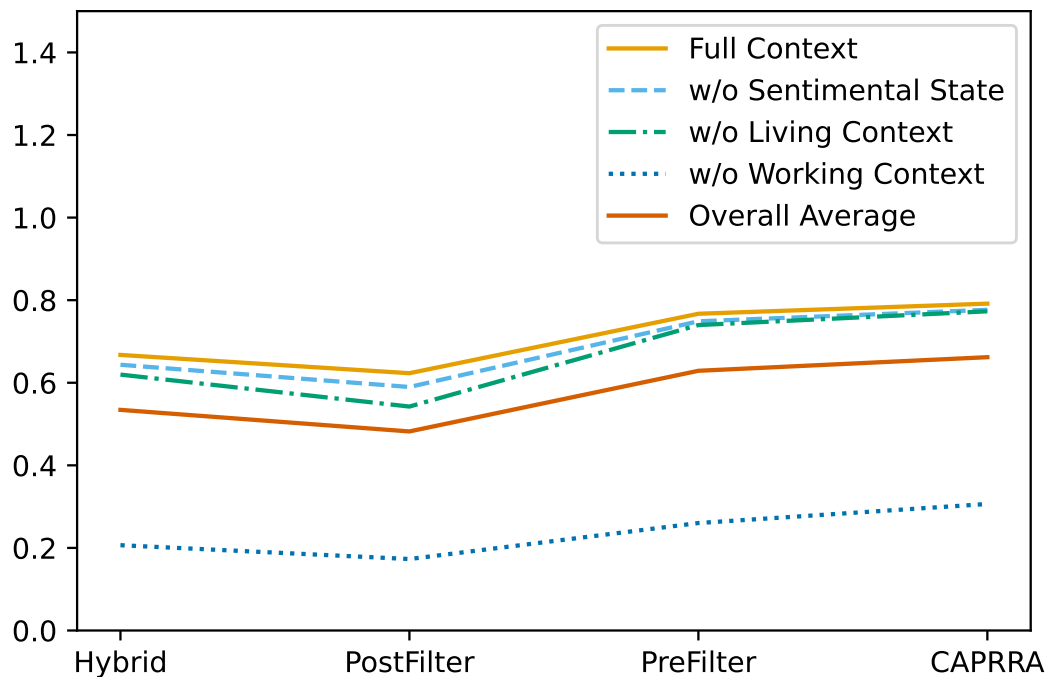


Figure 8.18: Average overall F1 scores obtained by the compared recommender approaches across all experiments.

CAPRRA or as effective in maintaining recommendation quality when specific contexts are eliminated.

Notably, the postfiltering approach tends to have the lowest average F1 scores across all experiments. This outcome implies that the postfiltering approach, which narrows down the recommendation list after an initial set of recommendations is generated, might struggle to achieve high levels of personalization and relevance, especially in the absence of certain context features.

Overall, the results emphasize CAPRRA's adaptability to diverse context variations and its potential to consistently provide valuable and contextually relevant pedagogical resource recommendations, followed closely by the prefiltering approach. These findings offer valuable insights for tailoring recommendation systems to the complex and multifaceted needs of teachers.

8.3.3 Experimentation with Peer Scenarios

While generating the list of teachers' peer recommendations, our main goal is to enforce various measures using semantic rules during the peer matching process to illustrate the impact of applying such rules on a teacher context. The sentimental state of teacher, expertise level, work and living environments, and field of science are the selected concepts in our approach, in addition to language and teaching style. The usage of SWRL rules to leverage these essential concepts proved to be effective. The following scenario illustrates the importance of the SWRL rules in enforcing the desired concepts.

Scenario: Let's assume that teacher T_A is a novice teacher with a negative mood while viewing a pedagogical resource and the teacher agrees to the

offered help, as shown in Figure 8.19. Algorithm 7 computes the list of contextual factors as well as their final weights W_{global} for this teacher's context. The semantic similarity calculated within *contextMatching*, indicates that the peer recommendation T_C obtains higher similarity score $Sim(A,C)$ than T_B despite the negative mood of this teacher while viewing the same resource. However, when the SWRL rules are applied to both recommendations, teacher T_B obtains higher score $Score_{SWRL}(A,B)$ than T_C by enforcing the positive sentimental state of teacher T_B as well as the teaching style. The *sortList* function offers the final list to the teacher in the following order: teacher T_B , teacher T_C .

$W_{global} = \{ \langle \text{Language}, 0.2 \rangle, \langle \text{Science}, 0.4 \rangle, \langle \text{Mood}, 0.3 \rangle, \langle \text{Working Environment}, 0.35 \rangle, \langle \text{Living Environment}, 0.1 \rangle \}$

$T_A = \{ \{ \text{Language 1}, \text{Language 2} \}, \text{Science 1}, \text{Negative}, \text{Small City}, \text{Small City} \}$

$T_B = \{ \{ \text{Language 1}, \text{Language 2}, \text{Language 3} \}, \text{Science 1}, \text{Positive}, \text{Large City}, \text{Small City} \}$

$T_C = \{ \{ \text{Language 2}, \text{Language 3} \}, \text{Science 1}, \text{Negative}, \text{Small City}, \text{Small City} \}$

$$Sim(A,B) = \frac{1*0.2+1*0.4+1*0.3+0*0.35+1*0.1}{0.2+0.4+0.5+0.35+0.1} = 0.74$$

$$Sim(A,C) = \frac{1*0.2+1*0.4+0*0.3+1*0.35+1*0.1}{0.2+0.4+0.5+0.35+0.1} = 0.78$$

$$Score_{SWRL}(A,B) = 8/10 = 0.8$$

$$Score_{SWRL}(A,C) = 6.5/10 = 0.65$$

The illustrated case can result in a different output if the contextual factors were selected without the mood enforcing approach. As previously stated, the sentimental analysis of the user has a remarkable impact on the provided results. In this thesis, we enforce the sentimental state of the teacher by enforcing the contextual factor f_{mood} in Algorithm 8 and by applying the SWRL rules in Table 4.1. If one of these two steps is escaped, the final recommendations are not of the desired quality. In Scenario 1, the mood contextual factor fails to prioritize the positive-mood teacher recommendation, but the SWRL rules readjusted the final recommendations list. Therefore, it can be said that this approach is more efficient regarding the sentimental state point-of-view.

Accordingly, the mentioned rules can be reordered or modified to obtain different results. For example, if Rule 12 is replaced with a different rule that matches the precise science field of both compared teachers, the list's order will be affected and accordingly the impact of the result on the teacher's sentimental state. Moreover, the rules can be increased to adjust the resulting list towards another approach, resulting in a flexible algorithm that can adapt to the different teachers' situations. TSCCO ontology and the SWRL rules provide a human understandable situation that can be easily evaluated using readable context as in Figure 8.19 and logical rules as in Table 4.1. The reviewed approaches target only a user in a unique situation without any generalization or problem abstraction. At the end, the provided results in these circumstances, cannot be evaluated except by the teacher. Hence, the teacher's evaluation of result list is used to recommend the same peer for other teachers with the same context which enhances the computing overhead for large datasets used by other approaches.

8.4 Results Overview

In this chapter, we have delved into the experimental results and findings of the various experiments conducted in this thesis, focusing on different phases and aspects. Firstly, the mood detection experiment explores the impact of sentimental state of teachers on their search and choice behavior during experiments. Through observations of teachers' search patterns, behavior, and the time spent on different tasks, the connection between mood and behavior was established. The experiment demonstrated that mood plays a significant role in shaping teachers' preferences and decisions, emphasizing the importance of integrating mood data in recommender systems.

Afterwards, the experiment of VIFSTC approach delves into the selection of appropriate features for representing and distinguishing teachers. The VIFSTC approach was evaluated against various approaches, demonstrating its effectiveness in retaining key contextual features. The use of classification methods showcased the robustness of the VIFSTC approach in feature selection, highlighting its ability to consider significant context features.

Subsequently, the experiments related to the recommender approaches focus on evaluating the performance of the context-aware pedagogical resources recommender approach (CAPRRA). Through a series of experiments, CAPRRA was compared against other approaches, assessing its effectiveness in providing relevant pedagogical resource recommendations to teachers under varying conditions. Different numbers of recommendations and variations of contextual features are explored to analyze CAPRRA's behavior and adaptability. Additionally, the impact of semantic rules on the peer recommendation process is examined to support the sentimental state of teachers through a peer teacher collaboration offering.

Overall, the experimental results highlighted the significance of contextual features, mood detection, feature selection, and the effectiveness of the CAPRRA approach in the context of pedagogical resource recommendation. The findings contribute to a deeper understanding of the interplay between teacher preferences, contextual features, and recommendation approaches, laying the groundwork for the conclusion of the study and insights for further research.

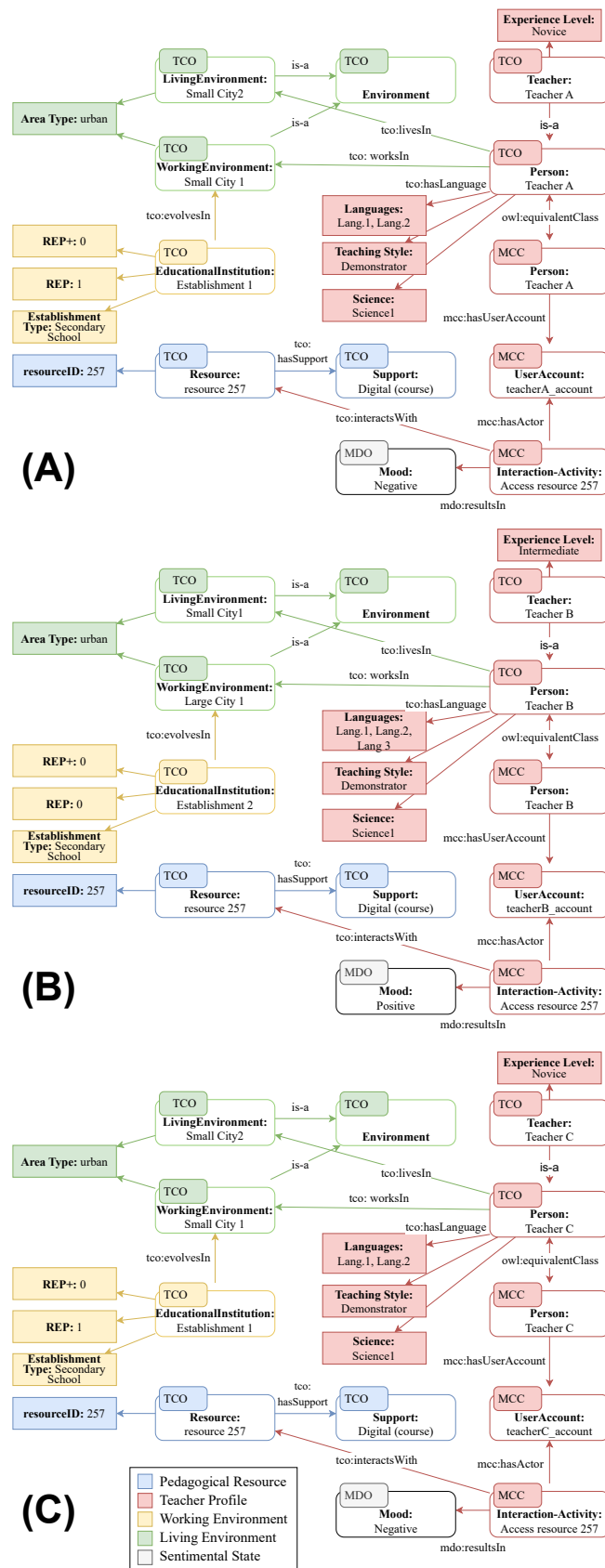


Figure 8.19: A-Box TSCCO instances of Teachers A, B and C used in peer scenario.

Part V

Conclusion and Perspectives

Chapter 9

Conclusion

Teachers face complex challenges during their teaching career. Some of these challenges are addressed in this thesis. These challenges span a broad range, encompassing the motivational dynamics, the precarious balance between professional and personal lives of teachers, the development of self-evaluation skills, enhancing the learner discipline and behavior, cultural differences, and learner within specialized areas. The interconnections among these challenges are thoroughly examined, unveiling their intricate influence on the teaching process.

One of the foundational findings of this research is the recognition that these struggles are not isolated but interconnected, implying that addressing one problem can have a reflected effect on other related challenges. This recognition paved the way for a comprehensive approach that provides cumulative solutions capable of simultaneously addressing the research questions.

1. How to represent the multiple contexts of teachers? And how to integrate data from these different contexts?
2. How taking account the multiple contexts of teachers to recommend suitable pedagogical resources to teachers using the most descriptive features?

In order to address these questions, the first question is answered by understanding and representing the multiple contexts of teachers. While the second question is addressed through context-aware pedagogical resource recommendations.

A critical first step was the understanding and representation of the multiple contexts of teachers. The multifaceted nature of teachers' professional and personal lives demanded a holistic approach. To this end, the Teacher Sentimental and Collaboration Context Ontology (TSCCO) was developed. TSCCO stands as a testament to the fusion of ontological modeling and semantic reasoning, providing a structured framework to encapsulate diverse teacher contexts. TSCCO consists of Teacher-Context Ontology (TCO), Mood-Detection Ontology (MDO), and the existing MEMORAE Collaboration Context (MCC) ontology. Each of these components plays a distinct yet interrelated role in the grand tapestry of teacher contexts.

TCO captures the essential dimensions of living environments, working environments, and the intricate interplay between these spheres for teachers. It is within the TCO that the contextual essence of teachers' experiences is encapsulated, allowing for a nuanced and comprehensive representation of their multifaceted lives. While MDO delves into the realm of sentiment analysis and emotional states. The MDO recognizes the profound influence of teachers' moods

on their interactions with pedagogical resources and their teaching styles. This component allows for the nuanced integration of sentimental context, enriching the recommendation process by considering not just the necessities of teachers but also their sentimental state. The already existing MCC ontology plays an integral role by incorporating the interactions of teacher with their sentimental state. Inclusion of the MCC ontology ensures that collaborative contexts are seamlessly woven into the fabric of the multiple contexts of teachers, offering a holistic connectivity of their contexts.

These comprehensive ontologies encompass various dimensions, including living environments, working environments, and sentimental states, offering a harmonious synthesis of information from disparate sources. The semantic reasoning mechanisms further ensure the integration of different data and prevent redundancy, facilitating a holistic understanding that underpinned subsequent stages of the pedagogical resource recommendation process.

Addressing the second research question, the heart of this research endeavor lays in the tailoring of context-aware pedagogical resource recommendations for teachers. Context-Aware Pedagogical Resource Recommender System (CAPRRS) represents the culmination of efforts to provide teachers with context-aware recommendations for pedagogical resources. CAPRRS harnesses the rich tapestry of the multiple contexts of teachers to tailor recommendations through the careful amalgamation of collaborative and content-based filtering techniques. The Variance-based Importance Feature Selection for Teacher Contexts (VIFSTC) methodology played a pivotal role in systematically identifying and weighing relevant features.

Feature selection is a fundamental pre-processing step that optimizes the pedagogical resources context-aware recommender system by reducing dimensionality, enhancing interpretability, improving computational efficiency, accommodating mixed data types, and filtering out noise. The VIFSTC approach is designed to systematically enhance the efficacy of feature selection within the pedagogical resources context-aware recommender system. The process commences with the extraction of ontological feature data. These features, encapsulating pertinent features of teachers, undergo a transformation from their original feature-value representation to a vectorized format optimized by a two-step feature selection process. VIFSTC incorporates two unsupervised filtering algorithms, information gain and variance-based, that obtains a more balanced selection of features that considers both the categorical and numerical features.

The task of providing context-aware recommendations in education is a challenging and complicated process, particularly when considering the multiple contexts of teachers. In order to address this challenge, the recommender approach introduces a methodology that aims to provide two types of recommendations: pedagogical resources and peer teachers recommendations. These two types address the potential solutions elicited through the second research question to offer help to teachers in the form of aiding pedagogical resources and collaboration with experienced teachers. This approach utilizes the integration of TSCCO ontology to represent multiple contexts of teachers such as living environment, working environment, and sentimental state within the collaboration management ontology, MCC. The approach first detects the mood of the teacher using MoodFlow@doubleYou and tracks their activity using the MEMORAe platform.

Then, the contextual features are extracted and selected carefully, keeping the mood-related features due to its importance. The contextual features are used at different phases of the context-aware recommendation process, but the approach starts with a hybrid 2D recommender approach. The 2D recommender approach captures the teacher-to-teacher similarity easily captured by user-based collaborative filtering recommender approach as well as the pedagogical resources suitability for a teacher excelled by content-based filtering recommender approach. Accordingly, the countability of both types of similarities combines the advantages of both approaches.

The context-aware recommender approach of this thesis introduces a pivotal framework for context-aware recommendations tailored to the multiple contexts of teachers. The approach commences by elucidating the intricate nature of pedagogical resource recommendation, encompassing both explicit feedback through ratings and implicit feedback through sentimental state of teachers. It leverages a hybridization two-dimensional approach that combines content-based and collaborative filtering techniques. Collaborative filtering leveraged historical interactions, while content-based filtering incorporated pedagogical resource data, resulting in comprehensive yet contextually nuanced recommendations. The consideration of teacher contexts becomes paramount, leading to the conceptualization of context-aware recommendations. Three distinct types of contextual filtering - prefiltering, postfiltering, and contextual modeling - are explored in the context of recommendation, offering insights into the integration of contextual features in the recommendation process. This exploration results in a combination of these different contextual filtering approaches called the Context-Aware Pedagogical Resources Recommender Approach (CAPRRA).

The experimental phase of this research journey provided valuable insights and findings. Mood detection emerged as a significant factor, shedding light on the impact of teachers' sentimental states on their behavior. The connection between mood and behavior underscored the importance of integrating mood data into recommender systems, offering a more personalized and effective recommendation process.

The VIFSTC approach, designed for feature selection, demonstrated its efficacy in retaining key contextual features. Utilizing various classification methods, the approach showcased robustness in feature selection, emphasizing its ability to consider significant context features.

CAPRRS, the context-aware recommendation system, underwent rigorous evaluation in various experiments. Comparative analyses against other approaches highlighted its effectiveness in providing relevant pedagogical resource recommendations to teachers under varying conditions. The experiments also delved into the impact of semantic rules on peer recommendations, which supported the sentimental state of teachers through peer teacher collaboration offerings.

In conclusion, this thesis embarked on a journey to unravel the multifaceted challenges faced by teachers in the realm of education. Through a comprehensive approach that involved understanding teacher contexts, developing an ontology, and crafting a context-aware recommendation system, this research has contributed valuable insights and potential solutions. The findings emphasize the importance of considering teacher contexts when providing pedagogical resource recommendations. By doing so, we can offer teachers the support they

need to navigate the intricate landscape of education, enhance their professional development, and ultimately enrich the quality of education for learners. This research underscores the significance of context-aware recommender systems in the field of education and opens the door to further exploration and innovation in this domain.

Chapter 10

Perspectives

This final chapter explores the potential avenues for future research and development in the context of the proposed context-aware pedagogical resource recommendation system and the broader field of educational technology. The research conducted in this thesis has paved the way for several short-term and long-term objectives that hold promise for enhancing the effectiveness of teacher support and the quality of education.

10.1 Short-Term Objectives

Short-term objectives focus on immediate and feasible advancements that can be pursued in the near future to build upon the current research.

Refinement of recommendation algorithms: While the Context-Aware Pedagogical Resource Recommender System (CAPRRS) has demonstrated its efficacy, further research can refine and enhance the recommendation algorithms. Experimentation with advanced machine learning techniques, such as deep learning and reinforcement learning, could improve the accuracy and personalization of resource recommendations.

Integration of real-time experiments: The incorporation of real-time experimentation into the recommendation process could lead to more dynamic and context-aware recommendations. This could include monitoring and analyzing teacher sentiment and learner engagement in real time to adapt recommendations on the fly, thereby addressing evolving needs.

User experience enhancement: Enhancing the user experience of the recommendation system is crucial. User interface design improvements, user feedback mechanisms, and user studies can provide valuable insights for optimizing the system's usability and effectiveness.

Scalability and deployment: In the short term, efforts should be directed towards making CAPRRS scalable and easy to deploy in various educational settings. This involves addressing issues related to system performance, compatibility, and accessibility.

Ethical and privacy considerations: As recommendation systems become more pervasive, addressing ethical and privacy concerns becomes paramount. Future research should focus on developing robust mechanisms for protecting user data and ensuring transparent and ethical recommendation practices.

10.2 Long-Term Objectives

Long-term objectives encompass more ambitious and forward-looking goals that may require substantial research and development efforts over an extended period.

Multimodal recommendation: The future of recommendation systems lies in multimodal data integration. Incorporating not only text-based and weblink pedagogical resources but also audio, video, and interactive content into the recommendation process can offer a richer learning experience.

Personalized learning paths: Long-term objectives should include the development of personalized learning paths for teachers. This involves creating adaptive curricula that consider individual teacher profiles, preferences, and progress, guiding them through continuous professional development.

Sentiment-aware recommendations: Building on the insights gained from mood detection, future work can delve deeper into sentiment-aware recommendations. Natural language processing and sentiment analysis can be integrated to tailor recommendations based on teachers' emotional states.

Cross-institutional collaboration: Long-term objectives should explore the possibilities of cross-institutional collaboration. Recommender systems that facilitate the sharing of best practices and resources among teachers from different educational institutions can foster a broader and more diverse community of practice.

Publications

International Publications

- Nashed, Nader N., Christine Lahoud, and Marie-Hélène Abel. "Contextual Ontology-based Feature Selection for Teachers." International Conference on Web-Based Learning. Cham: Springer International Publishing, 2022.
- Nashed, Nader N., Christine Lahoud, and Marie-Helene Abel. "Contextual and sentimental teachers' peer recommendations." 2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design (CSCWD). IEEE, 2022.
- Nashed, Nader, Christine Lahoud, and Marie-Hélène Abel. "Teacher Educational Resources Recommendation in the COVID-19 Context." 14th International Conference on Computer Supported Education (CSEDU 2022). Vol. 1. SCITEPRESS-Science and Technology Publications, 2022.
- Nashed, Nader N., Christine Lahoud, and Marie-Hélène Abel. "Ontology-based teacher-context data integration." 2022 IEEE/SICE International Symposium on System Integration (SII). IEEE, 2022.
- Nashed, Nader N., Christine Lahoud, and Marie-Hélène Abel. "Educational resources recommender system for teachers: why and how?." Advances in Deep Learning, Artificial Intelligence and Robotics: Proceedings of the 2nd International Conference on Deep Learning, Artificial Intelligence and Robotics, (ICDLAIR) 2020. Springer International Publishing, 2022.
- Nashed, Nader N., et al. "Mood detection ontology integration with teacher context." 2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE, 2021.
- Nashed, Nader N., Christine Lahoud, and Marie-Hélène Abel. "TCO: a teacher context ontology." 2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD). IEEE, 2021.
- Nashed, Nader N., Christine Lahoud, and Marie-Hélène Abel. "Towards personalized educational resources recommendations for teachers." International Conference on Deep Learning, Artificial Intelligence and Robotics. Cham: Springer International Publishing, 2021.

Bibliography

- Abel, Marie-Hélène (2022). “MEMORAE project: an approach and a platform for learning innovation”. In: *Multimedia Tools and Applications* 81.25, pp. 35555–35569.
- Ackoff, Russell L (1989). “From data to wisdom”. In: *Journal of applied systems analysis* 16.1, pp. 3–9.
- Adekitan, Aderibigbe Israel and Odunayo Salau (2019). “The impact of engineering students’ performance in the first three years on their graduation result using educational data mining”. In: *Heliyon* 5.2, e01250.
- Adomavicius, Gediminas and Alexander Tuzhilin (2005a). “Incorporating context into recommender systems using multidimensional rating estimation methods”. In: *The First International Workshop on Web Personalization, Recommender Systems and Intelligent User Interfaces*. Vol. 2. SCITEPRESS, pp. 3–13.
- (2005b). “Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions”. In: *IEEE transactions on knowledge and data engineering* 17.6, pp. 734–749.
- (2010). “Context-aware recommender systems”. In: *Recommender systems handbook*. Springer, pp. 217–253.
- Agarwal, Deepak, Bee-Chung Chen, and Bo Long (2011). “Localized factor models for multi-context recommendation”. In: *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 609–617.
- Aggarwal, Charu C et al. (2016). *Recommender systems*. Vol. 1. Springer.
- Ainley, John and Ralph Carstens (2018). “Teaching and learning international survey (TALIS) 2018 conceptual framework”. In.
- Akharraz, Laila, Ali El Mezouary, and Zouhir Mahani (2018). “To context-aware learner modeling based on ontology”. In: *2018 IEEE Global Engineering Education Conference (EDUCON)*. IEEE, pp. 1326–1334.
- Akkerman, Sanne F and Paulien C Meijer (2011). “A dialogical approach to conceptualizing teacher identity”. In: *Teaching and teacher education* 27.2, pp. 308–319.
- Al-Nafjan, Abeer et al. (2017). “Review and classification of emotion recognition based on EEG brain-computer interface system research: a systematic review”. In: *Applied Sciences* 7.12, p. 1239.
- Ali, Ben Ameer Mohamed et al. (2017). “Recommendation of pedagogical resources within a learning ecosystem”. In: *Proceedings of the 9th International Conference on Management of Digital EcoSystems*, pp. 14–21.
- Ali, Sadia et al. (2022). “Enabling recommendation system architecture in virtualized environment for e-learning”. In: *Egyptian Informatics Journal* 23.1, pp. 33–45.

- Alyahyan, Eyman and Dilek Düşteğör (2020). "Predicting academic success in higher education: literature review and best practices". In: *International Journal of Educational Technology in Higher Education* 17.1, pp. 1–21.
- Ananiadou, Katerina and Magdalen Claro (2009). "21st century skills and competences for new millennium learners in OECD countries". In: — (2009). *21st Century Skills: Preparing Young People for the Future of Work*. Paris: OECD Publishing.
- Andres, Frederic et al. (2021). "Overview of the 4th DECOR Workshop 2021". In: *37th IEEE International Conference on Data Engineering*.
- Ang, Jun Chin et al. (2015). "Supervised, unsupervised, and semi-supervised feature selection: a review on gene selection". In: *IEEE/ACM transactions on computational biology and bioinformatics* 13.5, pp. 971–989.
- Aroyo, Lora et al. (2006). "Interoperability in personalized adaptive learning". In: *Journal of Educational Technology & Society* 9.2, pp. 4–18.
- Atrash, Ala (2015). "Modeling a system of expertise capitalization to support organizational learning within small and medium-sized enterprises". PhD thesis. Université de Technologie de Compiègne.
- Avgerou, Chrisanthi (2001). "The significance of context in information systems and organizational change". In: *Information systems journal* 11.1, pp. 43–63.
- Awla, Hawkar Akram (2014). "Learning styles and their relation to teaching styles". In: *International Journal of Language and Linguistics* 2.3, pp. 241–245.
- Baltrunas, Linas, Bernd Ludwig, and Francesco Ricci (2011). "Matrix factorization techniques for context aware recommendation". In: *Proceedings of the fifth ACM conference on Recommender systems*, pp. 301–304.
- Baylor, Amy L and Donn Ritchie (2002). "What factors facilitate teacher skill, teacher morale, and perceived student learning in technology-using classrooms?" In: *Computers & education* 39.4, pp. 395–414.
- Bazire, Mary and Patrick Brézillon (2005). "Understanding context before using it". In: *Modeling and Using Context: 5th International and Interdisciplinary Conference CONTEXT 2005, Paris, France, July 5-8, 2005. Proceedings* 5. Springer, pp. 29–40.
- Beauchamp, Catherine and Lynn Thomas (2009). "Understanding teacher identity: An overview of issues in the literature and implications for teacher education". In: *Cambridge journal of education* 39.2, pp. 175–189.
- Beijaard, Douwe, Nico Verloop, and Jan D Vermunt (2000). "Teachers' perceptions of professional identity: An exploratory study from a personal knowledge perspective". In: *Teaching and teacher education* 16.7, pp. 749–764.
- Belmont, M et al. (1988). *Teacher as social context: A measure of student perceptions of teacher provision of involvement, structure, and autonomy support*. Tech. rep. Tech. rep.
- (1992). "Teacher as social context (TASC): Two measures of teacher provision of involvement, structure, and autonomy support". In: *Unpublished manuscript, Graduate School of Education and Human Development, Department of Psychology, University of Rochester*.
- Ben-Peretz, Miriam, Nili Mendelson, and Friedrich W Kron (2003). "How teachers in different educational contexts view their roles". In: *Teaching and Teacher Education* 19.2, pp. 277–290.

- Benayache, AHCÈNE (2005). "Construction d'une mémoire organisationnelle de formation et évaluation dans un contexte e-learning: le projet MEMORAE". PhD thesis. Compiègne.
- Bhattacharya, Arnab, Manoj Kumar Tiwari, and Jenny A Harding (2012). "A framework for ontology based decision support system for e-learning modules, business modeling and manufacturing systems". In: *Journal of Intelligent Manufacturing* 23.5, pp. 1763–1781.
- Bizer, Christian and Richard Cyganiak (2006). "D2r server-publishing relational databases on the semantic web". In: *Poster at the 5th international semantic web conference*. Vol. 175.
- Blanco-Fernández, Yolanda et al. (2008). "Exploiting synergies between semantic reasoning and personalization strategies in intelligent recommender systems: A case study". In: *Journal of Systems and Software* 81.12, pp. 2371–2385.
- Board, DCMI Usage (2018). *Dublin core metadata initiative (DCMI) metadata terms*. url: <http://dublincore.org/documents/dcmi-terms/> (visited on 07/07/2023).
- Bodily, Robert and Katrien Verbert (2017). "Review of research on student-facing learning analytics dashboards and educational recommender systems". In: *IEEE Transactions on Learning Technologies* 10.4, pp. 405–418.
- Borge-Holthoefer, Javier and Alex Arenas (2010). "Semantic networks: Structure and dynamics". In: *Entropy* 12.5, pp. 1264–1302.
- Bouchet, François et al. (2017). "Who wants to chat on a MOOC? Lessons from a peer recommender system". In: *European Conference on Massive Open Online Courses*. Springer, pp. 150–159.
- Boud, David, Ruth Cohen, and Jane Sampson (1999). "Peer learning and assessment". In: *Assessment & evaluation in higher education* 24.4, pp. 413–426.
- Bozo, Jorge, Rosa Alarcón, and Sebastian Iribarra (2010). "Recommending learning objects according to a teachers' context model". In: *European Conference on Technology Enhanced Learning*. Springer, pp. 470–475.
- Brachman, Ronald J and Hector J Levesque (1985). *Readings in knowledge representation*. Tech. rep. AT and T Bell Labs.
- Brewster, Christopher and Kieron O'Hara (2004). "Knowledge representation with ontologies: the present and future". In: *IEEE Intelligent Systems* 19.1, pp. 72–81.
- Burke, Robin (2000). "Knowledge-based recommender systems". In: *Encyclopedia of library and information systems* 69.Supplement 32, pp. 175–186.
- (2002). "Hybrid recommender systems: Survey and experiments". In: *User modeling and user-adapted interaction* 12, pp. 331–370.
- Bušljeta, Rona (2013). "Effective use of teaching and learning resources". In: *Czech-polish historical and pedagogical journal* 5.2.
- Cai, Jie et al. (2018). "Feature selection in machine learning: A new perspective". In: *Neurocomputing* 300, pp. 70–79.
- Campos, Pedro G et al. (2013). "Context-aware movie recommendations: an empirical comparison of pre-filtering, post-filtering and contextual modeling approaches". In: *E-Commerce and Web Technologies: 14th International Conference, EC-Web 2013, Prague, Czech Republic, August 27-28, 2013. Proceedings 14*. Springer, pp. 137–149.

- Carbonell, Jaime G, Ryszard S Michalski, and Tom M Mitchell (1983). "An overview of machine learning". In: *Machine learning*, pp. 3–23.
- Carton, Annie and Eric Fruchart (2014). "Sources of stress, coping strategies, emotional experience: Effects of the level of experience in primary school teachers in France". In: *Educational Review* 66.2, pp. 245–262.
- Carver, Liz and Murray Turoff (2007). "Human-computer interaction: the human and computer as a team in emergency management information systems". In: *Communications of the ACM* 50.3, pp. 33–38.
- Chandrasegaran, Senthil K et al. (2013). "The evolution, challenges, and future of knowledge representation in product design systems". In: *Computer-aided design* 45.2, pp. 204–228.
- Chang, Maiga et al. (2020). "Building Ontology-Driven Tutoring Models for Intelligent Tutoring Systems Using Data Mining". In: *IEEE Access* 8, pp. 48151–48162.
- Chen, Jizhi and Junzhong Gu (2021). "ADOL: a novel framework for automatic domain ontology learning". In: *The Journal of Supercomputing* 77, pp. 152–169.
- Christian, Hans, Mikhael Pramodana Agus, and Derwin Suhartono (2016). "Single document automatic text summarization using term frequency-inverse document frequency (TF-IDF)". In: *ComTech: Computer, Mathematics and Engineering Applications* 7.4, pp. 285–294.
- Clemente, Julia, Jaime Ramírez, and Angélica De Antonio (2011). "A proposal for student modeling based on ontologies and diagnosis rules". In: *Expert Systems with Applications* 38.7, pp. 8066–8078.
- Cobos, Carlos et al. (2013). "A hybrid system of pedagogical pattern recommendations based on singular value decomposition and variable data attributes". In: *Information Processing & Management* 49.3, pp. 607–625.
- Cooper, Karyn and Margaret R Olson (2020). "The multiple 'I's' of teacher identity". In: *Changing research and practice*. Routledge, pp. 78–89.
- Corson, Yves (2002). "Effects of positive, negative, and neutral moods on associative and semantic priming." In: *Cahiers de psychologie cognitive/Current Psychology of Cognition*.
- Cranton, Patricia and Ellen Carusetta (2002). "Reflecting on teaching: The influence of context". In: *The International Journal for Academic Development* 7.2, pp. 167–176.
- Cristani, Matteo and Roberta Cuel (2005). "A survey on ontology creation methodologies". In: *International Journal on Semantic Web and Information Systems (IJSWIS)* 1.2, pp. 49–69.
- Dalrymple, Prudence W (2011). "Data, information, knowledge: The emerging field of health informatics". In: *Bulletin of the American Society for Information Science and Technology* 37.5, pp. 41–44.
- Dang, Quoc-Viet (2018). "Implementing an individualized recommendation system using latent semantic analysis". In: *Proceedings of the 6th International Conference on Information and Education Technology*, pp. 239–243.
- Darling-Hammond, Linda (2007). "Race, inequality and educational accountability: The irony of 'No Child Left Behind'". In: *Race Ethnicity and Education* 10.3, pp. 245–260.

- Dascalu, MI et al. (2022). "An Ontology for Educational and Career Profiling based on the Romanian Occupation Classification Framework: Description and Scenarios of Utilisation". In: *ICERI2022 Proceedings*. IATED, pp. 7386–7395.
- Davies, John, Rudi Studer, and Paul Warren (2006). *Semantic Web technologies: trends and research in ontology-based systems*. John Wiley & Sons.
- Davis, Elizabeth A and Joseph S Krajcik (2005). "Designing educative curriculum materials to promote teacher learning". In: *Educational researcher* 34.3, pp. 3–14.
- De Medio, Carlo et al. (2020). "MoodleREC: A recommendation system for creating courses using the moodle e-learning platform". In: *Computers in Human Behavior* 104, p. 106168.
- De Meo, Pasquale et al. (2017). "Combining trust and skills evaluation to form e-Learning classes in online social networks". In: *Information Sciences* 405, pp. 107–122.
- De Santiago, Rafael and Andre Raabe (2010). "Architecture for learning objects sharing among learning institutions—LOP2P". In: *IEEE Transactions on Learning Technologies* 3.2, pp. 91–95.
- Deparis, Etienne (2013). "Création de nouvelles connaissances décisionnelles pour une organisation via ses ressources sociales et documentaires". PhD thesis.
- Dermeval, Diego et al. (2019). "GaTO: An ontological model to apply gamification in intelligent tutoring systems". In: *Frontiers in Artificial Intelligence* 2, p. 13.
- Deschênes, Michelle (2020). "Recommender systems to support learners' Agency in a Learning Context: a systematic review". In: *International Journal of Educational Technology in Higher Education* 17.1, pp. 1–23.
- Dey, Anind K (2001). "Understanding and using context". In: *Personal and ubiquitous computing* 5, pp. 4–7.
- Dib, Claudio Zaki (1988). "Formal, non-formal and informal education: concept/s/applicability". In: *AIP conference proceedings*. Vol. 173. 1. American Institute of Physics, pp. 300–315.
- Dietze, Stefan et al. (2013). "Interlinking educational resources and the web of data: A survey of challenges and approaches". In: *Program* 47.1, pp. 60–91.
- Dinçol, Sinem et al. (2011). "The effect of matching learning styles with teaching styles on success". In: *Procedia-Social and Behavioral Sciences* 15, pp. 854–858.
- Dong, Guozhu and Huan Liu (2018). *Feature engineering for machine learning and data analytics*. CRC Press.
- Dorman, Jeffrey, Joy Kennedy, and Janelle Young (2015). "The development, validation and use of the Rural and Remote Teaching, Working, Living and Learning Environment Survey (RRTWLLES)". In: *Learning Environments Research* 18.1, pp. 15–32.
- Drachsler, Hendrik, Hans GK Hummel, and Rob Koper (2008). "Personal recommender systems for learners in lifelong learning networks: the requirements, techniques and model". In: *International Journal of Learning Technology* 3.4, pp. 404–423.
- Duboue, Pablo (2020). *The art of feature engineering: essentials for machine learning*. Cambridge University Press.

- Dudek, Mark (2000). *Architecture of schools: The new learning environments*. Routledge.
- Dutt, Ashish, Maizatul Akmar Ismail, and Tutut Herawan (2017). "A systematic review on educational data mining". In: *Ieee Access* 5, pp. 15991–16005.
- Epstein, Jonathan, Ruth Miyuki Santo, and Francis Guillemin (2015). "A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus". In: *Journal of clinical epidemiology* 68.4, pp. 435–441.
- Fensel, Dieter (2005). *Spinning the Semantic Web: bringing the World Wide Web to its full potential*. MIT press.
- Fensel, Dieter and Dieter Fensel (2001). *Ontologies*. Springer.
- Frangeto, Carolina et al. (2021). "Monitoring Karting Pilot's Moodflow: a First Experience". In: *2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, pp. 1716–1719.
- Frenzel, Anne C (2014). "Teacher emotions." In.
- Fung, Kin Wah and Olivier Bodenreider (2019). "Knowledge representation and ontologies". In: *Clinical research informatics*, pp. 313–339.
- George, Gina and Anisha M Lal (2019). "Review of ontology-based recommender systems in e-learning". In: *Computers & Education* 142, p. 103642.
- Gnoth, Juergen et al. (2000). "Emotions, mood, flow and motivations to travel". In: *Journal of Travel & Tourism Marketing* 9.3, pp. 23–34.
- Goddard, Richard, Patrick O'brien, and Marion Goddard (2006). "Work environment predictors of beginning teacher burnout". In: *British educational research journal* 32.6, pp. 857–874.
- Golshan, Behzad et al. (2017). "Data integration: After the teenage years". In: *Proceedings of the 36th ACM SIGMOD-SIGACT-SIGAI symposium on principles of database systems*, pp. 101–106.
- Gomm, Roger (2008). *Social research methodology: A critical introduction*. Bloomsbury Publishing.
- Gonzalez, Gustavo et al. (2007). "Embedding emotional context in recommender systems". In: *2007 IEEE 23rd international conference on data engineering workshop*. IEEE, pp. 845–852.
- Gopal, Madan (2019). *Applied machine learning*. McGraw-Hill Education.
- Grabinger, R Scott and Joanna C Dunlap (1995). "Rich environments for active learning: A definition". In: *ALTJ* 3.2, pp. 5–34.
- Greenhow, Christine, Beth Robelia, and Joan E Hughes (2009). "Learning, teaching, and scholarship in a digital age: Web 2.0 and classroom research: What path should we take now?" In: *Educational researcher* 38.4, pp. 246–259.
- Gregor, Shirley (2006). "The nature of theory in information systems". In: *MIS quarterly*, pp. 611–642.
- Grimm, Stephan (2009). "Knowledge representation and ontologies". In: *Scientific data mining and knowledge discovery: principles and foundations*. Springer, pp. 111–137.
- Grivokostopoulou, Foteini et al. (2019). "An Ontology-based Approach for User Modelling and Personalization in E-Learning Systems". In: *2019 IEEE/ACIS 18th International Conference on Computer and Information Science (ICIS)*. IEEE Computer Society, pp. 1–6.

- Grubišić, Ani, Slavomir Stankov, and Ivan Peraić (2013). "Ontology based approach to Bayesian student model design". In: *Expert systems with applications* 40.13, pp. 5363–5371.
- Grüninger, Michael and Mark S Fox (1995). "The role of competency questions in enterprise engineering". In: *Benchmarking—Theory and practice*. Springer, pp. 22–31.
- Guarino, Nicola (1995). "Formal ontology, conceptual analysis and knowledge representation". In: *International journal of human-computer studies* 43.5-6, pp. 625–640.
- Gupta, Vishal and Shri Ram Pandey (2019). "Recommender Systems for Digital Libraries: A review of concepts and concerns". In: *Library Philosophy and Practice*, pp. 1–9.
- Guyon, Isabelle and André Elisseeff (2003). "An introduction to variable and feature selection". In: *Journal of machine learning research* 3.Mar, pp. 1157–1182.
- Han, Jiyang and Hongbiao Yin (2016). "Teacher motivation: Definition, research development and implications for teachers". In: *Cogent education* 3.1, p. 1217819.
- Hanushek, Eric A and Steven G Rivkin (2010). "Generalizations about using value-added measures of teacher quality". In: *American economic review* 100.2, pp. 267–271.
- Hartig, Olaf, Christian Bizer, and Johann-Christoph Freytag (2009). "Executing SPARQL queries over the web of linked data". In: *The Semantic Web-ISWC 2009: 8th International Semantic Web Conference, ISWC 2009, Chantilly, VA, USA, October 25-29, 2009. Proceedings 8*. Springer, pp. 293–309.
- Heath, Tom and Christian Bizer (2022). *Linked data: Evolving the web into a global data space*. Springer Nature.
- Helliwell, John F and Robert D Putnam (2004). "The social context of well-being". In: *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 359.1449, pp. 1435–1446.
- Henze, Nicola, Peter Dolog, and Wolfgang Nejdl (2004). "Reasoning and ontologies for personalized e-learning in the semantic web". In: *Journal of Educational Technology & Society* 7.4, pp. 82–97.
- Heutte, Jean et al. (2016). "The EduFlow model: A contribution toward the study of optimal learning environments". In: *Flow experience*. Springer, pp. 127–143.
- Horn, Ilana Seidel and Judith Warren Little (2010). "Attending to problems of practice: Routines and resources for professional learning in teachers' workplace interactions". In: *American educational research journal* 47.1, pp. 181–217.
- Horrocks, Ian et al. (2004). "SWRL: A semantic web rule language combining OWL and RuleML". In: *W3C Member submission* 21.79, pp. 1–31.
- Hotz, Lothar et al. (2014). "Configuration knowledge representation and reasoning". PhD thesis. Morgan Kaufmann Amsterdam.
- Hutchinson, Linda (2003). "Educational environment". In: *Bmj* 326.7393, pp. 810–812.
- Ilkou, Eleni et al. (2021). "Educor: An educational and career-oriented recommendation ontology". In: *International Semantic Web Conference*. Springer, pp. 546–562.

- Imran, Ali Shariq et al. (2021). "A systematic mapping review on MOOC recommender systems". In: *IEEE Access*.
- Iniesto, Francisco and Covadonga Rodrigo (2019). "YourMOOC4all: A recommender system for MOOCs based on collaborative filtering implementing UDL". In: *Transforming Learning with Meaningful Technologies: 14th European Conference on Technology Enhanced Learning, EC-TEL 2019, Delft, The Netherlands, September 16–19, 2019, Proceedings 14*. Springer, pp. 746–750.
- Jakus, Grega et al. (2013). *Knowledge representation*. Springer.
- Jannach, Dietmar et al. (2021). "A survey on conversational recommender systems". In: *ACM Computing Surveys (CSUR)* 54.5, pp. 1–36.
- Jensen, Jesper (2019). "A systematic literature review of the use of Semantic Web technologies in formal education". In: *British Journal of Educational Technology* 50.2, pp. 505–517.
- Johnson, Larry et al. (2016). *NMC horizon report: 2016 higher education edition*. The New Media Consortium.
- Jonassen, David (2003). "Using cognitive tools to represent problems". In: *Journal of research on Technology in Education* 35.3, pp. 362–381.
- Kamalov, Firuz (2021). "Orthogonal variance decomposition based feature selection". In: *Expert Systems with Applications* 182, p. 115191.
- Kamalov, Firuz et al. (2021). "Orthogonal variance-based feature selection for intrusion detection systems". In: *2021 International Symposium on Networks, Computers and Communications (ISNCC)*. IEEE, pp. 1–5.
- Katagall, Raghavendra et al. (2015). "Concept mapping in education and semantic knowledge representation: an illustrative survey". In: *Procedia Computer Science* 48, pp. 638–643.
- Kebede, Gashaw (2010). "Knowledge management: An information science perspective". In: *International journal of information management* 30.5, pp. 416–424.
- Kelchtermans, Geert (2005). "Teachers' emotions in educational reforms: Self-understanding, vulnerable commitment and micropolitical literacy". In: *Teaching and teacher education* 21.8, pp. 995–1006.
- Keßler, Carsten, Mathieu d'Aquin, and Stefan Dietze (2013). "Linked Data for science and education." In: *Semantic Web* 4.1, pp. 1–2.
- Khalid, Samina, Tehmina Khalil, and Shamila Nasreen (2014). "A survey of feature selection and feature extraction techniques in machine learning". In: *2014 science and information conference*. IEEE, pp. 372–378.
- Khan, Faisal, RM Yusoff, and Anwar Khan (2014). "Job demands, burnout and resources in teaching a conceptual review". In: *World Applied Sciences Journal* 30.1, pp. 20–28.
- Khosravi, Hassan (2017). "Recommendation in personalised peer-learning environments". In: *arXiv preprint arXiv:1712.03077*.
- Khosravi, Hassan, Kendra Cooper, and Kirsty Kitto (2017). "RiPLE: Recommendation in peer-learning environments based on knowledge gaps and interests". In: *arXiv preprint arXiv:1704.00556*.
- Klašnja-Milićević, Aleksandra, Boban Vesin, and Mirjana Ivanović (2018). "Social tagging strategy for enhancing e-learning experience". In: *Computers & Education* 118, pp. 166–181.

- Knoblauch, Dee and Anita Woolfolk Hoy (2008). ““Maybe I can teach those kids.” The influence of contextual factors on student teachers’ efficacy beliefs”. In: *Teaching and Teacher Education* 24.1, pp. 166–179.
- Ko, Hyeyoung et al. (2022). “A survey of recommendation systems: recommendation models, techniques, and application fields”. In: *Electronics* 11.1, p. 141.
- Koochi, Hamidreza and Kourosh Kiani (2016). “User based collaborative filtering using fuzzy C-means”. In: *Measurement* 91, pp. 134–139.
- Kraskov, Alexander, Harald Stögbauer, and Peter Grassberger (2004). “Estimating mutual information”. In: *Physical review E* 69.6, p. 066138.
- Krause, Andreas, Asim Smailagic, and Daniel P Siewiorek (2005). “Context-aware mobile computing: Learning context-dependent personal preferences from a wearable sensor array”. In: *IEEE Transactions on Mobile Computing* 5.2, pp. 113–127.
- Kuhn, Max and Kjell Johnson (2019). *Feature engineering and selection: A practical approach for predictive models*. CRC Press.
- Kuhn, Max, Kjell Johnson, et al. (2013). *Applied predictive modeling*. Vol. 26. Springer.
- Kulkarni, Saurabh and Sunil F Rodd (2020). “Context Aware Recommendation Systems: A review of the state of the art techniques”. In: *Computer Science Review* 37, p. 100255.
- Labarthe, Hugues et al. (2016). “Does a Peer Recommender Foster Students’ Engagement in MOOCs?.” In: *International Educational Data Mining Society*.
- Labib, A Ezzat, José H Canós, and M Carmen Penadés (2017). “On the way to learning style models integration: a Learner’s Characteristics Ontology”. In: *Computers in Human Behavior* 73, pp. 433–445.
- Lachowska, Bogusława H et al. (2018). “Conflict and facilitation between family and occupational roles versus work-related stress in the teachers’ group”. In: *Medical Studies/Studia Medyczne* 34.1, pp. 41–56.
- Lakiotaki, Kleantli, Nikolaos F Matsatsinis, and Alexis Tsoukias (2011). “Multicriteria user modeling in recommender systems”. In: *IEEE Intelligent Systems* 26.2, pp. 64–76.
- Lakshmi Padmaja, D and B Vishnuvardhan (2019). “Variance-Based Feature Selection for Enhanced Classification Performance”. In: *Information Systems Design and Intelligent Applications*. Springer, pp. 543–550.
- Lassila, Ora and Deborah McGuinness (2001). *The role of frame-based representation on the semantic web*. Linköping University Electronic Press.
- Lehdonvirta, Vili et al. (2021). “Social media, web, and panel surveys: using non-probability samples in social and policy research”. In: *Policy & internet* 13.1, pp. 134–155.
- Leidner, Dorothy E and Sirkka L Jarvenpaa (1995). “The use of information technology to enhance management school education: A theoretical view”. In: *MIS quarterly*, pp. 265–291.
- Levesque, Hector J (1986). “Knowledge representation and reasoning”. In: *Annual review of computer science* 1.1, pp. 255–287.
- Levesque, Hector J and Gerhard Lakemeyer (2001). *The logic of knowledge bases*. MIT Press.

- Li, Baoli and Liping Han (2013). "Distance weighted cosine similarity measure for text classification". In: *International conference on intelligent data engineering and automated learning*. Springer, pp. 611–618.
- Li, Jundong et al. (2017). "Feature selection: A data perspective". In: *ACM computing surveys (CSUR)* 50.6, pp. 1–45.
- Li, Siying (2021). "Context-aware recommender system for system of information systems". PhD thesis. Université de Technologie de Compiègne.
- Li, Siying, Marie-Hélène Abel, and Elsa Negre (2019). "Using User Contextual Profile for Recommendation in Collaborations". In: *The International Research & Innovation Forum*. Springer, pp. 199–209.
- Lika, Blerina, Kostas Kolomvatsos, and Stathes Hadjiefthymiades (2014). "Facing the cold start problem in recommender systems". In: *Expert systems with applications* 41.4, pp. 2065–2073.
- Lindblom-Ylänne, Sari et al. (2006). "How approaches to teaching are affected by discipline and teaching context". In: *Studies in Higher education* 31.03, pp. 285–298.
- Liu, Huawen et al. (2009). "Feature selection with dynamic mutual information". In: *Pattern Recognition* 42.7, pp. 1330–1339.
- Lizana, Pablo A et al. (2021). "Impact of the COVID-19 Pandemic on Teacher Quality of Life: A Longitudinal Study from before and during the Health Crisis". In: *International Journal of Environmental Research and Public Health* 18.7, p. 3764.
- Lomborg, Stine and Anja Bechmann (2014). "Using APIs for data collection on social media". In: *The Information Society* 30.4, pp. 256–265.
- Lops, Pasquale, Marco De Gemmis, and Giovanni Semeraro (2011). "Content-based recommender systems: State of the art and trends". In: *Recommender systems handbook*, pp. 73–105.
- Lü, Linyuan et al. (2012). "Recommender systems". In: *Physics reports* 519.1, pp. 1–49.
- Mabkhot, Mohammed M, Ali M Al-Samhan, and Lotfi Hidri (2019). "An ontology-enabled case-based reasoning decision support system for manufacturing process selection". In: *Advances in Materials Science and Engineering* 2019.
- Maguire, David J (1991). "An overview and definition of GIS". In: *Geographical information systems: Principles and applications* 1.1, pp. 9–20.
- Manouselis, Nikos, Riina Vuorikari, and Frans Van Assche (2010). "Collaborative recommendation of e-learning resources: an experimental investigation". In: *Journal of Computer Assisted Learning* 26.4, pp. 227–242.
- Manouselis, Nikos et al. (2012). *Recommender systems for learning*. Springer Science & Business Media.
- Manuti, Amelia et al. (2015). "Formal and informal learning in the workplace: A research review". In: *International journal of training and development* 19.1, pp. 1–17.
- Markman, Arthur B (2013). *Knowledge representation*. Psychology Press.
- Martinez-Cruz, Carmen et al. (2015). "A model to represent users trust in recommender systems using ontologies and fuzzy linguistic modeling". In: *Information Sciences* 311, pp. 102–118.
- McIntyre, Teresa Mendonça et al. (2017). "Towards a dynamic integrative theory of educator stress". In: *Educator Stress*. Springer, pp. 261–289.

- Mérida-López, Sergio and Natalio Extremera (2017). "Emotional intelligence and teacher burnout: A systematic review". In: *International Journal of Educational Research* 85, pp. 121–130.
- Monk, David H (2007). "Recruiting and retaining high-quality teachers in rural areas". In: *The future of children*, pp. 155–174.
- Moore, Philip, Zhili Zhao, and Hai Van Pham (2019). "Towards cloud-based personalised student-centric context-aware e-learning pedagogic systems". In: *Conference on Complex, Intelligent, and Software Intensive Systems*. Springer, pp. 331–342.
- Morton, K Ll et al. (2016). "The school environment and adolescent physical activity and sedentary behaviour: a mixed-studies systematic review". In: *Obesity reviews* 17.2, pp. 142–158.
- Mouromtsev, Dmitry et al. (2013). "Development of an ontology-based e-learning system". In: *International Conference on Knowledge Engineering and the Semantic Web*. Springer, pp. 273–280.
- Mukhtar, Khadijah et al. (2020). "Advantages, Limitations and Recommendations for online learning during COVID-19 pandemic era". In: *Pakistan journal of medical sciences* 36.COVID19-S4, S27.
- Mullis, Ina VS and Michael O Martin (2019). *PIRLS 2021 Assessment Frameworks*. ERIC.
- Mullis, Ina VS et al. (2020). "TIMSS 2019 international results in mathematics and science". In: Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/international-results>.
- Mumtaz, Shazia (2000). "Factors affecting teachers' use of information and communications technology: a review of the literature". In: *Journal of information technology for teacher education* 9.3, pp. 319–342.
- Mutingi, Michael and Charles Mbohwa (2017). "Grouping genetic algorithms". In: *Advances and Applications*. Switzerland: Springer International Publishing 243.
- Nashed, Nader, Christine Lahoud, and Marie-Hélène Abel (2022). "Teacher Educational Resources Recommendation in the COVID-19 Context". In: *14th International Conference on Computer Supported Education (CSEDU 2022)*. Vol. 1. SCITEPRESS-Science and Technology Publications, pp. 587–598.
- Nashed, Nader N, Christine Lahoud, and Marie-Hélène Abel (2021). "TCO: a Teacher Context Ontology". In: *2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*. IEEE, pp. 757–762.
- Nashed, Nader N et al. (2021). "Mood detection ontology integration with teacher context". In: *2021 IEEE 20th International Conference on Machine Learning and Applications*. IEEE.
- Natale, Jo Anna (1993). "Why Teachers Leave." In: *Executive Educator* 15.7, pp. 14–18.
- Negro, Alessandro (2021). *Graph-powered machine learning*. Simon and Schuster.
- Nickel, Maximilian et al. (2015). "A review of relational machine learning for knowledge graphs". In: *Proceedings of the IEEE* 104.1, pp. 11–33.

- Nilashi, Mehrbakhsh et al. (2020). "Intelligent recommender systems in the COVID-19 outbreak: the case of wearable healthcare devices". In: *Journal of Soft Computing and Decision Support Systems* 7.4, pp. 8–12.
- Noy, Natalya F, Deborah L McGuinness, et al. (2001). *Ontology development 101: A guide to creating your first ontology*.
- Noy, Natalya Fridman, Ray W Ferguson, and Mark A Musen (2000). "The knowledge model of Protege-2000: Combining interoperability and flexibility". In: *International Conference on Knowledge Engineering and Knowledge Management*. Springer, pp. 17–32.
- O'Banion, Terry (1997). *A learning college for the 21st century*. Greenwood Publishing Group.
- Oberle, Eva et al. (2020). "Do students notice stress in teachers? Associations between classroom teacher burnout and students' perceptions of teacher social-emotional competence". In: *Psychology in the Schools* 57.11, pp. 1741–1756.
- Ong, Shyue Ping et al. (2015). "The Materials Application Programming Interface (API): A simple, flexible and efficient API for materials data based on Representational State Transfer (REST) principles". In: *Computational Materials Science* 97, pp. 209–215.
- Opdenakker, Marie-Christine and Jan Van Damme (2006). "Teacher characteristics and teaching styles as effectiveness enhancing factors of classroom practice". In: *Teaching and teacher education* 22.1, pp. 1–21.
- Osguthorpe, Russell T and Charles R Graham (2003). "Blended learning environments: Definitions and directions". In: *Quarterly review of distance education* 4.3, pp. 227–33.
- Papamitsiou, Zacharoula and Anastasios A Economides (2014). "Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence". In: *Journal of Educational Technology & Society* 17.4, pp. 49–64.
- Paradarami, Tulasi K, Nathaniel D Bastian, and Jennifer L Wightman (2017). "A hybrid recommender system using artificial neural networks". In: *Expert Systems with Applications* 83, pp. 300–313.
- Park, Youngki et al. (2015). "Reversed CF: A fast collaborative filtering algorithm using a k-nearest neighbor graph". In: *Expert Systems with Applications* 42.8, pp. 4022–4028.
- Pasricha, Rajiv and Julian McAuley (2018). "Translation-based factorization machines for sequential recommendation". In: *Proceedings of the 12th ACM Conference on Recommender Systems*, pp. 63–71.
- Patel, Archana and Sarika Jain (2018). "Formalisms of representing knowledge". In: *Procedia Computer Science* 125, pp. 542–549.
- Pazzani, Michael J and Daniel Billsus (2007). "Content-based recommendation systems". In: *The adaptive web: methods and strategies of web personalization*. Springer, pp. 325–341.
- Peacock, Matthew (2001). "Match or mismatch? Learning styles and teaching styles in EFL". In: *International journal of applied linguistics* 11.1, pp. 1–20.
- Peimani, Nastaran and Hesam Kamalipour (2021). "Online education and the COVID-19 outbreak: A case study of online teaching during lockdown". In: *Education Sciences* 11.2, p. 72.

- Pence, Holly M and Ian K Macgillivray (2008). "The impact of an international field experience on preservice teachers". In: *Teaching and teacher education* 24.1, pp. 14–25.
- Pereira, Crystiam Kelle et al. (2017). "Linked data in education: A survey and a synthesis of actual research and future challenges". In: *IEEE Transactions on Learning Technologies* 11.3, pp. 400–412.
- Perifanou, Maria, Anastasios Economides, and Katerina Tzafilkou (2021). "Teachers' digital skills readiness during COVID-19 pandemic". In: *Teaching and Teacher Education* 113, pp. 1–12.
- Porter, Sarah (2015). *To MOOC or not to MOOC: How can online learning help to build the future of higher education?* Chandos Publishing.
- Potts, Boyd A et al. (2018). "Reciprocal peer recommendation for learning purposes". In: *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, pp. 226–235.
- Pratt, Sharon (2014). "Achieving symbiosis: Working through challenges found in co-teaching to achieve effective co-teaching relationships". In: *Teaching and Teacher Education* 41, pp. 1–12.
- Pratt, Sharon M et al. (2017). "Co-planning in co-teaching: A practical solution". In: *Intervention in School and Clinic* 52.4, pp. 243–249.
- Pursel, Bart et al. (2016). "Bbookx: Design of an automated web-based recommender system for the creation of open learning content". In: *Proceedings of the 25th International Conference Companion on World Wide Web*, pp. 929–933.
- Qiu, Yuanbo (2017). "The openness of open application programming interfaces". In: *Information, Communication & Society* 20.11, pp. 1720–1736.
- Ramaswami, M and R Bhaskaran (2009). "A study on feature selection techniques in educational data mining". In: *arXiv preprint arXiv:0912.3924*.
- Raza, Shaina and Chen Ding (2019). "Progress in context-aware recommender systems—An overview". In: *Computer Science Review* 31, pp. 84–97.
- Rendle, Steffen (2010). "Factorization machines". In: *2010 IEEE International conference on data mining*. IEEE, pp. 995–1000.
- Ricci, Francesco, Lior Rokach, and Bracha Shapira (2015). "Recommender systems: introduction and challenges". In: *Recommender systems handbook*, pp. 1–34.
- Rice, Jennifer King (2003). *Teacher quality: Understanding the effectiveness of teacher attributes*. ERIC.
- RISK, USEATYOW (2002). "Draft standard for learning object metadata". In: *IEEE standard* 1484.1.
- Roberts, Aedan GK, Daniel R Catchpoole, and Paul J Kennedy (2018). "Variance-based Feature Selection for Classification of Cancer Subtypes Using Gene Expression Data". In: *2018 International Joint Conference on Neural Networks (IJCNN)*. IEEE, pp. 1–8.
- Robin, Bernard R (2016). "The power of digital storytelling to support teaching and learning". In: *Digital Education Review* 30, pp. 17–29.
- Rodriguez, Agustin Canas et al. (2015). "A recommender system for non-traditional educational resources: a semantic approach". In: *Journal of Universal Computer Science* 21.2, pp. 306–325.

- Roeser, Robert W et al. (2012). "Mindfulness training and teachers' professional development: An emerging area of research and practice". In: *Child development perspectives* 6.2, pp. 167–173.
- Romero, Cristobal and Sebastian Ventura (2020). "Educational data mining and learning analytics: An updated survey". In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 10.3, e1355.
- Ross, Brian C (2014). "Mutual information between discrete and continuous data sets". In: *PloS one* 9.2, e87357.
- Rosse, Cornelius and José LV Mejino Jr (2003). "A reference ontology for biomedical informatics: the Foundational Model of Anatomy". In: *Journal of biomedical informatics* 36.6, pp. 478–500.
- Sabourin, Jennifer, Lucy Kosturko, and Scott McQuiggan (2015). "Where to next? A comparison of recommendation strategies for navigating a learning object repository". In: *User Modeling, Adaptation and Personalization: 23rd International Conference, UMAP 2015, Dublin, Ireland, June 29–July 3, 2015. Proceedings* 23. Springer, pp. 208–215.
- Sadeghyan, Saman (2018). "A new robust feature selection method using variance-based sensitivity analysis". In: *arXiv preprint arXiv:1804.05092*.
- Saleh, Majd (2018). "Digital ecosystem: towards a system of information systems". PhD thesis. Université de Technologie de Compiègne.
- Schafer, J Ben et al. (2007). "Collaborative filtering recommender systems". In: *The adaptive web: methods and strategies of web personalization*. Springer, pp. 291–324.
- Schneider, Silke L (2013). "The international standard classification of education 2011". In: *Class and stratification analysis*. Emerald Group Publishing Limited.
- Schwarz, Norbert and Seymour Sudman (2012). *Context effects in social and psychological research*. Springer Science & Business Media.
- Seger, Cedric (2018). *An investigation of categorical variable encoding techniques in machine learning: binary versus one-hot and feature hashing*.
- Sergis, Stylianos and Demetrios G Sampson (2015a). "Enhancing learning object recommendations for teachers using adaptive neighbor selection". In: *2015 IEEE 15th International Conference on Advanced Learning Technologies*. IEEE, pp. 391–393.
- (2015b). "Learning object recommendations for teachers based on elicited ICT competence profiles". In: *Ieee transactions on learning technologies* 9.1, pp. 67–80.
- Sergis, Stylianos, Panagiotis Zervas, and Demetrios G Sampson (2014). "Towards learning object recommendations based on teachers' ICT Competence Profiles". In: *2014 IEEE 14th International Conference on Advanced Learning Technologies*. IEEE, pp. 534–538.
- Shadbolt, Nigel, Tim Berners-Lee, and Wendy Hall (2006). "The semantic web revisited". In: *IEEE intelligent systems* 21.3, pp. 96–101.
- Shein, Khin Phyu Phyu and Thi Thi Soe Nyunt (2010). "Sentiment classification based on ontology and SVM classifier". In: *2010 Second International Conference on Communication Software and Networks*. IEEE, pp. 169–172.
- Sieg, Ahu, Bamshad Mobasher, and Robin Burke (2007). "Representing context in web search with ontological user profiles". In: *Modeling and Using Context:*

- 6th International and Interdisciplinary Conference, CONTEXT 2007, Roskilde, Denmark, August 20-24, 2007. *Proceedings 6*. Springer, pp. 439–452.
- Sikelis, Konstantinos, George E Tsekouras, and Konstantinos Kotis (2021). “Ontology-based feature selection: A survey”. In: *Future Internet* 13.6, p. 158.
- Silva, Felipe Leite da et al. (2023). “A systematic literature review on educational recommender systems for teaching and learning: research trends, limitations and opportunities”. In: *Education and Information Technologies* 28.3, pp. 3289–3328.
- Skaalvik, Einar M and Sidsel Skaalvik (2011). “Teacher job satisfaction and motivation to leave the teaching profession: Relations with school context, feeling of belonging, and emotional exhaustion”. In: *Teaching and teacher education* 27.6, pp. 1029–1038.
- Soemantri, Diantha, Cristian Herrera, and Arnaldo Riquelme (2010). “Measuring the educational environment in health professions studies: a systematic review”. In: *Medical teacher* 32.12, pp. 947–952.
- Stéfanou, Alexia (2017). “ÉDUCATION PRIORITAIRE”. In: *Éducation et Formations* 95, pp. 87–106.
- Tarus, John K, Zhendong Niu, and Ghulam Mustafa (2018). “Knowledge-based recommendation: a review of ontology-based recommender systems for e-learning”. In: *Artificial intelligence review* 50, pp. 21–48.
- Topping, Keith J (2005). “Trends in peer learning”. In: *Educational psychology* 25.6, pp. 631–645.
- Tuomi, Ilkka (2013). “Open educational resources and the transformation of education”. In: *European Journal of Education* 48.1, pp. 58–78.
- Urdaneta-Ponte, María Cora, Amaia Mendez-Zorrilla, and Ibon Oleagordia-Ruiz (2021). “Recommendation systems for education: systematic review”. In: *Electronics* 10.14, p. 1611.
- Vaidhehi, V and R Suchithra (2018). “A systematic review of recommender systems in education”. In: *International Journal of Engineering & Technology* 7.1.
- Van Harmelen, Frank, Vladimir Lifschitz, and Bruce Porter (2008). *Handbook of knowledge representation*. Elsevier.
- Varelas, Maria (2012). *Identity construction and science education research: Learning, teaching, and being in multiple contexts*. Vol. 35. Springer Science & Business Media.
- Veenman, Simon (1984). “Perceived problems of beginning teachers”. In: *Review of educational research* 54.2, pp. 143–178.
- Veisi, Hadi, Niloofar Aflaki, and Pouyan Parsafard (2020). “Variance-based features for keyword extraction in Persian and English text documents”. In: *Scientia Iranica* 27.3, pp. 1301–1315.
- Verbert, Katrien et al. (2012). “Context-aware recommender systems for learning: a survey and future challenges”. In: *IEEE transactions on learning technologies* 5.4, pp. 318–335.
- Villalba, Klinge et al. (2017). “A recommender system of open educational resources based on the purpose of learning”. In: *2017 Twelfth Latin American Conference on Learning Technologies (LACLO)*. IEEE, pp. 1–4.

- Villegas, Norha M et al. (2018). "Characterizing context-aware recommender systems: A systematic literature review". In: *Knowledge-Based Systems* 140, pp. 173–200.
- Voigt, Paul and Axel Von dem Bussche (2017). "The eu general data protection regulation (gdpr)". In: *A Practical Guide, 1st Ed., Cham: Springer International Publishing* 10.3152676, pp. 10–5555.
- Wang, Ning (2016). "Towards a competency recommender system from collaborative traces". PhD thesis.
- White, Piran CL et al. (2005). "Questionnaires in ecology: a review of past use and recommendations for best practice". In: *Journal of applied ecology* 42.3, pp. 421–430.
- Xiao, Jun et al. (2018). "A personalized recommendation system with combinational algorithm for online learning". In: *Journal of ambient intelligence and humanized computing* 9, pp. 667–677.
- Xu, Zhe et al. (2019). "Deep learning application for predicting soil organic matter content by VIS-NIR spectroscopy". In: *Computational Intelligence and Neuroscience* 2019, pp. 1–11.
- Yu, Hong Qing et al. (2011). "A linked data-driven & service-oriented architecture for sharing educational resources". In.
- Zaffar, Maryam et al. (2021). "A review on feature selection methods for improving the performance of classification in educational data mining". In: *International Journal of Information Technology and Management* 20.1-2, pp. 110–131.
- Zebari, Rizgar et al. (2020). "A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction". In: *Journal of Applied Science and Technology Trends* 1.2, pp. 56–70.
- Zhang, Qian, Jie Lu, and Guangquan Zhang (2021). "Recommender Systems in E-learning". In: *Journal of Smart Environments and Green Computing* 1.2, pp. 76–89.
- Zhao, Meijing et al. (2014). "A Knowledge-based teaching resources recommend model for primary and secondary school oriented distance-education teaching platform". In: *Frontier and Future Development of Information Technology in Medicine and Education: ITME 2013*. Springer, pp. 511–521.
- Zheng, Y., R. Burke, and B. Mobasher (2013a). "The role of emotions in context-aware recommendation". In: vol. 1050, pp. 21–28.
- Zheng, Yong, Robin Burke, and Bamshad Mobasher (2012). "Optimal feature selection for context-aware recommendation using differential relaxation". In: *Acm Recsys* 12.
- (2013b). "Recommendation with differential context weighting". In: *International Conference on User Modeling, Adaptation, and Personalization*. Springer, pp. 152–164.