APPROACH FOR OPTIMIZING COURSE RECOMMENDATION BASED ON INTEGRATING MODIFIED FELDER-SILVERMAN LEARNING STYLE MODEL WITH META-HEURISTICS ALGORITHM

ERUM ASHRAF

UNIVERSITI SAINS MALAYSIA

2023

APPROACH FOR OPTIMIZING COURSE RECOMMENDATION BASED ON INTEGRATING MODIFIED FELDER-SILVERMAN LEARNING STYLE MODEL WITH META-HEURISTICS ALGORITHM

by

ERUM ASHRAF

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

December 2023

ACKNOWLEDGEMENT

In the name of Allah, the most gracious and the most merciful, Whom I am thankful for giving me the strength, knowledge, ability, and opportunity to undertake this study and complete it satisfactorily. I would like to express my sincere gratitude to my supervisor, Assoc. Prof. **Dr. Selvakumar Manickam,** for his support and guidance throughout the completion of this thesis. His wide knowledge and understanding have been invaluable to me, especially in giving constructive comments and advice throughout this study. I am grateful to my co-supervisor **Dr. Shankar Karuppayah** for his valuable insights and contributions to the successful completion of this theses. I would like to convey my appreciation to all **academic staffs and colleagues** in National Advanced IPv6 Centre (NAv6), Universiti Sains Malaysia for their dedication and persistent support especially **Dr. Shams Ul Arfeen Laghari**. Besides that, the **administration and support staff** in Nav6 deserve a high mention for keeping everything running smoothly. My sincere heartfelt gratitude to my husband **Khurrum Mahmood,** who supported me both in my research and household matters that resulted in the success of this research. I dedicate my research to my father, **Ch. M. Ashraf,** who would surely be proud today for my achievement but unfortunately, he is not in this world today. I express my love to my dearest mother, **Ms. Khurshid Akhtar,** who has supported me throughout my life. I am thankful to my family, especially my brother **Zeeshan Ashraf** as they provided me encouragement during this study period. Finally, I am thankful to my colleague at Bahria University **Mahwish Pervaiz** who has helped and supported me over these years.

Thank you.

Erum Ashraf

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

LIST OF SYMBOLS

C	Centroid
dim	Dimension
ED	Euclidean Distance
ED_A	Euclidean Distance Accurate
ED_P	Euclidean Distance Predicted
f	Fitness Function
$fval$	Function Value
I	Set of Items
lb	Lower Bound
ΔLS	Difference in Learning Style
MV	Matching Value
$nvars$	Number of Variables
R	Set of Integers in Range
r_a	Average Rating
r_u	Average Similar User Rating
SS	Satisfaction Score
T	Threshold
U	Utility Function
ub	Upper Bound
$w_{a,u}$	Degree of Correlation

LIST OF ABBREVIATIONS

ADO.net	Active Data Objects.net
AI	Artificial Intelligence
ALS	Alternating Least Square
ARHR	Average Reciprocal Hit Rank
CBF	Content based Filtering
CBR	Case Based Reasoning
CF	Collaborative Filtering
COVID19	Corona Virus Disease 2019
CRS	Course Recommender System
CS	Computer Science
CSV	Comma Separated Values
DBN	Dynamic Bayesian Network
DSS	Decision Support System
ED	Euclidean Distance
EDM	Educational Data Mining
eDX	American Massive Open Online Courses
E-Learning	Electronic Learning
FP growth	Frequency Pattern Growth
FSLSM	Felder Silverman Learning Style Model
GA	Genetic Algorithm
HCT	High Confidence Tree
HF	Hybrid Filtering
ICT	Information and Communication Technology
IDE	Integrated Development Environment
IDSS	Intelligent Decision Support System

ILSQ	Index Learning Style Questionnaire
ILS	Index of Learning Styles
IM	Instruction Method
IR	Information Retrieval
IT	Information Technology
KBF	Knowledge based Filtering
LAN	Local Area Network
LMS	Learning Management System
LO	Learning Objects
LS	Learning Style
LSBCR	Learning Style based Course Recommendation
MOOC	Massive Open Online Course
MOODLE	Modular Object-Oriented Dynamic Learning Environment
NLP	Natural Language Processing
OWA	Ordered Weighted Average
RDBMS	Relational Database Management System
RMS	Root Mean Square
RO	Research Objective
RS	Recommender System
SLIM	Sparse Linear Method
SO	Surrogate Optimization
SQL	Structured Query Language
SVM	Support Vector Machine
TS	Teaching Strategies
URL	Uniform Resource Locator
WAN	Wide Area Network

LIST OF APPENDICES

PENDEKATAN MENGOPTIMUMKAN CADANGAN KURSUS BERDASARKAN INTEGRASI MODEL GAYA PEMBELAJARAN FELDER-SILVERMAN DENGAN PENGUBAHSUAIAN ALGORITMA META-HEURISTIK

ABSTRAK

Populariti e-pembelajaran meningkat disebabkan oleh teknologi yang membanjiri platform *Massive Open Online Course* (MOOC) dengan pelbagai kursus, menyebabkan kesaratan maklumat. Sistem pencadang boleh menyaring kursus-kursus tersebut tetapi menghadapi kesulitan dengan gaya pembelajaran kerana ketiadaan dataset piawai dan pendekatan pengukuran, yang menjadi penghalang terhadap pengumpulan data dalam institusi pendidikan yang mempunyai sumber yang terhad. Penyelidikan ini dapat menyelaraskan pemilihan kursus dengan gaya pembelajaran pelajar. Ia melibatkan pengenalpastian potensi gaya pembelajaran kursus, yang disahkan melalui algoritma genetik dan pengoptimuman penggantian meta-heuristik, serta menggunakan model Felder-Silverman untuk pengenalpastian gaya pembelajaran. Skema yang dicadangkan menyokong cadangan kursus diperibadikan berdasarkan kesesuaian gaya pembelajaran. Pengoptimuman skor kepuasan pelajar telah meningkatkan ramalan gaya pembelajaran kursus sebanyak 2% bagi Peratusan Min Ralat Mutlak (MAPE) dan 10.8% bagi Ralat Nilai Punca Min Kuasa Dua (RMSE). Setelah diaplikasikan kepada 129 kursus sains komputer MOOC, analisis kuantitatif mengesahkan kepentingan penemuan kajian ini, menekankan potensi sumbangan dalam bidang cadangan kursus e-pembelajaran. Dengan memadankan pelajar dengan kursus yang serasi dengan gaya pembelajaran,

ianya dapat meningkatkan keberkesanan pembelajaran, kepuasan pelajar, sambil mengurangkan kadar keguguran.

APPROACH FOR OPTIMIZING COURSE RECOMMENDATION BASED ON INTEGRATING MODIFIED FELDER-SILVERMAN LEARNING STYLE MODEL WITH META-HEURISTICS ALGORITHM

ABSTRACT

E-learning's popularity surges due to technology, flooding *Massive Open Online Course* (MOOC) platforms with courses, causing information overload. Recommender systems filter courses but struggle with learning styles due to lack of standardized datasets and measurement approaches, hindering data collection in resource-constrained educational institutions. This research streamlines course selection by matching it with learners' styles. It involves identifying potential courses learning style, validated through genetic and surrogate meta-heuristics optimization algorithms, and employing Felder-Silverman model for learning style identification. The proposed scheme supported the personalized course recommendations to students suitable with student learning style. Optimizing student satisfaction scores improves course learning style prediction by 2% Mean Absolute Percentage Error (MAPE) and 10.8% Root Mean Square Error (RMSE). Applied to 129 computer science MOOC courses, a quantitative analysis underscores the significance of the research findings, illuminating the potential theoretical, applied, and methodological contributions within the realm of e-learning course recommendation. This research holds promise for enhancing online learning experiences in multiple dimensions. By assisting learners in discovering courses aligned with their learning styles, it fosters improved learning outcomes, heightened learner satisfaction, and increased learner engagement, all while mitigating course abandonment rates.

CHAPTER 1

INTRODUCTION

1.1 Background and Introduction

The period of Education 1.0 began after the first industrial revolution that took place at the end of the 18th century. The education sector flourished with the introduction of machines and tools like the typewriter, ballpoint pen, paper-making machines, and mechanical printing. In this education period, the teacher was deemed the centre of education, and students had a passive role (Maria et al. 2018). The second industrial revolution coincided with Education 2.0 in the early 20th century. In this period, the fundamental source of education has now changed from teachers to open-source materials from libraries. New tools were introduced due to the technological advancements in Education 2.0, including electronic devices like calculators, computers, and printers. The role of the student was enhanced to being active, being the owner of the knowledge but the learning approach was still teacher-centred; however, peer assessment was encouraged. In this period, correspondence, education and broadcast education were also introduced (Tan et al. 2018).

Education 3.0 appeared at the end of the 20th century during the third industrial revolution. The advancement in Information Technology and other technological advancement mostly changes the learning process. Multimedia, virtual laboratories and other online tools supported the teaching. In this new communication era, the student and teacher transitioned to a vision in which they no longer needed to participate in a synchronous session for learning to happen (Nadiah et al. 2019).

Presently it is the period of Education 4.0 where the fourth industrial revolution combined technology and innovative pedagogical practices. The traditional learning method has been modified in Education 4.0. The use of technology and pedagogical approaches, strategies and styles are considered to provide more effective, accessible and flexible educational programs that are becoming popular in many universities (Sharma 2019). New programs adapt student-centred models where students have an active role in learning and are facilitated by the available and emerging technologies to enhance teaching and learning processes. As a result, innovative learning methods are emerging along with pedagogical approaches such as learning by doing, problem-based learning, challenge-based learning, and gamification-based learning.

1.1.1 Role of ICT in Education 4.0

In Education 3.0 and 4.0, Information and Communication Technology (ICT) has facilitated the education process. ICT encompasses the technological tools and resources, technologies and online platforms that facilitate information collection, access, and distribution. Computers and internet connections are employed to communicate and handle information for learning purposes, such as web-based learning and e-learning (Marín-Díaz et al. 2020). The utilization of ICTs in higher education has great significance as educational structures have transformed from a traditional teacher-centred model to a student-centred model. Machine Learning, Artificial Intelligence, High Data Processing, and Data Analytics are the technical implementations of ICT in Education 4.0 that support the pedagogical procedures (Peng et al. 2023). The use of ICT has transformed the formal teaching and learning processes by introducing Learning Management Systems (LMSs). LMS supply learning spaces where students and teachers work asynchronously or synchronously. Some of the most used LMSs are Blackboard, CANVAS, Google Classroom, Moodle, Sakai, and Edmondo. Other relevant implementations of tools and platforms include M-learning, web-based learning, intelligent tutoring systems, and virtual reality applications (Swerzenski 2021).

ICT has aided the evolution of Education 4.0, which offers various delivery modalities to deliver more accessible, relevant, personalized and flexible content like E-learning (Maria et al. 2018). The course delivery, interaction and facilitation in the E-learning program is provided through an information network- such as the internet, the local area network (LAN) or wide area network (WAN). E-learning influences current technological platforms to carry out remote processes through virtualization, digitalization and connectivity through synchronous and asynchronous activities. The key elements of Education 4.0 are depicted in Figure 1.1.

Figure 1.1 Key Elements of Education 4.0 Technologies

1.1.2 Decision Support Systems in Education (DSS)

ICT techniques are also applied to develop AI Engines to support decision support systems (Smith and Wong 2022). DSS has a practical application of great interest to the educational community. DSS is being used in the education industry to raise the standard of education and learning by incorporating various concepts such as educational data mining (EDM) (Shalabi 2020). The EDM is one of the essential methods where intelligent steps are implemented to extract data patterns in students' databases to discover key features. This new exploration field has developed exponentially and picked up prominence in the advanced instructive period because of its capability to improve the nature of instructive foundations and frameworks.

An intelligent decision support system (IDSS) is used in education to measure students' performance, predict the number of failures early in the learning process, or recommend courses (Alisan and Serin 2021). Implementing intelligent methods is essential for extracting data patterns and analytical information to discover hidden knowledge from student databases (Dasgupta 2019). In e-learning, LMS accumulates a vast amount of valuable information that can be explored to analyze students' behaviour and the effectiveness of course design. It will help to predict students' performance and their final mark, to group students according to their preferences, and in short, to improve the educational process (Swerzenski 2021). This new research area has become popular and has grown exponentially in the new era of modern education, the reason behind its potential and capacity building for familiarizing improvement in the quality of education infrastructure (Alexeyev and Solianyk 2020).

1.2 Motivation

There are two types of learning pedagogies: the first one being teacher-centric and the second one being learner centric. In teacher-centric pedagogy, the teacher controls the learner and decides the activities the learner should carry on during the learning process. On the other hand, in the learner-centric type, the learner himself decides the content he wants to learn and through what methodology can improve the learning performance (Shah 2020). E-learning has a personalization feature that can fulfil the requirement of different learners following their preferences like background, goals, personality and capability (Rawashdeh et al. 2021). Personalization means tailoring learners' needs to the technical needs and requirements, knowledge, learning style and other preferences of the learners'. This approach is valuable as it motivates the learners to enhance their performance (Elcullada Encarnacion et al. 2021).

The recommender platform manages information overloads in an E-learning system by filtering the data according to the learner's requirement. Recommender engines are being implemented in various fields such as E-commerce, Social Networks, E-health, Entertainment, Online Advertising, Travel and Hospitality, Education, News and Content Recommendation, Gaming, Finance and Investment, Food and Dining, Job and Career, and Healthcare., reflecting this social behaviour to support users in making wise choices with minimum effort (Sunitha and Kiran 2023). Recommender engines provide users with personalized assistance and screened information in daily life for decision-making. Personalization can guide the learners in the course recommendation process according to their interests and needs (Yildiz et al. 2023).

It is observed that more efforts are made to advance technology than to understand individual learners' needs, learning styles, and instructional design (Singh et al. 2021). Several recommender engines in E-learning have been introduced, but most focus on recommending learning material only (Sunitha and Kiran 2023; Tarus et al. 2017). Therefore, it encourages researchers to seek its benefits in the e-learning domain for course recommendation as they know recommender platforms have a high potential for achieving personalization at an advanced level in e-learning. These recommendations include accommodating multiple learning styles in assignments and other online education options. Learning style detection at the beginning of learning can help the learner to follow the learning more efficiently and improve the motivation to learn (El-Sabagh 2021).

Due to their unique characteristics, the recommendation mechanism has not been effectively applied in E-learning compared to other domains, such as e-commerce. The requirements of learners and the content of learning activities contain uncertain and vague information. For instance, the learners may not know which skills or courses they need; on the other hand, they are aware of what job they want to pursue. The context of learning is vital to learners, such as the course's purpose and learning style (Zhang et al. 2021). The above characteristics prompted the development of recommendation techniques and strategies for the e-learning context. It is expected that new e-learning mechanisms should appear to ensure personalization not only in terms of material but also by considering various factors that will make this domain complex. Depending on the items, recommender engines need different techniques, strategies, and architectural designs to fit the context of each scenario for automating the e-learning paradigm (Benhamdi et al. 2017).

1.3 Problem Statement

It is evident from the literature and current practices that the perceived effectiveness of E-learning among students has risen. E-learning has strived to unite learners, instructors, experts, practitioners, and various stakeholders within a single platform. Consequently, it has fostered a commendable culture of knowledge exchange facilitated by various online channels. This practice is significant in the current era where competition is rising in the growing world. Hence, prompt access to relevant information helps in the better growth of an individual. However, some shortcomings associated with E-learning are listed as follows:

i. A notable gap has been identified in the research concerning learning style-based course recommendations. This gap has led to an exploration of learning style-based challenges in e-learning, with a specific focus on selecting suitable courses from courseware based on learner requirements. This research area has received limited attention, potentially due to the complexity introduced by the distinct nature of learners and courses, warranting further investigation. Additionally, lack of influence of time on learning activities has been observed in existing learning style-based course recommendation system. Hence, there is a need for the development of a comprehensive course recommendation framework to address personalized learning styles.

ii. The lack of data is a major challenge for the development of learning style-based course recommendations. There is no single, universally accepted way to measure learning styles. This makes it difficult to collect data on learning styles in a consistent way. Students may be reluctant to share information about their learning styles, as they may perceive this information as being personal or sensitive. Educational institutions may not have the resources to collect and store data on learning styles. Numerous research papers have been found concerning movie recommendations, while there have been relatively few identified in the domains of health, tourism, and education-related recommender engines. This discrepancy can be attributed to the abundance of publicly available movie datasets. Consequently, there is a compelling need to create datasets in these other domains as well.

iii. The application of optimization techniques through meta-heuristic algorithms in e-learning course recommendation remains a relatively unexplored domain. While some initial steps have been made in developing adaptive applications based on learning styles, this area is still in its emerging stages, necessitating further research and development to reveal its full potential.

1.4 Research Questions

As mentioned earlier, there is a need to develop a novel framework for learning style-based course recommendations in an e-learning environment. Some of the research questions presented are:

RQ1: How can a course learning style (LS) identification model be designed to support an automatic learning style-based course recommendation model?

RQ2: What methods or techniques can be employed to accurately identify the LS associated with a given course?

RQ3: How can the effectiveness and accuracy of the course LS identification model be validated?

RQ4: How can a course recommendation model be developed that suggests courses based on the LS of the student?

RQ5: What algorithms or approaches can be utilized to match the LS of the student with the LS of available courses?

1.5 Research Objectives

To achieve the aim of this research study, the following specific objectives are defined that will resolve the research questions:

RO1: To design a course learning style identification model for supporting an automatic learning style-based course recommendation model.

RO2: To employ optimization techniques for the improvement and refinement of course LS identification models.

RO3: To develop a course recommendation model that recommends courses matching the student's learning style.

1.6 Research Scope

The focus of this research is to investigate and develop a methodology tailored specifically for E-learning courses. To accomplish this objective, data was extracted from Massive Open Online Courses (MOOC) platforms, allowing for the application of the proposed methodology to MOOC courses. This deliberate selection of MOOC courses aligns with the research aim, as it provides a relevant and representative sample for the evaluation and validation of the proposed methodology in the E-learning context. By utilizing MOOC courses as the basis for data extraction, this research aims to address the specific challenges and requirements associated with e-learning environments, ultimately contributing to the advancement of instructional design and delivery in online education.

This research presents a course recommendation model to be used in an E-learning domain. The model focuses solely on the user LS factor and consider LS identification of courses related to computer science field. The instructional methods and learning objects used to deliver the course contents are related to the support it provides for different LSs. The available instructional methods and learning objects data are used to identify the LS support of the computer science courses. The recommendation model matches the course LS support and the user LS. However, the user's LS is considered static and obtained through Felder and Silverman LS Model Questionnaire. One limitation of this research study is the absence of standard dataset availability to analyse proposed model's effectiveness. This problem has been tackled by performing an experiment and optimizing the model with experimental results to improve the accuracy in a real-world data set.

1.7 Research Methods

The methodology followed in this research thesis to obtain the objectives laid down in Section 1.5 consists of five phases described in Figure 1.2.

Phase 1: Problem Identification and Formulation of Research Objectives
- Understanding E-learning and role of recommendation mechanism - Identify shortcomings in existing recommendation engines in E-learning - Define research objectives
Phase 2: Literature Review
- Review of approaches used in existing course recommendation models - Study and select the LS model
Phase 3: Model Development Strategy
- Identify the course elements and their support for LSs - Selection technique to identify the LS of learners - Determine the approach to be used to develop the course recommender model
Phase 4: Development of Course Recommender Model
- Course recommender model development - Application on real-world data of experiment and optimization - Application of model in the E-learning domain
Phase 5: Results Analysis
- Evaluation of results - Suggestions for future research

Figure 1.2 Research Design

The first phase explores E-learning and recommendation mechanism' role in this domain. The various factors that are considered in course recommendation models are studied to identify their shortcomings. These shortcomings are translated into research objectives and the research scope to include to obtain these.

The second phase consists of a thorough literature review of the approaches used in course recommendation models, and their strengths and limitations are listed.

Various LS models available are studied to select the appropriate model to be used for this research study.

Based on the in-depth study of the literature and information explored and gathered in previous phases, a strategy to identify the course LS support is developed based on the available data in E-learning MOOC platforms using deductions from previous research studies. The selection of models and techniques is done to obtain the learners' LS. After determining the strategy to obtain course LS support and learners' LS, an appropriate approach to recommend courses based on LS is determined.

The fourth phase is to evaluate the performance of the proposed research model. A prototype model is developed in which the course LS identification, student LS and recommendation of course strategy are implemented. The results obtained from the prototype model are evaluated by the results obtained from recorded experimental values. The model is optimized with a fitness function that minimises deviation of the student feedback and evaluated result values. The finalized model is implemented in the E-learning domain for Computer Science, and course recommendation results are computed for random test cases comprising distinct LSs.

The last phase consists of analysis, discussion and evaluation of results obtained from the proposed model. Future improvements are also suggested to continue research in this area.

1.8 Research Contribution

The existing course recommender models employ various factors, including course rating, complexity, field, job interest, student feedback, course prerequisites, and instructor expertise. The characteristic of LS has only been used to filter the content that supports the learner's LS. This research study has introduced a model that will compute the LS support for the course available in the E-learning domain and recommend the courses matching the learners' LSs. The LS is a theoretical characteristic of a learner, and Felder and Silverman have quantified it in their learning model. Inspired by Felder and Silverman, the course LS support in the form of quantitative scores has been proposed in the model developed.

The first phase of research is focused on formulating a model that provides course LS support that is called Course LS Identification Model that is a fundamental part of the course recommender model. The lack of a standard data set is a limitation for validating the Course LS Identification Model. Therefore, a verification activity is performed that employs real-world data. During Covid19 universities were shifted to online learning mode and this provided an opportunity to verify the course LS identification model. Online study conducted in a local study was selected to obtain the data of students LSs, course content delivery and student satisfaction score that provided basis to test and optimize the Course LS Identification Model developed.

This research also depicts the methodology to implement the proposed course recommender model on an E-learning course dataset to recommend courses to learners per their LSs. The recommender model uses the Course LS Identification Model and identifies the course LS support for MOOC platform. To reduce computational expense, the k-means clustering algorithm is used to group courses with similar LS support. The demonstration of the recommender model is presented in this thesis in which data of 129 courses are scrapped and LS support for these courses is computed. Clustering is done to group similar courses with respect to the LSs in seven cluster groups. Test cases of learners are demonstrated, and a detailed explanation is provided to elaborate the working of the recommender model developed.

1.9 Chapter Organization

This thesis is organized into six chapters as follows.

Chapter 1 introduces the evolution of E-learning by discussing the development of formal education from Education 1.0 to Education 4.0. The significance and role of recommendation mechanism in E-learning are discussed along with their shortcomings to formulate the problem statement. The research methodology, scope and contributions are also presented. The research aim, questions and objectives are prepared to provide the road map for this research study.

Chapter 2 presents the literature review conducted in which recommender mechanism approaches and their application for course recommender models are discussed. The factors that are considered to recommend courses are also laid down. The LS that is the focus of the recommender model designed in this thesis is also studied and presented. Pros and Cons of different LS models developed in the past are listed. The justification for selection of the Felder and Silverman LS Model (FSLSM) is explained by highlighting its significance and ease of application. This chapter concludes by discussing the research gap in detail by conducting a survey of recommender engines using LSs in education.

Chapter 3 contains the research methodology followed to design the course recommender model based on the LS proposed for the E-learning domain. The procedure is explained sequentially explicitly with process flow. The relationship between the learning objects and instructional methods is associated with the course LS support identification. The data scrapping requirements to demonstrate the Course Recommender model are explained. The validation scheme to test and optimize the algorithm of Course LS Identification Model is discussed. Finally, the clustering scheme of MOOC courses data and approach to recommend courses to learners matching their LS is explained.

Chapter 4 discusses the implementation of the research methodology. The tools used are briefly introduced to develop the model. The dataset collection and formatting in the form feasible for the model are explained. The execution of the validation scheme for Couse LS Identification Model along with optimization of the threshold parameters that categorizes the LS in FSLSM style is laid down. This chapter is concluded by discussing the k-means clustering algorithm application and the method of matching of Course with student LS is recommender model developed for use in E-learning is explained.

Chapter 5 presents the results obtained through the validation of the Course LS Identification Model. The LS of students, Course LS and improvement in the Course LS Identification Model is presented. The results of the k-means clustering for the MOOC courses are presented. Finally, the demonstration of the course recommender model is done by recommending matching courses for test cases from the dataset of courses, obtained from MOOC platform for the field of computer science. The test cases are formed to represent the LS of distinct learners. This chapter concludes with analysis of the quantitative results obtained.

Chapter 6 summarizes the research thesis by discussing the achievement of the research objectives. The future course of study in this research area is suggested at the end to conclude this chapter.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter offers a comprehensive overview of course recommendations and the core concepts underlying this study. It reviews relevant research on various recommendation techniques and delves into learning style models. Furthermore, it highlights the limitations of each approach, which have driven the research in this area.

Section 2.2.1 explores relevant research in the area of course recommendation while also examining the methodologies used in recommendation systems. This evaluation is essential for understanding the current state of the art, identifying trends, and determining the strengths and limitations of existing systems. A thorough description of important LS models is provided in Section 2.2.2, including the Gregorc, Kolb, VAK (Visual-Auditory-Kinesthetic), Honey-Mumford, Hermann Brain Dominance Instrument (HBDI), 4MAT, and Felder-Silverman LS Model (FSLSM). Section 2.3 discusses the complex domain of recommendation systems, highlighting collaborative filtering, content-based filtering, knowledge-based filtering, and hybrid filtering strategies. These approaches serve as the foundation for the chapter's investigation of individualized LS-based course suggestions. Finally, Section 2.4 concludes the chapter with a review of research gaps, highlighting the lack of LS-based course recommendations in the e-learning environment and a concise summary is presented in Section 2.5.

To formulate a proper architecture for course recommender model, the literature review has been carried out as depicted in Figure 2.1.

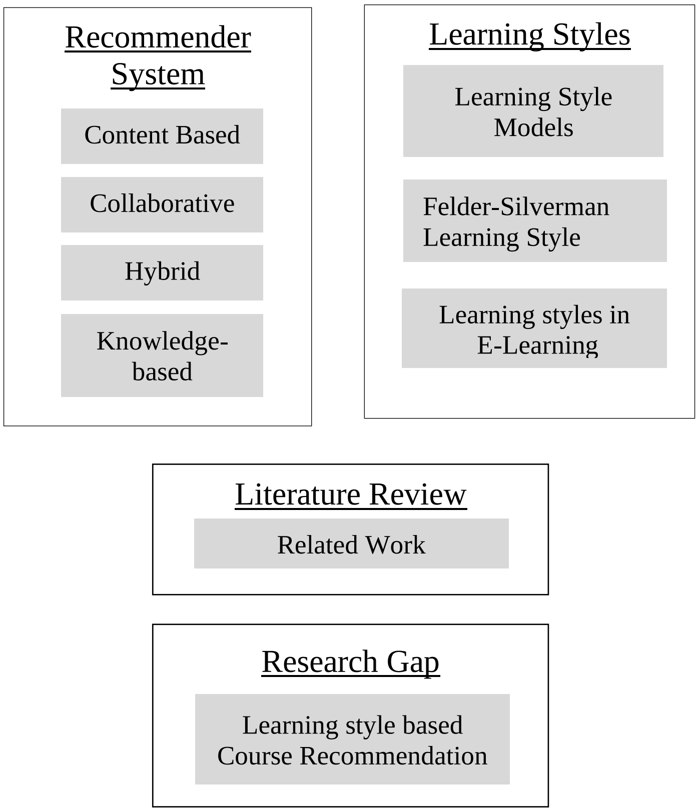

Figure 2.1 Outline of Literature Review

2.2 Background

In the background Section, an overview of recommendation system methodologies is presented, with an emphasis on their pivotal role in personalizing content delivery. LS models are also explored, clarifying their significance in shaping customized educational experiences. This Section establishes the foundation for the examination of the interaction between recommendation systems and LS models, a central focus of proposed study.

2.2.1 Recommender System

The development of Web 2.0 has caused information overload at an exceptional rate. An example of it can be seen as there are thousands of results obtained in a search engine when mobile or laptop is entered, which is outrageous for users. To handle the results, the significance of filters is evident, allowing users to obtain results per their specific interests. Information filtering systems tackle the information overload problem by filtering vital fragments from extensive dynamically generated information data (Agarwal et al. 2022). For information filtering, recommender systems (RS) aim to provide users with suitable and personalized content. The recommender system is defined as follows:

> *“A recommender system, or a recommendation system (sometimes replacing 'system' with a synonym such as platform or engine), is a subclass of information filtering system that provide suggestions for items that are most pertinent to a particular user.”*
>
> (Francesco et al. 2022)

A recommender system is essential for numerous domains such as products, movies, videos, books, music, CDs, news, and blogs. By employing these RS, users can save significant time and effort in searching for relevant information over the internet. Recommendation engines have an alternative methodology in which the users interact with each other, providing a richer experience. In e-commerce, online retailers and customers require RS to select a product for sale or purchase. The significance of recommendation systems in e-commerce is also quite evident, including providing users with videos, books, literature and content according to their personalized interests (Taneja and Arora 2018).

Recommendation systems have significance and application in various domains. Thus, they are an active research area that allows users to automatically obtain relevant information of their interest from a scenario of information overload.

There have been many improvements and advancements in these systems. Still, due to the enormous growth of the user population on social media, certain refinements are needed to be carried out on these systems. These refinements are required to make recommendation techniques applicable to a broader range of social media domains related to real-life applications. High-quality and exclusive RS must be built in various disciplines that will provide personalized recommendations to the users (Malik et al. 2020).

The mathematical expression representing a recommendation system is laid down as equation (2.1). In this equation, *U* represents the set of users, and I represent the set of items. The utility function f measures the relevance of the item to a particular user:

$$f: U * I \rightarrow R \tag{2.1}$$

In the above equation, *R* is a set of integers or real numbers within a specific range. The relevant items are selected for each user *U* that will maximize the utility function *f*.

$$\forall u \in U, i'_u = \max f(u, i) \quad where \;\; i \in I \tag{2.2}$$

The rating usually represents the utility of an item that indicates how much the user has given preference to that item. For example, if a rating of '9' is given to a movie by the user, it indicates a strong liking of a movie from that user. Table 2.1 enlists an example of a user-item matrix in which user 1 provides a rating for item 1 and item 2; however, for items 3 and 4, it is the recommendation system that should predict the ratings for the unrated items 3 and 4.

Table 2.1 User-Item Rating Matrix

Users/ Items	Item 1	Item 2	Item 3	Item 4
User 1	5	10	-	-
User 2	-	-	3	-
User 3	7	4	-	-

The foremost recommender system was created unknowingly by a librarian Grundy that used to recommend books to the user by forming stereotype groups. The users were interviewed to gather data about their preferences and match them with the data of the stereotyped group and recommend books accordingly (Rich 1979). The first formal development of such a system was Tapestry, developed by Xerox PARC. The users were allowed to comment on the documents they read, starting in the form of liking, or disliking it. This system used a collaborative filtering approach in which the users could filter the information according to their interests based on the notes and reviews of other users reading similar documents (Goldberg et al. 1992). In e-commerce, buyers were facilitated by providing personalized suggestions to purchase products. Malik et al. (2020) have presented their study about recommender techniques and their applications in various fields, such as e-tourism, e-learning, or e-commerce.

In this thesis, state-of-the-art methods related to RS have been explored. A comprehensive recommendation of taxonomy presenting a broader perspective encompassing the empirical literature is depicted in Figure 2.2, that lists the application areas, approaches and techniques employed in designing recommendation systems. RS are broadly divided into four categories, namely Collaborative Filtering (CF), Content-Based (CB), Knowledge-based (KB), and Hybrid. The upcoming Sections will elaborate on the recommendation system approaches and their pros and cons.

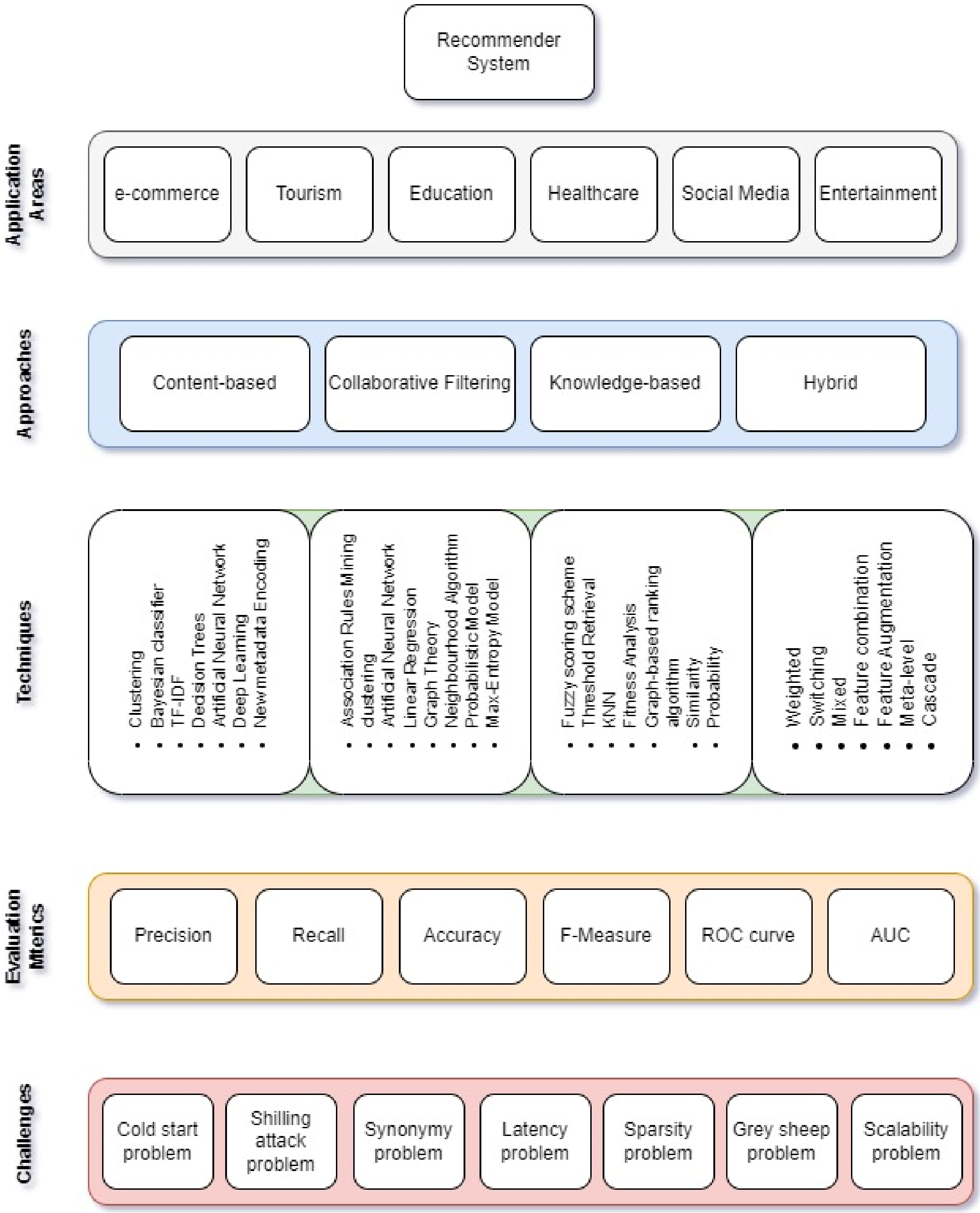

Figure 2.2 Taxonomy of Recommender Systems

The recommendation systems are categorized according to the approach it uses to recommend or filter the product. These approaches are elaborated on in this Section.

2.2.1(a) Collaborative Filtering (CF-RS)

A collaborative filtering recommendation system (CF-RS) uses a collaborative filtering technique in which the items are recommended based on the reactions of similar users. In this technique, a large group of users are searched to identify a smaller set with similar tastes to the user to recommend items. The database of a different identical set of users is maintained and updated with time. This technique assumes that the user's taste will not change with time. Figure 2.3 depicts the architecture of CF-RS.

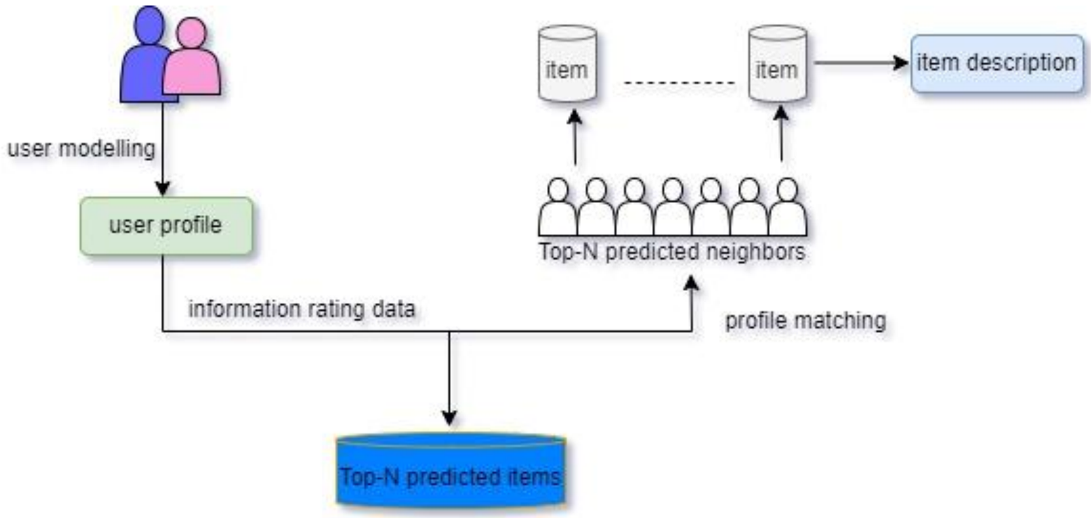

Figure 2.3 CF-RS Architecture

CF approach is the most popular and mature form of implementing recommendation systems. Broadly it has been divided into two categories:

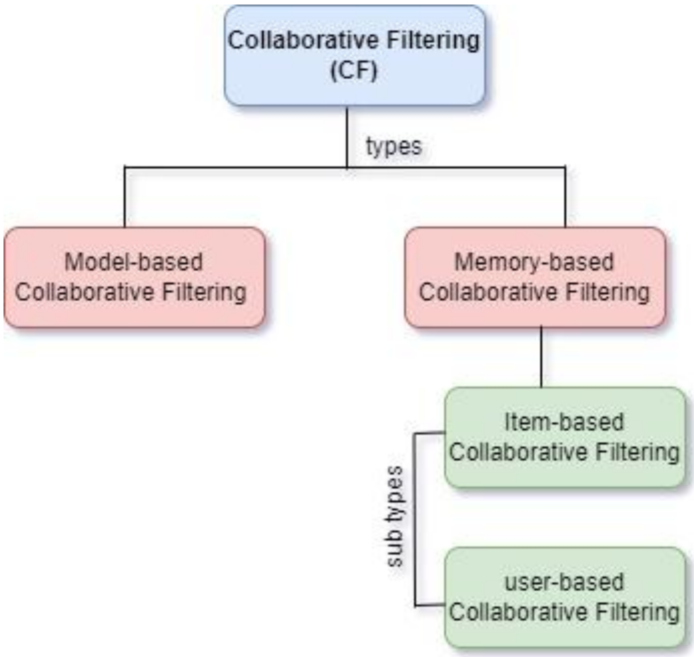

Figure 2.4 Types of Collaborative Filtering

Memory-based Approach

In the memory-based approach, the user-item rating matrix is used. A neighbourhood-based algorithm is applied that identifies the similarity among the items or users corresponding to the active user, and a recommendation is made accordingly. There are three steps in the neighbourhood algorithm; the first is the user-item rating matrix that computes the similarity between the users/items, the second is to predict the ratings, and the final step is the recommendation list generation that contains the top n-recommendations.

The similarity between users or items is obtained by computing the similarity weight between the users or items using the 'cosine similarity measures' or 'Pearson correlation coefficient'. The degree of similarity between the active and other users is computed by using the Pearson correlation coefficient (Resnick et al. 1994). The value of this coefficient varies from –1 to +1. If the value is closer to 1, it shows a strong correlation between users meaning they have similar tastes. If the value is