Fuzzy Systems and Soft Computing
ISSN : 1819-4362

REVIEW BASED RECOMMENDATION SYSTEM FOR E-LEARNING RESOURCES USING SENTIMENT ANALYSIS AND DEEP LEARNING

Anurag Chandna, Ph.D Scholar, Department of Computer Science & Engineering, Quantum University, Roorkee
Satender Kumar, Dean Academic, Quantum University, Roorkee, India

Abstract

In the rapidly growing e-learning domain, the vast amount of available resources poses a challenge for learners in selecting the most suitable material. Traditional recommendation systems often fall short in addressing individual learner preferences and sentiments. This paper reviews the application of sentiment analysis and deep learning techniques in developing a review-based recommendation system for e-learning resources. We explore various methodologies, their implementation, and the challenges faced, along with a discussion on future directions in this field. Deep learning techniques have significantly improved the accuracy and efficiency of these systems. However, there is a lack of literature regarding classification in systematic review papers that summarize the latest deep-learning techniques used in recommendation systems. Moreover, certain existing review papers have either overlooked state-of-the-art techniques or restricted their coverage to a narrow spectrum of domains. To address these research gaps, we present a systematic review paper that comprehensively analyzes the literature on deep learning techniques in recommendation systems, specifically using term classification. We analyzed relevant studies published between 2018 and 2023, examining the techniques, datasets, domains, and measurement metrics used in these studies, utilizing a thorough SLR strategy. Our review reveals that deep learning techniques, such as graph neural networks, convolutional neural networks, and recurrent neural networks, have been widely used in recommendation systems. Furthermore, our study highlights the emerging area of research in domain classification, which has shown promising results in applying deep learning techniques to domains such as social networks, e-commerce, and e-learning. Our review paper offers insights into the deep learning techniques used across different recommendation systems and provides suggestions for future research. Our review fills a critical research gap and offers a valuable resource for researchers and practitioners interested in deep learning techniques for recommendation systems.

1. Introduction

The e-learning industry has experienced significant growth, particularly with the advent of the COVID-19 pandemic, which necessitated a shift from traditional classroom-based education to online platforms. With this growth, there has been a proliferation of e-learning resources, making it increasingly difficult for learners to find resources that best fit their needs. Traditional recommendation systems, which often rely on collaborative filtering or content-based filtering, struggle to capture the nuanced preferences of individual learners. Sentiment analysis, combined with deep learning, offers a promising approach to enhance recommendation systems by analyzing learner reviews to gauge their sentiments and preferences. As data becomes more readily available, recommendation systems are gaining popularity in e-commerce. In today's business world, data-driven decisions are crucial, and many companies are incorporating recommendation system features into their websites and apps to enhance user experience and increase revenue [1], [2], [3], [4], [5]. The purpose of recommendation systems is to provide personalized and relevant suggestions to users based on their past behavior, preferences, and interests [6], [7], [8], [9], and to solve the problem of information overload in various domains such as e-commerce [10], [11], [12], e-learning [13], [14], [15], social networks [16], [17], [18], [19], [20], [21], and entertainment [22], [23], [24], [25]. However, recommendation systems face challenges such as data sparsity, cold start, and the need for collecting past user feedback [26], [27]. Researchers are developing more effective recommendation algorithms to overcome these challenges and improve accuracy and user satisfaction [28], [29]. As the amount of data being collected continues to increase due to technological advances, there is a growing need for techniques that can efficiently handle large amounts of data [30], [31].

Deep learning techniques based on artificial neural net- works have proven exceptionally effective in accurately predicting outcomes for big data applications, particularly for recommendation systems

[32], [33]. Deep learning models can learn and extract relevant features from raw data, making them highly effective in handling complex and high-dimensional datasets [34]. This capability is particularly relevant in recommendation systems, where personalized recommendations can be made to users based on their behavior and preferences [35], [36].
Given the considerations mentioned above, there is a need for a comprehensive and practical guide to deep learning-based recommendation systems. No systematic review has been conducted that focuses explicitly on deep learning-based recommendation systems using broad search terms that cover a wide range of recommendation domains. Furthermore, there is a lack of recent literature reviews that include the most advanced deep learning techniques developed within the last five years. Most importantly, no systematic study classifies deep learning techniques and domains into more appropriate categories based on the terms used in existing research articles. Such categorization would significantly expedite the summarization of future research in this field.

2. Sentiment Analysis in e-Learning

The field of sentiment analysis plays a crucial role in text mining and holds significant importance in recommendation systems [15]. For instance, Alatrash et al. [15] introduced a recommendation approach that combines sentiment analysis and genre-based similarity in collaborative filtering methods. Dang et al. [29] suggested the use of BERT for genre preprocessing and feature extraction, along with hybrid deep learning models for sentiment analysis of user reviews. Moreover, Liu et al. [95] proposed a novel multilingual review-aware deep recommendation model (MRRec) for rating prediction tasks. Sentiment analysis, or opinion mining, involves extracting and analyzing subjective information from text data. In the context of e-learning, sentiment analysis can be used to understand learners' opinions and feelings about specific resources. By analyzing reviews and feedback, sentiment analysis can help in identifying the strengths and weaknesses of e-learning materials.
Sentiment analysis, also known as opinion mining is a natural language processing technique (NLP) technique to identify the emotional tone behind a text Bansal et al., [96]. The technique has been widely used to extract user opinions on products and services from their reviews and create actionable knowledge for an entity Ligthart et al., [97]. This enables businesses to improve their strategies and gain insights into customers' feedback about their products and services. Opinion mining is a multi-disciplinary field that includes machine learning, NLP, sociology, and psychology to detect underlying customer or user opinions. Social media sites like Twitter, Facebook, and Instagram are significant sources of user opinions. The analysis of these sources to extract sentiment or opinions of users started a decade ago Tao et al., [98]; Zhou et al., [99].
Advancements in technology have transformed fields like health- care Tao et al., [100]; Shaik et al., [101] and education by adopting AI and NLP techniques. In the education domain, student feedback plays a vital role in evaluating and analyzing learning management systems, teaching, pedagogical procedures, and courses Elfeky et al., [37]. Educational institutions use student feedback surveys at the end of each semester to record their opinions on courses enrolled in the semester McKinney et al., [38]. The feedback comprises both qualitative and quantitative data which includes demographics of students, courses, ratings, and comments. The quantitative data can provide a statistical understanding of feedback on the courses, but the students' intent can be determined based on qualitative data analysis. The textual comments have to be preprocessed through NLP techniques such as feature extraction, and feature selection for further analysis Zhao et al., [39]. Qualitative data analysis enables listening to students' opinions on each course, content, and teaching.
In sentiment analysis, the initial step is to label text with emotional tags like positive, negative, or neutral which denotes students' emotional opinions on the services provided. The levels of the sentiment analysis differ based on application requirements Zhang et al., [40]. Few applications might need an overview of the student satisfaction report and few might need fine-grained analysis at the topic level to understand which aspect of the course delivery has negative reviews for improvisation. The manual annotation or labelling of the sentiment orientation is time-consuming Liu et al., [41] and requires many resources with pedagogical understanding in education. This challenge has been addressed by developing different sentiment anno-tation approaches using lexicons and corpus. These techniques act as unsupervised techniques for initial level understanding of the student feedback. AI's role in sentiment analysis is unavoidable as it assists to process and analyzing a large number of student

comments Zhu et al., [42]. AI methodologies like machine learning, deep learning, and transformers Acheampong et al., [43] are capable of learning student opinions with attention mechanisms and classifying or predicting their emotions for unlabeled student comments Kuleto et al., [44]. The unsupervised sentiment annotation techniques and the AI methodologies overcome the challenge of manual labelling to a certain extent.

Sentiment analysis has the potential to extract student opinions at the document level, sentence level, entity level, and aspect level with their sentiment orientation Ligthart et al., [97]. The document level analyses a comment and determines the overall sentiment of the comments towards a course is positive, negative, or neutral. At the sentence level, the sentiment extracts from each sentence and helps calculate a course's positives and negatives. Entity-level sentiment extraction combines entity and sentiment analysis to provide student opinion on an entity like tutor, course, and assignment. Aspect-based sentiment analysis is a fine-grained analysis of different data categories in a comment and identifies sentiment orientation on each data category. According to an education application, student feedback data will be analyzed at different levels. For example, a decision-making application would consider document-level sentiment analysis, and to understand student engagement, the analysis would be at the aspect level. (see Fig. 1).

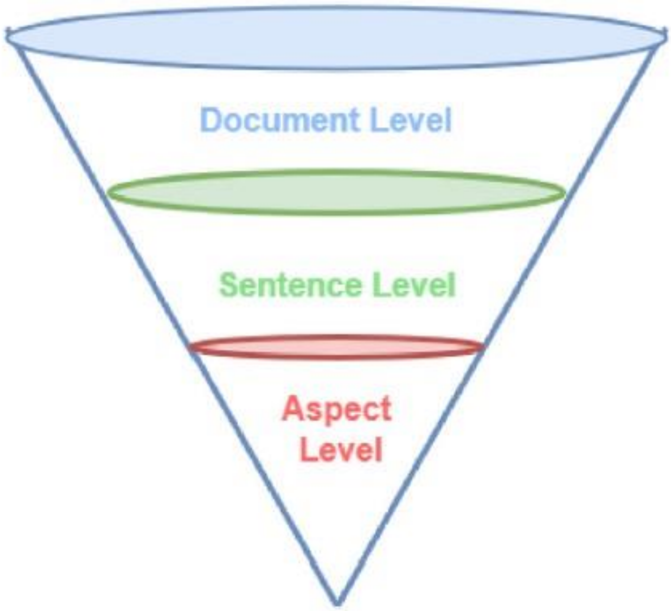

Fig. 1. Sentiment Analysis in Education.

2.1 Document level

In the education domain, developing sentiment extraction tools would require research resources. Hence, many institutions employ general sentiment extraction tools which are not domain specific. Dolianiti et al. [45] compared the performance of five commercial sentiment analysis tools IBM Watson Natural Language Understanding, Microsoft Azure Text Analytics API, OpinionFinder 2.0, Repustate, and Sentistrength with educational domain-specific tools in the document and sentence levels. In that study, two educational datasets were used which contain student forum posts of two courses across a semester in the learning management system. The sentiment orientation of the forum posts in datasets was manually annotated in two different versions: document level and sentence level. Four education domain tools, two for each course, were developed using SVM and k-fold cross-validation techniques. The study reported that educational-domain tools outper- formed commercial tools in one of the courses at both document and sentence levels. Faculty performance evaluation using document-level sentiment analysis was discussed by Ahmad et al. [46]. Two machine learning classifiers SVM and NB were trained on a preprocessed dataset of 5000 comments and achieved an accuracy of 72.80% and 81% respectively.

2.2 Sentence level

Sivakumar and Reddy [47] used the cosine similarity method to measure semantic relatedness between aspect words and student opinion sentences. The dataset used in that study was extracted from Twitter API, preprocessed, and classified the comments into seven aspects at the sentence level. Three machine learning algorithms decision trees, SVM, and NB were used to classify the sentences into different aspects. Subjective sentences were extracted using parts-of-speech (POS) tagging and their sentiment orientation was attributed using a lexicon-based approach SentiWordNet. Nikolić et al. [48] conducted a sentence- level analysis to extract one or more aspects in the sentences and classify their polarity into positive or negative sentiment. The authors used SVM, cascade classifier, and rule-based methods for aspect extraction and merged them. Similarly, the sentiment was detected using a merged component of SVM and a dictionary-based approach. Overall, the research was able to detect negative sentiment with an F-measure of 0.94 whereas positive sentiment was identified with an F-measure of 0.83 only.

2.3 Entity level

In an educational context, student feedback on entities such as teachers, learning management systems, or a specific concept of course content would help analyses their opinions. Yang [49] discussed automatic entity extraction in educational contexts such as content, teacher, lesson, and curriculum. Named Entity Recognition (NER) acts an important role in entity-level sentiment analysis which could a word or phrase that clearly detects a person, company or location Li et al., [49]. Ding et al. [50] developed an entity-level sentiment analysis tool, SentiSW, to extract sentiment and entity from comments in the form of the tuple ‘‘sentiment, entity)’’. The authors used TF– IDF and Doc2Vec to evaluate sentiment classification and achieved the best performance with Gradient Boosting Tree and Linear Support Vector machine. The results show that the proposed SentiSW tool outperformed existing tools like SentiStrength and SentiStrength-SE in positive and neutral comments classification. Li and Lu [51] pro- posed a novel approach to learning latent sentiment scopes at the entity level with named entities and sentiment is highlighted.

2.4 Aspect level

Aspect-based sentiment analysis (ABSA) is the most widely adopted approach for sentiment extraction in education. This approach provides a fine-grained analysis of educational data at phrase or sentence levels and extracts opinions or emotions at key aspects or entities. To perform ABSA, sentiment analysis techniques have to be ensembled with topic modelling techniques such as LDA, Latent Semantic Analysis (LSA), Non-negative matrix factorization (NMF), and Probabilistic latent semantic analysis (PLSA). A detailed definition and applications of the topic modelling techniques in educational data were discussed by Shaik et al. [52].

Rosalind and Suguna [53] proposed an ABSA system to extract student satisfaction on online courses in Coursera using a machine learning algorithm. Aspects in the student reviews were retrieved using unsupervised and semi-supervised LDA techniques. The actual reviews were segmented into sentences and then the aspects and their sentiment polarity were estimated. A customized lexicon was used to calculate the sentiment polarity of the sentences and it yield positive polarity for learning aspects, lab, job/career, and grade/test, and negative polarity for instructor, content, course, fee, and teaching. The maximum entropy classifier was trained and tested for the classification of multi-aspects and sentiments. The proposed model achieved an accuracy of 80.67% for aspect-based sentiment classification. Similarly, Edalati et al. [54] conducted aspect-based opinion mining on student comments from the Coursera platform to study the experience of conducting lectures and taking classes online. The authors deployed RF, SVM, decision tree, and deep learning models to identify teaching related-aspects and predict student opinions on the aspects. RF has the best performance with an F1 score of 98.01% and 99.43% in aspect identification and aspect sentiment classification respectively.

Wehbe et al. [55] proposed a Spatiotemporal sentiment analysis frame-work to analyze an education-related Twitter dataset using ASBA, sentiment analysis, and emotional analysis based on specific time and location. This framework used five machine learning classifiers decision tree, RF, multinomial, SVC, and gradient boosting to classify six emotions such as happiness, surprise, sad, anger, fear, and disgust, and four aspects educational rights, job security, financial security, safety, and death. In both aspects and emotions classification, RF had the best performance with an accuracy of 96.99% and 88.72% respectively. Aspect-level sentiment analysis can be applied to verbal speech to extract student emotions and predict their performance in collaborative learning. Dehbozorgi and Mohandoss [56] proposed a multi-class emotion analysis on students’ speech in teams and their performance. The authors classified emotions such as anger, happiness, sadness, surprise, and fear using Text2Emotion, a python package. Rule-based POS tagging was used to extract aspects, and then the k-nearest neighbour (KNN) algorithm was used to predict student performance by connecting extracted aspects and emotions. Kastrati et al. [57] proposed an aspect-level sentiment analysis framework to identify sentiment or opinion polarity towards a given aspect related to MOOCs. The pro- posed framework used weakly supervised annotation to identify aspect categories in unlabeled students' reviews. This reduces the need for manually annotated datasets in deep learning techniques. The authors used the convolutional neural networks (CNN) model for the prediction of aspects and sentiment classification. The proposed framework achieved an F1 score of 86.13% for aspect category identification and 82.10% for aspect sentiment classification. The research works adopted AI in different levels of sentiment analysis are presented in Table 1.

Table 1

Sentiment analysis in education.

SA Level	Algorithms	Application	Reference
Document Level	SVM, NB	Evaluate the documents' overall opinion towards a context is positive or negative.	Dolianiti et al. [45], Ahmad et al. [46]
Sentence Level	SVM, NB, cascade classifier, rule-based methods	Breakdown educational feedback and perform sentiment analysis at each sentence in a document	Sivakumar and Reddy [47], Nikolić et al. [48]
Entity Level	Gradient boosting tree, SVM, NER, TF–IDF, Doc2Vec	Analyse sentiment or opinion in feedback towards an entity in educational domain	Yang [49], Li et al. [58], Ding et al. [50], Li and Lu [51]
Aspect Level	LDA, RF, SVM, decision tree, CNN, KNN	Understand positive or negative aspects in educational practices.	Rosalind and Suguna [53], Edalati et al. [54], Wehbe et al. [55], Dehbozorgi and Mohandoss [56], Kastrati et al. [57]

Sentiment analysis can draw student opinions on different levels of feedback and provide insights to educational institutions for making informed decisions. However, the sentiment analysis can be performed only after labelling or annotating the text data with its sentiment orientation such as positive, negative, or neutral. There are supervised and unsupervised approaches to labelling or annotating student feedback. In the next section, sentiment annotation techniques that are adopted in education are discussed.

3. Techniques for Sentiment Analysis

Sentiment annotation is to label a document, sentence, or phrase with its semantic emotions that could be positive, negative, or neutral. The annotation can be fine-grained with micro-level analysis to extract precise emotions of users towards a service or product. This annotation mechanism can be applied to educational applications where student feedback towards learning management system, course content, learning–teaching practices, and instructor teaching abilities. Based on student feedback, the educational infrastructure can be stream- lined, can predict student performance, and support students with personalized learning. However, an increase in the number of student enrolments and their feedback made the manual sentiment annotation process next to impossible as it requires massive resources and time. Advancements in Artificial Intelligence and NLP methodologies led to a variety of tools for sentiment analysis. The tools are built on a lexicon-based approach, corpus-based approach, machine learning, deep learning, and transformer approaches.

Yeruva et al. [59] discussed differences between human annotators and machine annotators in sentiment analysis tasks. The authors presented a human-in-the-loop approach to explore human–machine collaboration for sentiment analysis. The sentiment annotation data obtained from 60 human annotators were compared with machine annotations extracted from six toolkits CoreNLP's sentiment annotator, Vader, TextBlob, Glove+LSTM, LIME, and RoBERTa. The authors con- ducted coefficient correlation analysis to understand the importance of features, words, and topics in human–machine collaboration. The results show that computational sentiment analysis has high performance and can be applied to text analysis.

3.1 Unsupervised annotation techniques

3.1.1 Lexicon-based approach

Sentiment and opinion words often act a vital role in the sentiment annotation of a document or sentence Catelli et al., [60]. Identifying the sentiment and opinion words can help to categories the sentiment in an unsupervised manner Dolianiti et al., [45]. In a lexicon-based approach, a sentiment dictionary with lexical units like words or phrases and their corresponding sentiment orientation like real values (eg: ranging from −1 to +1), classes (eg: positive, negative, or neutral), fine-grained classes like (eg: very positive to very negative). The sentiment orientation is based on the polarity of the content words like adjectives (Hatzivassiloglou and McKeown, [61]; Taboada et al., [62]), adverbs (Benamara et al., [63]), verbs (Vermeij, [64]), nouns (Neviarouskaya et al., [65]) and phrases in a sentence or document. Different lexicon-based approaches are developed for the English language based on the core idea of the sentiment dictionary. The approaches are SentiWordNet (Baccianella et al., [66]), Opinion Finder (Wilson et al., [67]), Bing Liu's Opinion Lexicon (Liu, 2012), MPQA subjectivity lexicon (Wilson et al., [68]), Harvard General Inquirer (Stone et al., [69]), AFINN (Nielsen, [70]), Senti- Ful (Neviarouskaya et al., [70]), Vader (Hutto and Gilbert, [72]), TextBlob and so on.

Tzacheva and Easwaran [73] used National Research Council (NRC) (Ortony and Turner, [74]; Mohammad and Turney, [75]) lexicon to label student feedback with fine-grained emotions such as joy, fear, trust, anger, sadness, disgust, and anticipation for teaching innovation assessment. The authors evaluated the impact of active learning methodologies like flipped classroom approach (Maher et al., [76]) and lightweight teams (Latulipe et al., [77]; MacNeil et al., [78]) implementation using emotion detection and sentiment analysis. The results show that the trust component increased from the time period of 2015 to 2020. Existing lexicon methods can be improved by modifying the strategies adopted to develop the method. Rosalind and Suguna [79] proposed steps to improve the existing Bing lexicon and develop a customized sentiment lexicon (CSL). The authors calculated sentiment polarity on academic course feedback texts using Bing sentiment Lexicon and CSL approaches. The process is to tokenize a sentence using the bag-of-words (BoW) method and calculate the polarity score of each word with Bing and CSL methods to output the cumulative polarity score of the sentence. The estimated cumulative score decides the sentiment orientation of the sentence. The CSL approach performed better than the existing Bing sentiment lexicon at the document level, detection of opinion words, and polarity scoring.

Lexicon-based approaches can perform unsupervised labelling of educational data and avoid manual labelling. The challenging part of the lexicon-based approach in sentiment analysis is domain or context understanding. The dictionary-based approach has a simple mapping of certain keywords with their sentiment. The words outside the listed keywords or opinions of students in the context are ignored. To avoid this, corpus-based sentiment analysis can be adopted.

3.1.2 Corpus-based approach

A corpus is a collection of texts on a specific topic or domain. The corpus-based approach in sentiment analysis is based on con- occurrence statistics and syntactic patterns of words in text corpora. This helps to enhance the sentiment lexicons with prior information about words across the semantic orientation of sentiment. The under- lying intuition of corpus-based techniques for sentiment lexicon is to calculate the semantic distance between a word and a set of positive or negative words to estimate the semantic polarity of the target word. This process could help to adapt the domain-independent sentiment lexicon to a domain-specific lexicon Alqasemi et al., [80].

3.2 Supervised annotation techniques

The lexicon and corpus-based approaches are capable to label text data in student feedback with sentiment orientation. However, processing and analyzing the opinions and emotions of multiple students' feedback or any other feedback would be a challenging task. AI methods have the potential to get trained on preprocessed NLP data and be able to classify or predict the sentiment orientation of the data.

3.2.1 Machine learning

A learning management system (LMS) plays an important role to provide access to course content for offline and online students and record their engagement. Ömer Osmanogˇlu et al. [81] analysed student feedback gathered from a university to assess the course materials. The authors used machine learning

techniques to classify the materials into positive, negative, or neutral sentiments and then improve the course materials with negative feedback for their upcoming semester. Six classifiers multinomial logistic regression, decision tree, multi-layer perceptron, XGBoost, support vector classifier, gaussian Naive Bayes, and k-nearest neighbours were used after preprocessing the student comments. Logistic regression was able to perform better than the other five classifiers. To implement a feedback analysis system, Lwin et al. [82] processed student textual comments along with quantitative ratings. The quantitative rating scores were clustered using the K-means clustering algorithm. The authors also used six classifiers support vector classifiers, logistic regression, multi-layer perceptron, and random forest to classify the clustered dataset. The textual comments were manually labelled as positive or negative as their sentiment. A naive Bayes classifier was used to train the labelled dataset to classify the comments into positive or negative. Faizi [83] proposed a machine learning approach to classify learners' feedback sentiment towards YouTube educational videos. Traditional machine learning methods such as random forest, logistic regression, Naive Bayes, and SVM were adopted to learn the feedback datasets and classify the learner comments with sentiment as positive or negative. The SVM algorithm was able to perform better than other models with an accuracy of 92.82% on a combination of unigrams and bigrams and an accuracy of 92.67% for associations of unigrams and trigrams. The algorithms discussed so far might have certain limitations individually. Kaur et al. [84] proposed a hybrid frame-work based on three classifiers random forest, logistic regression, and SVM models. The authors used the lexicon- based method sentiwordnet to label student comments with positive or negative sentiments. The proposed hybrid classifier was trained on the labelled dataset and compared its performance
with the SVM model at different test-training ratios. The hybrid classifier outperformed the SVM model in all classification metrics.

3.2.2 Deep learning

Early identification of students who are likely to fail a course can be predicted using sentiment analysis and deep learning techniques. Yu et al. [85] proposed a deep learning model convolutional neural networks (CNN) to learn structured data like attendance, grades and unstructured text feedback from 181 undergraduate students. The text feedback was manually annotated with positive, negative, or labelling based on the Self-Assessment Manikin rating scale. The authors trained support vector machines and CNN models with structured and unstructured feedback. The CNN model outperformed the SVM model with an F-measure of 0.78, 0.73, and 0.71 for the 5th week, 7th week, and 9th week of a semester. Similarly, the CNN model was used to evaluate lecturer effectiveness based on student feedback to a questionnaire by Sutoyo et al. [86]. The model was able to achieve accuracy, precision, recall, and F1-Score of 87.95%, 87%, 78%, and 81%, respectively. Deep learning models can be further enhanced by adding attention layers to identify sadness influence of words on emotion. Sangeetha and Prabha [87] proposed a multi-head attention fusion model for sentiment analysis of student feedback. The input sequences of sentences from feedback are processed in parallel across the multi-head attention layer with word and context embeddings like Glove and Cove. The outputs of the two multi-head attention layers with the embeddings are passed to the deep learning model LSTM. The dropouts of these layers are regulated to improve the accuracy of the model. The proposed fusion model was able to classify three sentiment orientations positive, negative, and neutral more accurately when compared to individual multi-head attention and LSTM.

3.2.3 Transformers

Dyulicheva and Bilashova [88] proposed a bidirectional encoder representation from transformers (BERT) model for sentiment detection of students, identifying top words describing positive and negative polarities. The authors used K-Means clustering and cosine similarity on 300 MOOCs titles from Udemy to extract 14 clusters and top words in each cluster. The BERT model was used to investigate the relationship between student–teacher, student-course, and description of issues while learning. The results show more negative sentiment towards courses when compared to teachers. Similarly, Li et al. [89] analyzed the sentiment of learning comments MOOCs using a shallow BERT-CNN model. The authors took advantage of deep learning not to depend on feature engineering and ensemble with the BERT model. The proposed BERT-CNN with a self-attention mechanism performed nearly equal to the traditional BERT model even after reducing the number of parameters to half. The proposed shallow BERT-CNN model with 6 layers outperformed all other lexicon-based methods, BERT variants with just 61 million parameters with an accuracy, F1-Score (Positive class), and F1-Score (Negative class) of 92.8&, 95.2%, and 81.3% respectively.

4. Deep Learning in Recommendation Systems

Deep learning adapts a multilayer approach to the hidden layers of the neural network. In traditional machine learning approaches, features are defined and extracted either manually or by making use of feature selection methods. However, in deep learning models, features are learned and extracted automatically, achieving better accuracy and performance. In general, the hyper parameters of classifier models are also measured automatically. Figure 1 shows the differences in sentiment polarity classification between the two approaches: traditional machine learning (Support Vector Machine (SVM), Bayesian networks, or decision trees) and deep learning. Artificial neural networks and deep learning currently provide the best solutions to many problems in the fields of image and speech recognition, as well as in natural language processing. Several types of deep learning techniques are discussed in this section.

4.1 Deep Neural Networks (DNN)

A deep neural network [90] is a neural network with more than two layers, some of which are hidden layers (Fig. 2.). Deep neural networks use sophisticated mathematical modeling to process data in many different ways. A neural network is an adjustable model of outputs as functions of inputs, which consists of several layers: an input layer, including input data; hidden layers, including processing nodes called neurons; and an output layer, including one or several neurons, whose outputs are the network outputs.

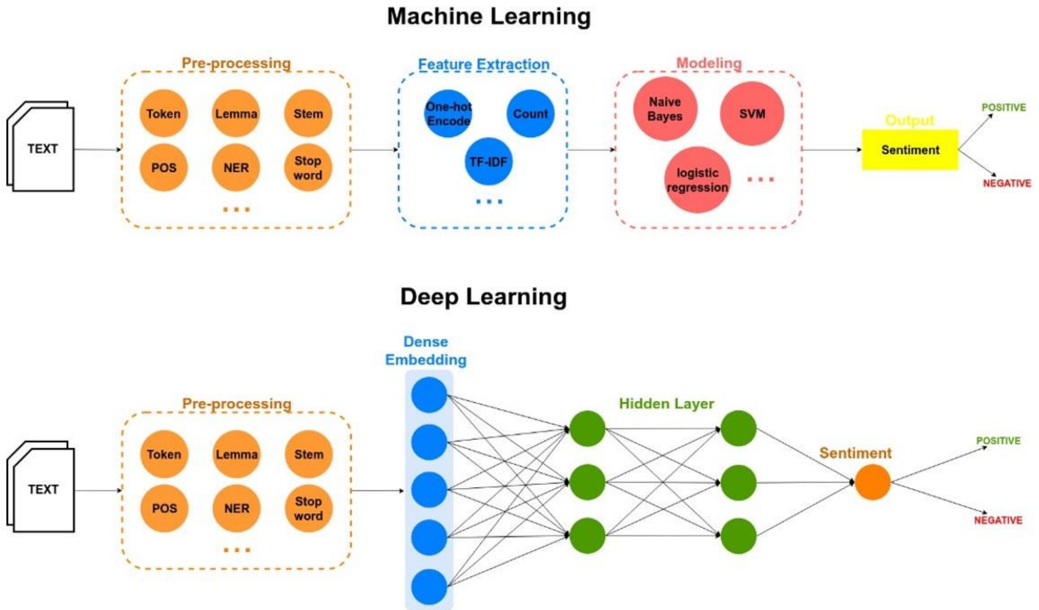

Fig. 2. Differences between two classification approaches of sentiment polarity, machine learning (top), and deep learning (bottom). Part of Speech (POS); Named Entity Recognition (NER); Term Frequency-Inverse Document Frequency (TF-IDF).

4.2 Convolutional Neural Networks (CNN)

A convolutional neural network is a special type of feed-forward neural network originally employed in areas such as computer vision, recommender systems, and natural language processing. It is a deep neural network architecture [91], typically composed of convolutional and pooling or subsampling layers to provide inputs to a fully-connected classification layer. Convolution layers filter their inputs to extract features; the outputs of multiple filters can be combined. Pooling or subsampling layers reduce the resolution of features, which can increase the CNN's robustness to noise and distortion. Fully connected layers perform classification tasks. An example of a CNN architecture can be seen in Fig. 3. The input data was preprocessed to reshape it for the embedding matrix. The figure shows an input embedding matrix processed by four convolution layers and two max pooling layers. The first two convolution layers have 64 and 32 filters, which are used to train different features; these are followed by a max pooling layer, which is used to reduce the complexity of the output and to prevent the overfitting of the data. The third and fourth convolution layers have 16 and 8 filters, respectively, which are also followed by a max pooling layer. The final layer is a fully connected layer that will reduce the vector of height 8 to an output vector of one, given that there are two classes to be predicted (Positive, Negative).

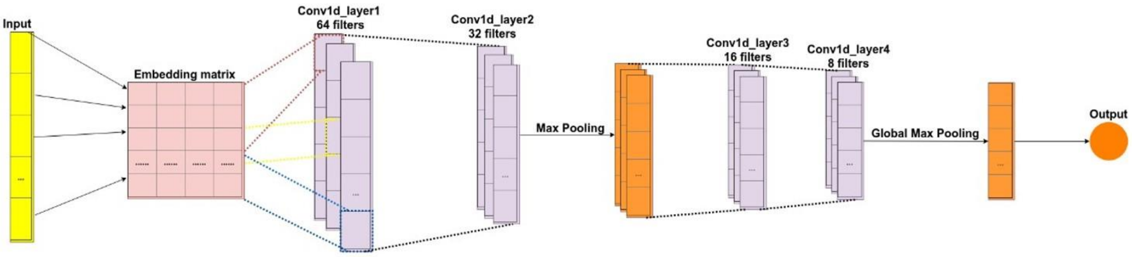

Fig. 3. A convolutional neural network.

4.3 Recurrent Neural Networks (RNN)

Recurrent neural networks [92] are a class of neural networks whose connections between neurons form a directed cycle, which creates feedback loops within the RNN. The main function of RNN is the processing of sequential information on the basis of the internal memory captured by the directed cycles. Unlike traditional neural networks, RNN can remember the previous computation of information and can reuse it by applying it to the next element in the sequence of inputs. A special type of RNN is long short-term memory (LSTM), which is capable of using long memory as the input of activation functions in the hidden layer. This was introduced by Hochreiter and Schmidhuber (1997) [93]. Fig. 4. illustrates an example of the LSTM architecture. The input data is preprocessed to reshape data for the embedding matrix (the process is similar to the one described for the CNN). The next layer is the LSTM, which includes 200 cells. The final layer is a fully connected layer, which includes 128 cells for text classification. The last layer uses the sigmoid activation function to reduce the vector of height 128 to an output vector of one, given that there are two classes to be predicted (positive, negative).

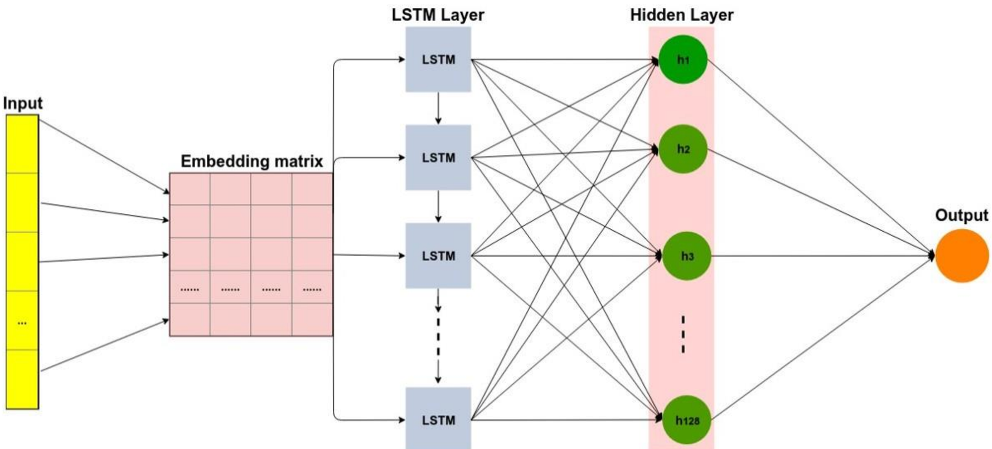

Fig. 4. A long short-term memory network. LSTM, long short-term memory.

5. Recommendation System Challenges

We reviewed several research studies on using deep learning techniques in recommendation systems. For example, De et al. [94] proposed using feed-forward neural networks as advanced frameworks for designing collaborative filtering engines, which consolidate and improve upon previous work. Dai et al. [95] introduced a personalized recommendation algorithm for online learning resources based on an improved backpropagation neural network algorithm. Dang et al. [96] discussed session-based recommendation, a recently pro- posed approach that reduces dependence on user profiles while maintaining high accuracy. The paper used real-world datasets and evaluation metrics to compare the performance of various session-based recommendation algorithms, including the deep learning approach named GRU4Rec. Additionally, Anantha et al. [2] discussed how deep learning techniques are used in various domains, including recommender systems, and compared the performance of traditional recommender systems with deep learning-based recommender systems.

Deep learning techniques are beneficial for developing recommendation systems, but organizations may face challenges such as data sparsity, the cold start problem [102], overfitting [103], scalability [104], and ethical concerns. Data sparsity occurs due to limited data available for the recommendation system to train on, and the cold start problem arises when insufficient data is available for accurate recommendations for new users or items [105]. Overfitting can occur when the model becomes too specialized to the training data and cannot generalize to new data.

5.1 Comparison

In comparison to Batmaz et al.'s review of thesis, journal, workshop, and conference papers from 2007 to 2017 [106] and Da'u et al.'s selection of 99 studies published from 2007 to 2018 [107], our review focuses exclusively on peer-reviewed journal articles published within the last five years (2018-2022), including some articles from 2023. This timeframe allows us to analyze the most current literature and provide insights into the latest developments in the field. Our review prioritizes peer-reviewed journal articles for their high-quality research and reliability, as they undergo rigorous evaluation by experts and provide detailed analyses.

A significant benefit of our approach is broad search keywords, enabling us to cast a wide net and identify a broad range of research studies that may have been overlooked in previous reviews. By utilizing this strategy, we can provide a comprehensive and up-to-date analysis of the latest research on deep learning-based recommendation systems.

Our review offers a valuable resource for researchers and practitioners interested in understanding the latest developments in this rapidly evolving deep learning-based recommendation systems field.

The following Table 1 displays detailed information for each article, including category, publication year, journal, domain, and technique, in addition to the comparison with our article.

TABLE 1. Comparison with related work.

Category	Reference	Year	Journal Index	Domain	Reviewed Technique	Comparison with our Review
Survey	[40]	2023	J1	General	Deep Reinforcement Learning	Diverse DL techniques
Survey	[41]	2023	J2	General	Deep Neural Networks	Various techniques beyond NN
Survey	[42]	2022	J3	General	Graph Neural Networks	Expands beyond GNN techniques
Survey	[49]	2022	J4	E-learning	Deep Learning	Broad E-Learning Spectrum
Survey	[50]	2022	J2	POI	Deep Learning	Wider range beyond POI
Survey	[43]	2022	J5	General	Deep Neural Networks	Not confined to NN techniques
Survey	[44]	2020	J6	Citation	Deep Learning	Focuses on diverse domains
Survey	[45]	2020	J7	General	Autoencoder-Based	Encompasses various DL techniques
Survey	[46]	2019	J3	General	Deep Learning	Focuses on recent 5 years
Survey	[47]	2018	J8	General	Deep Learning	Focuses on recent 5 years
Survey	[48]	2021	J9	Course	Deep Learning	Explores beyond a single domain
Survey	[58]	2022	J10	Trust-Aware	Deep Learning	Not confined to one type of RS
Study	[4]	2022	J11	Service	Neural Collaborative	Broad domain and techniques
Study	[51]	2021	J12	E-learning	Neural Networks	Covers diverse domain
Study	[52]	2020	J13	Session-Based	Deep Learning	Focuses on various application
Study	[2]	2018	J14	General	Deep Learning	Encompasses wider domain and includes recent articles
Review	[13]	2022	J15	E-learning	Deep Learning	Board domain and techniques
Review	[54]	2022	J1	General	Deep Knowledge Graph	Not confined to embedding techniques
Review	[55]	2022	J2	POI	Deep Learning	Explores beyond Point of Interest
Review	[56]	2019	J15	General	Deep Learning	Covers recent articles from the last 5 year

Journal Index: J1: Knowledge-Based Systems; J2: Neurocomputing; J3: ACM Computing Surveys; J4: Applied Sciences; J5: IEEE Transaction on Knowledge and data Engineering; J6: Expert System with Application; J7: Frontiers of Computer Science; J8: International Transaction Journal of Engineering Management & Applied Sciences & Technologies; J9: Pacific Asia Journal of the Association for Information System; J10: International Journal of Embedded System; J11: SN Computer Science; J12: Advance in Modeling and Analysis B; J13: Artificial Intelligence Review; J14: Data; J15: Electronics

6. Limitation and Future Work

This section discusses the limitations of our study and outlines avenues for future research. Despite the valuable insights gained, it’s essential to acknowledge the constraints within our approach.

6.1 Limitations in Our Study and Future Work

Our analysis offers a comprehensive examination of deep learning-based recommendation systems. However, it's important to consider some limitations that could affect the interpretation of our findings. These limitations create opportunities for future research and refinement of our methodology.

One limitation is that we only focused on peer-reviewed articles, meaning we may have missed out on state-of- the-art techniques published in other forms. It may be worth exploring non-traditional publication sources to gain additional insights. Furthermore, we selected articles based on title keywords, which could have limited their inclusion. Future studies could consider broader inclusion criteria. Additionally, our analysis provides a broad overview of the techniques' application, popularity, and categorization, rather than an in-depth evaluation of each technique. Future studies could delve deeper into each category of techniques and evaluate their performance in greater detail. Lastly, due to the large number of articles in the domain-based category, we extracted information from abstracts, titles, and keywords, potentially missing relevant data in datasets and metrics. Despite these limitations, our findings still provide significant value in understanding the employed techniques and dataset preferences, offering an overview of the field of recommendation systems. This information can serve as a foundation for further investigations and guide researchers and practitioners in their work. Therefore, even with the limitations mentioned above, this practice provides worthwhile insights for those involved in recommendation systems research and application.

6.2 Current Study Limitations and Suggestions

Deep learning-powered recommendation systems possess immense potential in different domains. However, they confront challenges that need addressing to boost their performance and reliability. These challenges extend beyond typical problems such as cold-start and data sparsity, which researchers have been dealing with. It is essential to recognize that enhancing performance requires more than just refining the techniques themselves.

Firstly, the selection of techniques should be tailored to the specific characteristics of the application domain. Researchers must carefully evaluate which techniques are most suitable for their particular problem and establish a clear classification strategy to ensure consistency and facilitate the discovery of related work.

Secondly, recommendation systems have been developed and tested in specific domains like e-commerce or entertainment. To enhance recommendation models, exploring their potential application in a broader range of domains and utilizing diverse data sources, including industry data, e-commerce data, publication sources, and contextual data like census data for point-of-interest recommendations, is necessary.

Lastly, while offline metrics like recall and precision are commonly used to evaluate recommendation systems, online metrics that measure real-time performance are necessary to reflect system performance in actual deployment settings. By addressing these challenges, we can improve the reliability and performance of deep learning-based recommendation systems across diverse domains.

On the other hand, based on our analysis, we suggest the following strategies to enhance the performance and reliability of deep learning-based recommendation systems:

1) **Clear Classification Criteria:** Develop precise classification criteria for recommendation systems based on hybridity and specific techniques to enhance consistency and clarity. A new classification system should cater to the evolving needs of recommendation systems research.

2) **Utilize Novel Techniques:** Explore innovative techniques within recommendation systems, including new neural network architectures and embedding methods, to push the boundaries and enhance performance. Adapt these architectures to diverse recommendation scenarios.

3) **Specialize Techniques:** Carefully select techniques for recommendation systems based on domain-specific characteristics to improve relevance and performance. Researchers should consider the appropriateness of techniques across various application domains.

4) **Expand Application Across Domains:** Extend the application of recommendation systems to diverse domains such as healthcare, finance, or education, broadening the impact of recommendation technologies.

5) **Utilize More Data:** Incorporate more data sources, including industry-specific datasets, to enhance recommendation accuracy and relevance, reflecting real-world usage more accurately.

6) **Develop Online Metrics:** Create online metrics that enable real-time assessment of recommendation system performance, providing more timely and relevant feedback on system behavior.

7) **Increase Interpretability:** Improve the interpretability of deep learning-based recommender systems by integrating attention mechanisms and visualization techniques, offering deeper insights into system operations.

8) **Address Bias and Transparency:** Mitigate biases in training data through data cleansing or augmentation methods and enhance transparency by providing explanations for recommendations.

9) **Improve Personalization:** Explore techniques for enhancing personalization in deep learning models by incorporating user feedback and advanced feature engineering methods, utilizing user and contextual information for tailored recommendations.

Implementing these strategies will contribute to developing more effective and adaptable recommendation systems and address the challenges currently in the field.

7. Conclusion

Our research focused on analyzing recommendation systems that utilize deep learning techniques. We found that neural collaborative filtering was the most commonly used technique across various domains, although the popularity of specific techniques varied depending on the domain. This highlights the importance of considering domain-specific traits when selecting an appropriate technique.

In our analysis, we discovered that matrix factorization was the most frequently utilized technique in the primary studies we examined, followed by graph neural networks and attention mechanisms. Other techniques, such as convolutional neural networks, deep reinforcement learning, knowledge graphs, and sequential recommendation, were also frequently employed. Among the different domains, social, session- based, and point-of-interest (POI) recommendation systems were the most popular, with numerous articles mentioning their use.

Our analysis made significant contributions to the field of deep learning-based recommendation systems. We also successfully analyzed and summarized the latest advanced deep learning techniques developed in the past five years, highlighting their applications in recommendation systems. Additionally, we categorized deep learning techniques and their application domains into meaningful groups based on study terminologies, facilitating better understanding and navigation. Our creation of a term classification system is a valuable resource, helping researchers effectively target specific terms within the field. Furthermore, we provided a summary of the datasets and metrics commonly used in the reviewed papers, serving as a reference for future researchers.

In conclusion, our research has played an important role in advancing the field of deep learning-based recommendation systems. Our work has enhanced our understanding of current trends, clarified domain-specific terminology, and provided valuable tools for researchers to navigate and contribute to this evolving and crucial study area.

REFERENCES

[1] G. Bathla, R. Rani, and H. Aggarwal, ''Improving recommendation techniques by deep learning and large scale graph partitioning,'' Int. J. Adv. Comput. Sci. Appl., vol. 9, no. 10, pp. 403–409, 2018.

[2] N. L. Anantha and B. Bathula, ''Comparative study on traditional recommender systems and deep learning based recommender systems,'' Adv. Model. Anal. B, vol. 61, no. 2, pp. 64–69, Jun. 2018.

[3] A. Akbar, P. Agarwal, and A. Obaid, ''Recommendation engines- neural embedding to graph-based: Techniques and evaluations,'' Int. J. Nonlinear Anal. Appl., vol. 13, no. 1, pp. 2411–2423, Mar. 2022.

[4] P. D. Rosa, M. Deriaz, M. D. Marco, and L. Laura, ''Service recommendations with deep learning: A study on neural collaborative engines,'' Pacific Asia J. Assoc. Inf. Syst., vol. 14, pp. 59–70, Jan. 2022.

[5] Z. Batmaz and C. Kaleli, ''AE-MCCF: An autoencoder-based multi- criteria recommendation algorithm,'' Arabian J. Sci. Eng., vol. 44, no. 11, pp. 9235–9247, Nov. 2019.

[6] Y. Geng, Y. Zhu, Y. Li, X. Sun, and B. Li, ''Multi-feature extension via semi-autoencoder for personalized recommendation,'' Appl. Sci., vol. 12, no. 23, p. 12408, Dec. 2022.

[7] N. Chizari, N. Shoeibi, and M. Moreno-García, ''A comparative analysis of bias amplification in graph neural network approaches for recommender systems,'' Electronics, vol. 11, no. 20, p. 3301, Oct. 2022.

[8] Z. Qiu, Y. Hu, and X. Wu, ‘‘Graph neural news recommendation with user existing and potential interest modeling,’’ ACM Trans. Knowl. Discovery Data, vol. 16, no. 5, pp. 1–17, Oct. 2022.
[9] Z. Huang, Y. Liu, C. Zhan, C. Lin, W. Cai, and Y. Chen, ‘‘A novel group recommendation model with two-stage deep learning,’’ IEEE Trans. Syst., Man, Cybern., Syst., vol. 52, no. 9, pp. 5853–5864, Sep. 2022.
[10] S. Bandyopadhyay and S. Thakur, ‘‘Product prediction and recommen- dation in e-commerce using collaborative filtering and artificial neural networks: A hybrid approach,’’ in Intelligent Computing Paradigm: Recent Trends, vol. 784. Singapore: Springer, 2020.
[11] P. Basile, C. Greco, A. Suglia, and G. Semeraro, ‘‘Deep learning and hierarchical reinforcement learning for modeling a conversational rec- ommender system,’’ Intelligenza Artificiale, vol. 12, no. 2, pp. 125–141, Jan. 2019.
[12] N. Hanafi, E. Pujastuti, A. Laksito, R. Hardi, R. Perwira, A. Arfriandi, and N. Asroni, ‘‘Handling sparse rating matrix for e-commerce recommender system using hybrid deep learning based on LSTM, SDAE and latent factor,’’ Int. J. Intell. Eng. Syst., vol. 15, no. 2, pp. 379–393, Jan. 2022.
[13] T. Liu, Q. Wu, L. Chang, and T. Gu, ‘‘A review of deep learning-based recommender system in e-learning environments,’’ Artif. Intell. Rev., vol. 55, no. 8, pp. 5953–5980, Dec. 2022.
[14] P. Balasamy and K. Athiyappagounder, ‘‘An optimized feature selection method for e-learning recommender system using deep neural network based on multilayer perceptron,’’ Int. J. Intell. Eng. Syst., vol. 15, no. 5, pp. 461–472, Jul. 2022.
[15] R. Alatrash, R. Priyadarshini, H. Ezaldeen, and A. Alhinnawi, ‘‘Aug- mented language model with deep learning adaptation on sentiment analysis for e-learning recommendation,’’ Cogn. Syst. Res., vol. 75, pp. 53–69, Sep. 2022.
[16] W. Chang, D. Sun, and Q. Du, ‘‘Intelligent sensors for POI recommen- dation model using deep learning in location-based social network big data,’’ Sensors, vol. 23, no. 2, p. 850, Jan. 2023.
[17] S. Safavi and M. Jalali, ‘‘DeePOF: A hybrid approach of deep convolutional neural network and friendship to Point-of-Interest (POI) recommendation system in location-based social networks,’’ Concur- rency Comput., Pract. Exper., vol. 34, no. 15, Jul. 2022, Art. no. e6981.
[18] L. Sun, ‘‘POI recommendation method based on multi-source information fusion using deep learning in location-based social networks,’’ J. Inf. Process. Syst., vol. 17, no. 2, pp. 352–368, 2021.
[19] Y. Liu and A.-B. Wu, ‘‘POI recommendation method using deep learning in location-based social networks,’’ Wireless Commun. Mobile Comput., vol. 2021, Jul. 2021, Art. no. 9120864.
[20] X. Lu and H. Zhang, ‘‘A content-aware POI recommendation method in location-based social networks based on deep CNN and multi- objective immune optimization,’’ J. Internet Technol., vol. 21, no. 6, pp. 1761–1772, 2021.
[21] L. Wan, F. Xia, X. Kong, C.-H. Hsu, R. Huang, and J. Ma, ‘‘Deep matrix factorization for trust-aware recommendation in social networks,’’ IEEE Trans. Netw. Sci. Eng., vol. 8, no. 1, pp. 511–528, Jan. 2021.
[22] C. Tang and J. Zhang, ‘‘An intelligent deep learning-enabled recommen- dation algorithm for teaching music students,’’ Soft Comput., vol. 26, no. 20, pp. 10591–10598, Oct. 2022.
[23] P. Magron and C. Févotte, ‘‘Neural content-aware collaborative filtering for cold-start music recommendation,’’ Data Mining Knowl. Discovery, vol. 36, no. 5, pp. 1971–2005, Sep. 2022.
[24] P. Linlin, ‘‘Tchaikovsky music recommendation algorithm based on deep learning,’’ Mobile Inf. Syst., vol. 2022, Sep. 2022, Art. no. 1265451.
[25] H. Xia, K. Huang, and Y. Liu, ‘‘Unexpected interest recommender system with graph neural network,’’ Complex Intell. Syst., vol. 9, no. 4, pp. 3819–3833, Aug. 2023.
[26] A. M. Al-Sbou and N. H. A. Rahim, ‘‘An improved hybrid semi-stacked autoencoder for item-features of recommendation system (iHSARS),’’ Indonesian J. Electr. Eng. Comput. Sci., vol. 30, no. 1, pp. 481–490, Apr. 2023.
[27] W. Jing, A. K. Sangaiah, L. Wei, L. Shaopeng, L. Lei, and L. Ruishi, ‘‘Multi-view fusion for recommendation with attentive deep neural network,’’ Evol. Intell., vol. 15, no. 4, pp. 2619–2629, Dec. 2022.
[28] Z. Lyu, Y. Wu, J. Lai, M. Yang, C. Li, and W. Zhou, ‘‘Knowledge enhanced graph neural networks for explainable recommendation,’’ IEEE Trans. Knowl. Data Eng., vol. 35, no. 5, pp. 4954–4968, May 2023.
[29] C. Dang, M. Moreno-García, and F. De la Prieta, ‘‘Using hybrid deep learning models of sentiment analysis and item genres in recommender systems for streaming services,’’ Electronics, vol. 10, no. 20, p. 2459, Oct. 2021.
[30] F. Yang, H. Wang, and J. Fu, ‘‘Improvement of recommendation algorithm based on collaborative deep learning and its parallelization on spark,’’ J. Parallel Distrib. Comput., vol. 148, pp. 58–68, Feb. 2021.
[31] Y. Huo, ‘‘Talent management recommendation technology based on deep learning,’’ Math. Problems Eng., vol. 2022, Sep. 2022, Art. no. 7697192.

[32] V. Tapaskar and M. M. Math, ‘‘Deep recurrent Gaussian Nesterovs recommendation using multi-agent in social networks,’’ Evolving Syst., vol. 13, no. 3, pp. 435–452, Jun. 2022.

[33] G. Qiu, C. Song, L. Jiang, and Y. Guo, ‘‘Multi-view hybrid recommen- dation model based on deep learning,’’ Intell. Data Anal., vol. 26, no. 4, pp. 977–992, Jul. 2022.

[34] Y. Li and X. Tong, ‘‘Trust recommendation based on deep deterministic strategy gradient algorithm,’’ IEEE Access, vol. 10, pp. 48274–48282, 2022.

[35] X. Huang and X. Liu, ‘‘Incorporating a topic model into a hypergraph neural network for searching-scenario oriented recommendations,’’ Appl. Sci., vol. 12, no. 15, p. 7387, Jul. 2022.

[36] D. H. Tran, Q. Z. Sheng, W. E. Zhang, N. H. Tran, and N. L. D. Khoa, ‘‘CupMar: A deep learning model for personalized news recommendation based on contextual user-profile and multi-aspect article representation,’’ World Wide Web, vol. 26, no. 2, pp. 713–732, Mar. 2023.

[37] Elfeky, A.I.M., Masadeh, T.S.Y., Elbyaly, M.Y.H., 2020. Advance organizers in flipped classroom via e-learning management system and the promotion of integrated science process skills. Think. Skills Creativity 35, 100622.

[38] McKinney, L., Burridge, A.B., Lee, M.M., Bourdeau, G.V., Miller-Waters, M., 2022. Incentivizing full-time enrollment at community colleges: What influences students’ decision to take more courses? Commun. College Rev. 50 (2), 144–170.

[39] Zhao, H., Liu, Z., Yao, X., Yang, Q., 2021. A machine learning-based sentiment analysis of online product reviews with a novel term weighting and feature selection approach. Inf. Process. Manage. 58 (5), 102656

[40] Zhang, Y., Wang, J., Zhang, X., 2021. Conciseness is better: Recurrent attention LSTM model for document-level sentiment analysis. Neurocomputing 462, 101–112.

[41] Liu, Z., Yang, C., Rüdian, S., Liu, S., Zhao, L., Wang, T., 2019. Temporal emotion aspect modeling for discovering what students are concerned about in online course forums. Interact. Learn. Environ. 27 (5–6), 598–627.

[42] Zhu, J.J., Chang, Y.-C., Ku, C.-H., Li, S.Y., Chen, C.-J., 2021. Online critical review classification in response strategy and service provider rating: Algorithms from heuristic processing, sentiment analysis to deep learning. J. Bus. Res. 129, 860–877.

[43] Acheampong, F.A., Nunoo-Mensah, H., Chen, W., 2021. Transformer models for text-based emotion detection: a review of BERT-based approaches. Artif. Intell. Rev. 54 (8), 5789–5829.

[44] Kuleto, V., Ilić, M., Dumangiu, M., Ranković, M., Martins, O.M.D., Pˇaun, D., Mihoreanu, L., 2021. Exploring opportunities and challenges of artificial intelligence and machine learning in higher education institutions. Sustainability 13 (18), 10424.

[45] Dolianiti, F.S., Iakovakis, D., Dias, S.B., Hadjileontiadou, S.J., Diniz, J.A., Natsiou, G., Tsitouridou, M., Bamidis, P.D., Hadjileontiadis, L.J., 2019. Sentiment analysis on educational datasets: a comparative evaluation of commercial tools. Educ. J. Univ. Patras UNESCO Chair 6,

[46] Ahmad, M., Aftab, S., Bashir, M.S., Hameed, N., 2018. Sentiment analysis using svm: a systematic literature review. Int. J. Adv. Comput. Sci. Appl. 9 (2).

[47] Sivakumar, M., Reddy, U.S., 2017. Aspect based sentiment analysis of students opinion using machine learning techniques. In: 2017 International Conference on Inventive Computing and Informatics. ICICI, IEEE,

[48] Nikolić, N., Grljević, O., Kovačević, A., 2020. Aspect-based sentiment analysis of reviews in the domain of higher education. Electron. Libr. 38 (1), 44–64.

[49] Yang, R., 2021. Machine learning and deep learning for sentiment analysis over students’ reviews: An overv Li, J., Sun, A., Han, J., Li, C., 2022b. A survey on deep learning for named entity recognition. IEEE Trans. Knowl. Data Eng. 34 (1), 50–70. iew study.

[50] Ding, J., Sun, H., Wang, X., Liu, X., 2018. Entity-level sentiment analysis of issue comments. In: Proceedings of the 3rd International Workshop on Emotion Awareness in Software Engineering. ACM,

[51] Li, H., Lu, W., 2017. Learning latent sentiment scopes for entity-level sentiment analysis. In: Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 31, (1),

[52] Shaik, T., Tao, X., Li, Y., Dann, C., McDonald, J., Redmond, P., Galli-gan, L., 2022b. A review of the trends and challenges in adopting natural language processing methods for education feedback analysis. IEEE Access 10, 56720–56739.

[53] Rosalind, J.M., Suguna, S., 2022. Predicting students’ satisfaction towards online courses using aspect-based sentiment analysis. In: Computer, Communication, and Signal Processing. Springer International Publishing, pp. 20–35.

[54] Edalati, M., Imran, A.S., Kastrati, Z., Daudpota, S.M., 2021. The potential of machine learning algorithms for sentiment classification of students’ feedback on MOOC. In: Lecture Notes in Networks and Systems. Springer International Publishing, pp. 11–22.

[55] Wehbe, D., Alhammadi, A., Almaskari, H., Alsereidi, K., Ismail, H., 2021. UAE e-learning sentiment analysis framework. In: ArabWIC 2021: The 7th Annual International Conference on Arab Women in Computing in Conjunction with the 2nd Forum of Women in Research. ACM, Sharjah, UAE, http://dx.doi.org/10.1145/3485557.3485570.

[56] Dehbozorgi, N., Mohandoss, D.P., 2021. Aspect-based emotion analysis on speech for predicting performance in collaborative learning. In: 2021 IEEE Frontiers in Education Conference. FIE, IEEE, http://dx.doi.org/10.1109/fie49875.2021.637330.

[57] Kastrati, Z., Imran, A.S., Kurti, A., 2020. Weakly supervised framework for aspect-based sentiment analysis on students' reviews of MOOCs. IEEE Access 8, 106799–106810. http://dx.doi.org/10.1109/access.2020.3000739.

[58] Li, J., Sun, A., Han, J., Li, C., 2022b. A survey on deep learning for named entity recognition. IEEE Trans. Knowl. Data Eng. 34 (1), 50–70. http://dx.doi.org/10.1109/tkde.2020.2981314.

[59] Yeruva, V.K., Chandrashekar, M., Lee, Y., Rydberg-Cox, J., Blanton, V., Oyler, N.A., 2020. Interpretation of sentiment analysis with human-in-the-loop. In: 2020 IEEE International Conference on Big Data (Big Data). IEEE, http://dx.doi.org/10.1109/bigdata50022.2020.9378221.

[60] Catelli, R., Pelosi, S., Esposito, M., 2022. Lexicon-based vs. bert-based sentiment analysis: A comparative study in italian. Electronics 11 (3), 374. http://dx.doi.org/10.3390/electronics11030374.

[61] Hatzivassiloglou, V., McKeown, K.R., 1997. Predicting the semantic orientation of adjectives. In: Proceedings of the 35th Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics, http://dx.doi.org/10.3115/976909.979640.

[62] Taboada, M., Gillies, M.A., McFetridge, P., 2006. Sentiment classification techniques for tracking literary reputation. In: LREC Workshop: To-Wards Computational Models of Literary Analysis. Citeseer, pp. 36–43.

[63] Benamara, F., Cesarano, C., Picariello, A., Recupero, D.R., Subrahmanian, V.S., 2007. Sentiment analysis: Adjectives and adverbs are better than adjectives alone. ICWSM 7, 203–206.

[64] Vermeij, M., 2005. The orientation of user opinions through adverbs, verbs and nouns. In: 3rd Twente Student Conference on IT, Enschede June. Citeseer.

[65] Neviarouskaya, A., Prendinger, H., Ishizuka, M., 2009a. Compositionality principle in recognition of fine-grained emotions from text. In: Proceedings of the International AAAI Conference on Web and Social Media. Vol. 3, pp. 278–281.

[66] Baccianella, S., Esuli, A., Sebastiani, F., 2010. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In: Calzolari, N., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Rosner, M., Tapias, D. (Eds.), LREC. European Language Resources Association, http://nmis.isti.cnr.it/sebastiani/Publications/LREC10.pdf.

[67] Wilson, T., Hoffmann, P., Somasundaran, S., Kessler, J., Wiebe, J., Choi, Y., Cardie, C., Riloff, E., Patwardhan, S., 2005. Opinionfinder: A system for subjectivity analysis. In: Proceedings of HLT/EMNLP 2005 Interactive Demonstrations, pp. 34–35.

[68] Wilson, T., Wiebe, J., Hoffmann, P., 2005. Recognizing contextual polarity in phraselevel sentiment analysis. In: Proceedings of Human Languagetechnology Conference and Conference on Empirical Methods in Natural Language Processing, pp. 347–354.

[69] Stone, P.J., Dunphy, D.C., Smith, M.S., 1966. The general inquirer: A computer approach to content analysis.

[70] Nielsen, F.Å, 2011. A new anew: Evaluation of a word list for sentiment analysis in microblogs. arXiv preprint arXiv:1103.2903.

[71] Neviarouskaya, A., Prendinger, H., Ishizuka, M., 2009a. Compositionality principle in recognition of fine-grained emotions from text. In: Proceedings of the International AAAI Conference on Web and Social Media. Vol. 3, pp. 278–281.

[72] Hutto, C.J., Gilbert, E., 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In: Adar, E., Resnick, P., Choudhury, M.D., Hogan, B., Oh, A.H. (Eds.), ICWSM. The AAAI Press, http://dblp.uni-trier.de/db/conf/icwsm/icwsm2014.html#HuttoG14.

[73] Tzacheva, A., Easwaran, A., 2021. Emotion detection and opinion mining from student comments for teaching innovation assessment. Int. J. Educ. (IJE) 09 (02), 21–32. http://dx.doi.org/10.5121/ije2021.9203.

[74] Ortony, A., Turner, T.J., 1990. What's basic about basic emotions? Psychol. Rev. 97 (3), 315.

[75] Mohammad, S.M., Kiritchenko, S., Zhu, X., 2013. Nrc-canada: Building the state-of-theart in sentiment analysis of tweets. http://dx.doi.org/10.48550/ARXIV.1308.6242, https://arxiv.org/abs/1308.6242.

[76] Maher, M.L., Latulipe, C., Lipford, H., Rorrer, A., 2015. Flipped classroom strategies for CS education. In: Proceedings of the 46th ACM Technical Symposium on Computer Science Education. ACM, http://dx.doi.org/10.1145/2676723.2677252.

[77] Latulipe, C., Long, N.B., Seminario, C.E., 2015. Structuring flipped classes with lightweight teams and gamification. In: Proceedings of the 46th ACM Technical Symposium on Computer Science Education. ACM, http://dx.doi.org/10.1145/2676723.2677240.

[78] MacNeil, S., Latulipe, C., Long, B., Yadav, A., 2016. Exploring lightweight teams in a distributed learning environment. In: Proceedings of the 47th ACM Technical Symposium on Computing Science Education. ACM, http://dx.doi.org/10.1145/2839509.2844577.

[79] Rosalind, J.M., Suguna, S., 2022. Predicting students' satisfaction towards online courses using aspect-based sentiment analysis. In: Computer, Communication, and Signal Processing. Springer International Publishing, pp. 20–35. http://dx.doi.org/10.1007/978-3-031-11633-9_3.

[80] Alqasemi, F., Abdelwahab, A., Abdelkader, H., 2017. Constructing automatic domainspecific sentiment lexicon using KNN search via terms discrimination vectors. Int. J. Comput. Appl. 41 (2), 129–139. http://dx.doi.org/10.1080/1206212x.2017.1409477.

[81] Ömer Osmanoˇglu, U., Atak, O.N., Çaˇglar, K., Kayhan, H., Can, T., 2020. Sentiment analysis for distance education course materials: A machine learning approach. J. Educ. Technol. Online Learn. 3 (1), 31–48. http://dx.doi.org/10.31681/jetol.663733.

[82] Lwin, H.H., Oo, S., Ye, K.Z., Lin, K.K., Aung, W.P., Ko, P.P., 2020. Feedback analysis in outcome base education using machine learning. In: 2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON). IEEE, http://dx.doi.org/10.1109/ecticon49241.2020.9158328.

[83] Faizi, R., 2022. A sentiment-based approach to predict learners' perceptions towards YouTube educational videos. In: Innovations in Bio-Inspired Computing and Applications. Springer International Publishing, pp. 549–556. http://dx.doi.org/10.1007/978-3-030-96299-9_52.

[84] Kaur, C., Boush, M.S.A., Hassen, S.M., Hakami, W.A., Ab-dalraheem, M.H.O., Galam, N.M., Hadi, N.A., Hadi, N.A., Benjeed, A.O.S., 2022. Incorporating sentimental analysis into development of a hybrid classification model. Int. J. Health Sci. 6 (S1), 1709–1720. http://dx.doi.org/10.53730/ijhs.v6ns1.4924.

[85] Yu, L., Lee, C., Pan, H., Chou, C., Chao, P., Chen, Z., Tseng, S., Chan, C., Lai, K., 2018. Improving early prediction of academic failure using sentiment analysis on self-evaluated comments. J. Comput. Assist. Learn. 34 (4), 358–365. http://dx.doi.org/10.1111/jcal.12247.

[86] Sutoyo, E., Almaarif, A., Yanto, I.T.R., 2021. Sentiment analysis of student evaluations of teaching using deep learning approach. In: International Conference on Emerging Applications and Technologies for Industry4.0 (EATI'2020). Springer International Publishing, pp. 272–281. http://dx.doi.org/10.1007/978-3-030-80216-5_20.

[87] Sangeetha, K., Prabha, D., 2020. RETRACTED ARTICLE: Sentiment analysis of student feedback using multi-head attention fusion model of word and context embedding for LSTM. J. Ambient Intell. Humaniz. Comput. 12 (3), 4117–4126. http://dx.doi.org/10.1007/s12652-020-01791-9.

[88] Dyulicheva, Y.Y., Bilashova, E.A., 2022. Learning analytics of moocs based on natural language processing.

[89] Li, X., Zhang, H., Ouyang, Y., Zhang, X., Rong, W., 2019. A shallow BERT- CNN model for sentiment analysis on MOOCs comments. In: 2019 IEEE International Conference on Engineering, Technology and Education. TALE, IEEE, http://dx.doi.org/10.1109/tale48000.2019.9225993.

[90] Aggarwal, C.C. Neural Networks and Deep Learning; Springer: Berlin, Germany, 2018.

[91] Zhang, L.; Wang, S.; Liu, B. Deep learning for sentiment analysis: A survey. WIREs Data Min. Knowl. Discov. 2018, 8, e1253.

[92] Britz, D. Recurrent Neural Networks Tutorial, Part 1–Introduction to Rnns. Available online: http://www.wildml.com/2015/09/recurrent-neural-networkstutorial-part-1-introduction-to-rnns/ (accessed on 12 March 2020).

[93] Hochreiter, S.; Schmidhuber, J. LSTM can solve hard long time lag problems. In Proceedings of the Advances in Neural Information Processing Systems, Denver, CO, USA, 2–5 December 1996; pp. 473–479.

[94] P. D. Rosa, M. Deriaz, M. D. Marco, and L. Laura, ''Service recommendations with deep learning: A study on neural collaborative engines,'' Pacific Asia J. Assoc. Inf. Syst., vol. 14, pp. 59–70, Jan. 2022.

[95] Y. Dai and J. Xu, ''Study of online learning resource recommendation based on improved BP neural network,'' Int. J. Embedded Syst., vol. 14, no. 2, p. 101, 2021.

[96] T. K. Dang, Q. P. Nguyen, and V. S. Nguyen, ''A study of deep learningbased approaches for session-based recommendation systems,'' Social Netw. Comput. Sci., vol. 1, no. 4, p. 216, Jul. 2020.

[97] Ligthart, A., Catal, C., Tekinerdogan, B., 2021. Systematic reviews in sentiment analysis a tertiary study. Artif. Intell. Rev. 54 (7), 4997–5053.

[98] Tao, X., Zhou, X., Zhang, J., Yong, J., 2016. Sentiment analysis for depression detection on social networks. In: Advanced Data Mining and Applications. Springer International Publishing, pp. 807–810.

[99] Zhou, X., Tao, X., Yong, J., Yang, Z., 2013. Sentiment analysis on tweets for social events. In: Proceedings of the 2013 IEEE 17th International Conference on Computer Supported Cooperative Work in Design. CSCWD, IEEE

[100] Tao, X., Shaik, T.B., Higgins, N., Gururajan, R., Zhou, X., 2021. Remote patient monitoring using radio frequency identification (RFID) technology and machine learning for early detection of suicidal behaviour in mental health facilities. Sensors 21 (3), 776.

[101] Shaik, T., Tao, X., Higgins, N., Gururajan, R., Li, Y., Zhou, X., Acharya, U.R., 2022a. FedStack: Personalized activity monitoring using stacked federated learning. Knowl.-Based Syst. 257, 109929

[102] S.-T. Zhong, L. Huang, C.-D. Wang, J.-H. Lai, and P. S. Yu, ''An autoencoder framework with attention mechanism for cross-domain recommendation,'' IEEE Trans. Cybern., vol. 52, no. 6, pp. 5229–5241, Jun. 2022.

[103] F. Merabet and D. Benmerzoug, ''QoS prediction for service selection and recommendation with a deep latent features autoencoder,'' Comput. Sci. Inf. Syst., vol. 19, no. 2, pp. 709–733, 2022.

[104] Q. Li and J. Kim, ''A deep learning-based course recommender system for sustainable development in education,'' Appl. Sci., vol. 11, no. 19, p. 8993, Sep. 2021.

[105] A. C. Hansel and A. Wibowo, ''Using movie genres in neural network based collaborative filtering movie recommendation system to reduce cold start problem,'' Int. J. Emerg. Technol. Adv. Eng., vol. 12, no. 3, pp. 63–73, Mar. 2022.

[106] Z. Batmaz, A. Yurekli, A. Bilge, and C. Kaleli, ''A review on deep learning for recommender systems: Challenges and remedies,'' Artif. Intell. Rev., vol. 52, no. 1, pp. 1–37, Jun. 2019.

[107] A. Da'u and N. Salim, ''Recommendation system based on deep learning methods: A systematic review and new directions,'' Artif. Intell. Rev., vol. 53, no. 4, pp. 2709–2748,