J. Mountain Res. P-ISSN: 0974-3030, E-ISSN: 2582-5011 DOI: https://doi.org/10.51220/jmr.v21-S2.5
Vol. 21(S2), (2026) 37-52

Review-Based Recommendation Systems for E-Learning: A Systematic Review of Deep Learning and Sentiment Analysis Approaches

Anurag Chandna • Vivek Kumar

Department of Computer Science and Engineering, Quantum University, Roorkee, India

*Corresponding Author's Email Id: anuragchandnarke@gmail.com

Received: 12.06.2025; Revised: 02.11.2025; Accepted: 19.11.2025

Abstract: E-learning has rapidly evolved, offering diverse digital content that often overwhelms learners when selecting appropriate materials. Traditional recommendation systems, typically driven by content or collaborative filtering, are limited in addressing the dynamic and emotional aspects of learner preferences. This study explores how integrating sentiment analysis with deep learning improves the personalization of e-learning recommendation systems. A systematic review of research published between 2018 and 2023 was undertaken to assess the key models, datasets, and evaluation measures adopted across multiple learning contexts. The findings indicate that deep learning architectures—such as convolutional, recurrent, and graph neural networks—enhance adaptability and accuracy in personalized recommendations. However, limited focus has been placed on classification frameworks, sentiment-based feedback mechanisms, and domain-specific taxonomies. This paper consolidates recent sentiment-driven strategies, identifies persistent challenges such as data sparsity, interpretability, and transferability, and highlights opportunities for domain-aware learning models. The study contributes a comprehensive understanding of deep learning–enabled sentiment analysis in e-learning and outlines future pathways for developing intelligent, learner-centric recommendation frameworks.

Keywords: Deep learning • Sentiment analysis • Review-based recommendation • Personalized learning • Artificial intelligence in education

Introduction

The rapid growth of digital education platforms has transformed how learners interact with knowledge resources, leading to the emergence of intelligent systems that can tailor learning experiences. Recommendation systems (RS) play a crucial role in this transformation by assisting learners in identifying suitable materials based on their preferences and historical behavior (Akbar et al., 2022). These systems traditionally rely on collaborative and content-based filtering models, which analyze similarities among users and learning materials to generate recommendations (Bandyopadhyay and Thakur, 2020). While effective in basic personalization, these models often overlook deeper psychological or emotional cues that influence learner engagement.

Recent advancements in artificial intelligence (AI) and natural language processing (NLP) have significantly strengthened the capabilities of RS. Sentiment analysis—also referred to as opinion mining—has emerged as a complementary component that interprets user emotions and feedback. By understanding the tone, polarity, and intensity of learners' opinions, RS can refine their suggestions beyond behavioral data alone. For instance, analyzing textual feedback on courses and instructors helps determine satisfaction levels, which can then inform future recommendations.

Deep learning (DL) has become an essential driver in this domain due to its capacity for automatic feature learning. Architectures such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and graph neural networks (GNNs) allow systems to process complex, nonlinear patterns in learner data. Unlike traditional models that depend on manually engineered features, DL models capture hierarchical and contextual representations, enabling a more nuanced understanding of learner preferences and interactions. These advances have led to notable improvements in the accuracy and interpretability of recommendation outcomes (Dai and Xu, 2021; Dang et al., 2020).

However, despite substantial progress, significant challenges persist. Traditional recommendation systems struggle with issues such as data sparsity, the cold start problem, and limited domain adaptability. Furthermore, most existing surveys and reviews focus primarily on algorithmic efficiency rather than exploring the interplay between sentiment analysis and deep learning. Few studies have systematically examined how integrating emotion-driven insights with DL-based frameworks can improve learner engagement and recommendation precision (Liu et al., 2019; Yang, 2021).

Therefore, this study aims to bridge this gap by providing a consolidated examination of sentiment-aware deep learning approaches for e-learning recommendation systems. Specifically, it reviews models that combine emotional feedback with advanced neural architectures, evaluates their performance across datasets and domains, and highlights existing limitations in classification frameworks and contextual interpretation.

The major contributions of this review are as follows:

- It synthesizes current research integrating deep learning with sentiment analysis for e-learning personalization.
- It identifies key methodological gaps, including inadequate domain categorization and limited interpretability of model outputs.
- It proposes potential directions for developing adaptive, emotion-aware recommendation systems that enhance learner satisfaction and engagement.

Through this investigation, the paper seeks to advance the understanding of how sentiment-driven deep learning models can foster more human-centric, intelligent, and scalable recommendation frameworks for digital learning environments.

Sentiment Analysis in E-Learning

Sentiment analysis has become a crucial element in improving the adaptability of recommendation systems, as it incorporates users' emotional responses into predictive modeling (Alatrash et al., 2022). Alatrash et al. integrated sentiment cues with genre-based similarity within collaborative filtering to enhance personalization. Dang et al. (2021) utilized BERT-based preprocessing combined with hybrid deep learning techniques to analyze and interpret user reviews. In a related contribution, Dai and Xu (2021) proposed a multilingual review-aware deep recommendation model (MRRec) that enhances rating predictions across varied linguistic datasets.

Sentiment analysis, often referred to as opinion mining, applies natural language processing (NLP) methods to determine the emotional orientation of textual data and infer user perspectives (Dang et al., 2020; Ligthart et al., 2021). In e-learning settings, it serves to interpret learners' reactions toward course materials, instructional strategies, and overall teaching quality through the examination of qualitative feedback and reviews. Detecting positive or negative sentiments within this feedback enables educators and platforms to adjust course design, enhance engagement, and strengthen learner satisfaction.

As an interdisciplinary domain, sentiment analysis integrates concepts from machine learning, linguistics, and behavioural science to extract patterns from extensive text data. While its early applications centred on social media and e-commerce for analysing consumer opinions, recent research has expanded its scope to education, where student reflections serve as indicators of engagement and instructional quality (Tao et al., 2016; Zhou et al., 2013).

Advances in artificial intelligence and NLP have revolutionized data-driven evaluation methods across multiple sectors, including education (Shaik et al., 2022; Zhong et al., 2022). Within academic environments, sentiment analysis enables institutions to assess learning management systems, teaching effectiveness, and overall course delivery by examining students' written feedback (Elfeky et al., 2020; McKinney et al., 2022). Beyond numerical scores, qualitative comments reveal emotional and cognitive dimensions of learning experiences, offering practical guidance for improving teaching methods and curriculum design.

In e-learning sentiment analysis, textual feedback is preprocessed through techniques like tokenization and feature extraction, then categorized as positive, negative, or neutral (Zhao et al., 2021; Zhang et al., 2021). Since manual emotion labeling demands time and expertise (Liu et al., 2019), researchers have introduced automated lexicon- and corpus-based annotation methods. Combined with AI-driven models, these approaches enable scalable interpretation of student feedback while reducing human involvement.

Artificial intelligence advances sentiment analysis by automating classification through machine learning, deep learning, and transformer models such as BERT (Acheampong et al., 2021; Zhu et al., 2021). These architectures capture subtle emotional cues and contextual meaning, supporting accurate estimation of learner satisfaction (Kuleto et al., 2021). Depending on the application, analysis may be conducted at the document, sentence, or aspect level to identify overall polarity or isolate elements—such as course content or instructional delivery—that require refinement (Ligthart et al., 2021). **Figure 1** depicts these analytical levels, from broad document sentiment to fine-grained aspect evaluation.

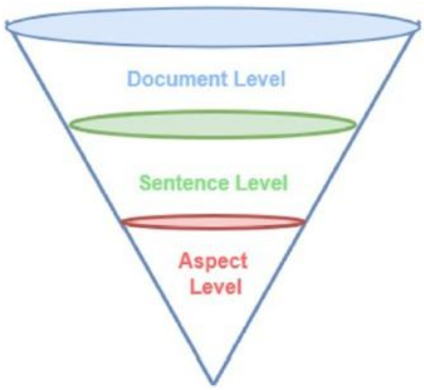

Figure 1. Different levels of sentiment analysis showing document-, sentence-, entity-, and aspect-level granularity used for identifying sentiment orientation in learner feedback.

A. Document Level

Research on sentiment extraction within education remains scarce, leading many institutions to rely on generic analytical platforms. Dolianiti et al. (Dai and Xu, 2021) compared five commercial tools—IBM Watson, Microsoft Azure, OpinionFinder 2.0, Repustate, and SentiStrength—with four education-specific models developed using support vector machines (SVM) and k-fold cross-validation on annotated student forum datasets. The tailored models produced higher accuracy in one course scenario. Likewise, Ahmad et al. (2018) utilized sentiment analysis to assess faculty performance, where SVM and Naïve Bayes classifiers trained on 5,000 comments achieved around 73% accuracy.

B. Sentence Level

Cosine similarity was utilized by Sivakumar and Reddy (2017) in order to evaluate the degree of semantic proximity between aspect words and student opinion phrases at the sentence level. After being cleaned up and

 DOI: https://doi.org/10.51220/jmr.v21-S2.5

separated into seven unique categories, the dataset that was collected through the Twitter API was prepared for analysis. The classification of phrases was accomplished through the utilization of decision trees, support vector machines (SVM), and Naive Bayes algorithms. Additionally, part-of-speech tagging was employed to extract subjective expressions. According to the SentiWordNet lexicon, the improvement in sentiment polarity was much more pronounced when it was combined with Gradient Boosting Tree and Linear Support Vector Machine models.

. This hybrid approach surpassed SentiStrength and SentiStrength-SE in identifying positive and neutral statements. Extending this work, Li and Lu (2017) introduced an entity-level model that associates named entities with their related sentiments to uncover implicit emotional cues within text.

C. Entity Level

Student input on teachers, course materials, and content is crucial for judging educational quality. Yang (2021) identified instructor, topic, lesson, and curriculum using automated entity extraction. Entity-level sentiment analysis relies on Named Entity Recognition (NER) to identify and classify feedback text phrases representing people, organizations, and education.

Ding et al. (2018) developed SentiSW, an entity-level model that produces sentiment–entity pairings utilizing TF-IDF and Doc2Vec representations, demonstrating significant classification efficacy. Edalati et al. (2021) used Random Forest (RF), Support Vector Machine (SVM), Decision Tree, and deep learning algorithms to look at Coursera student reviews and find teaching-related entities and predict feelings. RF did the best job, with F1-scores of 98.01% for entity recognition and 99.43% for sentiment detection. Wehbe et al. (2021) advanced this trajectory by introducing a spatiotemporal sentiment framework for education-related Twitter data, amalgamating Named Entity Recognition (NER) with Aspect-Based Sentiment Analysis (ABSA) to delineate entities alongside emotional patterns.

D. Aspect Level

Aspect-Based Sentiment Analysis (ABSA) is a method that is widely used in research on educational sentiment. This method gives a thorough look into how people feel or think about things like teachers, materials, and tests. ABSA accurately interprets learner feedback by combining sentiment classification with topic modelling techniques like Latent Dirichlet Allocation (LDA), Latent Semantic Analysis (LSA), Non-negative Matrix Factorization (NMF), and Probabilistic Latent Semantic Analysis (PLSA) (Shaik et al., 2022).

Rosalind and Suguna (2022) used Aspect-Based Sentiment Analysis to find out how happy students were with their Coursera courses by using machine learning and subject modelling. Divided the course reviews down into sentences and then used both supervised LDA algorithms that are either completely or partially self-sufficient to find the parts. A tailored vocabulary helped with polarity recognition by showing positive feelings about learning components, labs, and career goals, and negative feelings about teachers, materials, and course fees. The maximum entropy classifier they created was able to predict multi-aspect sentiment with an accuracy of 80.67%. In their study, Nikolić et al. (2020) conducted phrase-level automated sentence analysis (ABSA) by employing support vector machines (SVM), cascade classifiers, and rule-based extraction approaches. It was revealed that the F-measures for negative mood were 0.94, while the F-measures for positive classification were 0.83.

Table 1 summarizes key studies employing AI-driven sentiment analysis at document, sentence, entity, and aspect levels within educational contexts.

 DOI: https://doi.org/10.51220/jmr.v21-S2.5

TABLE 1: **Summary of studies applying sentiment analysis at different analytical levels in educational domains.**

Analytical Level	Representative Techniques	Primary Objective in the Educational Domain	Representative Studies
Document Level	Support Vector Machine (SVM), Naïve Bayes (NB)	Classify educational documents or feedback as expressing an overall positive or negative sentiment toward a given theme or context.	Dolianiti et al. (2019), Ahmad et al. (2018)
Sentence Level	SVM, NB, Cascade Classifier, Rule-Based Approaches	Perform fine-grained analysis by examining sentiment polarity at the sentence level within educational reviews or learning reflections.	Sivakumar and Reddy (2017), Nikolic´ et al. (2020)
Entity Level	Gradient Boosted Trees, SVM, Named Entity Recognition (NER), TF–IDF, Doc2Vec	Detect and interpret opinions associated with specific entities (e.g., instructors, courses, or tools) mentioned in learner feedback.	Yang (2021), Nikolic´ et al. (2020), Ding et al. (2018), Li and Lu (2017)
Aspect Level	Latent Dirichlet Allocation (LDA), Random Forest (RF), SVM, Decision Tree, Convolutional Neural Networks (CNN), K-Nearest Neighbors (KNN)	Identify and evaluate positive or negative aspects of educational practices, methods, or resources.	Rosalind and Suguna (2022), Edalati et al. (2021), Wehbe et al. (2021), Dehbozorgi and Mohandoss (2021), Dehbozorgi and Mohan- doss (2021)

As shown in Table 1, different algorithms are utilized across sentiment analysis levels, with deep learning and hybrid approaches increasingly demonstrating higher classification accuracy in educational datasets.

APPROACHES FOR SENTIMENT ANALYSIS

Sentiment annotation involves assigning emotional polarity—positive, negative, or neutral—to textual elements such as documents, sentences, or phrases. At a finer level, it can capture specific emotional cues reflecting users' attitudes toward a service or feature. Within educational contexts, this process helps interpret student feedback on learning management systems, instructional design, and teaching quality, enabling institutions to refine pedagogy, predict outcomes, and enhance personalization.

The human annotation of feelings becomes challenging due to time and resource limitations as the volume of data escalates. Consequently, advancements in Artificial Intelligence (AI) and Natural Language Processing (NLP) have enabled the implementation of automated procedures. These methodologies encompass lexicon-based, corpus-driven, machine learning, deep learning, and transformer-based techniques. Through the use of a human-in-the-loop system that included sixty annotators and six computational tools—CoreNLP, VADER, TextBlob, Glove+LSTM, LIME, and RoBERTa—Yeruva et al. (2020) conducted an evaluation of human and automated sentiment annotation. According to the findings of a correlation analysis, automated models often outperformed human annotators in terms of accuracy for large-scale text analysis. However, these models still benefited from minimal human oversight for contextual refinement.

A. Unsupervised Annotation Techniques

1) Lexicon-based Approach: Lexicon-based sentiment analysis employs predefined word lists—known as sentiment lexicons—that assign polarity scores (positive, negative, or neutral) to words or phrases (Catelli et al., 2022; Dolianiti et al., 2019). These methods analyze sentiment in an unsupervised way by

J. Mountain Res. P-ISSN: 0974-3030, E-ISSN: 2582-5011
Vol. 21(S2), (2026) 37-52
DOI: https://doi.org/10.51220/jmr.v21-S2.5

summing or averaging the polarity of key lexical elements such as adjectives (Hatzivassiloglou and McKeown, 1997; Taboada et al., 2006), adverbs (Benamara et al., 2007), verbs (Vermeij, 2005), and nouns (Neviarouskaya et al., 2009a).

Commonly used English sentiment lexicons include SentiWordNet (Baccianella et al., 2010), Opinion Finder (Wilson et al., 2005a), Bing Liu's Opinion Lexicon (Liu, 2012), MPQA Subjectivity Lexicon (Wilson et al., 2005b), Harvard General Inquirer (Stone et al., 1966), AFINN (Nielsen, 2011), SentiFul (Neviarouskaya et al., 2009b), VADER (Hutto and Gilbert, 2014), and TextBlob. These resources remain fundamental to rule-based sentiment analysis and serve as benchmarks for evaluating more advanced AI-driven techniques.

For the purpose of evaluating active learning models such as the flipped classroom (Maher et al., 2015) and lightweight team structures (Latulipe et al., 2015; MacNeil et al., 2016), Tzacheva and Easwaran (2021) utilized the NRC Emotion Lexicon (Ortony and Turner, 1990; Mohammad et al., 2013) to analyze student feedback across emotional dimensions. These dimensions included joy, trust, fear, anger, and anticipation. Their study of the longitudinal relationship between 2015 and 2020 found that trust was the most prominent positive feeling throughout that time period.

Roselind and Suguna (2022) created a Customized Sentiment Lexicon (CSL) that was modified for use in academic course assessments. This was done with the intention of improving upon previous lexicon-based methodologies. The researchers were able to reach a greater level of accuracy in determining the polarity of sentiment within educational reviews by comparing the polarity scores obtained from the Bing Lexicon and the CSL using a bag-of-words model.

Lexicon-based algorithms are useful for categorizing sentiments without supervision or manual annotation, which makes them good for vast amounts of educational data. However, their dependence on a certain domain sometimes limits how they may be understood in context. This is because preset dictionaries may not be able to keep up with new words or words that are specialized to a certain field. This limitation has led to the creation of corpus-based approaches, which get sentiment information directly from how words are used in domain-specific text corpora.

2) *Corpus-based Approach:* A corpus refers to a structured text collection representing a specific domain or topic. In sentiment analysis, corpus-based methods infer word polarity through **contextual and statistical analysis**, examining co-occurrence frequencies and grammatical dependencies within large textual datasets. This technique extends sentiment lexicons by identifying **semantic relationships** among words across varying contexts. The fundamental idea is to compute the **semantic proximity** between a target word and sets of positive or negative reference terms to determine its polarity. Corpus-based models are able to turn generic lexicons into domain-sensitive sentiment resources that more effectively capture the linguistic and emotional expressions that are present in educational data (Alqasemi et al., 2017). This is accomplished by exploiting contextual evidence.

B. Supervised Annotation Techniques

Although lexicon- and corpus-based approaches effectively categorize student opinions, handling extensive textual feedback remains complex. **Supervised learning** overcomes this challenge by training models on **annotated datasets** to recognize sentiment categories automatically. These models learn **linguistic and semantic patterns** from preprocessed text, allowing them to classify or predict learners' emotional orientations and attitudes in educational feedback with higher accuracy and scalability.

1) *Machine Learning:* To automate the interpretation of student feedback, machine learning models are essential to supervised

DOI: https://doi.org/10.51220/jmr.v21-S2.5

sentiment categorization. In order to distinguish between positive, neutral, and negative feelings, Osmanoğlu et al. (2020) assessed university course reviews using a variety of classifiers, including logistic regression, decision trees, multilayer perceptrons, XGBoost, support vector machines, Gaussian Naïve Bayes, and k-nearest neighbours. The most accurate of them was logistic regression, which allowed for efficient course material improvement in subsequent semesters.

Lwin et al. (2020) introduced a hybrid feedback analysis framework that merged quantitative student ratings with textual reviews. Ratings were clustered using the K-means algorithm, and multiple classifiers—SVM, logistic regression, multilayer perceptron, and random forest—were employed for sentiment prediction. A Naïve Bayes model, trained on manually annotated feedback, effectively differentiated positive and negative comments.

Building on this, Faizi (2022) applied machine learning to analyze learner responses to YouTube educational videos. Among the tested algorithms—random forest, logistic regression, Naïve Bayes, and SVM—the SVM model yielded the highest accuracy: 92.82% with unigram–bigram and 92.67% with unigram–trigram features.

Kaur et al. (2022) advanced conventional sentiment classification by developing a hybrid framework that integrated random forest, logistic regression, and SVM models with the SentiWordNet lexicon-based approach. Across varying training–testing splits, the hybrid classifier achieved superior precision and consistency, outperforming the individual SVM model in sentiment prediction.

2) *Deep Learning:* Deep learning models have significantly advanced sentiment analysis by capturing intricate linguistic and contextual cues in student feedback. The researchers Yu et al. (2018) used a Convolutional Neural Network (CNN) to identify students who were considered to be at danger of failing their classes. The CNN combined structured inputs, such as attendance and grades, with unstructured comments from 181 undergraduate students. The CNN model, trained on manually annotated data, achieved F-measures of 0.78, 0.73, and 0.71 across the 5th, 7th, and 9th weeks, outperforming the SVM baseline. Similarly, Sutoyo et al. (2021) applied a CNN-based framework to evaluate instructor performance, attaining 87.95% accuracy, 87% precision, 78% recall, and an F1-score of 81%. To improve emotional inference, Sangeetha and Prabha (2020) introduced a multi-head attention fusion model combining GloVe and CoVe embeddings within an LSTM network. Parallel attention heads processed sentence sequences while optimized dropout layers mitigated overfitting. The fusion model yielded higher accuracy in classifying positive, negative, and neutral sentiments than standalone attention or LSTM architectures.

3) *Transformer-Based Approach:* Recent transformer-based models, especially Bidirectional Encoder Representations from Transformers (BERT), have advanced sentiment analysis by learning contextual meaning in both directions of text. Dyulicheva and Bilashova (2022) employed a BERT-driven framework to examine student opinions and extract sentiment-bearing keywords. Through K-means clustering and cosine similarity applied to 300 Udemy MOOC titles, they organized the data into 14 clusters highlighting dominant emotional terms. The findings showed that students tended to express more negative perceptions of courses compared to instructors.

Li et al. (2019) enhanced transformer-based sentiment analysis through a lightweight BERT–CNN hybrid model applied to MOOC feedback. This framework merged BERT's contextual embeddings with convolutional and self-attention layers, achieving a 50% reduction in parameters without sacrificing accuracy. The model delivered 92.8%

accuracy, with F1-scores of 95.2% for positive and 81.3% for negative sentiments, surpassing traditional lexicon-driven techniques while using only 61 million parameters.

Deep Learning in Recommendation Systems

Deep learning models employ multilayered neural architectures that automatically learn complex feature hierarchies from data, removing the dependency on manual feature extraction typical in classical machine learning. Unlike traditional classifiers—such as Support Vector Machines (SVMs), Bayesian networks, or decision trees—which rely on predefined input features, deep networks derive abstract representations directly from raw data, improving adaptability and predictive precision. Automated hyperparameter tuning further refines model performance.

Figure 2 contrasts sentiment classification in traditional and deep learning approaches, highlighting how layered representations capture subtle emotional cues more effectively. Through artificial neural networks, deep learning now underpins key advancements in image, speech, and language processing, extending its influence to personalized recommendation systems.

1) Deep Neural Networks (DNN): A Deep Neural Network (DNN), as described by Aggarwal (2018), extends traditional neural architectures by integrating multiple hidden layers between the input and output stages. Each layer comprises interconnected neurons that process data through weighted links and nonlinear activation functions. This hierarchical configuration enables the model to learn increasingly abstract representations from raw input. A standard DNN consists of an input layer for feature reception, several hidden layers for transformation, and an output layer that delivers the final prediction or classification. *Figure 2* depicts the general framework of a deep neural network.

Convolutional Neural Networks (CNN): In the domains of computer vision, recommender systems, and natural language processing, a Convolutional Neural Network (CNN) is frequently employed as a deep feed-forward model (Zhang et al., 2018). Its architecture comprises successive convolutional and pooling layers, followed by fully connected layers for final classification. Convolutional layers extract localized features from input data, whereas pooling layers reduce dimensionality and enhance resistance to noise and overfitting. Figure 3 depicts a CNN framework in which the input embedding matrix passes through four convolutional and two max-pooling layers. The first pair, with 64 and 32 filters, capture essential patterns and are followed by pooling to minimize feature complexity. The subsequent layers, using 16 and 8 filters, refine these maps before reaching a fully connected layer that condenses the final feature representation into a single output node predicting two sentiment classes—positive or negative.

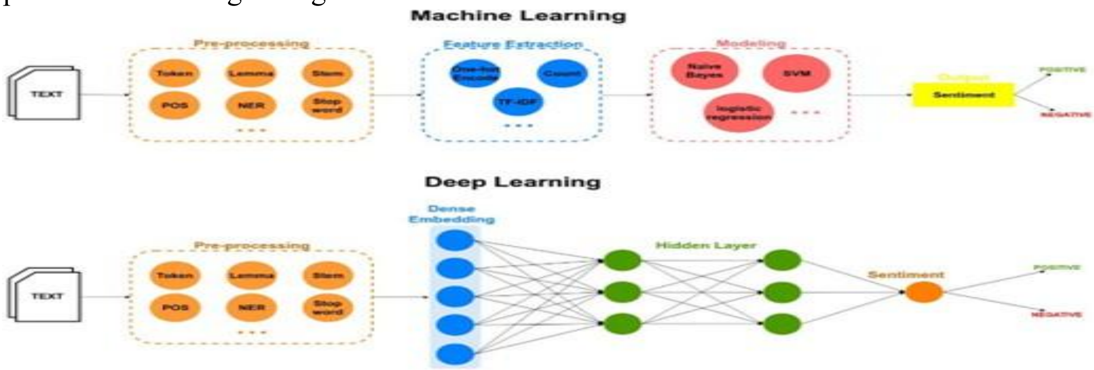

Fig. 2: Differences between two classification approaches of sentiment polarity: machine learning

(top) and deep learning (bottom). POS: Part of Speech; NER: Named Entity Recognition; TF-IDF: Term Frequency-Inverse Document Frequency.

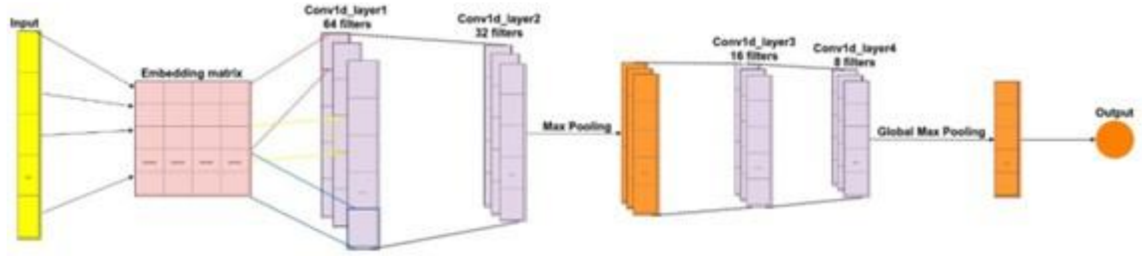

Fig. 3: A convolutional neural network (CNN).

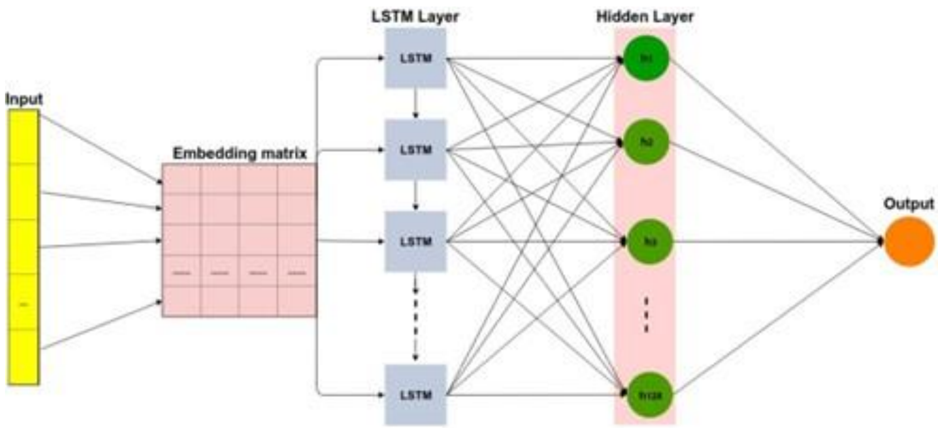

Fig. 4: A long short-term memory network (LSTM).

2) *Recurrent Neural Networks (RNN):* A Recurrent Neural Network (RNN) (Britz, 2015) is structured to process sequential or time-dependent data through feedback connections that enable the retention of information across iterations. Unlike standard feed-forward models, RNNs maintain internal states that influence later computations, making them suitable for capturing contextual or temporal patterns in text.

An advanced variant, the Long Short-Term Memory (LSTM) network (Hochreiter and Schmidhuber, 1996), addresses the vanishing gradient issue by incorporating gated memory cells that store information over long sequences. ***Figure 4*** presents a typical LSTM configuration in which input data are embedded into vector form and passed through an LSTM layer with 200 units. The resulting output feeds into a fully connected layer of 128 neurons, and a sigmoid activation generates a binary sentiment prediction—**positive** or **negative**.

Challenges of Recommendation Systems

Several researchers have examined how **deep learning enhances recommendation systems**, improving both **personalization and predictive accuracy**. Rosa et al. (2022) utilized **feed-forward neural networks** to strengthen collaborative filtering mechanisms beyond conventional architectures. Similarly, Dai and Xu (2021) designed a **personalized learning resource recommender** employing an optimized backpropagation neural model. Dang et al. (2020) explored **session-based recommendation frameworks** that reduce dependence on long-term user histories while maintaining accuracy, showcasing the effectiveness of **GRU4Rec** on real-world datasets. In addition, Anantha and Bathula (2018) performed a cross-domain comparison of classical and deep learning-based systems. They highlighted the improved adaptability and performance demonstrated by neural techniques.

Even with the substantial progress that has

J. Mountain Res. P-ISSN: 0974-3030, E-ISSN: 2582-5011
Vol. 21(S2), (2026) 37-52
DOI: https://doi.org/10.51220/jmr.v21-S2.5

been made, recommendation systems that are based on deep learning continue to face severe restrictions. One major issue is **data sparsity**, where a lack of sufficient user–item interactions undermines model accuracy. The **cold-start problem** similarly restricts predictions for new users or resources due to limited historical data (Zhong et al., 2022; Hansel and Wibowo, 2022). Another recurring challenge is **overfitting**, in which models adapt too closely to training data, impairing their ability to generalize to unseen inputs (Merabet and Benmerzoug, 2022). Moreover, scalability concerns arise when processing increasingly large datasets (Li and Kim, 2021), alongside growing **ethical considerations**, including algorithmic bias and privacy protection.

1) Comparison with Prior Reviews: For the purpose of ensuring that it accurately reflects the most current developments in deep learning–based recommendation systems, this review focuses on journal publications that have been subjected to peer review and published between the years 2018 and 2023. Earlier works, for example, **Batmaz and Kaleli (2019)**—which examined theses and conference papers from 2007–2017—and **Da'u and Salim (2020)**—who analyzed 99 studies from 2007–2018—provided broader yet less contemporary overviews. By focusing on a shorter and more recent timespan, this review emphasizes methodological rigor and analytical relevance. Furthermore, the use of **comprehensive search terms** enabled the inclusion of diverse studies across multiple domains. **Table II** summarizes key prior reviews and research works, outlining their publication venues, focus areas, employed deep learning approaches, and their correspondence with the present study.

Table 2: Comparative Summary of Surveys and Studies on Deep Learning–Based Recommendation Systems

Category	Representative Studies	Journal Index	Primary Application Domain	Key Deep Learning Focus	Distinctive Aspect of the Review
Survey	Zhang et al. (2021)	J1	General	Deep Reinforcement Learning	Covers a broader range of DL paradigms beyond DRL approaches
Survey	Liu et al. (2019)	J2	General	Deep Neural Networks	Expands the analysis to other network architectures and hybrid models
Survey	Zhu et al. (2021)	J3	General	Graph Neural Networks	Extends beyond GNN-based methods to include multiple DL families
Survey	Yang (2021)	J4	E-learning	Deep Learning	Addresses a wider educational recommendation spectrum
Survey	Ding et al. (2018)	J2	Point of Interest (POI)	Deep Learning	Broadens the discussion beyond POI-specific models
Survey	Acheampong et al. (2021)	J5	General	Deep Neural Networks	Encompasses DL strategies other than standard NN frameworks
Survey	Kuleto et al. (2021)	J6	Citation Analysis	Deep Learning	Incorporates diverse domain-oriented applications
Survey	Dolianiti et al. (2019)	J7	General	Autoencoder-Based	Synthesizes various DL architectures including autoencoders
Survey	Ahmad et al. (2018)	J3	General	Deep Learning	Focuses on major works within a five-year horizon
Survey	Sivakumar and Reddy (2017)	J8	General	Deep Learning	Highlights seminal contributions over a five-year period
Survey	Nikolic´ et al. (2020)	J9	Course	Deep Learning	Explores multi-domain educational recommender approaches
Study	Yeruva et al. (2020)	J10	Trust-Aware	Deep Learning	Integrates trust metrics with hybrid DL models
Study	Rosa et al. (2022)	J11	Service	Neural Collaborative	Provides a broader cross-domain analysis
Study	Li and Lu (2017)	J12	E-learning	Neural Networks	Extends coverage to diversified learning environments

DOI: https://doi.org/10.51220/jmr.v21-S2.5

Category	Representative Studies	Journal Index	Primary Application Domain	Key Deep Learning Focus	Distinctive Aspect of the Review
Study	Shaik et al. (2022)	J13	Session-Based	Deep Learning	Discusses several contextual and temporal use cases
Study	Anantha and Bathula (2018)	J14	General	Deep Learning	Includes multiple recent and domain-specific studies
Review	Liu et al. (2022)	J15	E-learning	Deep Learning	Synthesizes DL applications across educational contexts
Review	Edalati et al. (2021)	J1	General	Deep Knowledge Graph	Goes beyond embedding-based representation learning
Review	Wehbe et al. (2021)	J2	POI Recommendation	Deep Learning	Extends analysis beyond location-based recommendations
Review	Dehbozorgi and Mohandoss (2021)	J15	General	Deep Learning	Summarizes contemporary works from the last five years

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.

Conclusion

The purpose of this study was to investigate the development of recommendation systems that are driven by deep learning, with a particular focus on how contemporary designs improve customization and flexibility. Despite the fact that its performance frequently changed depending on the features of the domain in question, neural collaborative filtering emerged as the approach that was most widely used across all domains. There were also other methods that were utilized regularly, such as matrix factorization, graph neural networks, and attention-based mechanisms. Additionally, sophisticated models such as convolutional neural networks (CNNs), deep reinforcement learning, knowledge graphs, and sequential frameworks were utilized.

There was a significant amount of emphasis placed on social, session-based, and point-of-interest (POI) recommenders in the most recent research, which highlights the adaptability of these recommenders across a variety of applications.

After grouping deep learning approaches into groups and mapping their application across a variety of domains, this review brings together the findings that have been accumulated over the course of three years. In order to assist scholars in navigating the growing body of knowledge, this systematic taxonomy helps to clarify conceptual linkages. In addition, the incorporation of standard datasets and assessment criteria provides a point of reference for researchers who will conduct empirical research in the future.

This research, in general, contributes to a more in-depth knowledge of recent advancements in deep learning-based recommendation systems and highlights potential avenues for the construction of recommender models that are more predictive, scalable, and aware of their environment.

References

Acheampong F A, Nunoo-Mensah H and Chen W (2021). Transformer models for text-based emotion detection: a review of BERT-based approaches. *Artificial Intelligence Review*, 54(8):5789–5829.

Aggarwal C C (2018). *Neural networks and deep learning*. Springer International Publishing, Cham, Switzerland.

 DOI: https://doi.org/10.51220/jmr.v21-S2.5

Ahmad M, Aftab S, Bashir M S and Hameed N (2018). Sentiment analysis using SVM: a systematic literature review. *International Journal of Advanced Computer Science and Applications*, 9(2):135–142.

Akbar A, Agarwal P and Obaid A (2022). Recommendation engines – neural embedding to graph-based: techniques and evaluations. *International Journal of Nonlinear Analysis and Applications*, 13(1):2411–2423.

Alatrash R, Priyadarshini R, Ezaldeen H and Alhinnawi A (2022). Augmented language model with deep learning adaptation on sentiment analysis for e-learning recommendation. *Cognitive Systems Research*, 75:53–69.

Alqasemi F, Abdelwahab A and Abdelkader H (2017). Constructing automatic domain-specific sentiment lexicon using KNN search via terms discrimination vectors. *International Journal of Computer Applications*, 41(2):129–139.

Anantha N L and Bathula B (2018). Comparative study on traditional recommender systems and deep learning-based recommender systems. *Advances in Modelling and Analysis B*, 61(2):64–69.

Baccianella S, Esuli A and Sebastiani F (2010). SentiWordNet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In: *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*. European Language Resources Association (ELRA), Valletta, Malta.

Bandyopadhyay S and Thakur S (2020). Product prediction and recommendation in e-commerce using collaborative filtering and artificial neural networks: a hybrid approach. *International Journal of Computer Applications*, 176(2):784–789.

Benamara F, Cesarano C, Picariello A, Recupero D and Subrahmanian V (2007). Sentiment analysis: adjectives and adverbs are better than adjectives alone. In: *Proceedings of the International Conference on Weblogs and Social Media (ICWSM)*, Vol. 7, pp. 203–206.

Britz D (2015). Recurrent neural networks tutorial: part 1 – introduction to RNNs. Available at: http://www.wildml.com/2015/09/recurrent-neural-networkstutorial-part-1-introduction-to-rnns/ (Accessed: 12 March 2020).

Catelli R, Pelosi S and Esposito M (2022). Lexicon-based vs. BERT-based sentiment analysis: a comparative study in Italian. *Electronics*, 11(3):374.

Da'u A and Salim N (2020). Recommendation system based on deep learning methods: a systematic review and new directions. *Artificial Intelligence Review*, 53(4):2709–2748.

Dai Y and Xu J (2021). Study of online learning resource recommendation based on improved BP neural network. *International Journal of Embedded Systems*, 14(2):101.

Dang C, Moreno-García M and de la Prieta F (2021). Using hybrid deep learning models of sentiment analysis and item genres in recommender systems for streaming services. *Electronics*, 10(20):2459.

Dang T K, Nguyen Q P and Nguyen V S (2020). A study of deep learning-based approaches for session-based recommendation systems. *Social Network Computing Science*, 1(4):216.

Dehbozorgi N and Mohandoss D (2021). Aspect-based emotion analysis on speech for predicting performance in collaborative learning. In: *Proceedings of the IEEE Frontiers in Education Conference (FIE 2021)*. IEEE, San Diego, CA, USA.

Dolianiti F, Iakovakis D, Dias S, Hadjileontiadou S, Diniz J, Natsiou G, Tsitouridou M, Bamidis P and Hadjileontiadis L (2019). Sentiment

DOI: https://doi.org/10.51220/jmr.v21-S2.5

analysis on educational datasets: a comparative evaluation of commercial tools. *Educational Journal of the University of Patras UNESCO Chair*, 6(1).

Dyulicheva Y Y and Bilashova E A (2022). Learning analytics of MOOCs based on natural language processing. *Journal of Education and Information Technologies*, 27(3):4811–4825.

Edalati M, Imran A S, Kastrati Z and Daudpota S M (2021). The potential of machine learning algorithms for sentiment classification of students' feedback on MOOCs. In: *Proceedings of the International Conference on Data Mining and Machine Learning Applications (DMMLA 2021)*, pp. 11–22. IEEE, New York, USA.

Elfeky A I M, Masadeh T S Y and Elbyaly M Y H (2020). Advance organizers in flipped classrooms via e-learning management systems and the promotion of integrated science process skills. *Thinking Skills and Creativity*, 35:100622.

Faizi R (2022). A sentiment-based approach to predict learners' perceptions towards YouTube educational videos. In: *Proceedings of the 2022 International Conference on Advanced Learning Technologies (ICALT)*, pp. 549–556. IEEE.

Hansel A C and Wibowo A (2022). Using movie genres in neural network–based collaborative filtering for reducing the cold-start problem. *International Journal of Emerging Technology and Advanced Engineering*, 12(3):63–73.

Hatzivassiloglou V and McKeown K (1997). Predicting the semantic orientation of adjectives. In: *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics (ACL 1997)*, pp. 174–181. Association for Computational Linguistics, Madrid, Spain.

Hochreiter S and Schmidhuber J (1996). LSTM can solve hard long time-lag problems. In: *Advances in Neural Information Processing Systems (NIPS 1996)*, pp. 473–479. Denver, CO, USA.

Hutto CJ and Gilbert E (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. *In:* Proceedings of the International AAAI Conference on Web and Social Media (ICWSM 2014), pp. 216–225. AAAI Press, Palo Alto, CA, USA.

Kastrati Z, Imran A and Kurti A (2020). Weakly supervised framework for aspect-based sentiment analysis on students' reviews of MOOCs. *IEEE Access*, 8:106799–106810.

Kaur C, Boush M S A, Hassen S M, Hakami W A, Abdalraheem M H O, Galam N M, Hadi N A and Benjeed A O S (2022). Incorporating sentiment analysis into the development of a hybrid classification model. *International Journal of Health Sciences*, 6(S1):1709–1720.

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

Latulipe C, Long N B and Seminario C E (2015). Structuring flipped classes with lightweight teams and gamification. In: *Proceedings of the 46th ACM Technical Symposium on Computer Science Education (SIGCSE 2015)*, pp. 392–397. ACM, Kansas City, MO, USA.

Li H and Lu W (2017). Learning latent sentiment scopes for entity-level sentiment analysis. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 31, AAAI Press, San Francisco, CA, USA.

Li Q and Kim J (2021). A deep learning-based course recommender system for sustainable development in education. *Applied Sciences*, 11(19):8993.

 DOI: https://doi.org/10.51220/jmr.v21-S2.5

Li X, Zhang H, Ouyang Y, Zhang X and 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 2019)*, pp. 1–6. IEEE, Yogyakarta, Indonesia.

Ligthart A, Catal C and Tekinerdogan B (2021). Systematic reviews in sentiment analysis: A tertiary study. *Artificial Intelligence Review*, 54(7):4997–5053.

Liu T, Wu Q, Chang L and Gu T (2022). A review of deep learning-based recommender systems in e-learning environments. *Artificial Intelligence Review*, 55(8):5953–5980.

Liu Z, Yang C, Rudian S, Liu S, Zhao L and 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.

Lwin H H, Oo S, Ye K Z, Lin K K, Aung W P and Ko P P (2020). Feedback analysis in outcome-based education using machine learning. In: *Proceedings of the 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON 2020)*, pp. 272–276. IEEE, Phuket, Thailand.

MacNeil S, Latulipe C, Long B and Yadav A (2016). Exploring lightweight teams in a distributed learning environment. *In:* Proceedings of the 47th ACM Technical Symposium on Computer Science Education (SIGCSE 2016), pp. 158–163. ACM, Memphis, TN, USA.

Maher M L, Latulipe C, Lipford H and Rorrer A (2015). Flipped classroom strategies for computer science education. In: *Proceedings of the 46th ACM Technical Symposium on Computer Science Education (SIGCSE 2015)*, pp. 218–223. ACM, Kansas City, MO, USA.

McKinney L, Burridge A B, Lee M M, Bourdeau G V and Miller-Waters M (2022). Incentivizing full-time enrollment at community colleges: What influences students' decision to take more courses? *Community College Review*, 50(2):144–170.

Merabet F and Benmerzoug D (2022). QoS prediction for service selection and recommendation with a deep latent features autoencoder. *Computer Science and Information Systems*, 19(2):709–733.

Mohammad SM, Kiritchenko S and Zhu X (2013). NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets. *arXiv preprint*, arXiv:1308.6242.

Neviarouskaya A, Prendinger H and 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 (ICWSM 2009)*, Vol. 3, pp. 278–281. AAAI Press, San Jose, CA, USA.

Neviarouskaya A, Prendinger H and Ishizuka M (2009b). Compositionality principle in recognition of fine-grained emotions from text. In: *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM 2009)*, Vol. 3, pp. 278–281. AAAI Press, San Jose, CA, USA.

Nielsen FA (2011). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. *arXiv preprint*, arXiv:1103.2903.

Nikolić N, Grljević O and Kovačević A (2020). Aspect-based sentiment analysis of reviews in the domain of higher education. *The Electronic Library*, 38(1):44–64.

Ortony A and Turner TJ (1990). What's basic about basic emotions? *Psychological Review*, 97(3):315–331.

Rosa P D, Deriaz M, Di Marco M and Laura L (2022). Service recommendations with deep learning: A study on neural collaborative engines. *Pacific Asia Journal of the Association for*

DOI: https://doi.org/10.51220/jmr.v21-S2.5

Information Systems, 14:59–70.

Rosalind J M and Suguna S (2022). Predicting students' satisfaction towards online courses using aspect-based sentiment analysis. In: *Computer, Communication, and Signal Processing*, pp. 20–35. Springer International Publishing, Cham, Switzerland.

Sangeetha K and Prabha D (2020). Sentiment analysis of student feedback using multi-head attention fusion model of word and context embedding for LSTM. *Journal of Ambient Intelligence and Humanized Computing*, 12(3):4117–4126.

Shaik T, Tao X, Higgins N, Gururajan R, Li Y, Zhou X and Acharya U (2022). FedStack: Personalized activity monitoring using stacked federated learning. *Knowledge-Based Systems*, 257:109929.

Sivakumar M and Reddy U S (2017). Aspect-based sentiment analysis of students' opinion using machine learning techniques. In: *Proceedings of the 2017 International Conference on Inventive Computing and Informatics (ICICI)*, pp. 155–160. IEEE, Coimbatore, India.

Stone PJ, Dunphy DC and Smith MS (1966). *The General Inquirer: A Computer Approach to Content Analysis*. MIT Press, Cambridge, MA, USA.

Sutoyo E, Almaarif A and Yanto I T R (2021). Sentiment analysis of student evaluations of teaching using a deep learning approach. In: *Proceedings of the International Conference on Emerging Applications and Technologies for Industry 4.0 (EATI 2020)*, pp. 272–281. Springer International Publishing, Cham, Switzerland.

Taboada M, Gillies M and McFetridge P (2006). Sentiment classification techniques for tracking literary reputation. *In:* Proceedings of the LREC Workshop on Towards Computational Models of Literary Analysis, pp. 36–43. Citeseer, Genoa, Italy.

Tao X, Zhou X, Zhang J and Yong J (2016). Sentiment analysis for depression detection on social networks. In: *Proceedings of the International Conference on Advanced Data Mining and Applications (ADMA 2016)*, pp. 807–810. Springer International Publishing, Cham, Switzerland.

Tzacheva A and Easwaran A (2021). Emotion detection and opinion mining from student comments for teaching innovation assessment. *International Journal of Education (IJE)*, 9(2):21–32.

Vermeij M (2005). The orientation of user opinions through adverbs, verbs and nouns. *In:* Proceedings of the 3rd Twente Student Conference on IT, Enschede, The Netherlands. Citeseer.

Wilson T, Hoffmann P, Somasundaran S, Kessler J, Wiebe J, Choi Y, Cardie C, Riloff E and Patwardhan S (2005a). OpinionFinder: A system for subjectivity analysis. In: *Proceedings of HLT/EMNLP 2005 Interactive Demonstrations*, pp. 34–35. Association for Computational Linguistics, Vancouver, Canada.

Wilson T, Wiebe J and Hoffmann P (2005b). Recognizing contextual polarity in phrase-level sentiment analysis. In: *Proceedings of the Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP)*, pp. 347–354. Association for Computational Linguistics, Vancouver, Canada.

Yang R (2021). Machine learning and deep learning for sentiment analysis over students' reviews: An overview study. *Education and Information Technologies*, 26(2):2115–2135.

Yeruva V, Chandrashekar M, Lee Y, Rydberg-Cox J, Blanton V and Oyler N (2020). Interpretation of sentiment analysis with human-in-the-loop. In: *Proceedings of the 2020 IEEE International Conference on Big Data (Big Data)*, pp. 250–257. IEEE, Atlanta, GA, USA.

Yu L, Lee C, Pan H, Chou C, Chao P, Chen Z, Tseng S, Chan C and Lai K (2018). Improving early prediction of academic failure using sentiment analysis on self-evaluated comments. *Journal of Computer Assisted Learning*, 34(4):358–365.

Zhang L, Wang S and Liu B (2018). Deep learning for sentiment analysis: A survey. *WIREs Data Mining and Knowledge Discovery*, 8:e1253.

Zhang Y, Wang J and Zhang X (2021). Conciseness is better: Recurrent attention LSTM model for document-level sentiment analysis. *Neurocomputing*, 462:101–112.

Zhao H, Liu Z, Yao X and Yang Q (2021). A machine learning-based sentiment analysis of online product reviews with a novel term weighting and feature selection approach. *Information Processing and Management*, 58(5):102656.

Zhong S-T, Huang L, Wang C-D, Lai J-H and Yu P (2022). An autoencoder framework with attention mechanism for cross-domain recommendation. *IEEE Transactions on Cybernetics*, 52(6):5229–5241.

Zhou X, Tao X, Yong J and 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)*, pp. 557–562. IEEE, Whistler, Canada.

Zhu J, Chang Y-C, Ku C-H, Li S and Chen C-J (2021). Online critical review classification in response strategy and service provider rating: Heuristic processing, sentiment analysis to deep learning. *Journal of Business Research*, 129:860–877.