Sentiment analysis and opinion mining on educational data: A survey

Thanveer Shaik [a],*, Xiaohui Tao [a], Christopher Dann [b], Haoran Xie [c], Yan Li [a], Linda Galligan [a]

[a] *School on Mathematics, Physics, and Computing, University of Southern Queensland, Toowoomba 4350, Australia*
[b] *School of Education, University of Southern Queensland, Toowoomba 4350, Australia*
[c] *Department of Computing and Decision Sciences, Lingnan University, Tuen Mun, Hong Kong*

ARTICLE INFO

ABSTRACT

Sentiment analysis AKA opinion mining is one of the most widely used NLP applications to identify human intentions from their reviews. In the education sector, opinion mining is used to listen to student opinions and enhance their learning–teaching practices pedagogically. With advancements in sentiment annotation techniques and AI methodologies, student comments can be labelled with their sentiment orientation without much human intervention. In this review article, (1) we consider the role of emotional analysis in education from four levels: document level, sentence level, entity level, and aspect level, (2) sentiment annotation techniques including lexicon-based and corpus-based approaches for unsupervised annotations are explored, (3) the role of AI in sentiment analysis with methodologies like machine learning, deep learning, and transformers are discussed, (4) the impact of sentiment analysis on educational procedures to enhance pedagogy, decision-making, and evaluation are presented. Educational institutions have been widely invested to build sentiment analysis tools and process their student feedback to draw their opinions and insights. Applications built on sentiment analysis of student feedback are reviewed in this study. Challenges in sentiment analysis like multi-polarity, polysemous, negation words, and opinion spam detection are explored and their trends in the research space are discussed. The future directions of sentiment analysis in education are discussed.

Contents

* Corresponding author.
  E-mail address: Thanveer.Shaik@usq.edu.au (T. Shaik).

1. Introduction

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., 2022). 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., 2021). 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., 2016; Zhou et al., 2013).

Advancements in technology have transformed fields like healthcare (Tao et al., 2021; Shaik et al., 2022a) and education by adopting AI and NLP techniques. In the education domain, student feedback plays a vital role in evaluating and analysing learning management systems, teaching, pedagogical procedures, and courses (Elfeky et al., 2020). 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., 2022). 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., 2021). 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., 2021). 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., 2019) and requires many resources with pedagogical understanding in education. This challenge has been addressed by developing different sentiment annotation 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 analysing a large number of student comments (Zhu et al., 2021). AI methodologies like machine learning, deep learning, and transformers (Acheampong et al., 2021) are capable of learning student opinions with attention mechanisms and classifying or predicting their emotions for unlabelled student comments (Kuleto et al., 2021). The unsupervised sentiment annotation techniques and the AI methodologies overcome the challenge of manual labelling to a certain extent. In this review article, the following research questions were explored:

- What is the role of sentiment analysis and opinion mining in education?

- What are the different levels of sentiment analysis in the education domain?
- What are the current trends in sentiment annotation of educational data?
- What are the challenges of adapting sentiment analysis in the education domain?

This study explores sentiment analysis's role in enhancing educational procedures, different levels of sentiment analysis, and sentiment annotation techniques. The challenges that affect sentiment analysis in the education domain are discussed in detail. The contributions of this research are as follows:

- Enhancement in educational domain procedures by adopting the sentiment analysis.
- Overview on sentiment annotation techniques and role AI in sentiment analysis.
- Exploring trends and challenges in sentiment analysis that need to be addressed to be adopted in the education domain.

The remainder of the paper is organised into sections. Different levels of sentiment analysis adopted in education have been explored in Section 2. Sentiment annotation techniques and AI methodologies involved in sentiment analysis are reviewed in Section 3. Section 4 discusses the impact of sentiment analysis on educational procedures like evaluation, decision-making, and pedagogical concepts. Section 5 presents the challenges in sentiment analysis and its trends. In Section 6, the future directions of sentiment analysis in the educational domain are discussed. Finally, the article concludes in Section 7.

2. Sentiment analysis in education

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., 2021). 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 analysed 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).

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. (2019) 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
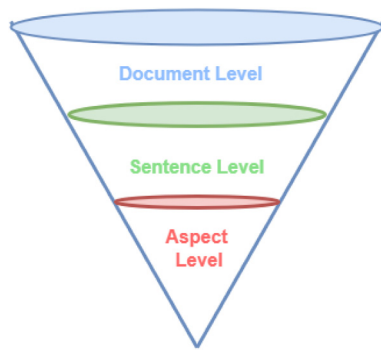
Fig. 1. Sentiment Analysis in Education.

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 outperformed 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. (2018). 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 (2017) 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. (2020) 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 analyse their opinions. Yang (2021) 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., 2022b). Ding et al. (2018) 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 (2017) proposed a novel approach to learning latent sentiment scopes at the entity level with named entities and sentiment is highlighted. The authors used graphical models to form nine different types of nodes at each word's position to determine before entity B, entity E, and after entity A. Based on their polarity, each word is annotated as positive, negative or neutral sentiment.

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 factorisation (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. (2022b).

Rosalind and Suguna (2022) 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 customised 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. (2021) 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. (2021) proposed a Spatiotemporal sentiment analysis frame-work to analyse 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, multinominal, 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 (2021) 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. (2020) proposed an aspect-level sentiment analysis framework to identify sentiment or opinion polarity towards a given aspect related to MOOCs. The proposed framework used weakly supervised annotation to identify aspect categories in unlabelled 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

Table 1
Sentiment analysis in education.

SA Level	Algorithms	Application	References
Document Level	SVM, NB	Evaluate the documents' overall opinion towards a context is positive or negative.	Dolianiti et al. (2019), Ahmad et al. (2018)
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 (2017), Nikolić et al. (2020)
Entity Level	Gradient boosting tree, SVM, NER, TF–IDF, Doc2Vec	Analyse sentiment or opinion in feedback towards an entity in educational domain	Yang (2021), Li et al. (2022b), Ding et al. (2018), Li and Lu (2017)
Aspect Level	LDA, RF, SVM, decision tree, CNN, KNN	Understand positive or negative aspects in educational practices.	Rosalind and Suguna (2022), Edalati et al. (2021), Wehbe et al. (2021), Dehbozorgi and Mohandoss (2021), Kastrati et al. (2020)

82.10% for aspect sentiment classification. The research works adopted AI in different levels of sentiment analysis are presented in Table 1.

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. Sentiment annotation techniques in education

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 streamlined, can predict student performance, and support students with personalised 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. (2020) 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 conducted 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., 2022). Identifying the sentiment and opinion words can help to categorise the sentiment in an unsupervised manner (Dolianiti et al., 2019). 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, 1997; Taboada et al., 2006), adverbs (Benamara et al., 2007), verbs (Vermeij, 2005), nouns (Neviarouskaya et al., 2009a) 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., 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), TextBlob (Loria, 2018) and so on.

Tzacheva and Easwaran (2021) used National Research Council (NRC) (Ortony and Turner, 1990; Mohammad and Turney, 2013) 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., 2015) and lightweight teams (Latulipe et al., 2015; MacNeil et al., 2016) 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 (2022) proposed steps to improve the existing Bing lexicon and develop a customised sentiment lexicon (CSL). The authors calculated sentiment polarity on academic course feedback texts using Bing sentiment Lexicon and CSL approaches. The process is to tokenise 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 co-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 underlying 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., 2017).

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 analysing 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 Osmanoğlu et al. (2020) 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. (2020) 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 naïve Bayes classifier was used to train the labelled dataset to classify the comments into positive or negative. Faizi (2022) proposed a machine learning approach to classify learners' feedback sentiment towards Youtube educational videos. Traditional machine learning methods such as random forest, logistic regression, Naïve 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. (2022) 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. (2018) 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 (Lang, 1980). 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. (2021). 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 (2020) 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 (2022) 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. (2019) analysed 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. Sentiment analysis impact on education

The research community has widely adopted sentiment analysis to extract student opinions toward teaching and learning practices in the education sector (Han et al., 2019). The opinion mining mechanism can be adapted to different educational domain applications such as understanding student engagement pedagogy, limitations of existing educational infrastructure, course and teaching evaluation, decision-making on educational policies, and so on. These applications can be achieved by understanding student feedback on a course in a semester, and their participation in online forums.

The role of sentiment analysis in the education domain is more toward student feedback analysis and enhancing learning–teaching practices (Zhou and min Ye, 2020). Sentiment analysis can also be adapted to personalise learning for individual students by understanding their issues, aspect-based, and entity-level opinions (D'Aniello et al., 2022). Text categorisation is to assign predefined topics in a comment. From a sentiment perspective, text can be categorised using sentence-level analysis. Emotional analysis summarises student feedback on a course into graded sentiment analysis with scales like very positive, positive, neutral, negative, and very negative (Sharma and Jain, 2020).

4.1. Learning and teaching systems evaluation

Teacher evaluation is part of education systems evaluations and it provides constructive feedback for teachers' professional growth. Tertiary education institutions conduct student feedback surveys on courses taught in that semester. The feedback has both qualitative and quantitative feedback on the courses. An ensemble learning of five machine learning algorithms such as Naive Bayes, logistic regression, support vector machine, decision tree, and random forest classifiers based on the majority voting principle was proposed by Lalata et al. (2019). The sentiment of student comments collected from each semester was manually labelled into positive, negative, and neutral comments. The authors performed individual classification and ensembled

the classifier output. The results show all model ensembling has achieved accuracy, F1-score, and recall of 90.32%, 93.80%, and 90.86%.

Roaring et al. (2022) analysed the relationship between the numerical rating of faculty performance and the actual opinions, feelings, and observations using text analytics. Association rule mining was used for pattern recognition based on rule-based metrics support, confidence, and lift. The faculty rating on a scale of 1–5 was evaluated in the study. Negative words were used in their feedback by students to attribute faculty with a rating of 1. Similarly, neutral words for faculty with a rating of 2, and positive words for faculty with a rating of 3. The feedback had positive observations and opinions for faculty with a rating of 4 and 5. Pramod et al. (2022) proposed a two-fold faculty evaluation system based on machine learning and text analytics. A structured questionnaire was used to record student and faculty characteristics, and qualitative and quantitative feedback from post-graduate students about their faculty. The qualitative feedback was analysed using sentiment analysis and converted the text feedback to polarity scores. The polarity of the words and sentences in the qualitative feedback varied across faculty and the time span of feedback. The engaging faculty members used case studies, imparting practical experiments, and real-time examples and they had high emotional engagement from students. The polarity scores of the qualitative feedback were combined with quantitative feedback to predict faculty effectiveness using ten different machine learning algorithms. Out of all, the random forest model achieved high accuracy, precision, and area under the curve of 98.87%, 97.71%, and 97.32% respectively.

Massive open online courses (MOOCs) deliver the learning content to participants breaking barriers like age, gender, race or geography. Onan (2020) evaluated the predictive performances of conventional machine learning, ensemble learning methods, and deep learning algorithms based on 66,000 student reviews for MOOCs. The conventional supervised methods and ensemble techniques used three-term weighting techniques: term presence, term frequency, and TF–IDF. Three word-embedding techniques word2vec, fast- Text, and GloVe were used in deep learning models. Out of all, the deep learning model LSTM with the GloVe word-embedding technique had the highest classification accuracy of 95.80%.

4.2. Enhance pedagogical concepts

Analysing student feedback each semester reveals issues in the pedagogical structure of teaching and learning practices. Student engagement is one of the key pedagogical concepts in education. Yan et al. (2019) evaluated learner engagement in Teach-Outs, which aims to drive meaningful discussions around topics of social urgency without any formal assessment. The authors used opinion-mining techniques including sentiment analysis, topic modelling, and similarity analysis to investigate the topics discussed by learners in the online forum. Latent Dirichlet Allocation (LDA) (Blei et al., 2003), a probabilistic topic modelling model technique, was used to extract topics from forum posts and the positive and negative emotions of each post were detected using Linguistic Inquiry and Word Count (LIWC) (Pennebaker et al., 2015). Li et al. (2022a) derived two distinct categories: knowledge-seeking and skill-seeking courses in MOOCs pedagogy by applying topical ontology of keywords and sentiment techniques to learners' comments on top-rated courses on Coursera.org. Knowledge-seeking courses stress learning concepts that provide insights, assessed with quizzes and final exams and they are driven by course design and materials. Skill-seeking courses require learning techniques through labs, projects, or assignments and they are driven by instructors. Sentence-level sentiment analysis was used for sentiment classification to detect the polarity of a sentence based on identified topics and to perform a fine-grained analysis of opposite sentiments on different topics. A learner review is chunked into multiple sentences and extracted the topic and sentiment for each sentence. To deliver pedagogical cognitive and affective skills, conversational agents based on reinforcement learning (RL) and sentiment analysis were proposed by Feidakis et al. (2019). In that study, the learning task is expressed in a goal-oriented

way to RL agents. The agents learn how to dialogue with students based on supervised learning of annotated datasets and the learning is through a trial-and-error method based on the incoming rewards for each action.

4.3. Decision making

Educational data mining assists educational institutions in measuring the teaching and learning process and improving their student recruitment and retention policies. Hussain et al. (2022) proposed a decision support system based on a multi-layered Aspect2Labels (A2L) approach. It is a three-layered topic modelling approach, the first layer extracts high-level aspects preserving semantic information, the second layer extracts low-level aspects associated with high-level aspects and the third layer was used for sentiment orientation. The authors annotated unlabelled student comments using the proposed A2L approach and achieved 91.3% accuracy. Two techniques Variable Stopwords Filtering (VSF) and Variable Global Feature Selection Scheme (VGFSS) proposed to improve the performance of the classifier. The authors used five machine learning models for sentiments analysis and Support Vector Machine with the linear kernel (SVM) and Artificial Neural Network (ANN) algorithms performed better than other models with an accuracy of 97% and 93% respectively.

4.4. Assessment evaluation

Distinctive understanding of a concept and the polarity of the words used by students shape the quality and presentation of a course assignment or assessment (Janda et al., 2019). In Janda et al. (2019), the authors proposed to use the semantic similarity of sentences of text along with their sentiments, and sentence syntax for automated essay evaluation. For example, if a student writing an argumentative and persuasive essay, he/she has to prove their view on the topic, check the tone, and the way of textual sentiment expression. The authors used Vader, a lexical approach to extract the polarity for each sentiment of the words used in the essay. Gkontzis et al. (2020) analysed the activity, polarity, and emotions of students and tutors to predict students' grades. The authors used six moodle datasets of postgraduate programs at the Hellenic Open University. The results show that extracting the polarity and emotions of students can provide a better prediction of students' performance. Furthermore, variables related to tutors also contribute to student performance. The study has compared actual and predicted grades and investigated which students had used third-party services to do their assignments. The research works outlined in Table 3 present the AI algorithms that can be adapted to education for its applications.

4.5. Sentiment analysis tools in education

Educational institutions have been working on building student feedback mining tools to analyse their students' opinions on existing learning infrastructures, teachers, and course contents. With the analysis, the institutions' intuition is to evaluate their teaching practices and learning management systems based on suggestions in student feedback. A few of those research and application works are enclosed in Table 2. AI algorithms are explicitly used in most of the works. Advanced AI methodologies like reinforcement learning would enable the education sector to personalise student learning experiences and support them.

5. Challenges

5.1. Negation handling

In sentiment analysis, the textual feedback received from students or instructors gets preprocessed to remove stopwords and prepare for sentiment labelling. The stopwords are low-level information that does not

Table 2
Sentiment analysis tools in educational Domain.

Tool	Application	Algorithms	Performance	Reference
Suggestion Extraction Tool	Topic extraction sentiment analysis suggestion extraction to improve instructor teaching	GLM, SVM, CIT, DT	DT with precision 0.802, recall 0.775, and F-score 0.781	Gottipati et al. (2018b), Pyasi et al. (2018)
Student Textual Feedback Analysis Tool	Topic extraction, sentiment analysis	LDA, TextBlob, Polarity Analyser	TextBlob improved version with F-Score 0.92	Gottipati et al. (2018a)
Automated Discussion Analysis (ADA) framework	Question Answering Systems, Recommendation Systems, Summarisation Systems, Opinion Mining Systems, Information retrieval systems	Clustering techniques		Gottipati et al. (2020)
SETSUM: Summarisation and Visualisation of Student Evaluations of Teaching	Aspect Annotation, Sentiment Analysis, Extractive Summarisation	LexRank, LDA, manual sentiment annotation		Hu et al. (2022)
Palaute tool	Topic modelling, emotional analysis	LDA, Syuzhet		Grönberg et al. (2020), Gronberg et al. (2021)
AI-enabled adaptive learning system	Generate personalised quizzes for each learner	Reinforcement Learning, LDA	Net Promoter Score 38	Tan et al. (2022)

Table 3
SA impact on education.

Educational Applications	References	Algorithms
Learning and Teaching Systems Evaluation	Lalata et al. (2019), Roaring et al. (2022), Pramod et al. (2022) Onan (2020)	Ensemble learning Association rule mining RF word2vec, fastText, GloVe, LSTM
Enhance Pedagogical Concepts	Yan et al. (2019) Li et al. (2022a) Feidakis et al. (2019)	LDA, LIWC Topic Ontology Reinforcement learning
Decision Making	Hussain et al. (2022)	VSF, VGFSS, SVM, ANN
Assessment Evaluation	Janda et al. (2019) Gkontzis et al. (2020)	Vader Additive regression, RF, SVN, KNN

add much value to sentiment analysis. By eliminating the stopwords, the analysis can focus on important information such as noun, verb, and adjective words of the textual feedback. However, the negation words such as never, not, none, do not, and so on in negative statements are preprocessed and these auxiliary verbs and adverbs are removed. For example, the negative statement "Lecturer never discuss practical knowledge concepts" passed through preprocessing step would remove "never" and the text analytic models get trained on inverse sentiment.

In the Aspect2Labels framework proposed by Hussain et al. (2022), the authors reversed the opinion score to −1 whenever negative words not, never, nothing, neither, or none are encountered in student feedback. The authors also dealt with blind negation words such as need, needed, require, and required by assigning a negative polarity score to the sentence. Singh and Paul (2021) proposed focusing on cue and scope could find the polarity shift in a sentence. A deep learning model LSTM was proposed to learn negation features of pre-annotated ConanDoyle corpus with negation information. The LSTM model identifies negation cues and BiLSTM extracts the relationship between the extracted cues and other words. The proposed approach was compared with traditional machine learning methods SVM, Hidden Markov Model, and CRF models. BiLSTM model outperformed other models with an F-Score of 93.34% in negation handling tasks. Similarly, Gupta and Joshi (2021) proposed a feature-based Twitter sentiment analysis (TSA) framework based on cues and the scope of sentences for negation handling. The research work is an improvement to the work proposed by Mohammad et al. (2013) in negation handling in which the authors

argue negation presence not always led to the negative polarity of a sentence.

5.2. Opinion spam detection

In the education domain, opinion spam detection is critical as students might use third-party services for assessment items and write fake feedback to enrolled courses at the end of the semester. The opinion spam detection concept has been widely explored in product reviews where people get incentives for writing fake reviews (Yuan et al., 2019).

Saumya and Singh (2018) discussed three challenges in the supervised approach to detect spam comments. They are manual labelling of comments into spam and non-spam, data imbalance problems due to scarcity of spam comments, and expensive computation to retrieve similarity among the spam comments. The authors proposed a system to overcome the labelling challenge and automate spam comments annotation. In that study, RF, SVM and Gradient Boosting classifiers were used to classify comments into spam or non-spam classes. Two oversampling techniques Synthetic Minority Over-sampling Technique (SMOTE) (Chawla et al., 2002) and Adaptive Synthetic (ADASYN) (He et al., 2008) to balance the datasets with the same number of spam and non-spam records. Deep learning methods were adopted for opinion spam detection by Bathla and Kumar (2021). Convolutional Neural Networks (CNN) model was used to classify the spam comments with the GloVe embedding technique. The results show that the deep learning method outperformed traditional machine learning methods such

Table 4
Sentiment analysis challenges in education.

Challenges	Trends	References
Negation Handling	Negative polarity scoring	Hussain et al. (2022), Singh and Paul (2021), Gupta and Joshi (2021), Mohammad et al. (2013)
Opinion Spam Detection	SMOTE, ADASYN, Deep Learning	Yuan et al. (2019), Saumya and Singh (2018), Bathla and Kumar (2021)
Multi-Polarity	Dictionary-based sentiment knowledge	Wei et al. (2020), Wang et al. (2018)
Polysemous Words	ACWE, Word Sense Disambiguation	Logacheva et al. (2020), Li et al. (2021), Jia et al. (2021), Yenicelik et al. (2020)

as SVM, KNN, Logistic Regression, Naive Bayes, MLP and other deep learning methods LSTM.

5.3. Multi-polarity

Multi-polarity refers to several contrasting ideas and emotions towards a product or service. In the education domain, student feedback might possess such multi-polarity challenges in which students might have positive emotions towards one aspect of a course and negative opinions on other aspects. In such situations, the sentiment analysis of the student feedback would be challenging as it might be ambiguous or misleading for sentiment annotation techniques to label the text.

Wei et al. (2020) proposed a BiLSTM model with multi-polarity orthogonal attention for implicit sentiment analysis. Dictionary-based sentiment knowledge such as neutral, positive, and negative embeddings are trained in the polarity layer with an orthogonal attention mechanism. Wang et al. (2018) present a semantic analysis model (SMA) framework to track the emotional tendencies of students based on comments, forum posts, and homework completion on the MOOC platforms. In that study, the SMA framework detects learners' multi-polarity sentiments by adopting the NRC lexicon for sentiment quantification.

5.4. Polysemous words

Polysemous refers to a word that has several meanings. This would cause word sense disambiguation (WSD) (Logacheva et al., 2020) to AI models and sentiment annotations while sentiment analysis. The word embedding concept in NLP represents each word by a vector to capture the semantic correlation between words. The research community tried to overcome the challenge by generating a small number of vectors for each word but it is difficult to estimate the number of meanings for each word.

Li et al. (2021) proposed a novel adaptive cross-contextual word embedding (ACWE) method to capture polysemous words in different contexts using topic modelling by defining a latent interpretable semantic space. A global word embedding was obtained based on an unsupervised cross-contextual probabilistic word embedding model and represents each word in the unified latent semantic space. An adaptive cross-contextual word embedding process is then defined to learn local word embedding for each polysemous word in different contexts. Jia et al. (2021) proposed an effective topical word embedding (TWE)-based WSD method to generate topical word vectors for each word under each topic. The authors used LDA to extract high-quality topic models, and then generate contextual vectors generated for an ambiguous word to exploit the context. Based on the topic models and contextual vectors, topical word vectors were generated. Yenicelik et al. (2020) used the BERT model to produce embedding vectors using quantitative analysis of linear separability and cluster organisation. The authors trained a simple linear classifier on top of the BERT embeddings to predict the semantic class and report accuracy. The challenges and trends in adopting sentiment analysis to education have been discussed and outlined in the research works in Table 4.

6. Future directions of SA in education

In this study, a comprehensive review of sentiment analysis in the educational context is conducted. The impact of sentiment analysis in the educational domain and the challenges in adopting sentiment analysis pave a path to future research directions. Sentiment analysis can assist educational institutions to evaluate their learning and teaching practices by extracting sentiment from student opinions. Advancements in transformers and deep learning techniques transform sentiment analysis applications in various fields.

The future direction should be focused on education-based sentiment annotation techniques. In current trends, the research community has adopted existing unsupervised sentiment annotation techniques such as corpus-based and lexicon-based approaches. However, there will be a huge semantic difference in student feedback on the education context and the language or words used in the unsupervised annotation techniques. Focusing on education-based unsupervised annotation techniques can overcome lexical ambiguity, and data labelling challenges and understand the core factors of educational context. To build an education-based annotation, a corpus of education documents has to be labelled manually in a dictionary-based approach. Based on the education-based annotation technique, the reinforcement learning process can be enhanced to provide improved support for goal-oriented learning tasks discussed in Feidakis et al. (2019).

Topic ontology discussed in Li et al. (2022a), Tao et al. (2008) can be further explored to build an educational knowledge base and focus on defining relationships among multimodal data collected in student feedback. With knowledge-based systems enhance the educational data mining process and overcome challenges such as sarcasm, negation handling, and ambiguity. Fine-tuning of aspect-based and entity sentiment analysis can be achieved by annotating entities and aspects that need to be emphasised in an educational domain.

7. Conclusions

In this review article, a comprehensive impact of sentiment analysis and opinion mining on educational data and challenges are explored. Though, sentiment analysis is the simple detection of the sentiment orientation of students towards course content or teaching, fine-grained analysis in education would require other NLP techniques like topic modelling. To analyse a large number of students' opinions, the usage of AI methodologies such as machine learning, deep learning, and transformers is inevitable.

This study reviews the role of sentiment analysis in the enhancement of educational procedures such as decision-making, pedagogical concepts, and evaluation systems. The pandemic situation expedites the process of online education all around the world and enlarges the MOOCs industry. This review article covered studies related to both offline and online modes of education. The research community has been working on aspect-level sentiment analysis for a fine-grained analysis at the topic level and to extract the sentiments of students. Dictionary and corpus-based techniques are unsupervised sentiment annotation techniques discussed in the study. Throughout the review, most of the

researchers used AI methods like traditional machine learning, deep learning, or BERT models to analyse, predict and classify the opinions of a large group of students. Yet, there are challenges in sentiment analysis that need attention to improvise the existing student feedback and mining techniques.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

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. http://dx.doi.org/10.1007/s10462-021-09958-2.

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).

Alqasemi, F., Abdelwahab, A., Abdelkader, H., 2017. Constructing automatic domain-specific 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.

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.

Bansal, M., Verma, S., Vig, K., Kakran, K., 2022. Opinion mining from student feedback data using supervised learning algorithms. In: Third International Conference on Image Processing and Capsule Networks. Springer International Publishing, pp. 411–418. http://dx.doi.org/10.1007/978-3-031-12413-6_32.

Bathla, G., Kumar, A., 2021. Opinion spam detection using deep learning. In: 2021 8th International Conference on Signal Processing and Integrated Networks. SPIN, IEEE, http://dx.doi.org/10.1109/spin52536.2021.9566048.

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.

Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent dirichlet allocation. J. Mach. Learn. Res. 3 (Jan), 993–1022.

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.

Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. SMOTE: Synthetic minority over-sampling technique. J. Artificial Intelligence Res. 16, 321–357. http://dx.doi.org/10.1613/jair.953.

D'Aniello, G., Gaeta, M., Rocca, I.L., 2022. Knowmis-ABSA: an overview and a reference model for applications of sentiment analysis and aspect-based sentiment analysis. Artif. Intell. Rev. 55 (7), 5543–5574. http://dx.doi.org/10.1007/s10462-021-10134-9.

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/fie49875.2021.9637330.

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, http://dx.doi.org/10.1145/3194932.3194935.

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, http://dx.doi.org/10.26220/UNE.2987, https://pasithee.library.upatras.gr/ejupUNESCOchair/article/view/2987.

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

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. http://dx.doi.org/10.1007/978-3-030-82199-9_2.

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. http://dx.doi.org/10.1016/j.tsc.2019.100622.

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.

Feidakis, M., Kasnesis, P., Giatraki, E., Giannousis, C., Patrikakis, C.Z., Monachelis, P., 2019. Building pedagogical conversational agents, affectively correct. In: CSEDU (1). pp. 100–107.

Gkontzis, A.F., Kotsiantis, S., Kalles, D., Panagiotakopoulos, C.T., Verykios, V.S., 2020. Polarity, emotions and online activity of students and tutors as features in predicting grades. Intell. Decis. Technol. 14 (3), 409–436. http://dx.doi.org/10.3233/idt-190137.

Gottipati, S., Shankararaman, V., Gokarn, M.N., 2020. Automated: discussion analysis - framework for knowledge analysis from class discussions. In: 2020 IEEE Frontiers in Education Conference. FIE, IEEE, http://dx.doi.org/10.1109/fie44824.2020.9273819.

Gottipati, S., Shankararaman, V., Lin, J., 2018a. Latent dirichlet allocation for textual student feedback analysis.

Gottipati, S., Shankararaman, V., Lin, J.R., 2018b. Text analytics approach to extract course improvement suggestions from students' feedback. Res. Pract. Technol. Enhanced Learn. 13 (1), http://dx.doi.org/10.1186/s41039-018-0073-0.

Grönberg, N., Knutas, A., Hynninen, T., Hujala, M., 2020. An online tool for analyzing written student feedback. In: Koli Calling '20: Proceedings of the 20th Koli Calling International Conference on Computing Education Research. ACM, http://dx.doi.org/10.1145/3428029.3428565.

Gronberg, N., Knutas, A., Hynninen, T., Hujala, M., 2021. Palaute: An online text mining tool for analyzing written student course feedback. IEEE Access 9, 134518–134529. http://dx.doi.org/10.1109/access.2021.3116425.

Gupta, I., Joshi, N., 2021. Feature-based twitter sentiment analysis with improved negation handling. IEEE Trans. Comput. Soc. Syst. 8 (4), 917–927. http://dx.doi.org/10.1109/tcss.2021.3069413.

Han, Z., Wu, J., Huang, C., Huang, Q., Zhao, M., 2019. A review on sentiment discovery and analysis of educational big-data. WIREs Data Min. Knowl. Discov. 10 (1), http://dx.doi.org/10.1002/widm.1328.

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.

He, H., Bai, Y., Garcia, E.A., Li, S., 2008. ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In: 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence). IEEE, http://dx.doi.org/10.1109/ijcnn.2008.4633969.

Hu, Y., Zhang, S., Sathy, V., Panter, A., Bansal, M., 2022. Setsum: Summarization and visualization of student evaluations of teaching. arXiv preprint arXiv:2207.03640.

Hussain, S., Ayoub, M., Jilani, G., Yu, Y., Khan, A., Wahid, J.A., Butt, M.F.A., Yang, G., Moller, D.P., Weiyan, H., 2022. Aspect2labels: A novelistic decision support system for higher educational institutions by using multi-layer topic modelling approach. Expert Syst. Appl. 209, 118119. http://dx.doi.org/10.1016/j.eswa.2022.118119.

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.

Janda, H.K., Pawar, A., Du, S., Mago, V., 2019. Syntactic, semantic and sentiment analysis: The joint effect on automated essay evaluation. IEEE Access 7, 108486–108503. http://dx.doi.org/10.1109/access.2019.2933354.

Jia, L., Tang, J., Li, M., You, J., Ding, J., Chen, Y., 2021. TWE-WSD: An effective topical word embedding based word sense disambiguation. CAAI Trans. Intell. Technol. 6 (1), 72–79. http://dx.doi.org/10.1049/cit2.12006.

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.

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.

Kuleto, V., Ilić, M., Dumangiu, M., Ranković, M., Martins, O.M.D., Păun, D., Mihoreanu, L., 2021. Exploring opportunities and challenges of artificial intelligence and machine learning in higher education institutions. Sustainability 13 (18), 10424. http://dx.doi.org/10.3390/su131810424.

Lalata, J.P., Gerardo, B., Medina, R., 2019. A sentiment analysis model for faculty comment evaluation using ensemble machine learning algorithms. In: Proceedings of the 2019 International Conference on Big Data Engineering. ACM, http://dx.doi.org/10.1145/3341620.3341638.

Lang, P., 1980. Behavioral treatment and bio-behavioral assessment: Computer applications. Technol. Mental Health Care Deliv. Syst. 119–137.

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.

Li, L., Johnson, J., Aarhus, W., Shah, D., 2022a. Key factors in MOOC pedagogy based on NLP sentiment analysis of learner reviews: What makes a hit. Comput. Educ. 176, 104354. http://dx.doi.org/10.1016/j.compedu.2021.104354.

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), http://dx.doi.org/10.1609/aaai.v31i1.11016.

Li, S., Pan, R., Luo, H., Liu, X., Zhao, G., 2021. Adaptive cross-contextual word embedding for word polysemy with unsupervised topic modeling. Knowl.-Based Syst. 218, 106827. http://dx.doi.org/10.1016/j.knosys.2021.106827.

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.

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.

Ligthart, A., Catal, C., Tekinerdogan, B., 2021. Systematic reviews in sentiment analysis: a tertiary study. Artif. Intell. Rev. 54 (7), 4997–5053. http://dx.doi.org/10.1007/s10462-021-09973-3.

Liu, B., 2012. Sentiment Analysis and Opinion Mining. Springer International Publishing, http://dx.doi.org/10.1007/978-3-031-02145-9.

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. http://dx.doi.org/10.1080/10494820.2019.1610449.

Logacheva, V., Teslenko, D., Shelmanov, A., Remus, S., Ustalov, D., Kutuzov, A., Artemova, E., Biemann, C., Ponzetto, S.P., Panchenko, A., 2020. Word sense disambiguation for 158 languages using word embeddings only. http://dx.doi.org/10.48550/ARXIV.2003.06651, https://arxiv.org/abs/2003.06651.

Loria, S., 2018. Textblob documentation. Release 0.15 2.

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/ecti-con49241.2020.9158328.

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.

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.

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. http://dx.doi.org/10.1177/00915521211061416.

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

Mohammad, S.M., Turney, P.D., 2013. Crowdsourcing a word–emotion association lexicon. Comput. Intell. 29 (3), 436–465.

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.

Neviarouskaya, A., Prendinger, H., Ishizuka, M., 2009b. Sentiful: Generating a reliable lexicon for sentiment analysis. In: 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops. IEEE, pp. 1–6.

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

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. http://dx.doi.org/10.1108/el-06-2019-0140.

Ömer Osmanoğlu, U., Atak, O.N., Çağlar, 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.

Onan, A., 2020. Sentiment analysis on massive open online course evaluations: A text mining and deep learning approach. Comput. Appl. Eng. Educ. 29 (3), 572–589. http://dx.doi.org/10.1002/cae.22253.

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

Pennebaker, J.W., Boyd, R.L., Jordan, K., Blackburn, K., 2015. The development and psychometric properties of liwc2015. Tech. rep..

Pramod, D., Bharathi, S.V., Raman, R., 2022. Faculty effectiveness prediction using machine learning and text analytics. In: 2022 IEEE Technology and Engineering Management Conference (TEMSCON EUROPE). IEEE, http://dx.doi.org/10.1109/temsconeurope54743.2022.9801997.

Pyasi, S., Gottipati, S., Shankararaman, V., 2018. SUFAT - an analytics tool for gaining insights from student feedback comments. In: 2018 IEEE Frontiers in Education Conference. FIE, IEEE, http://dx.doi.org/10.1109/fie.2018.8658457.

Roaring, B.F., Patacsil, F.F., Parrone, J.M., 2022. Analyzing pangasinan state university student's FacultyTeaching performance rating using text mining technique. Wseas Trans. Inf. Sci. Appl. 19, 161–170. http://dx.doi.org/10.37394/23209.2022.19.16.

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.

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.

Saumya, S., Singh, J.P., 2018. Detection of spam reviews: a sentiment analysis approach. CSI Trans. ICT 6 (2), 137–148. http://dx.doi.org/10.1007/s40012-018-0193-0.

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. http://dx.doi.org/10.1016/j.knosys.2022.109929.

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. http://dx.doi.org/10.1109/access.2022.3177752.

Sharma, N., Jain, V., 2020. Evaluation and summarization of student feedback using sentiment analysis. In: Advances in Intelligent Systems and Computing. Springer Singapore, pp. 385–396. http://dx.doi.org/10.1007/978-981-15-3383-9_35.

Singh, P.K., Paul, S., 2021. Deep learning approach for negation handling in sentiment analysis. IEEE Access 9, 102579–102592. http://dx.doi.org/10.1109/access.2021.3095412.

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, http://dx.doi.org/10.1109/icici.2017.8365231.

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

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.

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.

Tan, K.W., Lo, S.L., Ouh, E.L., Neo, W.L.L.W., 2022. Ai-enabled adaptive learning using automated topic alignment and doubt detection.

Tao, X., Li, Y., Nayak, R., 2008. A knowledge retrieval model using ontology mining and user profiling. Integr. Comput.-Aided Eng. 15 (4), 313–329. http://dx.doi.org/10.3233/ica-2008-15404.

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. http://dx.doi.org/10.3390/s21030776.

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. http://dx.doi.org/10.1007/978-3-319-49586-6_59.

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.

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

Wang, L., Hu, G., Zhou, T., 2018. Semantic analysis of learners' emotional tendencies on online MOOC education. Sustainability 10 (6), 1921. http://dx.doi.org/10.3390/su10061921.

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.

Wei, J., Liao, J., Yang, Z., Wang, S., Zhao, Q., 2020. Bilstm with multi-polarity orthogonal attention for implicit sentiment analysis. Neurocomputing 383, 165–173. http://dx.doi.org/10.1016/j.neucom.2019.11.

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.

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

Yan, W., Dowell, N., Holman, C., Welsh, S.S., Choi, H., Brooks, C., 2019. Exploring learner engagement patterns in teach-outs using topic, sentiment and on-topicness to reflect on pedagogy. In: Proceedings Ofthe 9th International Conference on Learning Analytics & Knowledge. ACM, http://dx.doi.org/10.1145/3303772.3303836.

Yang, R., 2021. Machine learning and deep learning for sentiment analysis over students' reviews: An overview study. http://dx.doi.org/10.20944/preprints202102.0108.v1.

Yenicelik, D., Schmidt, F., Kilcher, Y., 2020. How does BERT capture semantics? a closer look at polysemous words. In: Proceedings of the Third BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP. Association for Computational Linguistics, http://dx.doi.org/10.18653/v1/2020.blackboxnlp-1.15.

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.

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.

Yuan, C., Zhou, W., Ma, Q., Lv, S., Han, J., Hu, S., 2019. Learning review representations from user and product level information for spam detection. http://dx.doi.org/10.48550/ARXIV.1909.04455, https://arxiv.org/abs/1909.04455.

Zhang, Y., Wang, J., Zhang, X., 2021. Conciseness is better: Recurrent attention LSTM model for document-level sentiment analysis. Neurocomputing 462, 101–112. http://dx.doi.org/10.1016/j.neucom.2021.07.072.

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. http://dx.doi.org/10.1016/j.ipm.2021.102656.

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, http://dx.doi.org/10.1109/cscwd.2013.6581022.

Zhou, J., min Ye, J., 2020. Sentiment analysis in education research: a review of journal publications. Interact. Learn. Environ. 1–13. http://dx.doi.org/10.1080/10494820.2020.1826985.

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. http://dx.doi.org/10.1016/j.jbusres.2020.11.007.