International Journal of Data and Network Science 9 (2025) 253–266

Contents lists available at GrowingScience

International Journal of Data and Network Science

homepage: www.GrowingScience.com/ijds

Understanding students' sentiment from feedback with a new feature selection and semantics networks

Tran Anh Tuan[ab*], Dao Thi Thanh Loan[c] and Nichnan Kittiphattanabawon[a]

[a]*School of Informatics, Walailak University, Nakhon Si Thammarat, Thailand*
[b]*Informatics Innovation Center of Excellence (IICE), Walailak University, Nakhon Si Thammarat, Thailand*
[c]*Dak Lak College of Pedagogy, Dak Lak, Vietnam*

C H R O N I C L E

Article history:
Received: May 3, 2024
Received in revised format: May 28, 2024
Accepted: July 20, 2024
Available online: July 20, 2024

Keywords:
Students' sentiment
Students' feedback
Feature selection
Concatenated feature
Semantics network
Machine learning

A B S T R A C T

Sentiment analysis of students' feedback using machine learning algorithms has emerged as a valuable tool for understanding students' sentiments and improving educational outcomes. Currently, existing systems use frequency-based methods for feature selection (e.g., Term Frequency-Inverse Document Frequency (TF-IDF) and Bag of Words (BoW)) not to capture the subtleties of emotions expressed in student feedback and do not provide insights into the specific concerns of students via topics or themes. In this study, we propose the Student Sentiment from Feedback (SSF) framework, which includes four main procedures: pre-processing, feature selection, classification, and theme finding. The SSF framework classifies student sentiments and subsequently groups feedback into themes using semantic networks based on word co-occurrence. Our innovative feature selection approach combines TF-IDF with sentiment-based features derived from SentiWordNet and intensifiers, creating a robust feature vector that enhances the dataset's richness and improves classification accuracy and robustness. In the experiments, we utilize a public dataset from Kaggle, applying our proposed method and various machine learning models (e.g., k-nearest neighbor, decision tree, random forest, multilayer perceptron, support vector machine, gradient boosting, and extreme gradient boosting). The experimental results show that our concatenated features achieve the highest accuracy across all machine learning models (greater than 0.82). Our study demonstrates the efficacy of this hybrid feature selection method, contributing to better understanding and decision-making in educational settings.

1. Introduction

In the era of digital transformation, most domains, including education, have transitioned to digital platforms to support, drive, and balance business objectives. Consequently, understanding what internet users feel and expect about service and product quality has become crucial (Shenify, 2024; Xingyi & Adnan, 2024). In education, universities and institutions increasingly leverage technology to improve the learning experience and academic outcomes. A critical aspect of this transformation involves systematically collecting and analyzing students' feedback to understand their sentiments and perceptions regarding various aspects of their educational journey. Students' feedback, often captured through online surveys, course evaluations, and learning management systems, provides invaluable insights into their experiences, preferences, and challenges. Manually analyzing such feedback can be time-consuming and labor-intensive. However, advancements in artificial intelligence (AI), particularly machine learning (ML), enable the analysis of large-scale student feedback, helping educators gain a deep understanding of the sentiments expressed (Alamin et al., 2024; Louati et al., 2023; Pekrun et al., 2023; Sohel et al., 2023).

Sentiment analysis, defined as "the study of computation expressed in terms of opinions, sentiments, or emotions by text" (Bing, 2015; Hu & Liu, 2004), is essential for enhancing educational quality and fostering a supportive learning environment.

* Corresponding author.
E-mail address tran.tu@wu.ac.th (T. A. Tuan)

ISSN 2561-8156 (Online) - ISSN 2561-8148 (Print)

doi: 10.5267/j.ijdns.2024.7.010

Understanding students' sentiments from feedback is crucial for enhancing the quality of education and fostering a supportive, engaging learning environment. ML methods have proven effective in classifying text into various sentiment categories - positive, negative, and neutral - across multiple domains, including education (Abualhaj et al., 2024; Ahn & Kang, 2018; AL-Akhras et al., 2024; Alsubaie & Aldoukhi, 2024; Altun et al., 2022; Jain et al., 2022; Loan & Tuan, 2024; Mercha & Benbrahim, 2023; Nafea, 2018; Nasim et al., 2017; Pekrun et al., 2017; Ramasamy et al., 2021; Shenify, 2024; Surya & Subbulakshmi, 2019; Tran et al., 2021b; Tripathy et al., 2015; Ward et al., 2024; Xiao et al., 2023).

Achieving better performance is a primary goal of ML models, and feature selection methods are crucial in enhancing this performance. Previous studies have explored various feature selection techniques to improve ML model outcomes. Popular methods for feature selections in the previous works were Term Frequency-Inverse Document Frequency (TF-IDF) (Al-Dwish & Aljohani, 2024; Bensba et al., 2022; Bhardwaj & Srivastava, 2021; Esparza et al., 2018; Gutiérrez et al., 2018; Imran et al., 2022; Kathuria et al., 2023; Louati et al., 2023; Nasrulloh et al., 2019; Rakhmanov, 2020; Shenify, 2024; Sohel et al., 2023; Tamrakar & others, 2021) and Bag of Words (BoW) (Altrabsheh et al., 2015; Dsouza et al., 2019; Giang et al., 2020; Melba Rosalind & Suguna, 2021; Pacol & Palaoag, 2021; Rajesh & Suseendran, 2020; Ullah, 2016). There were a few studies conducted with other methods (e.g., lexicon-based, hybrid-based) (Dalal et al., 2015; Kabir et al., 2024; Lin et al., 2019; Nasim et al., 2017). However, these popular methods do not capture the nuanced emotions expressed in students' feedback. By leveraging the strengths of both frequency and sentiment information, our method aims to achieve a more precise understanding of students' feedback, ultimately leading to better insights and decision-making in educational settings. Hence, we propose a framework for student sentiment from feedback (SSF) comprising four main procedures: pre-processing, feature selection, classification, and theme finding. In the feature selection phase of the SSF framework, we propose a novel approach that concatenates both frequency- and sentiment-based features for enhanced classification performance. By combining TF-IDF, which quantifies the importance of words, with sentiment-based features derived using SentiWordNet and a list of intensifiers, we aim to capture a more comprehensive sentiment of student feedback. The concatenated feature vector enriches the dataset and improves the accuracy and robustness of classification models used in ML models. In this study, we demonstrate the versatility of our approach by employing various ML models, such as K-nearest neighbor (KNN), Decision Tree (DT), Random Forest (RF), Multilayer Perceptron (MLP), Support Vector Machine (SVM), Gradient Boosting, and Extreme Gradient Boosting (XGBoost). Finally, themes of three sentiment levels from students' feedback are determined by semantics networks based on word co-occurrence. The SSF framework could provide educators with deeper insights into students' emotional responses and help tailor educational strategies to better meet students' needs.

2. Related Work

In this section, previous studies related to sentiment classification with various ML and feature seletion methods are discussed.

TF-IDF is a statistical method used to assess the importance of a word in a document compared to a collection of documents. This paragraph discusses studies using TF-IDF and ML algorithms to classify sentiment. Gutiérrez et al. (2018) presented a model to analyze student reviews of teacher performance using SVM and TF-IDF from Twitter. Esparza et al. (2018) also suggested a model for analyzing student reviews of teacher performance using SVM and TF-IDF to classify comments as positive, negative, or neutral. Nasrulloh et al. (2019) introduced a framework for an educational feedback system that combined sentiment analysis and statistical methods to analyze student evaluations. The authors used TF-IDF with SVM to classify sentiment. Rakhmanov (2020) applied five ML models (e.g., NB, RF, SVM, ANN, and Gradient boosting) with TF-IDF to classify sentiment. With 5-fold cross-validation, the RF method achieved the highest prediction accuracy with optimal training time. Bhardwaj & Srivastava (2021) used ML models (e.g., NB, SVM, Logistic Regression (LR)) with TF-IDF to classify sentiment for product reviews. The experimental results showed that the NB method got the highest accuracy. Tamrakar et al. (2021) utilized LR, SVM, NB, and DT with two methods of feature selection (TF-IDF and BoW). The SVM model performed well with both methods. Bensba et al. (2022) suggested a model (SVM, KNN, and deep neural network) with the TF-IDF method to classify sentiment from online students, with SVM achieving the best accuracy. Imran et al. (2022) created an ensemble model of five ML models (e.g., DT, RF, LR, support vector classifier, and AdaBoost) with TF-IDF to classify the sentiment of online students during COVID-19, with the ensemble model outperforming individual ML models. Kathuria et al. (2023) introduced a model using various ML algorithms (e.g., SVM, NB, LR, RF, DT, and KNN) and TF-IDF to classify students' feedback, finding RF to have the best accuracy. Louati et al. (2023) suggested a model using the SVM algorithm and TF-IDF for classifying sentiment from student reviews studying Arabic courses. Sohel et al. (2023) classified the sentiment of reviews of Coursera courses using ML algorithms (e.g., SVM, NB, LR, DT, RF, AdaBoost) and TF-IDF, with LR achieving the highest accuracy. Al-Dwish & Aljohani (2024) incorporated TF-IDF and ML algorithms (SVM, LR, KNN, XGBoost) for classifying sentiment from online courses, with LR being the most accurate. Shenify (2024) used NB and SVM algorithms with TF-IDF for sentiment classification from Twitter in the e-commerce domain, with SVM achieving the best results.

BoW with uni-, bi-, and n-grams represents text by the frequency of words in a document, ignoring grammar and word order but maintaining multiplicity. This paragraph discusses studies using BoW and ML algorithms to classify sentiment. Altrabsheh et al. (2015)) employed uni-grams and ML models to detect the sentiment of sarcasm from students' feedback, with RB being the most accurate. Ullah (2016) utilized ML algorithms (e.g., NB, SVM, maximum entropy (ME)) and n-grams for analyzing students' sentiment from feedback, with SVM achieving the best results in tri-grams. Dsouza et al. (2019) used for sentimental analysis of student feedback using three machine learning algorithms with BoW, with NB achieving the best model. Giang et al. (2020) conducted sentiment analysis of student feedback using three machine learning algorithms with BoW, with NB achieving the best performance. Rajesh & Suseendran (2020) presented an SVM algorithm and n-grams to classify sentiment from e-learning reviews. Melba

Rosalind & Suguna (2021) created an ensemble model of five ML models (e.g., SVM, RF, and XGBoost) with uni-grams to classify the sentiment of online courses, with the ensemble model outperforming individual ML models. Pacol & Palaoag (2021) classified students' sentiments using various ML models (e.g., NB, SVM, and LR), with RF and n-grams achieving the best results.

This paragraph discusses studies using other feature selection methods (e.g., lexicon-based and hybrid approach) and ML algorithms to classify sentiment. Dalal et al. (2015) examined ML models (e.g., NB, ME, and SVM) with various feature selection methods (n-grams, lexicon-based) for classifying students' sentiments from feedback. NB with lexicon-based feature selection (SentiWordNet) achieved the best performance. Lin et al. (2019) developed a model using ML models (e.g., Gradient Boost and DT) with SentiWordNet for analyzing students' sentiments from feedback. Nasim et al. (2017) presented ML models (e.g., RF and SVM) with a hybrid approach integrating TF-IDF and lexicon-based methods to analyze students' sentiments, with RF emerging as the best model. Kabir et al. (2024) applied various ML models (e.g., NB, LR, SVM, and Gradient Boosting) and lexical features for evaluating students' feedback on teaching, with SVM performing the best.

In summary, there is a gap in research regarding capturing a more comprehensive sentiment of student feedback. Existing studies with frequency-based features, including TF-IDF and BoW, cannot adequately handle words with multiple meanings or identify the subtleties of students' feelings expressed in their feedback. Our study's contributions are as follows: 1) the proposed SSF framework to understand students' sentiment from feedback with themes using ML models and semantic networks, and 2) the proposed method of feature selection concatenated frequency- and sentiment-based features.

3. Proposed Method

The proposed framework for student sentiment from feedback (SSF) is illustrated in Fig. 1. The SSF framework comprises four main procedures: 1) pre-processing, 2) feature selection, 3) classification, and 4) theme finding. The input to the framework is students' feedback, while the output is the themes grouped to the student's sentiments, categorized into negative, neutral, and positive sentiments.

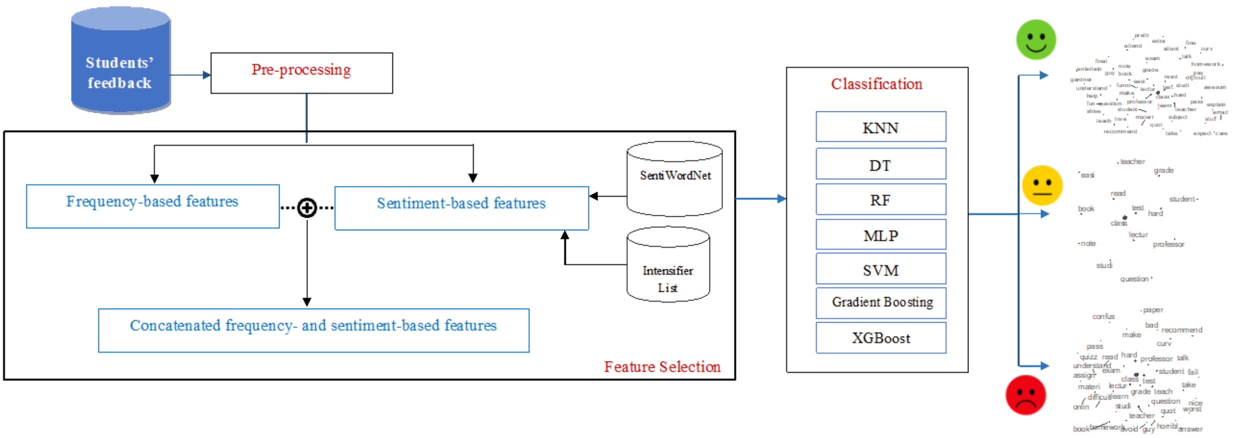

Fig. 1. An architecture of Student Sentiment from Feedback (SSF) framework

3.1. Pre-processing

The preprocessing procedure involves normalizing, segmenting words, removing stopwords, and tagging part-of-speech (POS) for each student feedback in the dataset. Additionally, this procedure determines the dependency among POS using the spaCy library (Spacy, 2024).

3.2. Feature Selection

The procedure aims to introduce a proposed hybrid approach of feature selection based on frequency and sentiment. The method will be thoroughly discussed in the upcoming sections.

3.2.1. Frequency-based features

Frequency-based feature selection methods aim to identify features (terms or words) most relevant to a particular feedback or class. The TF-IDF weights are extensively utilized and highly regarded techniques within the realm of text mining (Aizawa, 2003). We utilize the TF-IDF method to compute the importance of a term w, both in an individual document and in the whole dataset. The formula for TF-IDF can be referenced in Eq. (1).

Let m be the number of features based on frequency ($m \leq 500$).

$$tf - idf = tf \times idf \tag{1}$$

where, tf is term frequency, idf is inverse document frequency in Eq. (2)

$$idf = \log \frac{the\ number\ of\ total\ feedback}{the\ number\ of\ feedback\ contain\ a\ term\ w} \tag{2}$$

3.2.2. Sentiment-based features

Sentiment-based feature selection methods provide a deeper understanding of text data by capturing emotional nuances, thus improving relevance and interpretability and enabling personalized and data-driven decision-making processes. We employ SentiWordNet along with an intensifier list to derive scores for words in feedback (SentiWordNet, being a lexical resource, includes “positivity” and “negativity” scores of words (Baccianella et al., 2010)). Moreover, we also utilize sentiment intensity with intensifiers and values. An intensifier is a word (e.g., “very”, “really”, or “extremely”) that strengthens or intensifies the meaning of another word, often used to emphasize the degree or extent of an adjective, verb, or another adverb. For instance, given a sentence, “The examination was very difficult”. In this sentence, “very” is the intensifier for “difficult”. Hence, intensifiers are useful tools in language to convey stronger feelings, emphasize points, and provide more vivid descriptions.

The list of intensifiers with values retrieves from *WordNet*, the corpus *Brown* in Sk-learn library and sentiment libraries, e.g., VADER, TextBlob, Flair (VADER, TextBlob, Flair are sentiment libraries which are used to retrieve sentiment scores (Elbagir & Yang, 2019; Loria, 2018; Pedregosa et al., 2011)). The intensifiers and values are determined in Algorithm 1.

Algorithm 1: Determining intensifiers and values

Input: WordNet, Corpus, Frequency threshold γ
Output: The list of intensifiers and values L

1 L ← ∅ *// initialize the list of intensifiers and values*
2 *ti* ← initilize common intensifiers (e.g., “very”, “extremely”, “really”, “highly”)
3 *temp* ← ∅ *// initialize the list of synonyms*
4 for each intensifier iw_i in the initilized list *ti* do: *// retrieve synomyms from WordNet*
5 *syn* ← retrieve synonyms for iw_i from WordNet
6 add *syn* to *temp*
7 for each word w_j in the corpus do: *// calculate frquencies of the intensifier with Brown corpus*
8 if w_j is in *temp* then
9 increase one for w_j in the *temp*
10 for each intensifier iw_i in the *temp* do: *// choose an intensifier if its frequency* > threshold γ
11 if frequencies of the intensifier iw_i is greater than γ then
12 add iw_i to L
13 for each intensifier iw_i in L do: *// calculate intensifier values with the majority voting*
14 *vv* ← calculate a sentiment score for iw_i with VADER
15 *tv* ← calculate a sentiment score for iw_i with TextBlob
16 *fv* ← calculate a sentiment score for iw_i with Flair
17 *iv* ← calculate an intensifier value of iw_i with the majority voting method based on *vv*, *tv*, *fv*
18 add $<iw_i, iv_{iw_i}>$ to L
19 sort L in increasing order using iv_{iw} *// sort L based on intensifier values*
20 return the list of intensifiers and values L

The Determining intensifiers and values algorithm (Algorithm 1) is used to determine intensifiers and their values from WordNet, a corpus, and three sentiment libraries (VADER, TextBlob, and Flair). The list of intensifiers and values chosen by the algorithm is saved in L. Lines 1-3 initialize the list of intensifiers and values (L), the common intensifiers (*ti*), and the list of synonyms (*temp*). Lines 4-6 retrieve synonyms of intensifiers in the list *ti* from WordNet, saving these synonyms in *temp*. Lines 7-9 count the frequencies of each synonym in the list *ti* based on the Brown corpus, with the frequencies saved in *temp*. Lines 10-12 select intensifiers if the frequencies of each synonym are greater than the threshold γ, saving them in the list L. Lines 13-18 calculate intensifier values using the majority voting method with the three sentiment libraries (The majority voting method averages the scores of the most frequent polarity. If all three libraries return different results, the intensifier values of all three models are summed (Tuan, Nghia, et al., 2024)). The list of intensifiers with values is saved in the list L. In line 19, the list is sorted in increasing order iv_{iw}. In line 20, the algorithm returns the list of intensifiers and values L.

Let *iw* be an intensifier with a value listed in Table 1. Table 1 presents the list of intensifiers and their values as determined by Algorithm 1 (the threshold $\gamma = 0.5$). The “Intensifier” column contains intensifiers retrieved from WordNet and the Brown corpus. The “Value” column shows the values for the respective intensifiers, calculated using the majority voting method.

Table 1
Intensifiers and values

Intensifier	Value	Intensifier	Value	Intensifier	Value
awfully	-1.00	undoubtedly	0.00	highly	0.58
dreadfully	-0.86	somewhat	0.00	pretty	0.58
terribly	-0.85	altogether	0.00	reasonably	0.59
horribly	-0.84	quite	0.00	absolutely	0.60
violently	-0.79	entirely	0.00	purely	0.60
hopelessly	-0.75	critically	0.00	surprising	0.66
suspiciously	-0.70	moderately	0.00	tremendously	0.66
mildly	-0.61	simply	0.00	primarily	0.68
extremely	-0.55	exclusively	0.00	significantly	0.68
bitterly	-0.52	so	0.00	genuinely	0.70
largely	-0.51	almost	0.00	truly	0.72
partially	-0.49	predominantly	0.00	astonishing	0.75
roughly	-0.33	rather	0.00	strikingly	0.75
fairly	-0.22	nearly	0.35	supremely	0.79
enormously	0.00	very	0.40	amazingly	0.80
immensely	0.00	faintly	0.42	exceptionally	0.83
deeply	0.00	mainly	0.47	wonderfully	0.86
especially	0.00	slightly	0.48	remarkably	0.87
utterly	0.00	really	0.53	perfectly	0.88
totally	0.00	particularly	0.55	greatly	0.90
unquestionably	0.00	strongly	0.57	incredibly	0.95

From the students' feedback, which has been tagged with parts of speech (POS) and parsed for dependency in the previous phase with SpaCy library, and using the lexical sentiment scores from SentiWordNet, we propose eight features. Before exploring the method in detail, the subsequent definitions are presented:

Definition 1: ***Positive, negative, and object scores of one feedback*** (*ps, ns*, and *os*) are the total positive, negative, and object scores of all words in one feedback obtained from SentiWordNet and intensifier values, as shown in Eq. (3) – Eq. (5),

$$ps_f = \sum_{i=1}^{|w|} swnP(w_i) * \ [1 + IW(iw, dep(w_i, iw, f))] \tag{3}$$

$$ns_f = \sum_{i=1}^{|w|} swnN(w_i) * \ \left[1 + IW\left(iw, dep(w_i, iw, f)\right)\right] \tag{4}$$

$$os_f = \sum_{i=1}^{|w|} swnO(w_i) \tag{5}$$

where, $|w|$ is the number of words in one feedback. w_i is the word i^{th} in the feedback. $swnP(w_i)$, swnN(w_i), and swnO(w_i) are used to retrieve positive, negative, and object scores for one word from the SentiWordNet. $IW(iw, dep(w_i, iw, f))$ is used to retrieve an intensifier value of iw from Table 1 if there exists a dependency between the word w_i and the intensifier iw in the feedback f ($dep(w_i, iw, f)$ returns true).

Definition 2: ***Positive, negative, and object words of one feedback*** (*pw, nw*, and *ow*) are the number of positive, negative, and object words in one feedback, as shown in Eq. (6) – Eq. (8).

$$pw_f = \sum_{i=1}^{|w|} 1, if \ \ swnP(w_i) > swnN(w_i) \tag{6}$$

$$ns_f = \sum_{i=1}^{|w|} 1, if \ \ swnP(w_i) < swnN(w_i) \tag{7}$$

$$ow_f = \sum_{i=1}^{|w|} 1, if \ swnO(w_i) \ \geq (swnP(w_i) + swnN(w_i)) \tag{8}$$

Definition 3: ***Negation of opinion words of one feedback*** (*nno*) is the number of opinion words whose dependency associations are negative words in one feedback, as shown in Eq. (9).

$$nno_f = \sum_{i=1}^{|w|} 1\,, if\,(w_i\ is\ an\ opinion\ word) \wedge (neg(w_i, f) = true) \tag{9}$$

where, $neg(w_i, f)$ returns true if the word i^{th} (w_i) is an opinion word and has dependency association with negative words in the feedback *f*.

Definition 4: ***Intensifier of opinion words of one feedback*** (*niw*) is the number of opinion words whose dependency associations are intensifiers in one feedback, as shown in Eq. (10).

$$niw_f = \sum_{i=1}^{|w|} 1\,, if\ (w_i\ is\ an\ opinion\ word) \wedge (def(w_i, iw, f) = true) \tag{10}$$

Algorithm 2: Selecting features based on sentiment

Input: Dataset D, SentiWordNet, List of intensifiers and values L
Output: Sentiment-based feature vector of dataset VS
1 VS ← ∅ *// initialize the sentiment-based feature vector for dataset*
2 for each feedback f in D do:
3 $ps_f, ns_f, os_f, pw_f, nw_f, ow_f, nno_f, niw_f \leftarrow 0$
4 for each word w_i in f do:
5 $v_f \leftarrow \emptyset$ *// initialize feature vector for feedback*
6 $ps_f \leftarrow \sum_{i=1}^{|w|} swnP(w_i) * [1 + IW(iw, dep(w_i, iw, f))]$ *// calculate the positive score of feedback*
7 $ns_f \leftarrow \sum_{i=1}^{|w|} swnN(w_i) * [1 + IW(iw, dep(w_i, iw, f))]$ *// calculate the negative score of feedback*
8 $os_f \leftarrow \sum_{i=1}^{|w|} swnO(w_i)$ *// calculate the object score of feedback*
9 $pw_f \leftarrow \sum_{i=1}^{|w|} 1, if\ swnP(w_i) > swnN(w_i)$ *// calculate the positive words of feedback*
10 $nw_f \leftarrow \sum_{i=1}^{|w|} 1, if\ swnP(w_i) < swnN(w_i)$ *// calculate the negative words of feedback*
11 $ow_f \leftarrow \sum_{i=1}^{|w|} 1, if\ swnO(w_i) \geq (swnP(w_i) + swnN(w_i))$ *// calculate the object words of feedback*
12 $nno_f \leftarrow \sum_{i=1}^{|w|} 1, if(w_i\ is\ an\ opinion\ word) \wedge (neg(w_i, f) = true)$ *// calculate the negation of opinion word*
13 $niw_f \leftarrow \sum_{i=1}^{|w|} 1, if\ (w_i\ is\ an\ opinion\ word) \wedge (def(w_i, iw, f) = true)$ *// calculate the intensifier of opinion word*
14 add $< ps_f, ns_f, os_f, pw_f, nw_f, ow_f, nno_f, niw_f >$ to v_f *// save 8 features for feedback to v_f*
15 add v_f to VS *// save v_f of feedback f to VS*
16 return the sentiment-based feature vector of dataset VS

The selection of features based on sentiment algorithm (Algorithm 2) is used to select eight features for the dataset from SentiWordNet, dependency among POS, and the list of intensifiers and their values, L. The feature vector based on sentiment, chosen by the algorithm, is saved in VS. Line 1 initializes the sentiment-based feature vector of the dataset (VS). Lines 2-15 calculate values for variables $(ps_f, ns_f, os_f, pw_f, nw_f, ow_f, nno_f, niw_f)$ according to formulas (3)-(10) for all feedback in the dataset D. These values are selected as features and saved to VS. In line 16, the algorithm returns the sentiment-based feature vector VS.

For example, given the sentence "I absolutely hate his behavior in class", SpaCy parsed the sentence for POS, as shown in Fig. 2. Using Eq. (3)- Eq. (10), the values for $ps_f, ns_f, os_f, pw_f, nw_f, ow_f, nno_f$, and niw_f are 0.0, 1.2, 0.25, 0, 1, 3, 1, and 1, respectively.

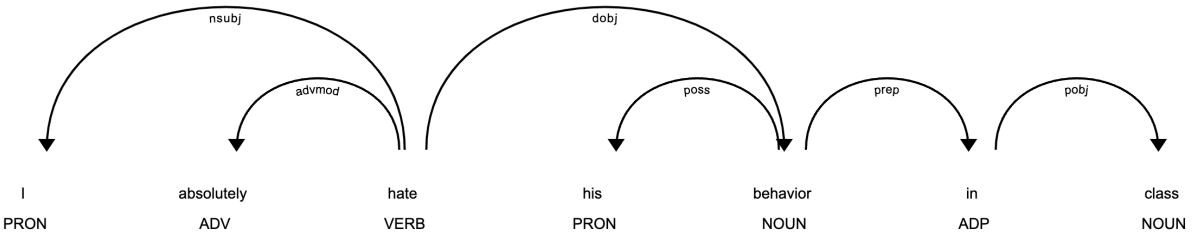

Fig. 2. An example of a sentence parsed by SpaCy

3.2.3. *Concatenated features based on frequency and sentiment*

In this study, we propose a novel approach that integrates both frequency- and sentiment-based features for enhanced classification performance. The frequency-based features are initially extracted using the TF-IDF method, which captures the importance of words in the feedback corpus. These features reflect the relevance and occurrence of terms within the dataset, providing a quantitative basis for text analysis. Concurrently, sentiment-based features are derived using a comprehensive approach that includes dependency associations among POS, SentiWordNet, and a list of intensifiers and their respective values. This step involves calculating sentiment scores that capture the emotional nuances of the feedback, thus enabling a profound understanding of the students' sentiments. These features include positive, negative, and objective sentiment scores, along with the impact of intensifiers on these sentiments. Once both sets of features are extracted, the sets are concatenated to assemble a comprehensive feature vector. By merging these two types of features, the classification models perform well in identifying and categorizing the sentiments expressed in student feedback, which is the aim of this study.

3.3. Classification Methods

There are seven ML algorithms, e.g., KNN, DT, RF, MLP, SVM, Gradient Boosting, and XGBoost, which have demonstrated strong performance in classifying various datasets (Aljrees, 2024; Batool & Byun, 2024; Chadaga et al., 2024; Nasir et al., 2024; Saha et al., 2024; Sarmiento Varón et al., 2023; Tuan, Trang, et al., 2024). The hyperparameters of these ML algorithms are depicted in Table 1. For KNN, '*n_neighbors*' determines the number of neighbors, the '*metric*' parameter calculates distance, and the '*p*' parameter specifies the power for the metric. For DT, '*criterion'* detects information gain, and '*splitter'* is used to choose the best split. For RF, '*n_estimators'* is the number of trees in the forest, and '*max_features*' is the number of features to consider for the best split. Similarly, in MLP, '*solver*' decides the optimization algorithm for weight adjustment during training, '*random_state*' parameter determines random number generation for weights and bias initialization, and '*hidden_layer_sizes*' denotes the neurons in each hidden layer. The number of neurons in each hidden layer are determined by the formula $nn = (2n^2 + 1)/(3n + 1)$, where n is the total input (Myo et al., 2019). In SVM, the '*C*' parameter is a regularization factor affecting decision boundary smoothness and accurate classification. Additionally, '*kernel*' defines the kernel function type, '*degree*' applies to polynomial kernel functions, and '*gamma*' controls single training example influence. For the Gradient Boosting, '*n_estimators*' sets the boosting stages, '*learning_rate*' shrinks the contribution of each tree, '*max_depth*' limits tree nodes, and '*random_state'* parameter controls the random seed given to each Tree estimator at each boosting iteration. In XGBoost, '*learning_rate*' shrinks each tree's contribution, '*subsample*' is the training instance subsample ratio, and '*colsample_bytree*' is the column subsample ratio when building each tree. Other hyperparameters of each ML algorithm are default values.

Table 2
ML algorithms with hyperparameters

ML algorithm	Hyperparameters
KNN	n_neighbors = 2, metric = 'minkowski', p = 2
DT	criterion = 'entropy', splitter = 'best'
RF	n_estimators=100, max_features = 'auto'
MLP	solver = 'lbfgs', hidden_layer_sizes = (*nn*,), random_state = 1
SVM	C = 1.0, kernel = 'linear', degree = 3, gamma = 'auto'
Gradient Boosting	n_estimators=100, learning_rate=1.0, max_depth=1, random_state=0
XGBoost	learning_rate = 0.3, subsample = 1.0, colsample_bytree = 1.0

3.4. Themes for sentiment-related feedback

Based on students' feedback and three sentiment levels (negative, neutral, positive), we employ a method to identify the primary themes within sentiment-related feedback. We create semantic networks for each sentiment level based on the co-occurrence frequencies of words. Words with high frequencies are selected, and a theme is emanated for each sentiment.

4. Results

In this study, we conducted our experiment using a dataset of students' feedback. The dataset released and published on Kaggle by Jayaprakashpondy (2022) consists of 2,345 feedback entries, categorized into three classes (negative/neutral/positive) with frequencies of 886, 381, and 1,078, respectively. The ratio for splitting data is 7:3 (70% dataset for training and 30% for testing). To evaluate our proposed method, we compared the results returned by the proposed method with other approaches using Precision (P), Recall (R), F1-score, and Accuracy measures as shown in Eq. (11)-Eq. (14).

$$\text{Precision (P)} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{11}$$

$$\text{Recall (R)} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{12}$$

$$\text{F1} - \text{score} = 2 \times \frac{\text{P} \times \text{R}}{\text{P} + \text{R}} \tag{13}$$

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \tag{14}$$

where TP (True Positive) occurs when an observation is correctly identified as positive. FP (False Positive) happens when an observation is negative but is predicted as positive. TN (True Negative) is when an observation is correctly identified as negative. FN (False Negative) occurs when an observation is positive but is incorrectly predicted as negative.

The results of the proposed feature (concatenated frequency- and sentiment-based) were compared to four feature selection methods: TF-IDF, BoW, Lexicon-based (Opinion Lexicon released by Hu & Liu (2004)), and SentiWordNet (SWN). The proposed feature achieved the highest accuracy across all ML algorithms (see Fig. 3 and Table 4 for more details). The highest accuracy scores were recorded by DT, MLP, SVM, and Gradient Boosting with 1.00, followed by KNN with 0.98. XGBoost scored the lowest accuracy with 0.82. We also experimented using the FastText model and compared it to our proposed feature. FastText is a library designed for efficiently learning word representations and performing sentence classification (Joulin et al., 2016). Our proposed feature demonstrated higher accuracy than the FastText model, which achieved an accuracy 0.75.

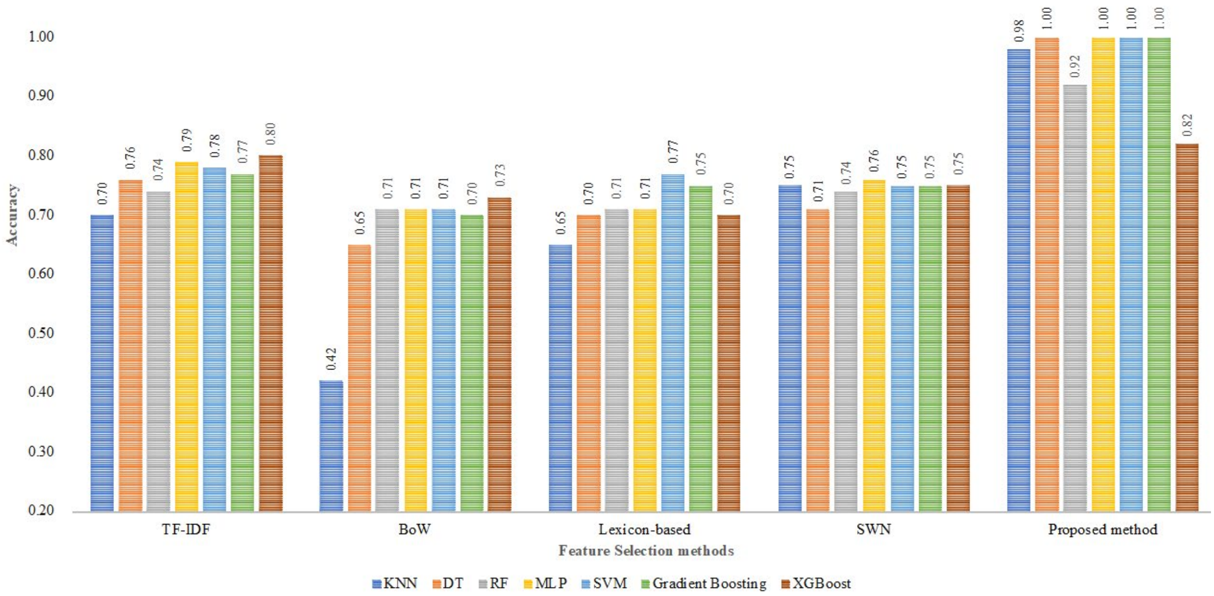

Fig. 3. A comparion of ML algorithms for classifying students' sentiment from feedback in terms of accuracy

Table 4
Results of the ML algorthims with various feature selection methods

	TF-IDF	BoW	Lexicon-based	SWN	Proposed method
KNN	0.70	0.42	0.65	0.75	0.98
DT	0.76	0.65	0.70	0.71	1.00
RF	0.74	0.71	0.71	0.74	0.92
MLP	0.79	0.71	0.71	0.76	1.00
SVM	0.78	0.71	0.77	0.75	1.00
Gradient Boosting	0.77	0.70	0.75	0.75	1.00
XGBoost	0.80	0.73	0.70	0.75	0.82

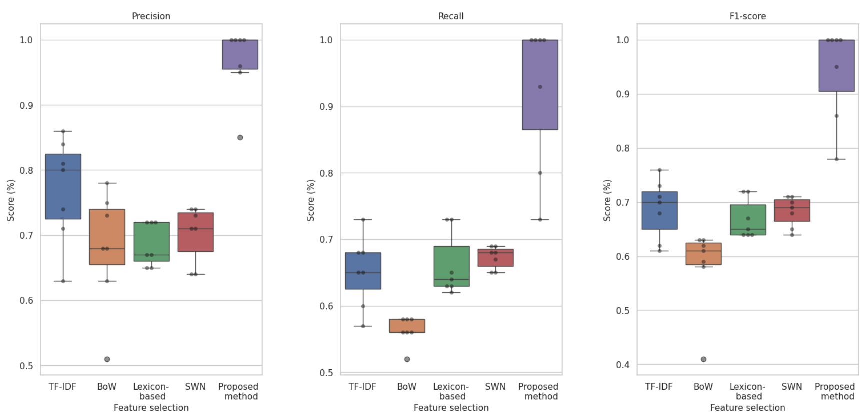

Fig. 4. A comparion among feature selection methods with the proposed method in terms of precision, recall, and F1-score

Among the five feature selection methods, the proposed method achieved the highest mean precision score of 0.97 [IQR: 0.96–1.00] with SE ± 0.02. The TF-IDF method followed with a mean precision score of 0.78 [IQR: 0.73–0.83] and SE ± 0.03. In terms of recall, the proposed method again led with a mean recall score of 0.93 [IQR: 0.87–1.00] with SE ± 0.04, followed by the SWN method, which had a mean recall score of 0.67 [IQR: 0.66–0.69] and SE ± 0.01. For the F1-score, the proposed method achieved the highest mean F1-score of 0.94 [IQR: 0.91–1.00] with SE ± 0.03, while the TF-IDF method had a mean F1-score of 0.69 [IQR: 0.65–0.72] and SE ± 0.02 (Fig. 4).

Using the method of word co-occurrence, semantics networks were generated for three sentiment levels of student feedback, as depicted in Fig. 5. In the positive sentiment-based feedback, the most frequent keywords were *class, hard, teacher, learn, lectur, expect, professor, grade, pass, take, recommend, understand, love, help, student, pretty, difficult, question*, and *test*. For the neutral sentiment-based feedback, the most frequent keywords were *class, test, lectur, hard, teacher, professor, studi, student, read, book, note, grade, easi*, and *question*. In the negative sentiment-based feedback, the most frequent keywords were *class, onlin, homework, professor, exam, studi, teacher, question, fail, recommend, quiz, hard, grade, worst, teach, assign, difficult, lectur, horribl, answer, pass, confus, bad*, and *paper*. These high-frequency words and their co-occurrence with other words enabled us to identify themes for each sentiment level. The identified themes were *"Classroom Experience and Academic Support"* for positive feedback, *"Academic Environment and Learning Resources"* for neutral feedback, and *"Challenges and Dissatisfaction in Academic Experience"* for negative feedback (Table 5).

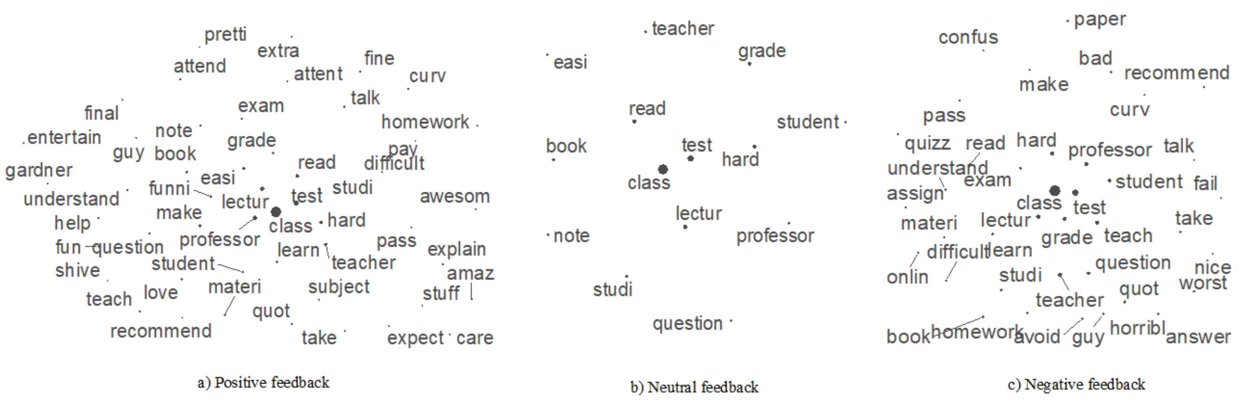

Fig. 5. Semantic networks based on sentiment of students' feedback

Table 5
Themes of each sentiment from sentiment-based feedback

Sentiment	***Theme of sentiment-based feedback*** **Key elements: details of words**
Positive	***Classroom Experience and Academic Support*** Class Experience and Expectations: "class", "learn", "expect", "understand" Teaching Quality: "teacher", "professor", "lectur" Academic Performance: "grade", "pass", "test" Student Engagement and Support: "student", "help", "question" Course Difficulty and Recommendations: "hard", "difficult", "recommend"
Neutral	***Academic Environment and Learning Resources*** Class Structure and Assessments: "class", "test", "lectur" Difficulty and Effort: "hard", "easi" Teaching Staff: "teacher", "professor" Study and Learning Materials: "studi", "read", "book", "note" Testing and Grading: "student", "question", "grade"
Negative	***Challenges and Dissatisfaction in Academic Experience*** Class and Teaching Quality: "class", "teacher", "professor", "teach", "lectur" Assignments and Homework: "homework", "assign", "paper" Exams and Quizzes: "studi", "confus" Student Performance and Grades: "fail", "grade", "pass" Class Difficulty and Workload: "hard", "difficult" Course Recommendations and Reviews: "recommend", "worst", "horribl", "bad"

5. Discussion

This research delves deeply into a framework for classifying students' sentiments from feedback using concatenated feature selection based on frequency and sentiment. Then, the framework represents students' feedback on three semantic networks. Our study highlights the significant advantages of the proposed feature combination method over other techniques such as TF-IDF, BoW, lexicon-based, and SentiWordNet (SWN). The concatenated frequency- and sentiment-based features consistently outperformed these conventional methods, demonstrating superior classification accuracy and enhanced performance metrics. The proposed approach's accuracy was higher across all machine learning algorithms, with a minimum accuracy of 0.82, underscoring its effectiveness in capturing subtle student sentiments that traditional methods often miss. This comparison underscores the innovative contribution of our method in sentiment analysis across different domains, such as customer feedback analysis, social media sentiment monitoring, and patient feedback in healthcare. This aligns with other research emphasizing that a hybrid method for feature selection in sentiment analysis achieved higher accuracy (Alamin et al., 2024; Harish et al., 2019; Kaur & Sharma, 2023; Khan et al., 2021; Shahi et al., 2022).

In our study, one crucial aspect is considering sentiment intensity, particularly the role of intensifiers in sentiment analysis. Intensifiers are words that amplify the sentiment of the feedback, providing a more precise differentiation of sentiment levels. By accurately capturing these intensifiers, our approach offers a deeper understanding of the emotional strength behind the feedback. This granularity enables more precise sentiment categorization and reveals the intensity of students' feelings, which is crucial for effectively addressing their emotional and academic needs. This is consistent with previous research that highlights the role of intensifiers in sentiment analysis (Asghar et al., 2019; Kabir et al., 2024; Mudgal & Khunteta, 2020; Tran et al., 2021a, 2021b).

Our analysis reveals a strong relationship between the themes identified in student feedback and their impact on student engagement and learning outcomes. Targeted interventions based on these insights can significantly enhance student participation and academic performance. Addressing themes related to classroom experience and academic support can foster a more engaging and supportive learning environment. By aligning teaching and support strategies with student feedback, educators can create conditions that promote active participation and improved academic achievement (Carless & Winstone, 2023; Li et al., 2023; Mandouit, 2018; Shaik et al., 2023).

While our study shows promising results, it also encounters certain limitations and challenges. The extensive preprocessing required for textual data can be time-consuming, and there is a risk of overfitting when using complex models. Additionally, accurately capturing subtle sentiment nuances remains challenging. Acknowledging these limitations, we suggest several ways to mitigate them in future research, such as employing advanced deep learning models and exploring alternative feature combination techniques. These steps can further refine the approach and enhance its robustness and applicability.

6. Conclusion

In this study, we proposed the student sentiment from feedback (SSF) framework with a novel feature selection method that combines frequency-based and sentiment-based features for analyzing student feedback. Our approach outperformed other techniques such as TF-IDF, BoW, lexicon-based methods, and SentiWordNet, achieving higher accuracy across various machine learning algorithms. By creating semantic networks, we identified key themes in positive, neutral, and negative feedback, providing insights that can enhance educational practices. This method's ability to capture nuanced sentiments underscores its potential for broader applications in sentiment analysis across different domains, including customer feedback and healthcare, highlighting the importance of integrating frequency and sentiment features for comprehensive sentiment understanding.

Looking ahead, future research could explore various directions to build on the current findings. Investigating other feature combination techniques, integrating additional linguistic features, and applying advanced deep learning models are potential avenues for further study. These future endeavors can enhance sentiment analysis accuracy and expand its applicability to more complex datasets and diverse contexts. By continually refining and evolving the approach, we can unlock new insights and applications, further advancing the field of sentiment analysis.

Acknowledgments

The authors would like to express their gratitude to the Informatics Innovation Center of Excellence (IICE), the School of Informatics, Research and Innovation Institute of Excellence (RIIE), Walailak University, Thailand for their support during this research.

Funding

This work was financially supported by Walailak University Personal Research Fund (Contract No. WU65231).

References

Abualhaj, M., Al-Zyoud, M., Hiari, M., Alrabanah, Y., Anbar, M., Amer, A., & Al-Allawee, A. (2024). A fine-tuning of decision tree classifier for ransomware detection based on memory data. *International Journal of Data and Network Science*, *8*(2), 733–742.

Ahn, E., & Kang, H. (2018). Introduction to systematic review and meta-analysis. *Kja*, *71*(2), 103–112. https://doi.org/10.4097/kjae.2018.71.2.103

Aizawa, A. (2003). An information-theoretic perspective of tf–idf measures. *Information Processing & Management*, *39*(1), 45–65.

AL-Akhras, M., Alghamdi, S., Omar, H., & Alshareef, H. (2024). A machine learning technique for Android malicious attacks detection based on API calls. *Decision Science Letters*, *13*(1), 29–44.

Alamin, Z., Lorosae, T. A., Ramadhan, S., & others. (2024). Improving Performance Sentiment Movie Review Classification Using Hybrid Feature TFIDF, N-Gram, Information Gain and Support Vector Machine. *Mathematical Modelling of Engineering Problems*, *11*(2).

Al-Dwish, G. M., & Aljohani, A. N. (2024). Sentiment Analysis of Saudi Arabian University Students by Using Ml Algorithms. *Advances and Applications in Discrete Mathematics*, *41*(5), 393–410.

Aljrees, T. (2024). Improving prediction of cervical cancer using KNN imputer and multi-model ensemble learning. *Plos One*, *19*(1), e0295632.

Alsubaie, F., & Aldoukhi, M. (2024). Using machine learning algorithms with improved accuracy to analyze and predict employee attrition. *Decision Science Letters*, *13*(1), 1–18.

Altrabsheh, N., Cocea, M., & Fallahkhair, S. (2015). Predicting learning-related emotions from students' textual classroom feedback via Twitter. *International Educational Data Mining Society*.

Altun, A., Köklü, M., & others. (2022). Optimizing the learning process of multi-layer perceptrons using a hybrid algorithm based on MVO and SA. *International Journal of Industrial Engineering Computations*, *13*(4), 617–640.

Asghar, M. Z., Khan, A., Zahra, S. R., Ahmad, S., & Kundi, F. M. (2019). Aspect-based opinion mining framework using heuristic patterns. *Cluster Computing*, *22*, 7181–7199.

Baccianella, S., Esuli, A., & Sebastiani, F. (2010). *SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining*.

Batool, A., & Byun, Y.-C. (2024). Towards Improving Breast Cancer Classification using an Adaptive Voting Ensemble Learning Algorithm. *IEEE Access*.

Bensba, A., Ahmim, N., Zakaria, C., & Bousbia, N. (2022). Analysis of Students' Emotions in an Online Learning Environment. *2022 International Conference on Advanced Aspects of Software Engineering (ICAASE)*, 1–7.

Bhardwaj, A., & Srivastava, P. (2021). A machine learning approach to sentiment analysis on web based feedback. *Applications of Artificial Intelligence and Machine Learning: Select Proceedings of ICAAAIML 2020*, 127–139.

Bing, L. (2015). *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press.

Carless, D., & Winstone, N. (2023). Teacher feedback literacy and its interplay with student feedback literacy. *Teaching in Higher Education*, *28*(1), 150–163.

Chadaga, K., Prabhu, S., Sampathila, N., Chadaga, R., Umakanth, S., Bhat, D., & GS, S. K. (2024). Explainable artificial intelligence approaches for COVID-19 prognosis prediction using clinical markers. *Scientific Reports*, *14*(1), 1783.

Dalal, R., Safhath, I., Piryani, R., Kappara, D. R., & Singh, V. K. (2015). A lexicon pooled machine learning classifier for opinion mining from course feedbacks. *Advances in Intelligent Informatics*, 419–428.

Dsouza, D. D., Deepika, D. P. N., Machado, E. J., Adesh, N. D., & others. (2019). Sentimental analysis of student feedback using machine learning techniques. *International Journal of Recent Technology and Engineering*, *8*(14), 986–991.

Elbagir, S., & Yang, J. (2019). Twitter sentiment analysis using natural language toolkit and VADER sentiment. *Proceedings of the International Multiconference of Engineers and Computer Scientists*, *122*(16).

Esparza, G. G., de-Luna, A., Zezzatti, A. O., Hernandez, A., Ponce, J., Álvarez, M., Cossio, E., & de Jesus Nava, J. (2018). A sentiment analysis model to analyze students reviews of teacher performance using support vector machines. *Distributed Computing and Artificial Intelligence, 14th International Conference*, 157–164.

Giang, N. T. P., Dien, T. T., & Khoa, T. T. M. (2020). Sentiment analysis for university students' feedback. *Advances in Information and Communication: Proceedings of the 2020 Future of Information and Communication Conference (FICC), Volume 2*, 55–66.

Gutiérrez, G., Ponce, J., Ochoa, A., & Álvarez, M. (2018). Analyzing Students Reviews of Teacher Performance Using Support Vector Machines by a Proposed Model. *Intelligent Computing Systems: Second International Symposium, ISICS 2018, Merida, Mexico, March 21-23, 2018, Proceedings 2*, 113–122.

Harish, B. S., Kumar, K., & Darshan, H. K. (2019). *Sentiment analysis on IMDb movie reviews using hybrid feature extraction method*.

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. *Proceedings of the 10th International Conference on Knowledge Discovery and Data Mining (SIGKDD)*, 168–177.

Imran, M., Hina, S., & Baig, M. M. (2022). Analysis of Learner's Sentiments to Evaluate Sustainability of Online Education System during COVID-19 Pandemic. *Sustainability*, *14*(8), 4529.

Jain, R., Bawa, S., & Sharma, S. (2022). Sentiment analysis of COVID-19 tweets by machine learning and deep learning classifiers. *Advances in Data and Information Sciences: Proceedings of ICDIS 2021*, 329–339.

Jayaprakashpondy. (2022). *Student Feedback dataset*. Accessed: March. 13, 2024. https://www.kaggle.com/datasets/jayaprakashpondy/student-feedback/data

Joulin, A., Grave, E., Bojanowski, P., Douze, M., Jégou, H., & Mikolov, T. (2016). FastText.zip: Compressing text classification models. *ArXiv Preprint ArXiv:1612.03651*.

Kabir, M. M., Othman, Z. A., & Yaakub, M. R. (2024). Enhanced Lexicon based Hybrid Method for Slang and Punctuation Scoring for Aspect Based Sentiment Analysis. *2024 6th International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT)*, 1418–1422.

Kathuria, A., Gupta, A., & Singla, R. K. (2023). AOH-Senti: Aspect-Oriented Hybrid Approach to Sentiment Analysis of Students' Feedback. *SN Computer Science*, *4*(2), 152.

Kaur, G., & Sharma, A. (2023). A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis. *Journal of Big Data*, *10*(1), 5.

Khan, J., Alam, A., & Lee, Y. (2021). Intelligent hybrid feature selection for textual sentiment classification. *IEEE Access*, *9*, 140590–140608.

Li, S., Xie, Z., Chiu, D. K. W., & Ho, K. K. W. (2023). Sentiment analysis and topic modeling regarding online classes on the Reddit Platform: educators versus learners. *Applied Sciences*, *13*(4), 2250.

Lin, Q., Zhu, Y., Zhang, S., Shi, P., Guo, Q., & Niu, Z. (2019). Lexical based automated teaching evaluation via students' short reviews. *Computer Applications in Engineering Education*, *27*(1), 194–205.

Loan, D. T. T., & Tuan, T. A. (2024). Towards visual sentiment summary to understand customers' satisfaction. *Bulletin of Electrical Engineering and Informatics*, *13*(4), 2774–2783.

Loria, S. (2018). Textblob Documentation. *Release 0.15*, *2*(8), 269.

Louati, A., Louati, H., Kariri, E., Alaskar, F., & Alotaibi, A. (2023). Sentiment analysis of Arabic course reviews of a Saudi university using support vector machine. *Applied Sciences*, *13*(23), 12539.

Mandouit, L. (2018). Using student feedback to improve teaching. *Educational Action Research*, *26*(5), 755–769.

Melba Rosalind, J., & Suguna, S. (2021). Ensemble Model for Classifying Sentiments of Online Course Reviews using blended Feature Selection methods. *Journal of Education: Rabindrabharati University*.

Mercha, E. M., & Benbrahim, H. (2023). Machine learning and deep learning for sentiment analysis across languages: A survey. *Neurocomputing*, *531*, 195–216.

Mudgal, P., & Khunteta, A. (2020). Handling double intensifiers in feature-level sentiment analysis based on movie reviews. *International Conference on Artificial Intelligence: Advances and Applications 2019: Proceedings of ICAIAA 2019*, 383–392.

Myo, W. W., Wettayaprasit, W., & Aiyarak, P. (2019). Designing classifier for human activity recognition using artificial neural network. *2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS)*, 81–85.

Nafea, I. T. (2018). Machine learning in educational technology. *Machine Learning-Advanced Techniques and Emerging Applications*, 175–183.

Nasim, Z., Rajput, Q., & Haider, S. (2017). Sentiment analysis of student feedback using machine learning and lexicon based approaches. *2017 International Conference on Research and Innovation in Information Systems (ICRIIS)*, 1–6.

Nasir, A., Makki, S. V. A.-D., & Al-Sabbagh, A. (2024). Pandemia Prediction Using Machine Learning. *PRZEGLĄD ELEKTROTECHNICZNY*, *5*, 211–214.

Nasrulloh, S., Permanasari, A., & Kusumawardani, S. (2019). A Framework of Educational Feedback System with Statistical Method and Sentiment Analysis. *6th International Conference on Educational Research and Innovation (ICERI 2018)*, 310–316.

Pacol, C. A., & Palaoag, T. D. (2021). Enhancing sentiment analysis of textual feedback in the student-faculty evaluation using machine learning techniques. *European Journal of Engineering Science and Technology*, *4*(1), 27–34.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., & others. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, *12*, 2825–2830.

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child Development*, *88*(5), 1653–1670.

Pekrun, R., Marsh, H. W., Elliot, A. J., Stockinger, K., Perry, R. P., Vogl, E., Goetz, T., Van Tilburg, W. A. P., Lüdtke, O., & Vispoel, W. P. (2023). A three-dimensional taxonomy of achievement emotions. *Journal of Personality and Social Psychology*, *124*(1), 145.

Rajesh, P., & Suseendran, G. (2020). Prediction of N-gram language models using sentiment analysis on E-learning reviews. *2020 International Conference on Intelligent Engineering and Management (ICIEM)*, 510–514.

Rakhmanov, O. (2020). On validity of sentiment analysis scores and development of classification model for student-lecturer comments using weight-based approach and deep learning. *Proceedings of the 21st Annual Conference on Information Technology Education*, 174–179.

Ramasamy, L. K., Kadry, S., Nam, Y., & Meqdad, M. N. (2021). Performance analysis of sentiments in Twitter dataset using SVM models. *International Journal of Electrical and Computer Engineering (IJECE)*, *11*(3), 2275–2284.

Saha, R., Malviya, L., Jadhav, A., & Dangi, R. (2024). Early stage HIV diagnosis using optimized ensemble learning technique. *Biomedical Signal Processing and Control*, *89*, 105787.

Sarmiento Varón, L., González-Puelma, J., Medina-Ortiz, D., Aldridge, J., Alvarez-Saravia, D., Uribe-Paredes, R., & Navarrete, M. A. (2023). The role of machine learning in health policies during the COVID-19 pandemic and in long COVID management. *Frontiers in Public Health*, *11*, 1140353.

Shahi, T. B., Sitaula, C., & Paudel, N. (2022). A hybrid feature extraction method for Nepali COVID-19-related tweets classification. *Computational Intelligence and Neuroscience*, *2022*(1), 5681574.

Shaik, T., Tao, X., Dann, C., Xie, H., Li, Y., & Galligan, L. (2023). Sentiment analysis and opinion mining on educational data: A survey. *Natural Language Processing Journal*, *2*, 100003.

Shenify, M. (2024). Sentiment analysis of Saudi e-commerce using na\"\ive bayes algorithm and support vector machine. *International Journal of Data and Network Science*, *8*(3), 1607–1612.

Sohel, A., Hossain, M. R., Mostofa, Z. B., Hasan, M. U., Das, U. C., & Parvin, S. K. (2023). Sentiment Analysis Based on Online Course Feedback Using Textblob and Machine Learning Techniques. *2023 26th International Conference on Computer and Information Technology (ICCIT)*, 1–6.

Spacy. (2024). spaCy library. In *Accessed: March. 13, 2024*. https://spacy.io/

Surya, P. P. M., & Subbulakshmi, B. (2019). Sentimental analysis using Naive Bayes classifier. *2019 International Conference on Vision towards Emerging Trends in Communication and Networking (ViTECoN)*, 1–5.

Tamrakar, M. L., & others. (2021). An Analytical Study Of Feature Extraction Techniques For Student Sentiment Analysis. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *12*(11), 2900–2908.

Tran, T. A., Duangsuwan, J., & Wettayaprasit, W. (2021a). A new approach for extracting and scoring aspect using SentiWordNet. *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, *22*(3), 1731–1738.

Tran, T. A., Duangsuwan, J., & Wettayaprasit, W. (2021b). A Novel framework for aspect knowledge base generated automatically from social media using pattern rules. *Computer Science*, *22*.

Tripathy, A., Agrawal, A., & Rath, S. K. (2015). Classification of sentimental reviews using machine learning techniques. *Procedia Computer Science*, *57*, 821–829.

Tuan, T. A., Nghia, N. H., An, T. D., & Loan, D. T. T. (2024). Exploring Mental Stress Expressions in Online Communities: A Subreddit Analysis. *Journal of Human, Earth and Future*, *5*(2), 131–150.

Tuan, T. A., Trang, L. T. T., An, T. D., Nghia, N. H., & Loan, D. T. T. (2024). An Analysis of Machine Learning for Detecting Depression, Anxiety, and Stress of Recovered COVID-19 Patients. *Journal of Human, Earth, and Future*, *5*(1), 1–18.

Ullah, M. A. (2016). Sentiment analysis of students feedback: A study towards optimal tools. *2016 International Workshop on Computational Intelligence (IWCI)*, 175–180.

Ward, T., Jenab, K., & Ortega-Moody, J. (2024). Machine learning models for condition-based maintenance with regular truncated signals. *Decision Science Letters*, *13*(1), 197–210.

Xiao, J., Liu, G., Huang, M., Yin, Z., & Gao, Z. (2023). A kernel-free L1 norm regularized ν-support vector machine model with application. *International Journal of Industrial Engineering Computations*, *14*(4), 691–706.

Xingyi, G., & Adnan, H. (2024). Potential cyberbullying detection in social media platforms based on a multi-task learning framework. *International Journal of Data and Network Science*, *8*(1), 25–34.