

Multifaceted Sentiment Detection System (MSDS) to Avoid Dropout in Virtual Learning Environment using Multi-class Classifiers

Ananthi Claral Mary. T¹, Arul Leena Rose. P. J^{2*}

Department of Computer Science-College of Science and Humanities,
SRM Institute of Science and Technology, Kattankulathur - 603 203, Chengalpattu, Tamil Nadu, India

Abstract—Sentiment analysis with machine learning plays a vital role in Higher Educational Institutions (HEI) for decision making. Technology-enabled interactions can only be successful when a strong student-teacher link is established, and the emotions of students are clearly comprehended. The paper aims at proposing Multifaceted Sentiment Detection System (MSDS) for detecting sentiments of higher education students participating in virtual learning and to classify the comments posted by them using Machine Learning (ML) algorithms. Present research evaluated a total of n=1590 students' comments with the presence of three specific multifaceted characteristics each providing 530 comments to perform Sentiment Analysis (SA) for monitoring their sentiments, opinions that facilitate predicting dropout in virtual learning environment (VLE). This begins with the phrase extraction; then data pre-processing techniques namely digits, punctuation marks and stop-words removal, spelling correction, tokenization, lemmatization, n-grams, and POS (Part of Speech) are applied. Texts are vectorized using two feature extraction techniques with count vectorization and TF-IDF metrics and classified with four multiclass supervised ML techniques namely Random Forest, Linear SVC, Multinomial Naive Bayes, and Logistic Regression for multifaceted sentiment classification. Analyzing students' feedback using sentiment analysis techniques classifies their positive, negative, or even more refined emotions that enables dropout prediction. Experimental results reveal that the highest mean accuracy result for device efficiency, cognitive behavior, technological expertise with cloud learning platform usage were achieved by Logistic Regression with 98.49%, Linear SVC with 93.58% and Linear SVC with 92.08% respectively. Practically, results confirm feasibility for detecting students' multifaceted behavioral patterns and risk of dropout in VLE.

Keywords—Sentiment analysis; opinions; TF-IDF; n-gram; virtual learning; machine learning; NLTK; text pre-processing

I. INTRODUCTION

The education system is crucially dependent on students' academic progress. The tremendous volume of data in educational databases has made it increasingly difficult to predict student performance. Low performance students may encounter several challenges, such as delayed graduation and dropping out. To quickly assist students who are performing poorly, educational institutions should regularly monitor the academic development of their students. One way to accomplish that is to use students' academic achievement prediction. The proposed method is based on hybrid approach

of sentiment analysis to achieve quality education [1]. With the explosive growth of internet, digital technologies, IT infrastructure cloud-based online learning is growing at a quicker pace. Cloud computing technologies facilitate virtual platforms and assist students in favorable paths despite the barriers. The platform inherits countless benefits namely reducing installation cost, high storage possibility, virtualization, security, easy access, etc. T. Zarra et al. states that academic institutions suffer from general budget cuts and the growing number of students. Implementation of cloud-based e-learning platforms is very much motivated. The application of cloud community assists members of education institution to work together. The research aims to propose a web service which analyzes exchange of information between learners from various universities connected on cloud using sentiment analysis [2].

The key contribution of this research is to analyze and deal with a huge amount of online data during the virtual learning course for predicting dropouts with the help of sentiment analysis of learners' comments. SA is receiving much attention in HEI for predicting student dropout. By applying SA in the context of virtual learning, provides better insights of understanding learners' attitudes, emotional reactions towards entities, events, attributes, and the notion of sentiment predictors of dropout as this reveals specific patterns in learners' behaviors that can be of practical importance for course designers and instructors. The learning process in virtual courses dramatically differs from face-to-face courses, as the courses are taken exclusively online hence, in this proposed system, it is appropriate to detect students' text sentiments composed of multifaceted characteristics that helps in identifying their risk of college dropout. One of the formidable tasks is systematical surveying of students' opinions and perceptions, as there is huge volume of opinionated text and sentiment detection of real-world data is full of challenges. It's difficult for humans to extract sentences with sentiments, read, review, and classify them into usable formats. Automated sentiment discovery and summarization system are thus required. SA is classified as positive, negative, or neutral, that employs automatic process of text classification for opinion mining and finding wide array of sentiments expressed by learners explaining their personal attitudes and evaluation regarding provided services that determines high and low sentiment scores of their multi-dimensional characteristic phenomena. Thus, this research

aims at investigation of sentiment analysis based on learners' comments from VLE to avoid dropouts. The remnant paper is organized as follows. Section II discusses the background of SA with related research works. Section III specifies the explanation of research methodology used in this research. Section IV explains MSDS architecture in detail. Section V represents the experimental outcomes. Section VI discusses conclusion and future scope of research.

II. LITERATURE REVIEW

MOOC based teaching is implemented in traditional curriculum and educational practices. Sentiment analysis approaches can be used for course content evaluations that deliver extensive intimations to course designers and educators for evaluating courses periodically and introduce probable enhancements [3]. SASys (Sentiment Analyzes System) framework built on lexical approach and polarized frame network is proposed. Its primary objective is identifying early students' risk of dropout by emotional state detection. Text sentiment is essential for defining learners' motivational profile that depends on their activities in VLE. Students' class engagement can be identified by analyzing their frequency of access data and interactions [4].

Students' motivation to learn is greatly influenced by emotional elements, and in online education, emotions can be inferred from textual discussion while giving responses. Natural Language Processing methods can be used for emotional classification of students by extracting information from Whatsapp and organizing them into corresponding categories. RNN algorithm increases the accuracy by 75% for students' emotion analysis in online learning [5]. Emotion mining and SA were performed by gathering Arabic tweets regarding online education during pandemic. Results disclose that the proposed method performs efficiently in identifying peoples' opinion on online learning in context of the pandemic using SVM with greatest accuracy of 89.6%. By considering emotion analysis, anger is the top emotion. Most significant reasons behind negative sentiments were lack of face-to-face interaction, breakdown of network, ambiguity, and games [6]. Academic issues are one of the factors which cause stress or depression among students. For example, a drop in grades, a fear of failing, and challenging competitions. Most students are young, and when they encounter difficulties, they occasionally lack wisdom and may cause harm to themselves. Hence, advice and support from professionals, family members, experts, and others are crucial for preventing adolescents from engaging in risky behavior. Sentiment analysis using Naive Bayes algorithm classifies students into stress and depression [7].

In recent years, due to generation of voluminous data, technologies have been developed for storing and data processing effortlessly. A huge amount of data can be obtained, mined, and realized for sentiment analysis. This helps in making better policies for the education sector and the target users are teachers, students, and educational organizations. In every aspect of teaching-learning approach, education sectors can incorporate sentiment analysis extensively [8]. Learners are more interested in "Engineering and Technology" but hold negative thinking towards "Life

Science and medicine", this research infers that it is an emergency to explore SA systems for cross-domain that accelerates SA application in multiple learning domains. SA is used for enhancing the learning process; additionally, it provides valuable information for educational institutions. Students receive results of SA through visualization tools such as word clouds, dashboards, virtual agents, etc [9]. Technological advancements like Blockchain, IoT, Cloud Computing and Big Data have broadened applications of SA permitting this to be utilized in any discipline [10]. Moreover, in Machine Learning and Natural Language Processing, SA has turned out to be a hot trend and is being accepted extensively all over the globe [11].

The primary objective of implementing SA approaches on students' feedback in an online learning system is to identify learners' emotions, feelings, participation and evaluate educators' performance [12]. From student learning and achievement, emotions are inseparable. Responses were collected from students to discover their emotional experience around test and an online quiz. Greater positive level emotions and lesser negative level emotions were experienced in online quizzes than tests. Future guidelines of research must integrate a complex relation between cognitive, emotional, and motivational aspects of learning [13]. Sentiment analysis accurately portrays students' learning circumstances in the online learning community. To improve quality of teaching practice, the proposed model can recognize students' sentiment tendencies. SSM (Sentiment Score Matrix) is formed to compute sentiment scores. This can efficiently identify sentiment tendencies of learners for enhancing information services quality in education practice [14].

HEI are increasingly looking for best ways to understand the learning experience of their students. Sentiment analysis helps to investigate emotions and attitudes of students regarding their course experience. To analyze sentiment text was fed into Google's Cloud-based Natural Language Processing. Results presented that students' sentiment in online interaction during two online courses is more positive than in face-to-face courses [15]. During COVID-19 pandemic, to support remote and distance learning useful learning resources are lecture recordings. Students with illness, learning disabilities, and work commitments have narrated that availability of lecture recordings has shaped an inclusive education setting. Sentiment analysis was conducted using Microsoft Azure cognitive services text analytics API. With a large text dataset, machine learning was employed that were labeled for sentiments along 0 and 1. Findings depicted that lecture recordings serve as an additional resource for preparing notes or exams [16]. Therefore, to solve real world problems through design and development of smart learning environment, we can use Azure cognitive services, a text analytics API that leverages natural language processing capabilities for deploying high-quality AI models [17].

Sentiment Analysis is the most crucial area in text mining, as the thoughts of several individuals are analyzed and compiled as a single dataset. E-learning is an educational attempt to deliver knowledge through computers. SA helps users to easily classify their emotional input information. Students' anti-course feelings can be tracked that serves as

feedback from online learning sites [18]. In the teaching-learning process virtual learning environments (VLE) deliver a set of communication and interaction tools utilized by learners and educators. Researchers presented the SentiEduc framework that uses Multi-Agent System (MAS) to gather and analyze opinion of texts posted by learners in VLE. SenticNet tool was used to analyze sentiments automatically. Educators with tutoring experience utilized the framework with real data for verifying the efficiency; as a result accuracy obtained was 73.88% [19]. From an organization, probable text reviews collected on 270 training programmes by 2688 participants were analyzed. RapidMiner Text Mining package was used to track tokenization, removal of stop words, stemming, and token filtering. Authors suggested that instead of content delivery and faculty expertise, the proposed approach can further be expanded to establish sentiment expressed over several aspects like internet connection, hospitality [20].

Researchers have developed a web-application system that uses text analytics and SA for providing educators with a deeper analysis of learners' feedback to assess the course they have taught that will enhance the students learning experience. Feedback was grouped into positive, negative, and neutral. The result depicted a larger number of neutral sentiments. Their system implementation was successful, and it significantly benefits students, lecturers, and administrators [21]. Twitter is the popular free social networking channel. In Anadolu University open and distance education system, sentiment analysis for learners was performed by fetching tweets. 400 tweets were used for validation and classification outcomes are presented. The negative feelings and student complaints can be concentrated by institution managers [22]. Numerous educational institutions utilize online education as media learning where each piece of media, maybe audio, video, or text, can accept learners' feedback. The lecture intends to realize emotions which learners experience when they access media, namely happiness, unhappiness, or disappointment and educators intend to recognize their delightfulness. The study developed a utility cellular for emotion detection from column comments in online media. Accuracy of mobile application utility is 70% for emotion detection [23]. Humans are easily prone to errors in interpreting text-based emotions. Four supervised ML classification techniques namely MNB, SVM, DT, and KNN were applied for analyzing basic emotions. The best performance was resulted by Multinomial Naive Bayes classifier with an average accuracy of 64.08% [24]. In the form of tweets, blogs, and updates of thoughts about interest, a lot of data is being produced. On a variety of subjects, including products, movies, politics, education, news, and more, people express their thoughts and ideas. Data analysis is useful to comprehend observations, sentiments, and attitudes of society. Additionally, decision-making would benefit more from such analysis. Naive Bayes, RF, Tailored RF and enhanced XGBoost were employed. Enhanced XGBoost achieved better accuracy of 72.26% [25].

With meticulous analysis of previous existing research, the limitations are they do not consider hidden structural features namely internet connection, mixed-emotional elements, and unstructured data. To address these limitations, the existing

methods can be operationalized and extended by considering specific factors namely device efficiency, cognitive behaviors, and technical familiarity with cloud platform usage. Moreover, they infer that there is an emergency to investigate SA systems for cross-domain which accelerates application of SA in multiple learning domains. Furthermore, they suggest that future researchers must integrate relationships between cognitive, emotional, and motivational learning aspects. In this context the present research bridges gap in students' sentiment detection, giving a new definition of multifaceted characteristics of concept to gain a deeper understanding of how sentiments provide an indication to educators for identifying whether a student is motivated or discouraged with virtual learning and has an intention to dropout. Thus, in this research sentiments are explored as a process that truly reflects students' learning circumstances by considering necessary features in VLE across multiple disciplines. Hence, in higher education context, sentiments are linked to students' cognitive, emotions, psychological and learning factors in students' behavior. The model is explored with four multiclass classification algorithms, to perform sentiment analysis on text and key phrases. After experimentation, the proposed MSDS system successfully outperforms in each case compared to other methodologies. The best performance for mean accuracy rate was achieved by Logistic Regression for device efficiency with 98.49%, Linear SVC for cognitive behavior and technological expertise with 93.58% and 92.08%.

III. RESEARCH METHODOLOGY

The main objective of this research is detecting the students' intention of dropping out from virtual learning courses by considering their text sentiments. There are several methods for dropout detection. These methods are based on ML techniques and require students' activity records for training and creating predictive models depending on the features extracted from raw data. In this research, four machine learning algorithms are implemented, and its performance metrics are evaluated. The subsequent research questions are considered:

- 1) To develop a Multifaceted Sentiment Detection System (MSDS) architecture for predicting dropout students in VLE.
- 2) To estimate the evaluation steps for measuring the proposed architectures' efficiency in terms of mean accuracy rate and standard deviation using four multi-class classification algorithms namely RF, Linear SVC, MNB, LR.
- 3) To investigate the most suited ML algorithm for classifying students' sentiments depending on their multifaceted characteristics and visualize the results to educators as an early intervention.

IV. MULTIFACETED SENTIMENT DETECTION SYSTEM (MSDS) ARCHITECTURE

The MSDS framework detects the students' sentiments with information gathered from VLE interactions. MSDS architecture attempts to address sentiment analysis approach for characterizing the sentiments across multifaceted subjects namely device efficiency, cognitive behavior, technological expertise with cloud platform usage. The research flow begins with data extraction from comments posted by learners of

VLE. The system reads the data stored with .csv (Comma Separated Values) format. Next, data pre-processing techniques are implemented to texts, namely removal of punctuation marks, stop words, etc. Feature extraction methods like count vectorizer, Term Frequency and Inverse Document Frequency (TF-IDF) are employed. To identify the

real opinions of learners, sentiment classification is performed based on machine learning approaches. Then the flow ends with detecting students' sentiments with three polarities: positive, negative, and neutral. The model is assessed with various evaluation metrics namely accuracy, recall, precision, and F1-Score. Fig. 1 depicts proposed (MSDS) model.

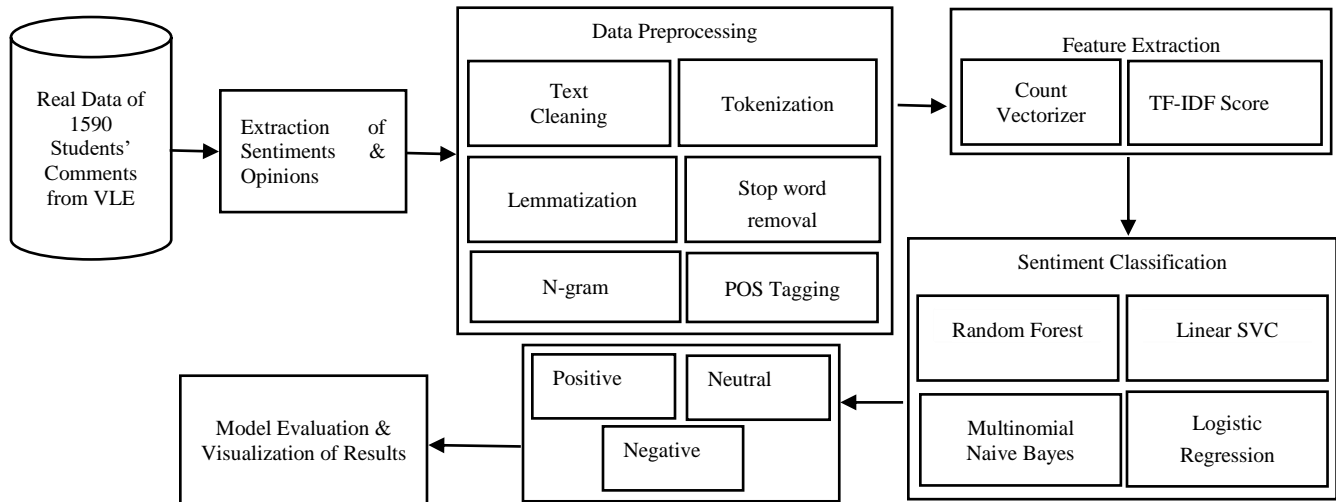


Fig. 1. Framework of proposed (MSDS) research model.

A. Data Collection from Participants

SA can be integrated into a virtual learning environment that realizes real-time analysis of learners' feedback. Real data was collected from students belonging to various disciplines of numerous HEI throughout India. The data sources comprise dataset that were collected by educators through online questionnaire that was broadly classified into various factors namely demographic features, device usage characteristics, self-efficacy which was estimated with 5-point Likert scale, familiarity with cloud platforms using 4-point range for identifying sentiments based on their interactions with VLE. Totally, the dataset contains n=1590 comments, opinions and feedbacks posted by students of virtual learning classes. Depending on their multifaceted characteristics, three datasets are grouped each containing 530 comments for conducting experiments. The data instances are branched into multiple classes. The sentiments are labelled with score range from 0-2, representing negative, neutral, and positive.

B. Data Pre-processing

Text pre-processing is the process of cleaning and preparing text data, as machines can use that processed text to perform various tasks like analysis, prediction, etc. These procedures help to reduce the volume of data and processing time. The comments posted by students are in the form of natural English language and this must be converted into machine readable format. With text data, there are many challenges as it contains a lot of noisy, semi or unstructured data, punctuations, numbers, special characters, spelling mistakes, etc that require to be processed with NLP techniques. These are the pre-processing steps carried out for improving prediction performance.

1) *Removing punctuations and numbers, converting all characters into lowercase*: The basic pre-processing method is punctuations removal from textual data. This process helps to treat each text similarly. As numbers do not hold any vital information in the text it has to be removed. Then the input text is converted into same casing format namely lowercase.

2) *Tokenization*: This is a method of breaking the sentence into meaningful words and phrases. This process is achieved by using delimiters like white spaces and punctuation. Inbuilt NLTK (Natural Language Tool Kit) libraries have tokenization function capability to divide into words.

3) *Lemmatization*: Lemmatization is a text pre-processing technique in NLP models to break a word to its root meaning for identifying similarities. Text normalization changes the words down to a base form. WordNetLemmatizer() is assigned to a variable which is used to improve the algorithms' performance and facilitates to focus on meaning of the words.

4) *Stop word removal*: Next step is to eliminate stop words as they are commonly used words yet not useful for analysis and generally eliminated from text. This technique is applied for removing stop words namely I, am, you, she, he, the, a, an, so, what, etc. These words are not required for sentiment classification as it increases the dataset size and is considered irrelevant for results set. It's always better to remove these words for improving the model accuracy, reducing computational processes and complexity of data.

5) *N-gram*: In NLP this is a continuous series of n items created from a given text sample where items can be characters or words. This language model predicts the probability within any sequence of n words in the language.

There are multiple n-grams like unigram, bigram, trigram etc. Unigram is simplest n-gram where only one word is considered. Bigram takes two words at a time. Unigrams and bigrams are used in this research, where ngram_range = (1,2) is considered for feature processing.

6) *Part of Speech (POS) Tagging*: This process labels each word in text format for a particular POS depending on its context and definition. This reads the text from a language and assigns some token (POS) for every word. `nlk.pos_tag(tokenized_text)` is applied. Parts of Speech Adjective, Verb, Noun, and Adverb are considered.

C. Feature Extraction Techniques

In text data classification feature extraction plays a dominant role in reduction of feature space and increasing classifier's accuracy. This process converts data into features applied for ML model, as ML algorithms are programmed with numbers. Textual data is converted into vector form using two feature extraction techniques namely Count Vector, Term Frequency, and Inverse Document Frequency (TF-IDF) vector. Fig. 2 represents the cleaned text results achieved after data pre-processing and feature extraction with cognitive behavior data.

| | text | label | clean_text |
|---|---|-------|---|
| 0 | I am Undecided to manage to solve difficult pr... | 2 | [undecided, manage, solve, difficult, problem,... |
| 1 | I am Agree to manage to solve difficult proble... | 2 | [agree, manage, solve, difficult, problem, agr... |
| 2 | I am Agree to manage to solve difficult proble... | 2 | [agree, manage, solve, difficult, problem, agr... |
| 3 | I am Agree to manage to solve difficult proble... | 2 | [agree, manage, solve, difficult, problem, dis... |
| 4 | I am Undecided to manage to solve difficult pr... | 2 | [undecided, manage, solve, difficult, problem,... |

Fig. 2. Sample processed text for cognitive behavior.

D. Machine Learning Techniques for Sentiment Analysis

In the current situation, feedback is provided through grading methods. Although this grading method masks students' genuine feelings, the textual response gives them a chance to emphasize qualities. Three different ML algorithms SVM, MNB, and RF were implemented. Experimental outcomes suggest that MNB with 80% accuracy performs better than other classifiers [26]. The study applies SA to self-evaluation comments, a form of unstructured data that provides valuable information representing students' learning status over course duration for identifying at-risk students. SVM and Convolutional Neural Networks (CNN) were

1) *Count vectorizer*: CountVectorizer transforms a given text into a vector depending on frequency (count) of each word which occurs in entire text. This is beneficial when there are multiple such texts, and for converting each word in each text into vectors. These functionalities make it an extremely adaptable feature description module for text.

2) *TF-IDF score*: This is a weighting measure that quantifies the string relevance representations namely words, phrases, lemmas in a given document. Term Frequency (TF) describes in a document how frequently a term occurs against the total number of words in document, as in (1). IDF is measurement of selected term's weight in the document. This is given in (2).

$$tf = \frac{\text{number of occurrences of a term in document}}{\text{total number of terms in the document}} \quad (1)$$

$$idf = \log \left(\frac{\text{total number of given documents}}{\text{number of document with existing selected word}} \right) \quad (2)$$

applied to predict student performance. The proposed model provides an effectiveness represented by F-measure of 0.66 (SVM) and 0.78 (CNN). Best effectiveness was presented by CNN, achieving an F-measure value of 0.78. Experimental results demonstrated that applying sentiment analysis to unstructured data can significantly improve accuracy of early-stage predictions [27]. In this research, NLP techniques are employed to pre-process and vectorize the text data. Vectorized data is then applied for training various ML models. Following are the list of steps required to train the model for sentiment analysis:

Algorithm: Sentiment Analysis Using Machine Learning

1. Collect dataset from VLE to train and test machine learning model classifier.
 2. Pre-process the data for subsequent processing.
 3. Convert textual data into vector form using NLP techniques namely Count Vectorizer and TF-IDF.
 4. Divide the dataset into training and testing groups.
 5. Train the ML classifier with training data. Apply algorithms such as RF, Linear SVC, MNB, LR.
 6. Predict the polarity of test data.
 7. Evaluate the ML model using various metrics namely accuracy, precision, recall and F1-Score.
 8. Select the best algorithm for multifaceted text characteristics using sentiment multiclass classification.
-

1) *Random forest*: SA is used for analyzing unstructured text data for extracting positive or negative sentiments contained in student advisor's notes to predict college student dropout using RF model. The authors have quoted that their study is the first to apply NLP techniques for dropout prediction. RF classifier achieved 73% accuracy when compared to SVM, LR and CART [28]. M. A. Fauzi stated that due to the development of social media and online website reviews, SA is an efficient way for text classification. Experimental results confirmed that RF gives excellent performance with an average OOB score of 0.829 [29]. For increasing predictive power of Random Forest, this research utilizes hyperparameters $n_estimators=100$ and $max_depth=5$. $n_estimators$ is the number of trees the algorithm creates before taking maximum voting. A higher number of trees enhances the performance and makes predictions more stable. The hyperparameter max_depth represents maximum depth of each decision tree in the forest.

2) *Linear Support Vector Classifier (SVC)*: [30] suggests an opinion analysis system for amazon review for identifying comments received from UCI Website either positive or negative. The proposed approach is applied to review dataset and obtained an accuracy of 91% with Linear SVC over Naive Bayes and Voting. Authors have proposed that Linear SVC takes lesser execution time of about 0.0972s to test samples and provides output which is better than other classifiers namely Naive Bayes, Logistic Regression, Decision Tree [31]. The Linear SVC method is a faster implementation of Support Vector Classification that applies a linear kernel function for performing classification. This performs well for NLP based text classification tasks with a large number of samples. Linear SVC in scikit-learn library does not provide `predict_proba` function, instead `decision_function` is used that predicts the confidence scores for samples, which is the signed distance of that sample to hyperplane.

3) *Multinomial Naive Bayes (MNB)*: This is a type of Naive Bayes (NB) classifier that finds probabilities of classes assigned to text that uses joint probabilities of words and classes and is often used as a baseline for sentiment analysis. MNB achieves significant results for text categorization with 90% accuracy. MNB algorithm is a fast, easy-to-implement, and modern text categorization algorithm [32]. Social networking has developed into a tool that is useful for gathering vital information about individuals. The sentiment of the user can be determined by extracting user comments via an API and feeding them to an algorithm that detects whether they are positive or negative. Results obtained show that Multinomial Naive Bayes performs good with classification accuracy of 85% when compared to SVM, Random Forest and Decision Tree [33]. MNB is specifically beneficial for problems that involve text data with discrete features, namely word frequency counts. This works on the principle of Bayes theorem and assumes that features are conditionally independent given the class variables. The computation is

performed by adding logarithms of probabilities, as in (3). The class with the highest log probability score is most feasible.

$$C_{NB} = \mathop{\text{arg}}_{c \in C} \max [\log P(c) + \sum_{1 \leq k \leq n_d} \log P(t_k | c)] \quad (3)$$

4) *Logistic regression*: In a variety of situations finding the polarity of reviews is useful. NLP techniques can be applied for analyzing the reviews and optimizing the strategic decision making. TFIDF is used for feature selection and Logistic Regression for classification as it uses sigmoid activation function. The LR with grid search model classifies the text accurately with 94% [34]. Reviews and feedback are deciding factors for understanding the opinions of users. SA is a method of information extraction for improving the work by review analysis. The users' review text is cleaned by Count Vectorizer and TFIDF. Sentiment prediction is done using various classifiers for reviews without ratings. Logistic Regression accuracy is 93% and more compared to NB, MNB, Bernoulli classifiers [35]. Logistic Regression is a simple classification algorithm that can be generalized to multiple classes. The LR uses sigmoid function, which is given by (4),

$$\text{sig}(t) = \frac{1}{1+e^{-t}} \quad (4)$$

These popular classifiers do not directly support multiclass classification problems. There are certain heuristic methods available which can split multiclass classification datasets into several binary classification datasets. The binary classifiers are then trained on each binary classification problem and predictions are made operating the model which is most confident. To implement this method for multi-class classification, the `OneVsRestClassifier` method is used.

V. EXPERIMENTAL RESULTS

In this research, implementation is done with efficient NLTK, Scikit-learn libraries by applying the proposed machine learning techniques. NLTK is used to perform regular expression patterns & tokenization to parse text, lemmatization, stop word removal, n-grams, and POS tagging. Vectorization and classification were accomplished by Scikit-learn. Data was preprocessed, vectorized with TF-IDF, and classified with four multiclass machine learning algorithms. Dataset was split into 75% training set and 25% test set. 5-fold Cross-Validation was used for training classification model. RF, Linear SVC, MNB and LR were applied as they are the most popular machine learning classifiers used to analyze students' opinions. Depending on the classification task different metrics were used to measure the classifiers' performance.

A. Multi-class Classification Positive/Neutral/Negative

The performance of MSDS framework is evaluated using various metrics. Accuracy is one of the popular metrics for multi-class classification. This is the ratio of the number of accurately classified instances to total number of instances. Macro-average precision calculates precision for all classes individually and then averages them. Weighted average precision calculates precision per class but considers no. of

samples of each class in data. Macro-average recall score computes the arithmetic mean of all recall scores of different classes. Weighted-average recall computes recall per class but considers the number of samples of each class in the data. Macro average F1score computes arithmetic mean of all per-class F1 scores. Weighted average F1 score calculates mean of all per-class F1 scores by considering each class's support.

$$Accuracy = \frac{(TP+TN)}{(TP +TN+FP+FN)} \quad (5)$$

$$Macro - average Precision = \frac{Precision_1+Precision_2+Precision_3}{3} \quad (6)$$

$$Weighted - average Precision = \frac{w_1*Precision_1+w_2*Precision_2+w_3*Precision_3}{Total\ number\ of\ samples} \quad (7)$$

$$Macro - average Recall = \frac{Recall_1+Recall_2+Recall_3}{3} \quad (8)$$

$$Weighted - average Recall = \frac{w_1*Recall_1+w_2*Recall_2+w_3*Recall_3}{Total\ number\ of\ samples} \quad (9)$$

$$Macro - average F1 - Score = \frac{F1-Score_1+F1-Score_2+F1-Score_3}{3} \quad (10)$$

$$Weighted - average F1 - Score = \frac{w_1*F1-Score_1+w_2*F1-Score_2+w_3*F1-Score_3}{Total\ number\ of\ samples} \quad (11)$$

1) *Device efficiency*: The first perspective analyzes the students' device usage characteristics as VLE increases the portability of learning processes through smart devices. To investigate new educational opportunities that result from expanding device access to VLE, enables users to establish more fleeting links to virtual campus for providing instructional procedures built on a model that is significantly less time and space consuming. The parameters used for configuring each of the algorithms implemented are the type of smart device, mode of device availability, device connectivity, number of hours the device is connected online. The optimized smart device environment will provide services to the educational community, regardless of their functional and cognitive elements. N. A. S. Remali with other researchers, states that its' a challenging task for educational institutions to identify learners' opinions and difficulties during online education as few students may have poor internet connection problem, lack of bandwidth and the environment makes learners not to concentrate on class. Thus, SA is utilized to assess students' opinions (positive, neutral, and negative) of virtual learning [36]. Considering device efficiency features, Table I shows the evaluation metrics of four multi-class classifiers with LR model showing high accuracy rate of 98%. The cross-validation method, with cv=5, was performed to cross-validate baseline models with feature extractors of TF-IDF and CV. Mean accuracy and standard deviation for each fold validates performance of

models, depicted in Table II. The highest mean accuracy of Logistic Regression is 98.49%. Fig. 3 shows mean accuracy with 5-fold cross validation. Results of sentiment distribution are visualized in Fig. 4. Fig. 5 displays classification metrics of device efficiency text characteristics.

TABLE I. CLASSIFICATION OF DEVICE EFFICIENCY

| Evaluation Metrics of Classifiers | RF | Linear SVC | MNB | LR |
|-----------------------------------|------|------------|------|------|
| Accuracy | 0.96 | 0.96 | 0.97 | 0.98 |
| Precision-macro avg | 0.58 | 0.58 | 0.60 | 0.61 |
| Precision-weighted avg | 0.95 | 0.95 | 0.96 | 0.97 |
| Recall-macro avg | 0.66 | 0.66 | 0.66 | 0.66 |
| Recall-weighted avg | 0.96 | 0.96 | 0.97 | 0.97 |
| F1-score macro avg | 0.61 | 0.61 | 0.63 | 0.64 |
| F1-score weighted avg | 0.96 | 0.96 | 0.96 | 0.97 |

TABLE II. MEAN ACCURACY AND SD

| ML Models | Mean Accuracy | Standard Deviation |
|-------------------------|---------------|--------------------|
| Random Forest | 0.981132 | 0.011554 |
| Linear SVC | 0.983019 | 0.013993 |
| Multinomial Naive Bayes | 0.981132 | 0.006671 |
| Logistic Regression | 0.984906 | 0.010756 |

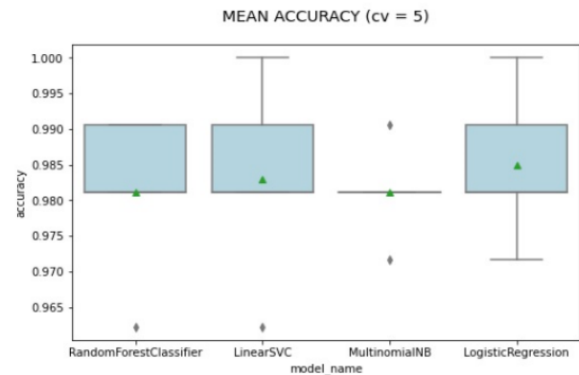


Fig. 3. Mean accuracy for device efficiency.

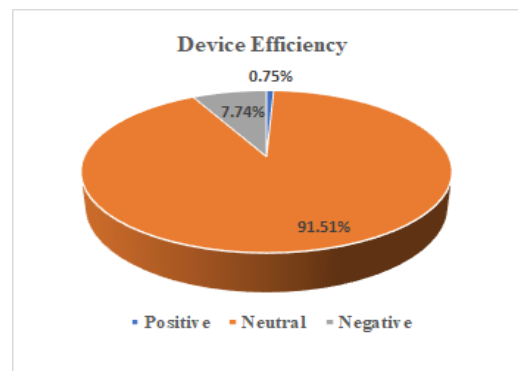


Fig. 4. Device efficiency sentiment polarity distribution.

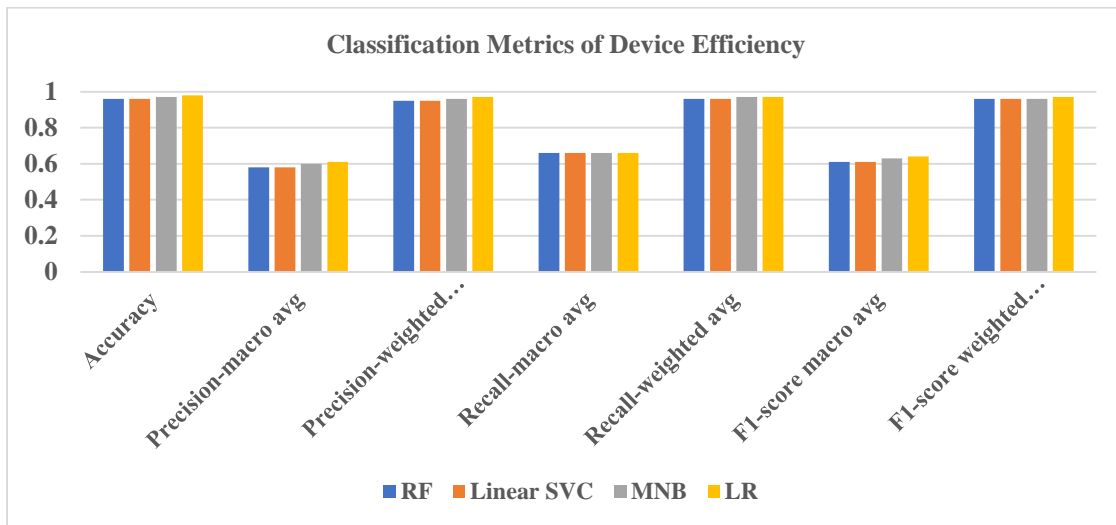


Fig. 5. Classification metrics of device efficiency text characteristics.

2) *Cognitive behavior*: The encompassing perspective analyzes the students’ psychological cognitive behavioral model through self-efficacy theory stated by Psychologist Albert Bandura. The parameters used for configuring each of the algorithms implemented are manageability, finding means & ways, stick to aims & accomplishments, handling unforeseen situations, investing effort, finding several solutions, handling whatever situations in online learning. In addition, the researchers in [37] explored that predictive power and feature generalization of cognitive skill score estimates possibility of learners’ success or failure in higher education course which provides suitable intervention to facilitate learners. Cognitive skill score proves to be efficient in identifying students’ performance when exact metrics correlated to learning activities and students’ social behavior are unavailable. Table III shows evaluation metrics of four multi-class classifiers for self-efficacy with Linear SVC showing a high accuracy rate of 93%. With cv=5, mean accuracy and standard deviation for each fold validates the models’ performance depicted in Table IV. The mean accuracy of Linear SVC is higher with 93.58% when compared to RF, MNB and LR. Fig. 6 presents mean accuracy with cv=5 for cognitive behavioral features. Sentiment analysis outcomes are visualized in Fig. 7. Classification metrics of self-efficacy text characteristics are displayed in Fig. 8.

| | | | | |
|-----------------------|------|------|------|------|
| F1-score macro avg | 0.61 | 0.62 | 0.57 | 0.59 |
| F1-score weighted avg | 0.91 | 0.93 | 0.87 | 0.90 |

TABLE IV. MEAN ACCURACY AND SD

| ML Models | Mean Accuracy | Standard Deviation |
|-------------------------|---------------|--------------------|
| Random Forest | 0.932075 | 0.018147 |
| Linear SVC | 0.935849 | 0.010334 |
| Multinomial Naive Bayes | 0.900000 | 0.023679 |
| Logistic Regression | 0.932075 | 0.012300 |

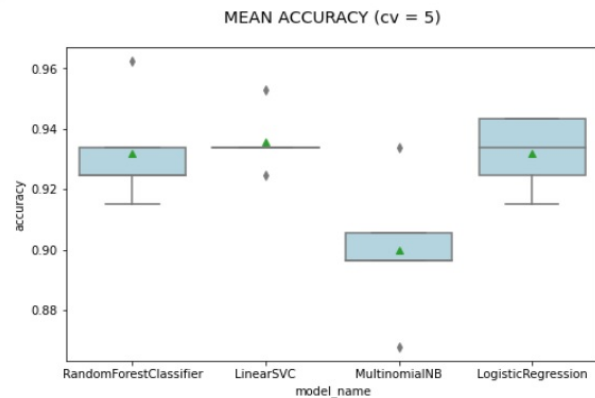


Fig. 6. Mean accuracy for behavioral features.

TABLE III. CLASSIFICATION OF SELF-EFFICACY

| Evaluation Metrics of Classifiers | RF | Linear SVC | MNB | LR |
|-----------------------------------|------|------------|------|------|
| Accuracy | 0.92 | 0.93 | 0.88 | 0.90 |
| Precision-macro avg | 0.61 | 0.62 | 0.59 | 0.60 |
| Precision-weighted avg | 0.91 | 0.92 | 0.88 | 0.90 |
| Recall-macro avg | 0.61 | 0.62 | 0.56 | 0.58 |
| Recall-weighted avg | 0.92 | 0.93 | 0.88 | 0.90 |

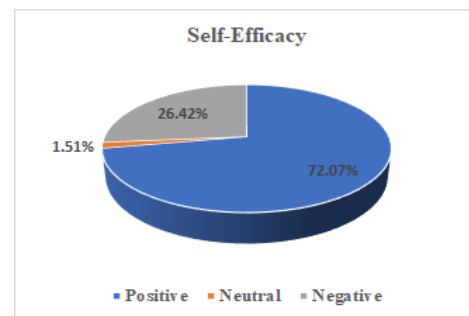


Fig. 7. Self-efficacy sentiment polarity distribution.

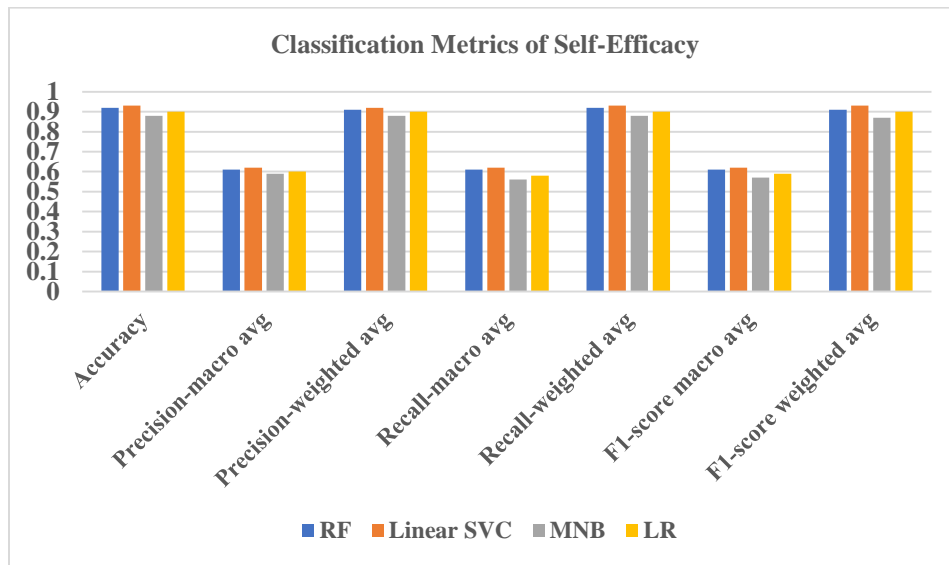


Fig. 8. Classification metrics of self-efficacy text characteristics.

3) *Technological expertise of cloud learning platform usage*: Outside the classroom, cloud platforms prepare students to make reasonable study schedules, thereby promoting self-directed learning, innovation, collaboration, and ease of accessibility. The parameters used emphasize on the platform learners are familiar with accessing online teaching materials namely Google Classroom, Google Meet, Zoom, Facebook Classroom, Twitter etc. Moreover, J. Zhang in [38] suggested that educators could discuss with learners through the cloud class even after online class and they can set time-limited online test for enhancing results. With students' feedback the course assistants can track their problems in online teaching. Using bullet screen educators can communicate useful course contents to learners in time, whereby the emotional communication between them is deepened. Table V illustrates the evaluation metrics of four classification algorithms for technological expertise with Linear SVC depicting a high accuracy rate of 92%. Table VI depicts mean accuracy and standard deviation for each fold to validate models' performance. The highest mean accuracy score of Linear SVC is 92.08%. Fig. 9 visualizes mean accuracy with $cv=5$ for technological familiarity with cloud platforms. The sentiment analysis results are presented in Fig. 10. Fig. 11 shows classification metrics of technological expertise text characteristics.

| | | | | |
|-----------------------|------|------|------|------|
| F1-score macro avg | 0.55 | 0.61 | 0.48 | 0.48 |
| F1-score weighted avg | 0.88 | 0.89 | 0.87 | 0.87 |

TABLE VI. MEAN ACCURACY AND SD

| ML Models | Mean Accuracy | Standard Deviation |
|-------------------------|---------------|--------------------|
| Random Forest | 0.918868 | 0.010756 |
| Linear SVC | 0.920755 | 0.008438 |
| Multinomial Naive Bayes | 0.916981 | 0.004219 |
| Logistic Regression | 0.916981 | 0.004219 |

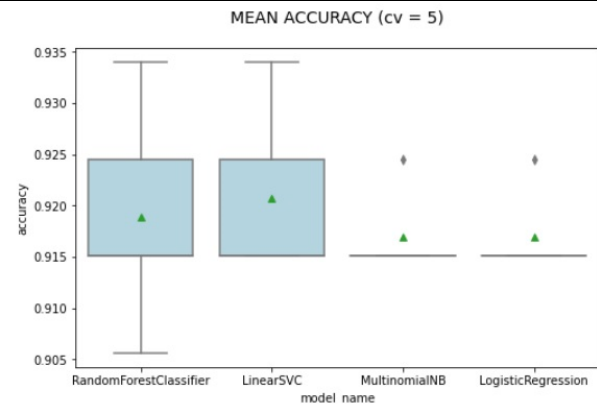


Fig. 9. Mean accuracy for technological expertise.

| Evaluation Metrics of Classifiers | RF | Linear SVC | MNB | LR |
|-----------------------------------|------|------------|------|------|
| Accuracy | 0.91 | 0.92 | 0.91 | 0.91 |
| Precision-macro avg | 0.71 | 0.79 | 0.45 | 0.45 |
| Precision-weighted avg | 0.88 | 0.90 | 0.83 | 0.83 |
| Recall-macro avg | 0.54 | 0.58 | 0.50 | 0.50 |
| Recall-weighted avg | 0.91 | 0.92 | 0.91 | 0.91 |

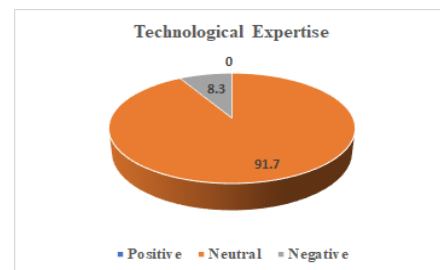


Fig. 10. Technological expertise sentiment polarity distribution.

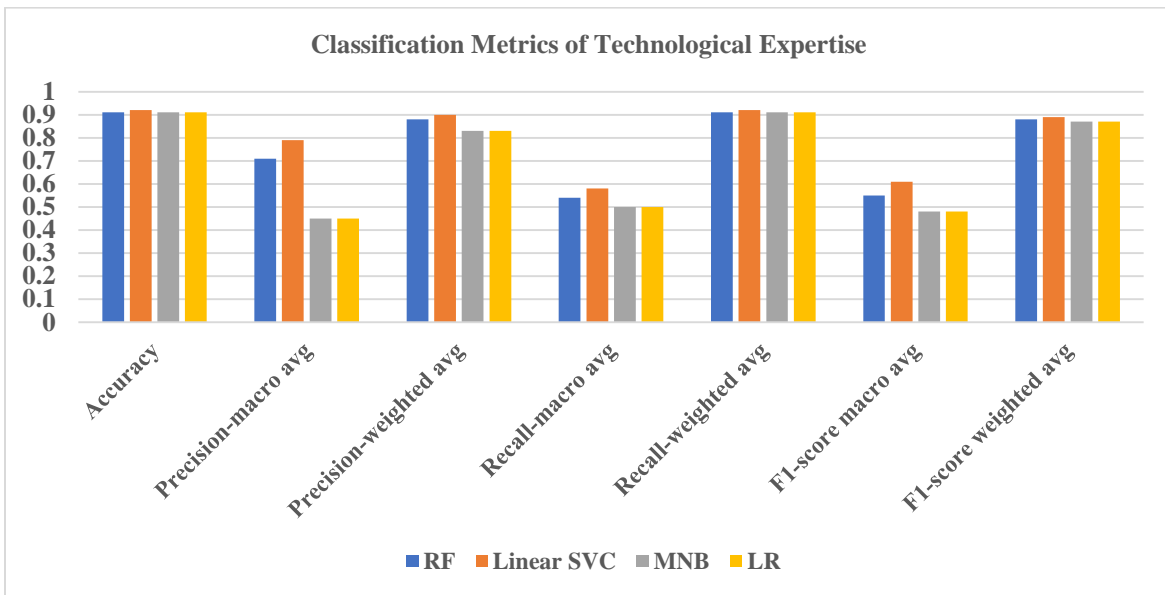


Fig. 11. Classification metrics of technological expertise text characteristics.

B. Performance Evaluation using ROC and AUC

To analyze the performance of MSDS architecture, ROC curve can be used as it measures the classifier’s predictive quality. Tradeoff between classifier’s sensitivity and specificity can be visualized by the user using a ROCAUC (Receiver Operating Characteristic/Area Under the Curve). ROC curve exhibits true positive rate on Y axis and false positive rate on X axis when plotted. Ultimate point is the top-left corner of plot: false positives are 0 and true positives are 1, which directs to another metric, AUC. The higher AUC represents a better model. However, this is vital to examine “steepness” of curve, as this illustrates maximization of sensitivity while minimizing specificity. ROC curve is extensively used in this research to describe diagnostic accuracy and for finding the best cut-off value for a model trained through multi-class machine learning techniques. ROC curves for cognitive behavior through self-efficacy characteristics using RF, Linear SVC, MNB, LR are visualized in Fig. 12, Fig. 13, Fig. 14, Fig. 15 respectively as this is main framework emphasizing the learners’ psychological cognitive behavioral patterns for risk prediction. The performance comparisons of ROC curves show that Linear SVC model performs the best for self-efficacy data.

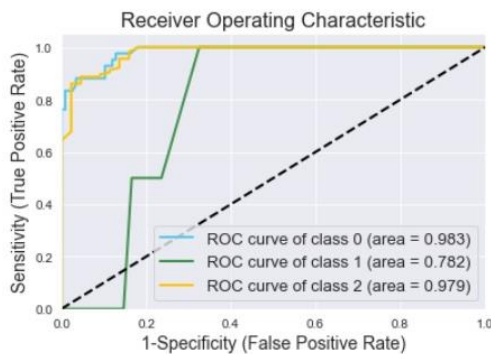


Fig. 12. ROC for random forest.

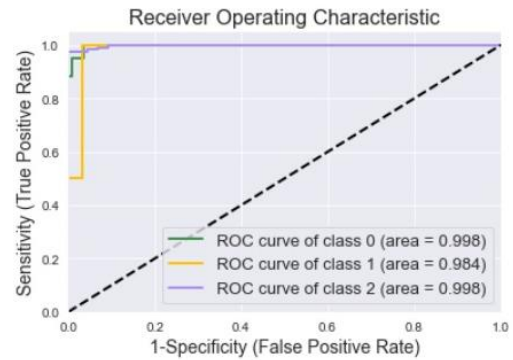


Fig. 13. ROC for linear SVC.

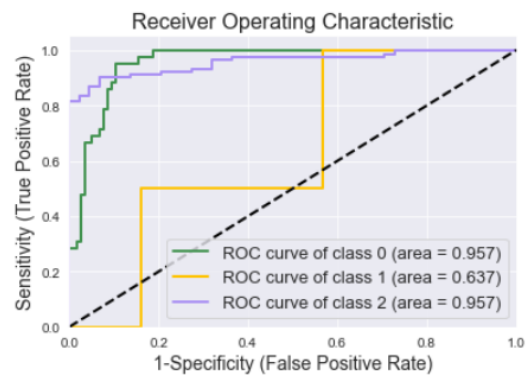


Fig. 14. ROC for MNB.

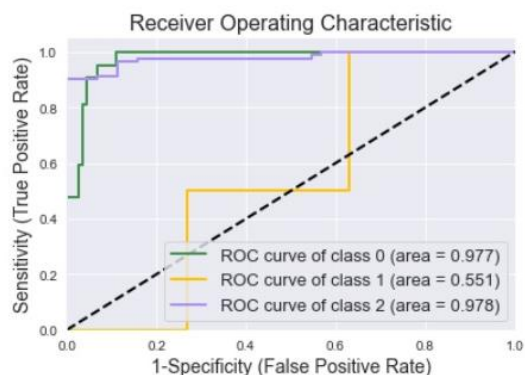


Fig. 15. ROC for logistic regression.

C. Data Visualization

Sentiment analysis encompasses a variety of SA tasks, including subjectivity detection and emotion analysis. This reflects the wide range of user tasks and data domains used in SA research and applications, which include everything from social media and news monitoring to theoretical linguistic research and NLP. This implies the use of numerous visual channels and interpretations. The visual representation involving polarity data includes word clouds [39]. Word cloud represents a powerful textual data visualization technique that enables quickly identifying the words that are most often used within a particular body of text. These are frequently employed communication tools for processing, analyzing qualitative sentiment data. When educators need to visualize the opinions of learners, word cloud can be used for identifying the messages posted by them. This visual representation gives a clear way for educators to easily interpret the messages. Fig. 16 represents the word cloud visualization of multifaceted factors.



Fig. 16. Word clouds of device efficiency, cognitive behavior, technological expertise of cloud platform Usage.

VI. CONCLUSION

Multifaceted Sentiment Detection System (MSDS) is proposed in this research to predict students' dropout using multiclass classification algorithms. For addressing this goal, the research analyzes higher education students' comments

and reviews while attending virtual classes through VLE that have been classified and analyzed using machine learning algorithms and word clouds. The results obtained from the research explore multideterminant characteristic features of respondents' device efficiency, psychological cognitive behavior, and technological knowledge of cloud platform usage. This finding mainly contributes to improving the classifiers' predictive ability of college student dropout through text classification. The proposed method obtained the highest mean accuracy results for device efficiency using Logistic Regression with 98.49%, Linear SVC for cognitive behavior with 93.58% and technological expertise with 92.08%. Thus, this novel assessment mechanism MSDS, aims in exploring efficient sentiment classification by proposing multiclass machine learning algorithms to avoid dropouts. Furthermore, experimental outcomes demonstrate that proposed system has obtained better accuracy results when comparing to previous methods. For future work, the algorithms can be validated and applied against other computational methods such as deep learning algorithms for superior performance.

REFERENCES

- [1] Pooja and R. Bhalla, "A Review Paper on the Role of Sentiment Analysis in Quality Education," *SN Comput. Sci.*, vol. 3, no. 6, pp. 1–9, 2022, doi: 10.1007/s42979-022-01366-9.
- [2] T. Zarra, R. Chiheb, R. Faizi, and A. El Afia, "Cloud computing and sentiment analysis in E-learning systems," *Proc. 2016 Int. Conf. Cloud Comput. Technol. Appl. CloudTech 2016*, pp. 171–176, 2017, doi: 10.1109/CloudTech.2016.7847695.
- [3] F. Dalipi, K. Zdravkova, and F. Ahlgren, "Sentiment Analysis of Students' Feedback in MOOCs: A Systematic Literature Review," *Front. Artif. Intell.*, vol. 4, pp. 1–13, 2021.
- [4] F. Campos, V. Stroele, R. Braga, and T. T. Torrent, "Using Sentiment Analysis to Identify Student Emotional State to Avoid Dropout in E-Learning," *International Journal of Distance Education Technologies*, vol. 20, no. 1, pp. 1–24, 2022, doi: 10.4018/IJDET.305237.
- [5] N. H. Harani and C. Prianto, "Sentiment Analysis of Student Emotion During Online Learning Using Recurrent Neural Networks (RNN)," *International Journal of Information System & Technology*, vol. 5, no. 3, pp. 299–307, 2021.
- [6] M. M. Ali, "Arabic sentiment analysis about online learning to mitigate covid-19," *J. Intell. Syst.*, vol. 30, no. 1, pp. 524–540, 2021, doi: 10.1515/jisys-2020-0115.
- [7] N. S. Izzati Samsari, M. Mohamad, and A. Selamat, "Sentiment Analysis on Students Stress and Depression Due to Online Distance Learning During the COVID-19 Pandemic," *Math. Sci. Informatics J.*, vol. 3, no. 1, pp. 66–74, 2022, doi: 10.24191/mij.v3i1.18273.
- [8] R. Archana and B. Kishore, "Role of Sentiment Analysis in Education Sector in the Era of Big Data: a Survey," *International Journal of Latest Trends in Engineering and Technology*, pp. 022–024, 2017.
- [9] J. Zhou and J. M. Ye, "Sentiment analysis in education research: a review of journal publications," *Interactive learning environments*, pp. 1-13, 2020, doi: 10.1080/10494820.2020.1826985.
- [10] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artificial Intelligence Review*, vol. 55, no. 7, pp. 5731-5780, 2022.
- [11] A. Raut and R. K. Pandey, "Sentiment Analysis using Optimized Feature Sets in Different Twitter Dataset Domains," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 11, pp. 3035-3039, 2019, doi: 10.35940/ijitee.k2195.0981119.
- [12] S. Ulfa, R. Bringula, C. Kurniawan, and M. Fadhli, "Student Feedback on Online Learning by Using Sentiment Analysis: A Literature Review," *Proc. - 2020 6th Int. Conf. Educ. Technol. ICET 2020*, no. 1, pp. 53–58, 2020, doi: 10.1109/ICET51153.2020.9276578.

- [13] K. Riegel and T. Evans, "Student achievement emotions: Examining the role of frequent online assessment," *Australas. J. Educ. Technol.*, vol. 37, no. 6, pp. 75–87, 2021, doi: 10.14742/ajet.6516.
- [14] K. Wang and Y. Zhang, "Topic sentiment analysis in online learning community from college students," *Journal of Data and Information Science*, vol. 5, no. 2, pp. 1–33, 2020, doi: 10.2478/jdis-2020-0009.
- [15] T. D. Pham *et al.*, "Natural language processing for analysis of student online sentiment in a postgraduate program," *Pacific J. Technol. Enhanc. Learn.*, vol. 2, no. 2, pp. 15–30, 2020, doi: 10.24135/pjtel.v2i2.4.
- [16] L. M. Nkomo and B. K. Daniel, "Sentiment Analysis of Student Engagement with Lecture Recording," *TechTrends*, vol. 65, no. 2, pp. 213–224, 2021, doi: 10.1007/s11528-020-00563-8.
- [17] S. Pande, "An overview of sentiment analysis using Azure Cognitive Services," *International Journal of Management, IT & Engineering*, vol. 9, no. 10, pp. 123–127, 2019.
- [18] P. Rajesh and D. Akila, "Sentimental Analysis on E-learning Videos for Learner's Opinion Using Machine Learning Methodology - Support Vector Machine," *Lecture Notes in Networks and Systems*, 2023, Volume 467, pp. 191–200, doi: 10.1007/978-981-19-2538-2_18.
- [19] M. A. dos Santos Alencar, J. F. de Magalhaes Netto, and F. de Morais, "A Sentiment Analysis Framework for Virtual Learning Environment," *Appl. Artif. Intell.*, vol. 35, no. 7, pp. 520–536, 2021, doi: 10.1080/08839514.2021.1904594.
- [20] K. Ravi, V. Siddeshwar, V. Ravi, and L. Mohan, "Sentiment analysis applied to Educational Sector," *2015 IEEE Int. Conf. Comput. Intell. Comput. Res. ICCIC 2015*, no. February 2017, 2016, doi: 10.1109/ICCIC.2015.7435667.
- [21] O. Abiodun Ayeni, A. Mercy, T. A.F, and M. A.S, "Web-Based Student Opinion Mining System Using Sentiment Analysis," *Int. J. Inf. Eng. Electron. Bus.*, vol. 12, no. 5, pp. 33–46, 2020, doi: 10.5815/ijieeb.2020.05.04.
- [22] Z. Kamisli Ozturk, Z. I. Erzurum Cicek, and Z. Ergul, "Sentiment analysis: An application to Anadolu University," *Acta Phys. Pol. A*, vol. 132, no. 3, pp. 753–755, 2017, doi: 10.12693/APhysPolA.132.753.
- [23] I. D. Wahyono *et al.*, "Emotion Detection based on Column Comments in Material of Online Learning using Artificial Intelligence," *International Journal of Interactive Mobile Technologies*, vol. 16, no. 3, pp. 82–91, 2022, doi: 10.3991/IJIM.V16I03.28963.
- [24] A. F. Ab Nasir *et al.*, "Text-based emotion prediction system using machine learning approach," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 769, no. 1, 2020, doi: 10.1088/1757-899X/769/1/012022.
- [25] B. N. Supriya and C. B. Akki, "Sentiment prediction using enhanced xgboost and tailored random forest," *Int. J. Comput. Digit. Syst.*, vol. 10, no. 1, pp. 191–199, 2021, doi: 10.12785/ijcds/100119.
- [26] D. D. Dsouza, Deepika, D. P. Nayak, E. J. Machado, and N. D. Adesh, "Sentimental analysis of student feedback using machine learning techniques," *Int. J. Recent Technol. Eng.*, vol. 8, no. 1 Special Issue 4, pp. 986–991, 2019.
- [27] L. C. Yu *et al.*, "Improving early prediction of academic failure using sentiment analysis on self-evaluated comments," *J. Comput. Assist. Learn.*, vol. 34, no. 4, pp. 358–365, 2018, doi: 10.1111/jcal.12247.
- [28] J. Jayaraman, "Predicting student dropout by mining advisor notes," *Educationaldatamining.Org*, no. Edm, pp. 629–632, 2020, https://educationaldatamining.org/files/conferences/EDM2020/papers/paper_233.pdf.
- [29] M. A. Fauzi, "Random Forest Approach for sentiment analysis in Indonesian language," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 12, no. 1, pp. 46–50, 2018, doi: 10.11591/ijeecs.v12.i1.pp46-50.
- [30] S. S. Sikarwar and N. Tiwari, "Analysis The Sentiments Of Amazon Reviews Dataset By Using Linear SVC And Voting Classifier," *International Journal of Scientific & Technology Research*, vol. 9, no. 06, pp. 461–465, 2020.
- [31] N. Sultana, P. Kumar, M. R. Patra, S. Chandra and S.K. Safikul Alam, "Sentiment Analysis System for Product Review: A Survey: A Survey," *ICTACT Journal on Soft Computing*, vol. 09, no. 03, pp. 268–278, 2019.
- [32] M. Abbas, K. Ali Memon, and A. Aleem Jamali, "Multinomial Naive Bayes Classification Model for Sentiment Analysis," *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, vol. 19, no. 3, pp. 62–67, 2019.
- [33] S. Bachhety, S. Dhingra, R. Jain, and N. Jain, "Improved Multinomial Naive Bayes Approach for Sentiment Analysis on Social Media," *Proceedings of 4th International Conference on Computers and Management (ICCM) 2018*, pp. 121–127, 2018.
- [34] P. S. Reddy, D. R. Sri, C. S. Reddy, and S. Shaik, "Sentimental Analysis using Logistic Regression," *International Journal of Engineering Research and Applications*, vol. 11, no. 7, pp. 36–40, 2021, doi: 10.9790/9622-1107023640.
- [35] S. P. Sheela, "Sentiment Analysis and Prediction of Online Reviews with Empty Ratings," *Int. J. Appl. Eng. Res.*, vol. 13, no. 14, pp. 11532–11539, 2018.
- [36] N. A. S. Remali, M. R. Shamsuddin, and S. Abdul-Rahman, "Sentiment Analysis on Online Learning for Higher Education During Covid-19," *2022 3rd Int. Conf. Artif. Intell. Data Sci. Championing Innov. Artif. Intell. Data Sci. Sustain. Futur. AiDAS 2022 - Proc.*, pp. 142–147, 2022, doi: 10.1109/AiDAS56890.2022.9918788.
- [37] S. Md and S. Krishnamoorthy, "Student performance prediction, risk analysis, and feedback based on context-bound cognitive skill scores," *Educ. Inf. Technol.*, vol. 27, no. 3, pp. 3981–4005, 2022, doi: 10.1007/s10639-021-10738-2.
- [38] J. Zhang, "Research on Sentiment Analysis and Satisfaction Evaluation of Online Teaching in Universities During Epidemic Prevention," *Frontiers in Psychology*, vol. 12, 2021, doi: 10.3389/fpsyg.2021.738776.
- [39] K. Kucher, C. Paradis, and A. Kerren, "The state of the art in sentiment visualization," *Comput. Graph. Forum*, vol. 37, no. 1, pp. 71–96, 2018, doi: 10.1111/cgf.13217.