International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Sentiment Analysis of Students Feedback Using Lexicon Based Method and Hybrid Machine Learning Method

[1]Ms. Shital A. Patil, [2]Dr. Krishnakant P. Adhiya, [3]Dr. Girishkumar K. Patnaik

Submitted: 19/11/2023 **Revised**: 31/12/2023 **Accepted**: 11/01/2024

Abstract: Sentiment classification and opinion mining are the most widely used NLP applications for detection of human intention and reviews. The use of social media is currently in high demand for the purpose of sharing information and interaction. This includes the textual and visual dissemination of feedback, emotions, and ideas. The most widely used social networks, such as Facebook, Twitter, and Instagram have become significant means for the exchange of knowledge and interaction among users who share a common interest group. The form of communication platform among these users may consist of large texts. In this day and age of advanced technological advances, people like to keep in touch with those they care about most by using a variety of social networking platforms. Many users are able to communicate their thoughts and feedback through on specific platform. In this paper we proposed a sentiment classification on the student feedback dataset using a hybrid machine learning algorithm in this work. The 9 different feature extraction methods, such as TF-IDF, N-Gram, NLP-based dependency features, and lexicon-based techniques, are used. The different machine learning algorithms are Naïve Bayes, ANN, SVM and HML classification algorithm. In a comparative analysis of the proposed system, the HML obtained higher classification accuracy of 98.20% with the lexicon-based method. The proposed model shows around 3-4% higher accuracy than the existing sentiment classification methodologies.

Keywords: *sentiment classification, feedback analysis, lexicon-based method, machine learning, feature extraction, feature selection, polarity detection.*

1. Introduction

Feedback is a message regarding an entity's previous conduct that the entity may use to assess its present and future behavior to produce the desired outcome. Feedback is a procedure that aids the company in keeping track of, assessing, and controlling the general working environment. Effective feedback practices provide the organization information that helps it enhance the teaching and learning process. The form may be categorized as textual or grading (Likert-scale based score) depending on the response provided by the students. Students are given questions with a Likert-scale based score and asked to respond to those questions using a rating-based scale. This method does not accurately reflect the students' feelings since it primarily focuses on a question that is linked to the same subject.

The textual feedback approach is utilized to determine the precise emotion of the pupils. Students are given a series of questions in this text form, and they must provide complete sentences in their responses. In order to resolve problems with their organizational structure, it is beneficial for both the academic administration and the teacher. Google Forms are used in this study to get student responses with a range of opinions. The goal is to identify opinion expressions and categorize them as neutral, positive, or negative using machine learning methods.

Sentimental analysis is a technique for figuring out the emotion conveyed in texts. Given the current circumstances that people throughout the globe are facing, the necessity for text sentiment analysis has become increasingly critical. There are typically three methods used in sentimental analysis. They use a hybrid method, machine learning, and lexicon-based techniques. Using a sentiment lexicon is one of the steps involved in the lexicon-based method, which is used to determine the polarity of the text's sentiment. The machine learning method is used to label the training data set using a sentiment lexicon. Then, using this model, testing data is assessed. The rest of the paper describes below; section II discusses the literature survey on various methods used for sentiment classification and different feature extraction, feature selection, and classification techniques. Section III details a description of the proposed model with research methodology. It also discusses the architecture of proposed systems and modules of entire executions. Section VI elaborates on an algorithm used for the proposed implementation. Section V shows the proposed system results and

[1]*Research Scholar, Department of Computer Engineering, SSBT, COET, Jalgaon,*
shital.a.patil2014@gmail.com
[2]*Professor, Department of Computer Engineering, SSBT, COET, Jalgaon,*
adhiyakp@gmail.com
[3]*Professor, Department of Computer Engineering, SSBT, COET, Jalgaon,*
patnaik.girish@gmail.com

comparative analysis with various state-of-the-art systems. Finally, section VI discusses about conclusion and future work of the proposed system.

2. Literature Review

In recent years, sentiment analysis and mined opinions have drawn a lot of interest from researchers in the field of education. Krishnaveni K S. et al. [1] proposed a system of student feedback analysis for faculty rating. The student's emotions are classified using supervised machine learning classification algorithms. This system helps identify the standard of faculties and improve the teaching standards—a data mining process to rank the faculties of a school according to certain criteria. The greatest method to learn about a faculty is to read the comments that students have left. This system can also categorize the faculties into the appropriate groups by analyzing the student suggestions supplied using the approach involving fundamental text analysis and the Naive Bayes classifier. To address the issues with conventional feedback systems, the system also developed the notions of student feedback importance and the ideal learner.

Anna Koufakou et al. [2] proposed a system of student comments evaluation for undergraduate students. They used a small dataset of student course assessment comments to test the viability of text-mining approaches in assessing the student's open-ended remarks. The four different supervised machine learning classification algorithms are used to detect sentiment. The first findings demonstrate the viability of text mining as a tool for unstructured data analysis, sentiment analysis, and opinion summarization by identifying relevant phrases in survey responses.

Andamlak Terkik et al. [3] proposed a system for gender bias analysis using students' evaluation—the faculty performance is evaluated based on students' comments. To assess the total impact of each feedback, they used sentiment analysis methods to collect words. After that, investigate if there is bias when individuals describe their professors using this information together with other characteristics. Students' vocabulary used to describe their trainers and classroom experiences is significantly influenced by the gender of the assessed teacher, even though this does not appear to alter their declared level of contentment with their training.

Zenun Kastrati et al. [4] proposed a system of aspect-based opinion mining based on student reviews for online courses. This system describes a supervised model for aspect-based opinion mining on students' comments on online courses offered in Coursera. The model may anticipate the direction of opinion toward certain course-related characteristics discussed in the textual evaluations. A dataset including a significant number of manually commented student evaluations is gathered to validate the model. Evaluating a 1D-CNN model employing different word embeddings and four traditional machine-learning classifiers yielded encouraging results. It also only discovered and characterized a few factors connected to how well online courses are taught, so in our follow-up study, we want to look at some other course-related factors with an emphasis on factors that affect how well courses are taught in conventional settings.

Karunya K. et al. [5] proposed a system to responsibility of reading every student's comments is challenging. Every student's input is examined using sentiment analysis to address this problem. Knowing whether or not the teaching and learning session was fruitful is the major goal of this estimate. The suggestion system will give instructors a hint as to which areas of change are required based on the study. The fact that this effort will provide an automated suggestion system once the opinion is created is a significant addition. If opinion mining classified the feature as bad, the plan for suggestion would then suggest changes that should be made to that specific element. The benefit of such systems is that rapid comparisons can be made overall, even with many student comments.

Muhammad Zubair Asghar et al. [6] proposed a system for analyzing student satisfaction using fuzzy-based sentiment classification techniques. This work seeks to construct a fuzzy-based evaluation of a sentiment system for evaluating feedback from students and happiness by giving opinion words and polarity shifters in the input reviews the appropriate sentiment scores. The emotional score of feedback from student evaluations is calculated using this method, and the fuzzy-logic component is then used to assess and quantify student satisfaction on an extremely fine-grained level. The study results show that the work has done better than the previous studies and the available machine learning classifiers.

Lamiaa Mostafa [7] proposed Student Sentiment Analysis Using Gamification for Education Context. Three classifiers—NB, SVM, and decision tree—were employed in the sentiment analysis classifier, which goes through text processing, choosing features, and artificial intelligence classification. The findings indicated that the NB has the highest classifier accuracy rates. Additionally, when a test involving 1000 students was conducted, the agree group performed better than the disagree group, demonstrating that augmented reality will improve student learning outcomes. The following is a list of the emotion classifier's limitations.

Samuel Cunningham-Nelson et al. [8] text analysis using student feedback and opinion dataset. To illustrate student satisfaction remarks, several approaches that utilized machine learning and methods for studying texts were used. The latent Dirichlet allocating statistical approach was applied to pinpoint elements of student evaluation of a course. Additionally, the tone of the student responses was determined. Then, in a case study that provides instances of these representations, this knowledge was displayed graphically for instructors.

M. Sivakumar and Dr U. Srinivasulu Reddy [9] proposed a system aspect-based sentiment classification on student feedback datasets using machine learning techniques. The Twitter API was used for acquiring student input, which was then analyzed using the k-mean clustering and naive Bayes classification algorithms. They employed the R emotion package to determine the opposite polarity among the sentences and the semantic match among the opinion statement and a certain aspect word. These findings were utilized to assign an aspect term to each phrase. They had excellent results on recall, accuracy, and F-score tests.

Zhi Liu et al. [10] proposed a system detection of negative emotions of academic students for generated online courses. The system describes an ETJM, or automated joint sentiment and topic modelling, as a method for exploring unfavourable subjects in SGCs. By adding an emotion layer and a subject layer, ETJM could accurately identify pairs of words that express emotion at the context level. Then, to identify successful negative subjects, extract the high-NED phrases from SGCs and create a vocabulary of negative emotions. The experimental findings demonstrate that determining negative themes has been performed as intended. The accuracy rate for identifying negative subjects is 88%, which is respectable. Additionally, based on the findings of the research on emo-topics, learners are generally worried about the degree of complexity in educational resources, assignments, and obtaining certifications.

3. Research Methodology

In proposed sentiment classification we collect various students' feedback data. The collected data has a significant amount of noise. A number of computations are performed on the data before the algorithms may begin working with it in order to prepare it for their use. Data normalization is a important pre-processing procedure. The process of clearing the data begins with the detection of punctuation and stop - word, which is proceeded by the elimination of these elements from the data. After that, the data are tokenized and turned into lower case by invoking a function [9] in order to correct the imbalance that existed between them. The length of the dataset will be reduced as a result of this process, which will eliminate data that is not necessary.

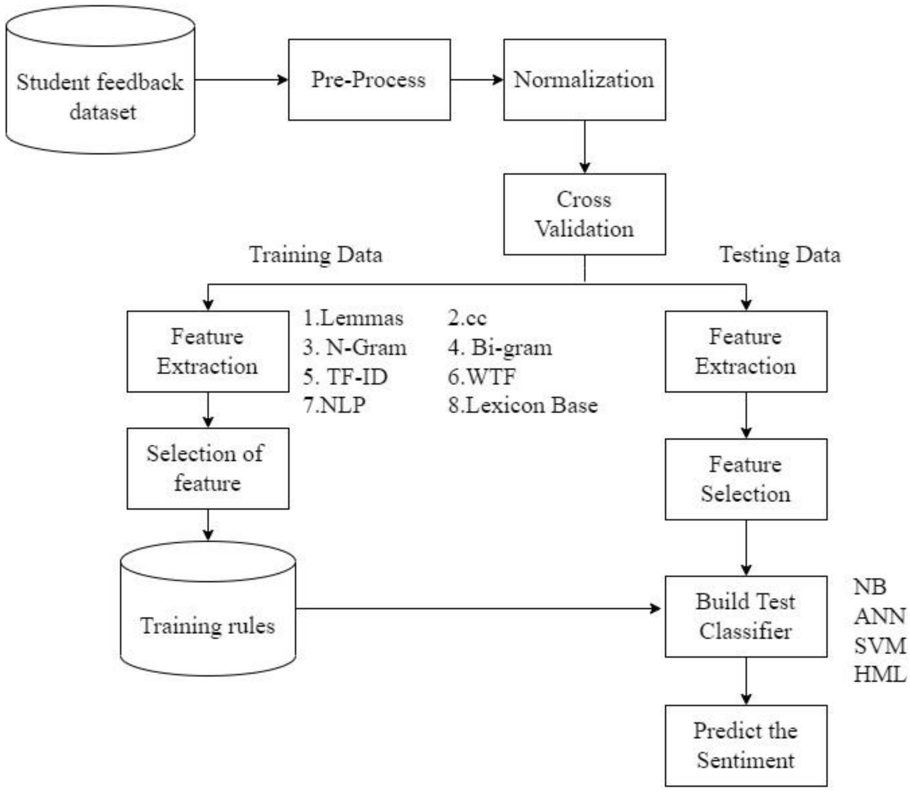

Fig 1: Proposed system architecture for sentiment classification using various feature extraction and machine learning techniques

Stop Words Removal : Stop words are words that enhance the meaning of other words or describe the relationship between two or more words. They may also contain conjunctions, pronoun, prepositions, and auxiliaries, among other grammatical components. Before the data is used as input for the classifiers, it is imperative that these stop words be removed from the dataset because our dataset contains a variety of articles. Words like "a," "an," "other," "neither," "but," "or," "forward with," "yet," "in," and so on are examples of such instances. Following their removal from the data corpus that we are working with, the result contains fewer different words [9].

Stemming : Converting the tokens to their corresponding fundamental or root terms is the next step in the standardization of the text. This step takes place after the tokens have been identified. The term "Stemming" refers to this particular procedure. It is used to decrease the number of different word forms that are present in data. This is accomplished through a shift in the word's fix position. The Snowball Stemmer method has been used in this model because, in comparison to the portal stemmer, it provides superior results. The word "minister" is changed to "minist" in the data collection, and it translates words like "extreme" and "extremely" to "extreme." The word "secretory" appeared the most frequently in the dataset; hence, the method was utilized most frequently with respect to this word.

Feature Extraction and Selection: During the process of feature extraction, a variety of features are extracted from the dataset that is being used for training, which results in a hybrid feature set. The following characteristics were gathered from the data:

Term Frequency Based Features: In text information, terms that appear frequently provide essential details about the context. The frequency of a phrase is used as a criterion in the selection of features in many different system designs. When we selective specific features such as TF count is higher or equal to *w* (*w* is weight such as 3,4,5…) those features considered as Weighted Term Frequency (WTF).

N-gram features: Both text mining and natural language processing operations make extensive use of this characteristic. N-grams are only a group of words that appear together in the same window at the same frequency. These are put to use in a variety of contexts. These concepts are not only utilized to construct unigram models, but also bigram as well as trigram models. In addition, they are used to generate unigram models.

Bi-tagged features: Using a POS tagger to tag phrases is the first step in the process of extracting bi-tagged characteristics. The elements of a sentence that are regarded as "bitagged" are those that have a substantial connection between two words that come immediately after one another in the phrase.

POS tag-based features: Many natural language processing applications, such as overview generation, query to response, as well as SA, depend on POS tagging of textual data. Words that are composed of nouns, actions, adverbs, and adjectives convey information that is significant to the reader about the context. Text mining systems can therefore benefit significantly from such key features.

Co-relation Coefficient Features: Correlation coefficient features refer to a type of feature engineering technique used in data analysis and machine learning to quantify the degree of linear relationship between two or more variables. The correlation coefficient indicates how well the variables move together and can range from -1 to 1. Correlation coefficient features are just one tool among many for understanding and analyzing relationships in data. Additionally, correlation coefficients are most suitable for capturing linear relationships; they might not capture more complex or nonlinear associations between variables.

Lexicon based Features: Lexicon-based features refer to linguistic features that are derived from a lexicon or a collection of words and their associated attributes. Lexicons are databases or lists of words or phrases, often accompanied by additional information such as part-of-speech, sentiment, frequency, and other linguistic properties. These features are widely used in natural language processing (NLP) and text analysis tasks to extract meaningful information from text data. Lexicon-based features can be useful for various NLP tasks, including sentiment analysis, text classification, topic modeling, and more. They provide a structured way to extract meaningful information from text data without requiring extensive training on large datasets. However, they may have limitations in capturing context and nuances in language that more advanced machine learning models can address.

Classification

The hybrid classification method is used for the classification process in the suggested system.

Support Vector Machine: When applied to linear data, the support vector machine method yields reliable classifications. If the input data has a non-linear form, we may use the kernel approach to convert it into a linear model without resorting to dimensionality reduction techniques. This method generates a hyper plane in N-dimensional space (precise position dependent on input feed dimension) during the model's training phase. This hyper plane is used to establish the boundary between the different classes identified in the dataset under consideration. In order to aid in the process of categorizing the new data, this line was selected to have the largest distance as from data points linked with each category.

Naïve Bayes: since of the properties of the dataset used, this classifier has been developed since it performs best when used with contextual data. It is built on the solid ground of the Bayes theorem. Using the file's word frequency distribution, the multinomial naive bayes classifier was used to determine the text's category and provide a prediction. This was done to ensure the highest degree of precision in our final conclusion.

HML: This algorithm having collaboration of ANN and SVM. Before sending input to a neural network, it must go through a series of steps to transform it into a format the network can understand, known as "preprocessing." In artificial neural networks, the news and information must first be translated into an acceptable format using a bag of words model since a lengthy 1-D vector of a specific length must be carried. This model accounts for the occurrence of each word in a statement, producing a vector whose bits reflect the number of occurrences of each word in the statement. Each layer of an artificial neural network is connected to the next through a weighted link. The precompiled news vector in a certain dimension is fed into the network at the first, or input, layer, which consists of several nodes. The second layer, called the hidden layer, is composed of the neurons that are responsible for carrying out the calculation necessary to produce the required output, which is the bitwise variable indicating class label. The output layer, which is the third layer, is responsible for showing the audience the findings from the concealed layer.

4. Results and Discussion

Dataset Description

The dataset has collected from student feedback which contains two major attributes as user comment and class label. The four different class are considered as final sentiment for each comment. The data contents classes such as poor, awful, average and awesome class labels. The below Table 1 shows the details about dataset. This data contains much noise as well as few emoji contents. During the preprocessing we eliminates such noisy information and recovering the null attributes.

Table 1: Dataset information

No.	Attribute Name	Values
1	Id	Numeric
2	Comment	Text
3	Class label	Text

The dataset contains around 3000 instances that we split into 70–30 patterns using the 5, 10, and 15-fold cross-validation techniques.

Table 2: Performance evaluation of proposed model with various feature extraction and machine learning based classification models

Feature extraction Method	Classification algorithm	Accuracy
N-Gram	Naive Bayes	**86.50**
Bi-Gram		**82.40**
TF-IDF		**88.95**
WTF		**91.30**
NLP (All Relationship)		**90.10**
NLP (selective Relationship)		**92.30**
Lemmas		**94.70**
Co-relation Coefficient		**93.60**
Lexicon based		**97.70**
N-Gram	ANN	**72.60**
Bi-Gram		**75.40**
TF-IDF		**86.70**
WTF		**89.40**
NLP (All Relationship)		**88.80**
NLP (selective Relationship)		**90.60**

Lemmas		**91.90**
Co-relation Coefficient		**93.70**
Lexicon based		**96.40**
N-Gram	SVM	**88.30**
Bi-Gram		**89.70**
TF-IDF		**90.50**
WTF		**92.80**
NLP (All Relationship)		**94.10**
NLP (selective Relationship)		**95.80**
Lemmas		**96.10**
Co-relation Coefficient		**95.95**
Lexicon based		**97.40**
N-Gram	HML	**91.20**
Bi-Gram		**93.60**
TF-IDF		**93.80**
WTF		**94.75**
NLP (All Relationship)		**95.20**
NLP (selective Relationship)		**96.70**
Lemmas		**97.10**
Co-relation Coefficient		**96.90**
Lexicon based		**98.20**

The above Table 2 describes 8 different feature extraction methods and 4 different supervised classification algorithms are used. The Lexicon based method with hybrid machine learning obtains highest results as 98.20%. The N-Gram methods obtains low accuracy with HML as 91.20%.

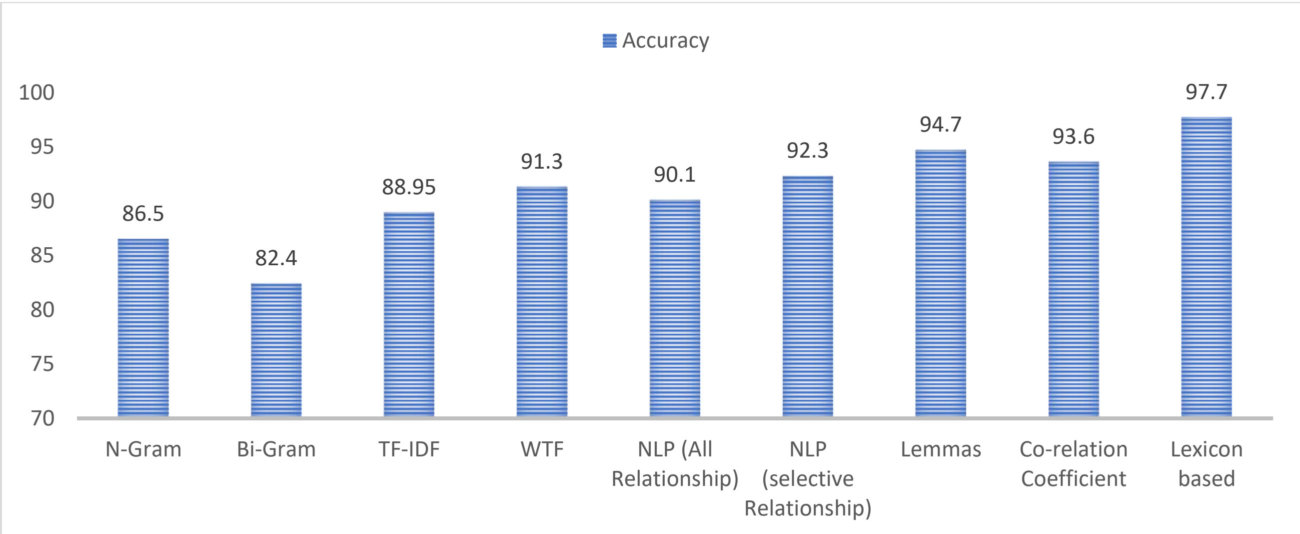

Fig 2 : Sentiment classification accuracy using Naïve Bayes with various feature extraction methods

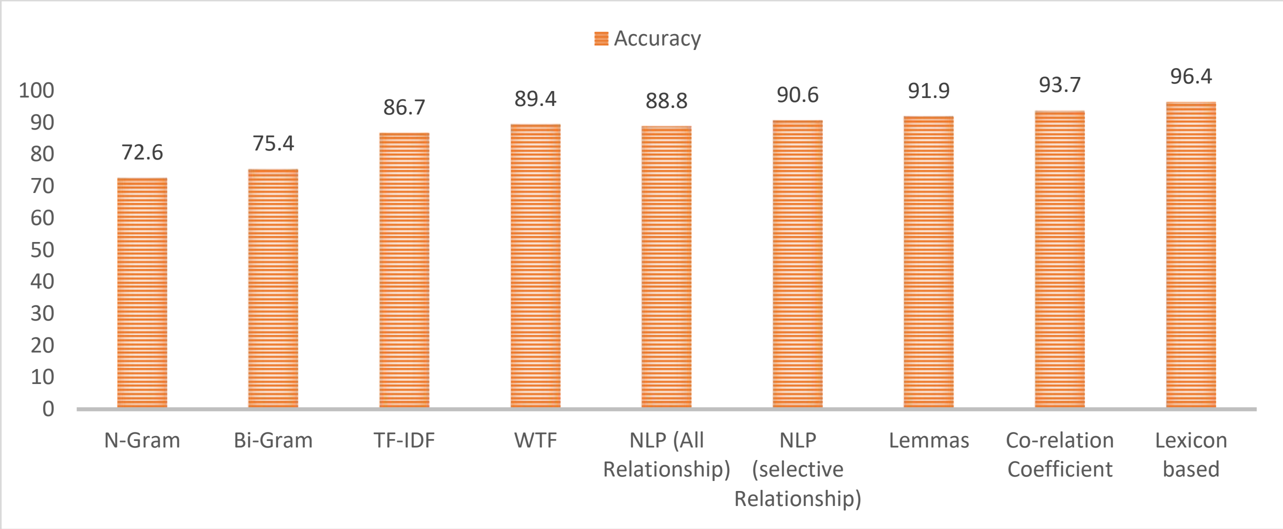

Fig 3 : Sentiment classification accuracy using ANN with various feature extraction methods

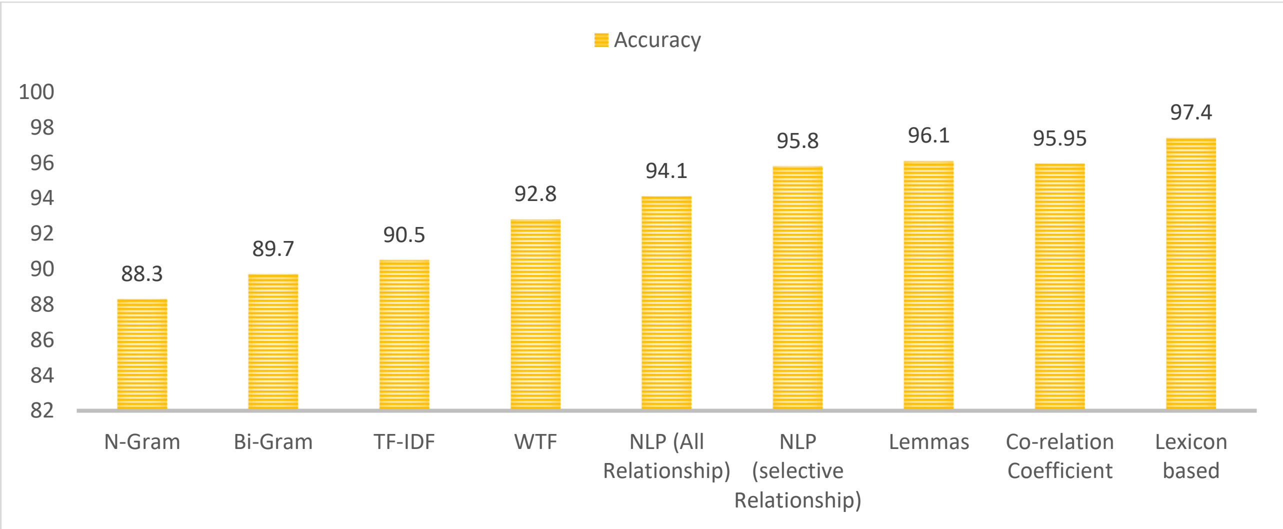

Fig 4 : Sentiment classification accuracy using SVM with various feature extraction methods

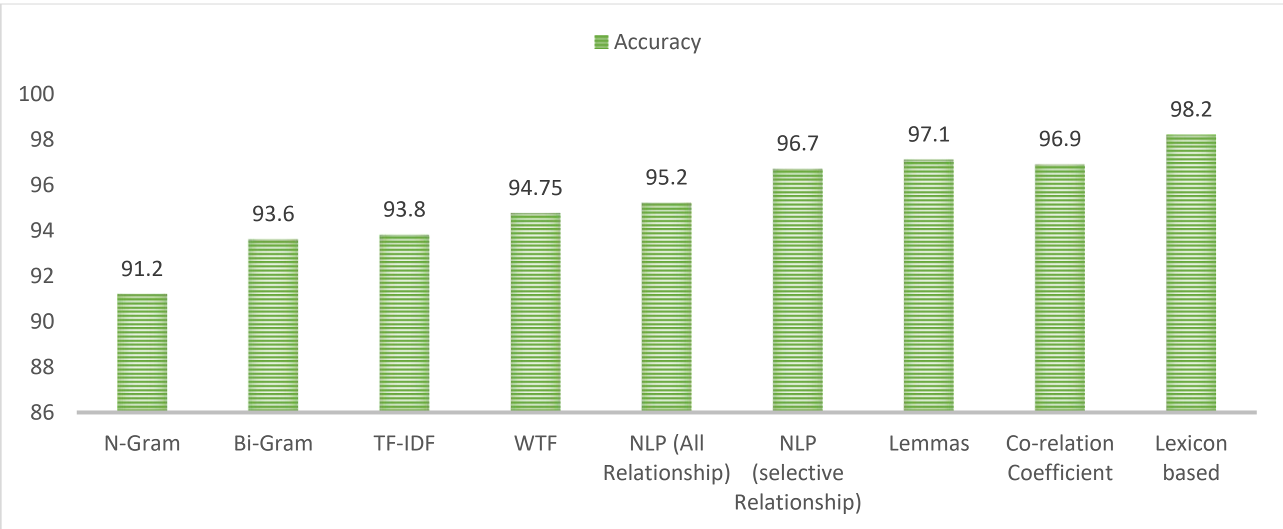

Fig 5 : Sentiment classification accuracy using HML with various feature extraction methods

The above Figure 5 shows HML based classification methods with nine different feature e extraction and selection techniques. The Lexicon based method achieves higher accuracy over the all-existing methods.

The Lexicon based features and HML algorithm obtains higher accuracy than other feature extraction and classification methodologies.

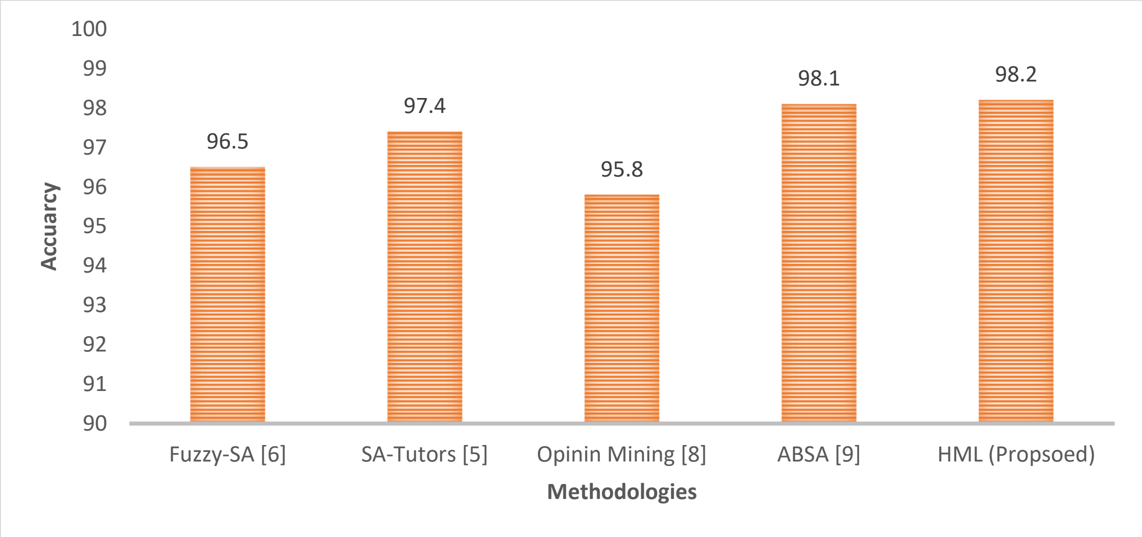

Fig 6 : comparative analysis for student feedback analysis with various ML with proposed HML

The above Figure 6 describes a comparative analysis of student feedback analysis using various ML models, including a proposed hybrid model. The HML models can often provide better results by leveraging the strengths of different techniques, especially when dealing with complex and diverse data sources like student feedback. As results we conclude the proposed HML provides around 3% higher accuracy than other machine learning techniques.

5. Conclusion

The conclusion of a study we describe a sentiment analysis method for analyzing student feedback using a lexicon-based approach and a hybrid machine learning method. In this study, we investigated and compared two different approaches for sentiment analysis of student feedback: a lexicon-based method and a hybrid machine learning method. Our goal was to determine the effectiveness and suitability of each approach in capturing the sentiments expressed in student feedback data.

Our analysis using the lexicon-based approach involved using pre-defined sentiment lexicons to assign sentiment scores to words and phrases within the student feedback. This method offered simplicity and efficiency in terms of implementation, as it did not require extensive training on large datasets. However, the hybrid machine learning method combined the strengths of machine learning techniques with the insights provided by lexicon-based analysis. By training a model on a labeled dataset of student feedback, we aimed to improve the accuracy and context-awareness of sentiment classification. In conclusion, both the lexicon-based method and the hybrid machine learning method have their merits and limitations when applied to sentiment analysis of student feedback. The lexicon-based approach is simple to implement but struggles with context and adaptability, while the hybrid machine learning method offers more accurate and adaptable sentiment analysis. However, it requires substantial data for training and potential model updates to remain effective in dynamic environments. The choice between these methods should depend on the specific goals, available resources, and the desired level of accuracy and context-awareness required for sentiment analysis in student feedback. Future research could focus on refining the hybrid model's architecture, experimenting with different lexicons, and exploring techniques to further enhance the accuracy and interpretability of sentiment analysis in educational settings.

References

[1] Krishnaveni, K.; Pai, R.R.; Iyer, V. Faculty rating system based on student feedbacks using sentimental analysis. In Proceedings of the 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Udupi, India, 13–16 September 2017; pp. 1648–1653.

[2] Koufakou, A.; Gosselin, J.; Guo, D. Using data mining to extract knowledge from student evaluation comments in undergraduate courses. In Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016; pp. 3138–3142.

[3] Terkik, A.; Prud'hommeaux, E.; Alm, C.O.; Homan, C.; Franklin, S. Analyzing gender bias in student evaluations. In Proceedings of the COLING 2016, the 26th International Conference on

Computational Linguistics: Technical Papers, Osaka, Japan, 11–16 December 2016; pp. 868–876.
[4] Kastrati, Z.; Arifaj, B.; Lubishtani, A.; Gashi, F.; Nishliu, E. Aspect-Based Opinion Mining of Students' Reviews on Online Courses. In Proceedings of the 2020 6th International Conference on Computing and Artificial Intelligence, Tianjin, China, 23–26 April 2020; pp. 510–514.
[5] Karunya, K.; Aarthy, S.; Karthika, R.; Deborah, L.J. Analysis of Student Feedback and Recommendation to Tutors. In Proceedings of the 2020 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 28–30 July 2020; pp. 1579–1583
[6] Asghar, M.Z.; Ullah, I.; Shamshirband, S.; Kundi, F.M.; Habib, A. Fuzzy-based sentiment analysis system for analyzing student feedback and satisfaction. Preprints 2019.
[7] Mostafa, L. Student sentiment analysis using gamification for education context. In International Conference on Advanced Intelligent Systems and Informatics; Springer: Berlin/Heidelberg, Germany, 2019; pp. 329–339.
[8] Cunningham-Nelson, S.; Baktashmotlagh, M.; Boles, W. Visualizing student opinion through text analysis. IEEE Trans. Educ. 2019, 62, 305–311.
[9] Sivakumar, M.; Reddy, U.S. Aspect based sentiment analysis of students opinion using machine learning techniques. In Proceedings of the 2017 International Conference on Inventive Computing and Informatics (ICICI), Coimbatore, India, 23–24 November 2017; pp. 726–731.
[10] Iram, A. Sentiment Analysis of Student's Facebook Posts. In International Conference on Intelligent Technologies and Applications; Springer: Berlin/Heidelberg, Germany, 2018; pp. 86–97.
[11] Reddy, B.R.S., Saxena, A.K., Pandey, B.K., Gupta, S., Gurpur, S., Dari, S.S., Dhabliya, D. Machine learning application for evidence image enhancement (2023) Handbook of Research on Thrust Technologies? Effect on Image Processing, pp. 25-38.

Ms. Shital Abhimanyu Patil, completed B.E from Godavari college Of Engineering, Jalgaon. and ME Computer Science & Engineering from SSBT COET Bambhori. She is working as Assistant Professor in SSBT's College of Engineering and Technology since 2006.She is pursuing PhD in Computer Science & Engineering from Kavayitri Bahinabai Chaudhari North Maharashtra University, Jalgaon. Her areas of interest are Machine Learning, Data Analytics and Natural Language Processing.

Dr. Krishnakant Prabhudas Adhiya has completed PhD degree in Computer Science and Engineering from North Maharashtra University Jalgaon. Currently he is working as Professor, Department of Computer Engineering, SSBT's College of Engineering & Technology, Jalgaon and recognized PhD Guide in Kavayitri Bahinabai Chaudhari North Maharashtra University, Jalgaon. He has 44 research papers in reputed peer reviewed journals in addition to 14 papers in International Conferences to his credit. He is Professional Member in ACM, Life member of ISTE. His research interests are Wireless Sensor Networks and Security, Machine Learning, Block Chain and Natural Language Processing

Dr. Girishkumar K. Patnaik, is working as Professor & Principal of SSBT's College of Engineering & Technology, Jalgaon. He is recognized PhD Guide in Kavayitri Bahinabai Chaudhari North Maharashtra University, Jalgaon. Professional Member in ACM, Life member of ISTE and CSI. His research interests are Wireless Sensor Networks and Security, Machine Learning, Blockchain and Natural Language Processing. He has published 30 research papers in reputed peer reviewed journals in addition to 12 papers in International Conferences to his credit. He is Senior Member in IEEE.