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

Streamlined Teacher Evaluations: Leveraging ML with SMOTE and Streamlit Approach for Real-Time Sentiment Analysis of Student Feedback

[1]Dr. Manisha M. More, [2]Dr. Vinaya Keskar, [3]Dr. Richa Purohit, [4]Prof. Rashmi Dharwadkar, [5]Dr. Pallavi Yarde, [6]Dr. Vijaya Kumbhar

Submitted: 03/05/2024 **Revised**: 16/06/2024 **Accepted**: 23/06/2024

Abstract: In this research, we collect student feedback on teachers through Google Forms using open-ended questions, allowing students to freely express their feelings, opinions, thoughts, and evaluations of teaching styles and pedagogies. The collected textual data is analyzed using the NLP – a natural language processing technique to predict the student feedback sentiments. We employ a Machine Learning (ML) model, specifically a Support Vector Machine (SVM) enhanced with SMOTE to address data imbalance. After rigorous evaluation and validation, the SVM model, achieving 86% accuracy, was selected for sentiment prediction. This pre-trained model is deployed through a Streamlit application, designed with HTML, CSS, and Python. The novelty of this research lies in the Streamlit app, enabling students to submit feedback and receive sentiment analysis results in real-time. This system aids in continuous feedback collection, facilitating daily, weekly, monthly, and semester-wise analysis. Such insights are valuable for teachers to understand student satisfaction, for higher authorities to monitor overall performance, and for HR during the appraisal process. Additionally, this feedback mechanism supports the NAAC accreditation process, making feedback collection timely and efficient. With 75% of the implementation completed, the remaining will be finalized soon, enhancing the system's capability to forecast feedback trends and support decision-making processes

Keywords: *Student feedback analytics, NLP - Natural Language Processing, ML - Machine Learning, sentiment prediction, SVM - support vector machine, SMOTE – synthetic minority over-sampling technique, Streamlit Application*

I. Introduction

In the realm of education, student feedback is indispensable for assessing the effectiveness of teaching methods, identifying areas for improvement, and fostering a responsive learning environment. Feedback from students provides educators with valuable insights into their teaching practices, highlighting strengths and pinpointing weaknesses. Traditional feedback collection methods, typically conducted at the end of a semester, are often time-consuming and may not reflect the immediate impacts of teaching strategies. The delaying of the student feedback can cause timely adjustments and improvements. Addressing these challenges, our proposed research aims to develop an automated advanced real time sentiment analyzer for student feedback, which utilizes the advanced technology such as NLP and ML. Researcher has integrated NLP, ML with Streamlit approach is the the pivotal in our research. Sentiment analyzer leverages the SVM Model with SMOTE to handle our imbalanced student feedback with improved accuracy in sentiment prediction. Our automated innovative platform not only help in feedback sentiment prediction but also provides an actionable insight.

The proposed research novelty has its capability to continuously collect the student feedback and automatically analyze and predict the sentiments. Throughout the semester it helps to the educators to adjust the teaching strategies. This automated system also helps to meets the educational requirements such as the faculty appraisal process, and NAAC accreditation process by providing meaningful insights. The researcher's motivation behind this automated system is to enhance the educational setup and improve the teaching methodologies.

During the research researcher gone through the extensive literature survey related to the proposed work in the context to the educational context. Teachers sentiments analysis and prediction, evaluating the teachers performance with ML and NLP. Bhowmik et. al. (2023), in their research categorizes the sentiments into three categories i.e. positive sentiments, negative sentiments

[1]Sr. Assistant Professor, School of Computer Studies, Sri Balaji University, Pune
Drmanishamore8@gmail.com 9511775185
[2]Head dept of computer science ATSS College of business studies and Computer Application Chinchwad Pune vasanti.keskar@gmail.com
[3]Associate Professor School of Computer Studies Sri Balaji University Pune Richapurohit81@gmail.com
[4]Sr Teaching Assistant School of Computer Studies Sri Balaji University Pune rashmidharwadkar1@gmail.com
[5]Balaji Institute of Technology & Management, Sri Balaji University, Pune pallavi.yarde@bitmpune.edu.in
[6]Sr. Asstt. Prof.School of Computer Studies, Sri Balaji University, Pune. veejeya.kumbhar@gmail.com

and neutral sentiments employing the aspect-based analysis of sentiments approach and providing the teaching aspects. Faizi (2023) in research study, used lexicon approach and predicted the student's feedback. Grimalt Alvaro et. al. (2023) in their automated sentiment analysis system used the variables such as gender and cultural context. Fargues et. al. (2023), the researcher developed NLP application to analyze textual student feedback and detected sentiments through it. It was helpful in enhancing the teaching course delivery. Jagtap et. al. (2014), the study focused on the automation for teacher feedback assessment. They used SVM model and HMM hybrid approach to highlight the education sector challenges. Su et. al. (2023), the researcher introduced the mechanism of hierarchical attention for sentiment analysis for the online courses. The researchers aimed to improve the accuracy over the traditional models. Bhowmik et. al. (2023), in their study explored the aspect based student feedback analysis on large data sets, which was focusing its efficiency in automated analysis of feedback. Beral et. al. (2023), the researchers investigated different patterns and sentiments in school of mathematics feedback, and analyzed the impact of punctuation marks on teacher sentiments. Ren et al. (2023) employed aspect-level sentiment analysis to evaluate teaching performance, achieving high precision and F1 values. Tian et al. (2022) applied the mini-Xception framework for the sentiments analysis on real-time students responses. in classroom settings, aiming to provide timely feedback to enhance teaching effectiveness. One of the researcher Usart et. al. (2023), in their research study, proposed a sentiment analysis for gender-sensitive method to characterize the emotions climates for the teachers who teaches online, highlighting gender differences in sentiment expression. Khanam (2023) reviewed sentiment analysis methods in online learning environments, discussing their potential in improving teaching methods and learning experiences. Mamidted and Maulana (2023) explored students' perceptions of teachers' online teaching performance, identifying key factors influencing student satisfaction and learning outcomes. The researcher Rajput et. al. (2016) integrated the tect analytics for analyzing the open ended students feedback, providing insights into teacher performance beyond Likert scale scores. Ortigosa e.t al. (2014) introduced a hybrid sentiment analysis approach in Facebook for e-learning contexts, demonstrating high accuracy. Mary et. al. (2023) proposed the sentiments detection and prediction using multifaceted sentiment approach to find out dropouts, predict dropout risks in online learning environments, utilizing machine learning techniques. One another researcher Kastrati et. al. (2021) conducted a sentiment analysis with the systematic mapping study in education, highligts the growth and also the various challenges when apply the NLP, DL, and ML technology to analyze student feedback. Sumers et. al. (2021), the researchers developed a framework and used aspect-based approach to analyze the sentiments with NLP, and AI agents that demonstrate automatically learn human feedback.

II. Research Gap and Significance of Proposed Research

The existing research studies on sentiment analysis on the student's feedback is neglecting the challenges in imbalanced datasets, and also not overlooked the automated Streamlit App which will work on real-time students feedback. The existing researchers has focused primarily on structured data and only on the traditional sentiment analysis and was failing to address the imbalance of sentiment classes and user frindely model deployment.

The proposed research fills all the gaps by applying SMOTE model with SVM to address the imbalance the datasets. This approach ensures accurate sentiment classification and prediction. The model deployment approach with the Streamlit library facilitates the real-time feedback submition and automate the sentiment prediction. This automated system enhances the overall effectiveness of teaching strategies and also improves the student satisfaction with the NLP, ML, SMOTE and Streamlit approach.

III. Methodology Used

The researcher used primary data which is collected from BCA and MCA students at Sri Balaji University, Pune through the well prepared Google Form with all open ended questions. The data was collected in textual form and then downloaded from Google Form to the Excel file. This file then uploaded to the Jupyter Notebook editor for preprocess and model building purpose.

The students feedback dataset with two columns viz. Teacher and Feedback with 340 rows for 7 seven teachers. For NLP and ML we used Python programming language with techniques such as TFIDFVectorizer and TextBlob, ML technique such as SVM with SMOTE for imbalanced data to perform sentiment analysis. For model deployment we employed Streamlit approach to facilitate the real-time data submission and sentiment prediction.

A. Feature Analysis

Initially we loaded the student feedback data to Jupyter notebook and then converted into the DataFrame, with 340 rows and 2 columns such as Teacher and Feedback. For data preprocessing we employed NLTK library to perform lemmatization and find and remove the stop-words with the help of regular expression to clean the text. The preprocessing included removing the non-alphabetic characters and numbers, converting the text to lowe case, splitting text into individual words. The lemmatization helps to reduce the words into the base forms and then remove the stop words. This process resulted into one new feature i.e. filtered_Feedback, which contained the cleaned data and ready for further analysis. The Fig. 1 shows the filtered summary of student's feedback about their teacher.

	Teacher	Feedback	filtered_Feedback
0	Prof. A	Excellent view of teaching	excellent view teaching
1	Prof. A	Mam have very good teaching method	mam good teaching method
2	Prof. A	Maam teaches very professionally and always co...	maam teach professionally always cool headed. ...
3	Prof. A	She teaches DBMS as per as first sem.\nAnd I w...	teach dbms per first sem. would like give star...
4	Prof. A	My genuine opinion is that mrs richa purohit i...	genuine opinion mr richa purohit beautiful con...
...	...	...	...
335	Prof. G	The teacher does not accommodate different lea...	teacher accommodate different learning styles....
336	Prof. G	Classroom management is poor, leading to frequ...	classroom management poor, leading frequent di...
337	Prof. G	The teacher does not use technology effectivel...	teacher use technology effectively teaching. o...
338	Prof. G	The teacher is frequently late to class, which...	teacher frequently late class, disrupts learni...
339	Prof. G	he teacher does not provide sufficient resourc...	teacher provide sufficient resource study mate...

340 rows × 3 columns

Fig. 1 Filtered_Feedback

B. Exploratory Data Analysis

In this phase researcher utilized TFIDFVectorizer from the Sklearn library and transformed the filtered_Feedback into the vectors. This process allowed one another DataFrame for further analysis. The text data is used to generate the WordClouds to highlights the most frequent terms from student feedback. The general WordClouds are generated on entire data set for positive, negative, ngrams. These WordClouds are generated by focusing on first three entries to understand the text distribution.

At the last we identified the unique terms by employing the DET- Determiner which helps in further refining the features in different model stages. The comprehensive EDA process provides the valuable insights from the textual data and sentiment distribution, and development of sentiment analysis model. Fig. 2 shows the WordCloud on student feedback.

Fig. 2 WordCloud

Calculating Subjectivity and Polarity

To calculate the subjectivity and polarity weutilized the TextBlob library. The subjectivity and polarity calculated from the filtered_Feedback. The subjectivity indicates the degree of personal opinions from the feedback, and it is ranging from 0 – objective to 1- subjective. The score of polarity is ranging from (-1) indicated as negative and (1) indicated as positive polarity. 0 represents the neutral. This process helps in sentiment classification and model training. Fig. 3 gives the demonstration of subjectivity and polarity in student feedback.

	Teacher	Feedback	filtered_Feedback	textblob_sentiment	Unique Terms	Subjectivity	Polarity
0	Prof. A	Excellent view of teaching	excellent view teaching	1.000000	[excellent, view, teaching]	1.000000	1.000000
1	Prof. A	Mam have very good teaching method	mam good teaching method	0.700000	[mam, good, method]	0.600000	0.700000
2	Prof. A	Maam teaches very professionally and always co...	maam teach professionally always cool headed. ...	0.212500	[maam, teach, professionally, always, cool, he...	0.445833	0.212500
3	Prof. A	She teaches DBMS as per as first sem.\nAnd i w...	teach dbms per first sem. would like give star...	0.250000	[dbms, per, first, sem., would, like, give, st...	0.333333	0.250000
4	Prof. A	My genuine opinion is that mrs richa purohit i...	genuine opinion mr richa purohit beautiful con...	0.583333	[genuine, opinion, mr, richa, purohit, beautif...	0.777778	0.583333
...	...	...	...	...	...	...	...
335	Prof. G	The teacher does not accommodate different lea...	teacher accommodate different learning styles....	0.000000	NaN	0.600000	0.000000
336	Prof. G	Classroom management is poor, leading to frequ...	classroom management poor, leading frequent di...	-0.150000	[poor,, leading, frequent, disruptions., hinde...	0.450000	-0.150000
337	Prof. G	The teacher does not use technology effectivel...	teacher use technology effectively teaching. o...	0.200000	[lesson, engaging.]	0.711111	0.200000
338	Prof. G	The teacher is frequently late to class, which...	teacher frequently late class, disrupts learni...	-0.300000	[frequently, disrupts, schedule., students']	0.600000	-0.300000
339	Prof. G	he teacher does not provide sufficient resourc...	teacher provide sufficient resource study mate...	-0.500000	NaN	1.000000	-0.500000

340 rows × 7 columns

Fig.3 Subjectivity and Polarity

Next, to calculate the sentiment score, we utilized an affinity score approach. The afinn2.txt file loaded into a dataframe and obtained word-affinity pairs. The Spacy library with en_core_sm open source language model for sentiment analysis. Created a custom function to calculate the sentiment scores, accessed the lemmatized words from the student feedback using affinity scores. With this approach we calculated the sentiment values in students feedback. Fig. 4 shows the sentiment score in student feedback.

	Teacher	Feedback	filtered_Feedback	textblob_sentiment	Unique Terms	Subjectivity	Polarity	Preprocessed_Feedback	sentiment_value
0	Prof. A	Excellent view of teaching	excellent view teaching	1.000000	[excellent, view, teaching]	1.000000	1.000000	excellent view teaching	3
1	Prof. A	Mam have very good teaching method	mam good teaching method	0.700000	[mam, good, method]	0.600000	0.700000	mam good teaching method	3
2	Prof. A	Maam teaches very professionally and always co...	maam teach professionally always cool headed. ...	0.212500	[maam, teach, professionally, always, cool, he...	0.445833	0.212500	maam teach professionally always cool headed m...	1
3	Prof. A	She teaches DBMS as per as first sem.\nAnd i w...	teach dbms per first sem. would like give star...	0.250000	[dbms, per, first, sem., would, like, give, st...	0.333333	0.250000	teach dbms per first sem would like give star ...	2
4	Prof. A	My genuine opinion is that mrs richa purohit i...	genuine opinion mr richa purohit beautiful con...	0.583333	[genuine, opinion, mr, richa, purohit, beautif...	0.777778	0.583333	genuine opinion mr richa purohit beautiful con...	5

Fig. 4 Sentiment Score

We visualized the distribution of sentiment values using a Seaborn distplot, which allowed us to observe the sentiment score spread across the feedback data. This plot provided a clear depiction of the overall sentiment tendencies within the dataset. Fig. 5 shows the distribution chart on sentiment value of student data.

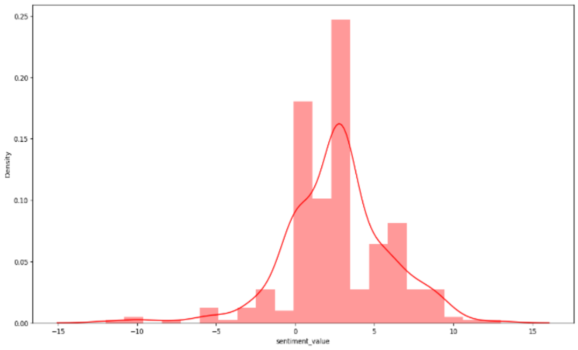

Fig. 5 Distribution of Sentiment Value

The density plot shows that most student feedback is slightly positive, with sentiment values clustering around 2-3. There are fewer instances of highly negative or highly positive sentiments, indicating a general trend towards mildly positive feedback.

The histogram shows that most feedback entries are short, peaking around 5 to 10 words, with fewer longer feedback entries. This graph was created by using Seaborn to combine a histogram with a kernel density estimate (KDE) for a clear view of the distribution. The following Fig. 6 illustrates the distribution of words count on student feedback.

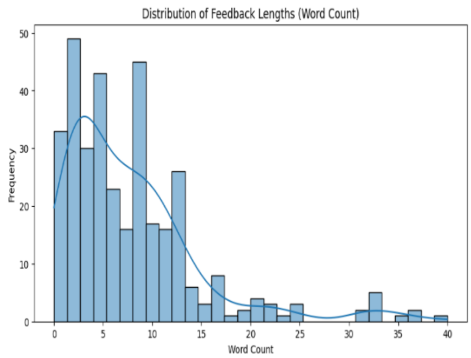

Fig. 6 Word Count in Student Feedback

A violin plot displays the distribution of feedback lengths by teacher, showing variations in word counts for each professor. Seaborn's violin plot technique combines KDE and box plot elements, effectively illustrating the spread and density of feedback lengths across different teachers. Fig. 7 shows the distribution of feedback length.

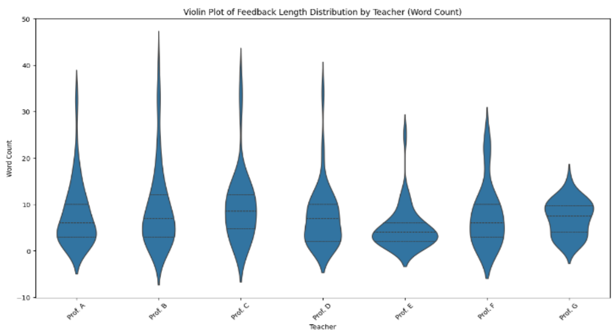

Fig. 7 Distribution of Feedback Length.

The DataFrame displays the emotion scores for various feedback entries, with columns representing different emotions such as anger, joy, and trust. NRCLex, a library for emotion detection, was used to extract these emotion scores from the textual feedback, resulting in a comprehensive overview of the emotional tone of each entry. Fig. 8 shows the overall emotion score on student feedback data.

	anger	anticipation	disgust	fear	joy	negative	positive	sadness	surprise	trust
0	0	0	0	0	1	0	1	0	0	1
1	0	1	0	0	1	0	1	0	1	1
2	0	0	0	1	1	0	4	0	1	2
3	1	1	0	1	2	1	2	1	1	2
4	0	0	0	0	2	0	4	0	0	3
...	...	...	...	...	...	...	...	...	...	...
335	0	0	0	0	0	0	2	0	0	1
336	0	0	0	0	0	0	3	0	0	2
337	0	1	0	0	1	0	4	0	0	3
338	0	2	0	0	1	2	3	1	0	3
339	0	0	0	1	0	0	3	0	0	2

340 rows × 10 columns

Fig. 8 Emotion Score

The figure categorizes teacher feedback into Negative, Neutral, and Positive sentiment classes using pandas' cut function. Sentiment values are segmented into bins: -53 to -1 for Negative, -1 to 0 for Neutral, and 0 to 78 for Positive. This technique simplifies continuous sentiment values into discrete categories, aiding analysis. The resulting DataFrame includes feedback, filtered feedback, sentiment score, and sentiment class, using pandas for effective data categorization and interpretation. The fig. 9 shows the sentiment score and related sentiment class on student feedback.

	Teacher	Feedback	filtered_Feedback	sentiment_value	Sentiment_Class
0	Prof. A	Excellent view of teaching	excellent view teaching	3	Positive
1	Prof. A	Mam have very good teaching method	mam good teaching method	3	Positive
2	Prof. A	Maam teaches very professionally and always co...	maam teach professionally always cool headed. ...	1	Positive
3	Prof. A	She teaches DBMS as per as first sem.\nAnd i w...	teach dbms per first sem. would like give star...	2	Positive
4	Prof. A	My genuine opinion is that mrs richa purohit i...	genuine opinion mr richa purohit beautiful con...	5	Positive
...	...	...	...	...	...
335	Prof. G	The teacher does not accommodate different lea...	teacher accommodate different learning styles....	1	Positive
336	Prof. G	Classroom management is poor, leading to frequ...	classroom management poor, leading frequent di...	-2	Negative
337	Prof. G	The teacher does not use technology effectivel...	teacher use technology effectively teaching. o...	2	Positive
338	Prof. G	The teacher is frequently late to class, which...	teacher frequently late class, disrupts learni...	-2	Negative
339	Prof. G	he teacher does not provide sufficient resourc...	teacher provide sufficient resource study mate...	-1	Negative

340 rows × 5 columns

Fig. 9 Sentiment Score and Sentiment Class

The pie chart shows the distribution of sentiment classes among teacher feedback: 252 Positive (73.7%), 58 Neutral (17.1%), and 30 Negative (8.8%). This indicates a predominance of positive feedback. The chart was generated using pandas for data manipulation, matplotlib for plotting, and seaborn for color palette customization. Fig. 10 shows the distribution of sentiment class on student feedback data.

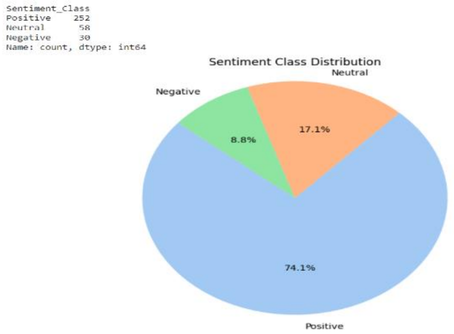

Fig. 10 Sentiment Score and Sentiment Class Distribution

The analysis reveals the average sentiment and emotion scores for each teacher, highlighting their perceived strengths and areas for improvement. Sentiment distribution charts show the balance of positive, neutral, and negative feedback. Key technologies used include Pandas for data manipulation, Matplotlib and Seaborn for visualization, and NRClex for extracting emotions from text. This approach provides clear insights into teacher performance based on student feedback. Fig. 11 shows teacher wise distribution of emotions

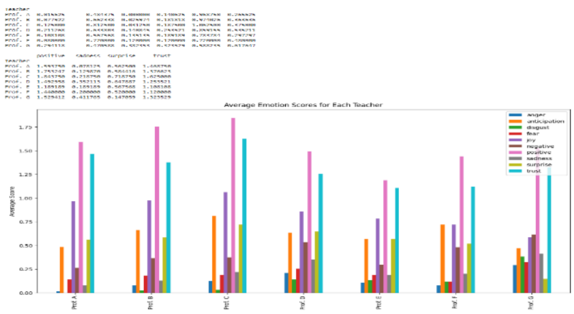

Fig. 11 Teacher Wise Emotion Score Distribution

The correlation matrix reveals the relationships between different emotions in student feedback. Notably, 'anger' has a strong positive correlation with 'disgust' (0.837) and 'sadness' (0.738), while 'joy' and 'trust' are highly correlated (0.741). These insights were generated using Pandas for data manipulation, Matplotlib and Seaborn for visualization. The matrix helps identify how different emotions are interconnected, aiding in understanding the emotional landscape of student feedback. The Fig. 12 shows Correlation matrix on student feedback.

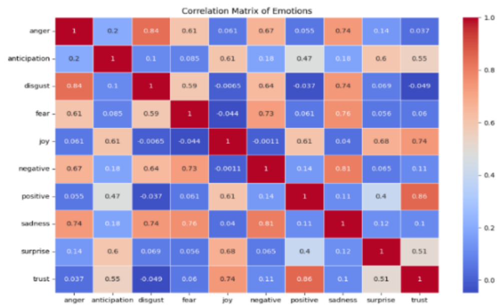

Fig. 12 Correlation Matrix on Student Feedback

The chart displays sentiment trends over time for each teacher, generated by applying TextBlob to calculate sentiment scores from feedback. Each teacher's sentiment score is plotted day-wise, showing fluctuations and trends. This analysis uses Pandas for data manipulation, TextBlob for sentiment analysis, and Matplotlib for visualization, highlighting how student feedback sentiment varies over time for different teachers. The Fig. 13 shows the sentiment trend over the time.

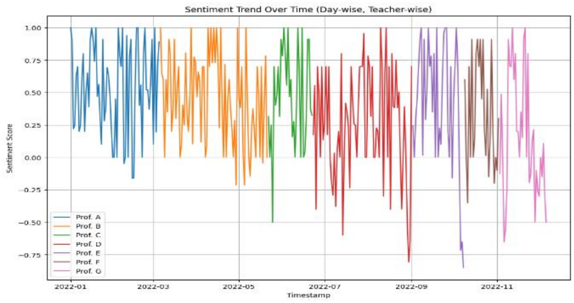

Fig. 13 Sentiment Trend Over the Time(Teacher Wise)

C. Model Building and Testing

The student feedback dataset contains open-ended feedback from students about teachers. The data was processed to extract sentiment scores and classified into sentiment classes (Positive, Neutral, Negative). The feedback text was filtered to remove noise, and sentiment values were calculated. The data was then split into features (filtered feedback) and target variables (sentiment classes) for analysis.

Machine Learning Models for Imbalanced Data

Initially, we applied Decision Trees, Random Forest, and Logistic Regression to predict sentiment classes. These models did not perform well due to class imbalance, leading to biased predictions towards the majority class. To address this, we have implemented SVM with a minority class balancing technique of ML i.e. SMOTE. SVM effectively handled all the imbalanced data by maximizing the margin between classes. The SMOTE helps in oversample the minority classes to balance in the training set, and this resulted in improved performance of SVM model with 86% accuracy for sentiment prediction. Table 1 shows the mathematical model for SMOTE

Table 1. SMOTE Model for Imbalanced Student Feedback

Step	Description
Initial Data Distribution	- Positive (P): 252 - Neutral (N): 58 - Negative (NG): 30
Identify Minority Classes	- Neutral samples: $X_{neutral}$ - Negative samples: $X_{negative}$

Synthetic Sample Generation	For each x_i in $X_{neutral}$ and $X_{negative}$ 1. Find k-nearest neighbors. 2. Create synthetic sample x^: $x\hat{}=x_i+\lambda \cdot (x_{i,\,nn} - x\_i)$ where, $\lambda \sim U\ (0,1)$
Final Balanced Dataset	Combine original and synthetic samples: $X_{balanced} = X_{positive} \cup X_{neutral} \cup X_{negative} \cup X_{synthetic}$

The table 1 outlines the process of SMOTE, how the synthetic samples which are generated on minority classes i.e. neutral class, and the negative class for balancing in the student feedback dataset. This process identifies the minority samples and find the nearest neighbors. It interpolates the new samples to achieve the more class distributions.

The code of the model begins with splitting the student feedback into the training as well as testing sets which is utilized with the Sklearn Library function train_test_split. The filtered_Feedback and Sentiment_Class are then transformed into the TFIDF feature with TFIDFVectorizer. The imbalance class is tackled with Synthetic Minority Over-Sampling Technique i.e. SMOTE to ensure all three classes are distributed equally.

We employed SVM model with RBF kernel to train the balanced data. The TFIDFVectorizer and the trained SVM models are saved by using joblib.dump library for model deployment purpose. The performance of this SVM model is evaluated and assessed through metrics such as the accuracy of the model and the classification report generated to provide the model report including precision, recall, F1 score on each sentiment class. This helps in evaluation of the performance of model on each class. Furthermore, the confusion matrix breaks down the models performance with true versus predicted classification. The confusion matrix highlights the misclassification. The model is trained and tested effectively or not is ensured by this approach. The evaluation guarantees the reliability in prediction of sentiment classes in student feedback.

D. Model Deployment

The pretrained SVM model is used to deploy it with Streamlit approach. To develop interactive interface we utilized the Streamlit library. The Streamlit App name is Student Feedback Sentiment Analyzer, designed with Python, HTML, CSS and Streamlit. The pretrained SVM model and TFIDFVectorizer are loaded with joblib and NLTK. For tokenizing feedback into the sentences. The Streamlit App is divided into three parts. First part enable students to submit the feedback about their teachers, second part shows the sentiments for te teachers, wordcloud, third part shows the sentiment class distribution on the student feedback. Overall in the Streamlit App process we employed joblib, pandas, matplotlib, NLTK, WordCloud techniques to design the user friendly interface which gives detailed analysis of student feedback. The model saved with following code,

Joblib.dump(svm_model, "Teachertrained_model_svm.sav") and the TFIDFVectorizer saved witk pickle library in .pkl format, joblib.dump(tfidf_vectorizer, "tfidf_vectorizer.pkl")

Fig 14 is the screenshot of the Streamlit system which shows the user friendly interface enable to the students to enter the feedback and can see the sentiments immediately. Students just has to input the feedback about their teacher and our system immediately predict the sentiment from given feedback just by clicking on "Predict" button. Streamlit app use the SVM pretrained model and TFIDFVectorizer to predict the result by using the labels such as Positive label, Negative label and Neutral label. It also displays the WordCloud and Piechart for the sentiments class score distribution.

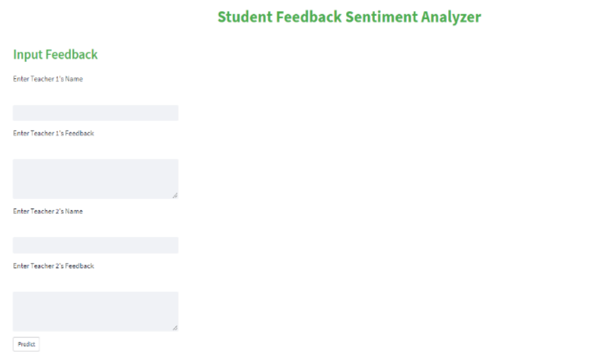

Fig. 14 Streamlit Interface

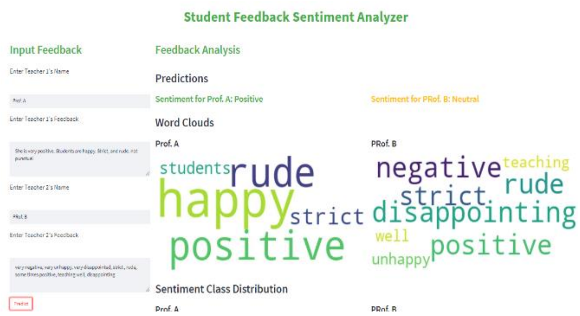

Fig. 15 Prediction on Students Feedback

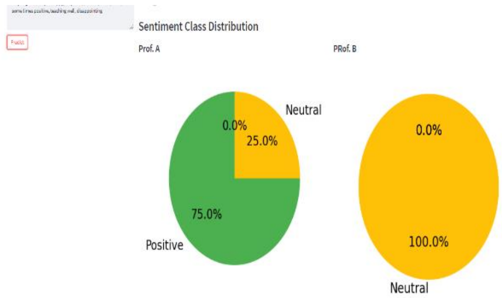

Fig. 16 Streamlit- Distribution of Sentiments

IV. Results And Discussion

To improve sentiment analysis of student feedback, we experimented with Decision Tree, Random Forest, CNN, and XGBoost models. Each model was trained using a combination of techniques, including TF-IDF vectorization for text feature extraction and SMOTE for balancing the dataset.

a. Decision Tree

We used a Decision Tree classifier to capture the hierarchical structure of feedback sentiments. Despite using TF-IDF and SMOTE to balance the data, the Decision Tree model achieved 68% model accuracy and 0.56 F1 score which indicates its limited capability to generalize sentiment patterns effectively.

b. Random Forest

Implementing a Random Forest classifier aimed to improve prediction by averaging multiple decision trees. The model was trained with TF-IDF features and balanced using SMOTE. However, it only marginally improved the results, achieving 72% model accuracy and a 0.70 F1 score, it is showing slight improvement but still not optimal.

c. Convolutional Neural Network (CNN)

We designed a CNN model to capture spatial patterns in text data using word embeddings. Despite extensive training and hyperparameter tuning, the CNN model achieving 74% model accuracy and a 0.72 F1 score. It is better than tree-based models, it struggled with text classification due to the imbalanced data.

d. XGBoost

Finally, we used XGBoost, a powerful gradient boosting algorithm, for classification. Trained on TF-IDF features and balanced with SMOTE, XGBoost performed best among the traditional models but still fell short, 76% model accuracy and 0.74 F1 score, indicating that it handled imbalance better but not sufficiently.

e. SVM with SMOTE

The sentiment analysis model achieved an overall accuracy of 86.76% on a dataset of 340 data points, splits into the training and testing sets having data points 272 and 68 respectively. Before applying SMOTE to address class imbalance, the training data had varying counts across sentiment classes (Positive: 202, Neutral: 46, Negative: 24). Post-SMOTE, the dataset was balanced with 202 samples per class. The model, based on a SVM with a RBF kernel and class weighting, achieved high precision and recall for positive (87% and 96%, respectively) and negative sentiments (100% precision and 67% recall). However, neutral sentiments exhibited lower precision (78%) and recall (58%), with misclassifications primarily as negative sentiments. The use of TfidfVectorizer to convert text data into TF-IDF features, SMOTE for balancing, and joblib for saving models were crucial technologies in achieving these results.

Table 2 Sentiment Class Distribution Before SMOTE

Class distribution before SMOTE:

	Sentiment Class	Count
0	Positive	202
1	Neutral	46
2	Negative	24

Table 3. Sentiment Class Distribution After SMOTE

Class distribution after SMOTE:

	Sentiment Class	Count
0	Negative	202
1	Neutral	202
2	Positive	202

SMOTE balanced the dataset by increasing the number of samples in each sentiment class to 202, ensuring equal representation for negative, neutral, and positive sentiments from the trained data.

Table 4. Model Evaluation Matrix

Confusion Matrix:

	Predicted: Positive	Predicted: Neutral	Predicted: Negative
Actual: Positive	4	0	2
Actual: Neutral	0	7	5
Actual: Negative	0	2	48

The table 4 demonstrates the evaluation matrix of the model is correctly predicting 4 out of 6 positive sentiments, 7 out of 12 neutral sentiments, and 48 out of 50 negative sentiments. The confusion matrix indicates the model effectively predicted negative sentiments but struggled with neutral sentiments.

The overall accuracy of a classification model is typically calculated using the formula:

Accuracy = Correct Predictions / Total Predictions

In the context of our student sentiment analysis model (SVM with SMOTE):

Correct Predictions: This includes the summation of True Positives (TP) and the summation of True Negatives (TN).

So total Predictions: is the sum of all predictions made by the model.

TP (Positive): 4,

TN (Neutral + Negative): 7 (Neutral) + 48 (Negative) = 55

Correct Predictions: TP + TN = 4 + 55 = 59

Total Predictions: Sum of all elements in the confusion matrix = 4 + 0 + 2 + 0 + 7 + 5 + 0 + 2 + 48 = 68

Therefore, the overall accuracy of our student sentiment analysis model (SVM with SMOTE) is:

Accuracy

= 59/68 ≈ 0.8676

So, the overall accuracy of your model is approximately 86.76%.

Table 5. Classification Report of SVM

Test Accuracy:
0.8676470588235294

Classification Report:

	precision	recall	f1-score	support
Negative	1	0.666667	0.8	6
Neutral	0.777778	0.583333	0.666667	12
Positive	0.872727	0.96	0.914286	50
accuracy	0.867647	0.867647	0.867647	0.867647
macro avg	0.883502	0.736667	0.793651	68
weighted avg	0.867201	0.867647	0.860504	68

The sentiment analysis SVM with SMOTE model achieved an overall accuracy of 86.76%. It showed high precision and recall for positive (87.27%, 96%) and negative (100%, 67%) sentiments, but lower precision and recall for neutral (77.78%, 58.33%) sentiments.

Our model has achieved the overall accuracy approximately 86.76% and showing the effectiveness in predicting the positive and negative sentiments, and dueto the SMOTE it also able to classify neutral class.

The researcher employed the integration of SVM with the SMOTE to balance the class, and the Streamlit approach for model deployment leveraging the promising potential in enhancing automated Student Feedback Sentiment Analyzer System.

V. Conclusions

This research has demonstrated the effectiveness of combining SVM, SMOTE, and Streamlit in enhancing an automated feedback analyzer system for sentiment analysis. The SVM model achieved an overall accuracy of 86.76% on a balanced dataset using SMOTE, effectively predicting positive and negative sentiments, though it showed challenges in accurately classifying neutral sentiments. The integration of Streamlit facilitated the deployment of this model into an interactive web application, enabling real-time feedback sentiment analysis across various domains such as customer service and social media monitoring. The use of TfidfVectorizer for text feature extraction and SMOTE for class balancing proved essential in improving model performance and handling class imbalance in sentiment analysis. This system provides actionable insights from user feedback and highlights the potential of SVM, SMOTE, and Streamlit in developing efficient and scalable automated feedback analyzer systems. Moving forward, further research could focus on optimizing neutral sentiment classification and exploring advanced NLP technology to enhancing the model accuracy and robustness of the sentiments analysis on the diverse datasets. Overall, our research contributes in the latest advancements in sentiment analysis methods, emphasizing the importance of addressing class imbalance and optimizing feature extraction methods for more accurate sentiment classification.

VI. Suggestions And Future Work

In the future, our system will enable real-time analysis of teacher feedback, predicting sentiment instantly upon submission. We plan to expand its accessibility to the public via the internet, allowing students to submit feedback after each session. Teachers can access comprehensive feedback reports over time, and higher authorities can use this data for overall performance assessment and HR appraisals. Currently, we focus solely on textual data; future enhancements will incorporate quantitative data, utilizing various parameters to collect real-time student feedback and generate comprehensive performance reports.

Acknowledge

During the entire research, the support and guidance of my colleagues and mentors were crucial in the completion of this research work and I am always thankful to them.

References:

[1] Bhowmik A, Nur, Miah, M. S. U., Karmekar, D., Aspect Based sentiment analysis model for evaluating teachers performance from student feedback, AIUB Journal of Science and Engineering, 2023, DOI: https://doi.org/10.53799/AJSE.V22I3.921.

[2] Faizi, R., Using sentiment analysis to explore student feedback: A lexicon approach, International Journal of Emerging Technologies in Learning, 2023.

[3] Grimalt Alvaro, C., Usrat M, Sentiment analysis for formative assessment in higher education: A systematic literature review, 2023, Journal of Open Access.

[4] Fargues M., Kadry S., Lawal I. A., S., H. T., Automated analysis of open-ended students feedback using sentiment, emotion and cognition classifications, Applied Sciences (Switzerland), 13(4), Article 2061, https://doi.org/10.3390/app13042061.
[5] Jagtap, B., Dhotre, V., SVM and HMM based hybrid approach of sentiment analysis for teacher feedback assessment, 2014, IJETTCS.
[6] Su, B., Peng, J., Sentiment analysis of comment texts on online courses based on hierarchical attention mechanism, Applied Sciences (Switzerland), 2023, Article 4204, https://doi.org/10.3390/app13074204.
[7] Bhowmik, A., Mohd Noor N., Saeff Ullah, M., Miah, M. Mazid-UI-Haque, M., Karmaker D., A comprehensive dataset for aspect based sentiment analysis in evaluating teacher performance, AIUB Journal of Science and Engineering, 22(2), https://doi.org/10.53799/AJSE.V22I2.862.
[8] Baral, S., Botelho, A. F., Santhanam, A., Gurung, A., Erickson, J, Hefferman, N. T., Inverstigating patterns of tone and sentiments in teacher written feedback messages, Communications in computer and Information Sciences, 2023, https://doi.org/10.1007/978-3-36336-8_53.
[9] Ren P, Yang L, Luo F., Automatic scoring of student feedback for teaching evaluation based on aspect level sentiment analysis, 2023, Education and Information Technologies. https://doi.org/10.1007/s10639-022-11151-z.
[10] Tian X, Tang S., Zhu H., Zia D., Real time sentiment analysis of student based on mini Xception architecture for wisdom classroom, concurancy and consumption: Practice and experience, https://doi.org/10.1002/cpe.7059.
[11] Usrat M, Grimalt Alvaro C., Iglesias Estrade A. M., Gender sensitive sentment analysis for estimating the emotional climate in online teacher education, Learning Environments Research., https://doi.org/10.1007/s10984-022-09405-1.
[12] Khanam Z., Sentiment analysis of user reviews in an online learning environment: Analyzing the methods and future prospects, European Journal of Education and Pedagogy, 2023, 4(2), 531-545, https://doi.org/10.24018/ejedu.2023.4.2.531.
[13] Mamidted A. D. Maulana S. S., The teaching performance of the teachers in online classes: A sentiment analysis of the students in a state university in the Philippines, Randwick International of Education and Linguistics Science Journal, 2023,
[14] Rajput Q., Haider S., Ghani S., Lexcon based sentiment analysis of teachers evaluation, Applied Computational Intelligence and Soft Computing, 2016, https://doi.org/10.1155/2016/2385429.
[15] Ortigosa A., Martin J. M., Carro R. M., Sentiment analysis in facebook and its application to e-learning, Computers in Human Behavior, 2014, 30, 625-633, https://doi.org/10.1016/j.chb.2013.05.024.
[16] Mary T. A. C., Rose P. J. L., Multifaceted sentiment detection system (MSDS) to avoid dropouts in a virtual learning environment using multi-class classifiers, International Journal of Advanced Computer Science and Applications.
[17] Kastrati Z., Dalipi F., Imran A., Nuci K. P., Wani M. A., Sentiment analysis of students feedback with NLP and deep learning: A systematic mapping study, Applied Sciences (Switzerland), 2021, 11(9), article 3986, https://doi.org/10.3390/app11093986.
[18] Sumers T., Hawkins R. D., Narasimhan K., Griffiths T. L, Learning rewards from linguistic feedback, In proceedings of 5th AAAI conference on Artificial Intelligence(AAAI 2021), https://doi.org/10.1609/aaai.v35i7.16749.
[19] Chamorro Atalaya, Sobrino Chunga L., Guerrero Carranza R., Vargas- Diaz, Poma Garcia C., Student satisfaction in the context of hybrid learning through sentiment analysis, International Journal of Evaluation and Research in Education, 2024, https://doi.org/10.11591/ijere,v13i2.26717 .Pong Inwong C., Songpan W., Sentiment analysis in teaching evaluations using sentiment phrase pattern matching (SPPM) based on association mining, International journal of Machine Learning and Cybernetics, 2019, https://doi.org/10.1007/s13042-018-0800-2.