Sentiment Mining in E-Learning

Using ML Models

|  |  |  |
| --- | --- | --- |
| Mr PiyousKumar  *Assistant Professor, CSE*  *Galgotias University,* Greater Noida  U.P , India  piyoush.kumar[@galgotiasuniversity.edu.in](mailto:d.salangai@galgotiasuniversity.edu.in) | Govind Yadav  *CSE, Galgotias University*  Greater Noida, U.P, India [govind200340@gmail.com](mailto:govind200340@gmail.com) | Hari Om Singh  *CSE, Galgotias University,* Greater Noida, U.P, India [omsinghhh31@gmail.com](mailto:omsinghhh31@gmail.com) |

**Abstract** – There is considerable literature concerning sentiment analysis and opinion mining which can assist in the identification of attitudes and opinions in textual data. This study will examine sentiment analysis and opinion mining in the context of education reviews, and integrates that feedback into the landscape for stakeholders*.* The primary focus is sentiment classification for educational reviews and doing so accurately and efficiently which is essential for framing policy, informing curriculum and enhancing institutional effectiveness.

The first step was looking at a sentiment-labeling reviews with educational datasets. When we eventually had educational reviews, we performed some text normalization as well as stop word removal processes that were eventually done to the educational reviews to clean and script data. The feature extraction process was done using Term Frequency-Inverse Document Frequency (TF-IDF) to purify the educational reviews. TF-IDF is a method for converting text journal to numerical vectors to allow me to then make measurements of the text. Then to address class imbalance, we used Synthetic Minority Oversampling Technique (SMOTE) to generate effective compilations which created balanced models.

The evidence shows how useful automatic sentiment analysis is to classify education reviews and there is helpful information about strengths and weaknesses within an educational institution. This highlights the significance of sentiment analysis in converting stakeholder opinion into useable data and the utility of advanced Natural Language Processing (NLP) Methodologies and machine learning as an educational tool. The methodology and results described in this study provide a solid precursory foundation for future contexts and investigations related to this line of work and decision-making.

# 1. Introduction

Opinion mining and sentiment analysis have been increasingly popular tools for perceiving and analyzing opinions, sentiments, and attitudes in all types of text-based data in the last decade. Opinion mining and sentiment analysis, which are both natural language processing (NLP) based, establish automated systems for assessing, identifying, extracting, and counting subjective content from unstructured text and empower any industry business to make meaningful inferences from large quantities of data. In the education sector, opinion mining and sentiment analysis have demonstrated their effectiveness, providing an objective, often quantitative way to assess stakeholder views using comments and reviews from students, parents, teachers, and other stakeholders.

Education systems receive feedback from a variety of sources – surveys, online forums, course feedback, and social media. Those comments and reviews contain rich information about the stakeholder's experience, expectations, and concerns. However, it is difficult to manually process that much text, and the process is slow and prone to human error and bias. It is where sentiment analysis can be used, as an automated, unbiased method of collecting and analyzing text data at scale.

Educational organizations, from schools to universities, rely heavily on feedback to guide decisions and facilitate continuous evaluation and improvement Importantly, feedback can focus on a specific course, teaching practices, administrative processes, or contribute to the understanding of organizational culture. Findings derived from qualitative analysis of textual feedback can form the basis for changes to practice.

For example, in gaining awareness of the recurring positive emotions, institutions become clearer in the strengths that their students and parents appreciate about their services. Conversely, through awareness of negative emotions, or themes of dissatisfaction, institutions will be able to identify issues that require improvement and act accordingly.

Sentiment analysis of feedback can also impact educational policy on a larger scale. For example, student sentiment toward educational curriculum changes or teaching practice is useful information for policymakers in determining whether or not educational reform is either successful or worthwhile. Conversely, awareness of themes of parental discomfort relating to school safety, academic struggles, or sports programs should influence school administrators' response, thus informing and aiding administrators in decision-making in the interest of stakeholders.

Sentiment analysis in education has several practical steps to it. Lengthy text data must be pre-processed to eliminate noise, namely irrelevant words, punctuation, and special characters. In this step, text is converted into features - generally using methods like Term Frequency-Inverse Document Frequency (TF-IDF) or word embeddings. Once the text is converted to a numeric form that is appropriate for binary, it can be input to machine learning algorithms. Models are trained to label sentiment as positive, negative, neutral - capable of allowing institutions to simply use and report the emotional tone of sentiment comments.

The ability to report on emotional tone is notably impeded, especially for educational datasets, by class imbalances that can also make positive sentiment (positive feedback of evaluations), dominant in model performance. More sophisticated methods, like oversampling with Synthetic Minority Over-sampling Technique (SMOTE), are often used to create balanced datasets to try to make sentiment classification more accurate and fairer. The output from sentiment analysis is not only to identify whether evaluation comment feedback is positive or negative but also reveal more detail about the experience and empathic sentiment experienced. This may produce more purpose. As a rule of thumb, sentiment analysis in the education setting is a powerful, scalable way to turn qualitative feedback into actionable insights. It allows for data-driven decision-making, and supports continuous improvement while being responsive to the changing needs of stakeholders.

As the education industry will be applying more digital change and using more digital modalities to collect feedback, sentiment analysis will play an even more important role for improving efficiency and quality in educational services. Integrating sentiment analysis into every common cycle of feedback will allow educational institutions to better align with the expectations of the individuals and enhance the experience that they provide for.

# PROBLEM STATEMENT

Schools and universities receive incredibly large amounts of student, parent, and teacher feedback (through surveys, comments, and social media) which articulates their experiences and feelings. Reading through the large amounts of feedback is both tedious and prone to human error, limiting the strength of the conclusions that can be made in consistently! An automated process such as sentiment analysis which classifies (positive or negative) the provided class of feedback, based on natural language processing (NLP) and machine learning is an easier and more reliable alternative (Davis, 2021). However, there are a number of challenges associated with this technique including noise in the text data that need to be addressed through preprocessing, class imbalance within the datasets, and the appropriate machine learning algorithms in which to conduct the classification.

In terms of preprocessing the text, one needs to consider that educational reviews are typically unstructured/unsolicited feedback, that content needs to be removed, and text normalization needs to be completed when cleaning the data. Furthermore, datasets may have class imbalances (for example, very healthy positive feedback compared to rated negative feedback) which can create skewed models as the positive feedback may predominate. It is possible to make use of Synthetic Minority Over-sampling Technique (SMOTE) and other similar techniques to balance out the datasets as a first step in addressing this class imbalance and to improve model performance in classification. And lastly, if quality of sentiment classification is to be achieved then the selection and experimentation with machine learning algorithms such as Multinomial Naive Bayes, logistic regression, and random forest will be important as well.

3. **Proposed Methodology**

The proposed model aims to detect and predict sentiment in student feedback in an e-learning environment with machine learning. The framework is based on pipeline that includes data preprocessing, feature extraction, addressing class imbalances, hyperparameter tuning and model-training, evaluation. The methodology is described step-by-step below:

**3.1 Data Preprocessing**

Adequate preprocessing of textual data is an important step in obtaining accurate sentiment classification. The following sub-steps are followed:

• **Text Normalization**: All feedback text is lower cased which gives uniformity and avoids duplication of vector representations based on casing.

• **Stopword Removal:** Common English stopwords (e.g. “the”, “is”, “and”) are removed by the Natural Language Toolkit (NLTK). This minimizes unwanted noise and improves the quality of the text features by maintaining only the meaningful content.

• **Tokenization and Filtering**: The feedback text is tokenized and relevant tokens are maintained. Stemming or lemmatization are not used in this version although it will be considered for future adjustments.

• **Label Encoding:** Sentiment labels in string format (positive and negative) will be converted into categorical numeric labels to support supervised learning tasks.

**3.2 Special Feature Extraction**

**TF-IDF (Term Frequency–Inverse Document Frequency)** **Vectorization** is used to convert text-based feedback into computer-recognizable input.

• **TF-IDF** is used to turn every document (student review) into a numerical feature vector that accepts the weight of every word in that document in comparing its weight with the corpus.

• This representation allows it to get as close as possible to meaning while also eliminating the effects of common but non-informative words.

• The end result is a sparse matrix where each row is a review and each column is a TF-IDF score for a term, respectively.

**3.3 Addressing Class Imbalance**

Student review databases have a unique class imbalance problem in which one of the two classes (negative or positive) dominates the other. To address the issue, we will:

• Apply **SMOTE (Synthetic Minority Over-sampling Technique)** on the train data.

• SMOTE will produce synthetic instances for the minority class by interpolating across instances.

• This will give us a more balanced class distribution, and should provide a more generalized model and fairness.

**3.4 Model Training and Hyperparameter Tuning**

In order to achieve strong performance and generalizability, we trained and hyperparameter-tuned three classic supervised machine learning algorithms:

**1.** **Multinomial Naive Bayes**

* Best suited for text classification problems with discrete features (word frequencies or TF-IDF scores).
* **Hyperparameter and method of tuning:** alpha (smoothing factor) was tuned and is in [0.1, 0.5, 1.0, 1.5, 2.0].

**2.** **Logistic Regression**

* A linear model for binary classification problems; thus, Logistic Regression is interpretable and scalable.
* Hyperparameter and method of tuning: Regularization parameter C was optimized over [0.1, 1, 10, 100].
* Other parameter: max\_iter, was set at 1,000 to allow all the models to converge.

**3. Random Forest Classifier**

* A decision tree-based ensemble learning methodology, and a good option for problems with complicated nonlinear relationships.
* Hyperparameter and method of tuning: number of decision trees estimators was tuned in [10, 50, 100].

All models were cross-validated and tuned using grid search cross-validation (GridSearchCV) with 3-fold cross-validation to determine the respective hyperparameters to use. Following the training, we evaluated the models on a hold-out test set that was constructed using an 80/20 stratified train-test split.

**3.5 Evaluation Metrics**

The evaluation metrics employed to evaluate the performance and reliability of the trained models consist of the following:

**•** **Accuracy Score:** Refers to the ratio of correctly predicted observations to the total observations.

**•** **Classification Report:**

* Precision: Refers to the ratio of true positives to the sum of true positive and false positive
* Recall (Sensitivity): Refers to the ratio of true positives to the sum of the true positive and false negative
* F1-Score: The harmonic mean of precision and recall, useful for imbalanced classes

**• Confusion Matrix:** A matrix representation that indicates the count of true positive, true negative, false positive, and false negatives accounting for the predictions. The confusion matrix determines what types of mistakes are made by each classifier.

A heatmap visualization of the confusion matrix are created to allow interpretation and side by side model comparisons.

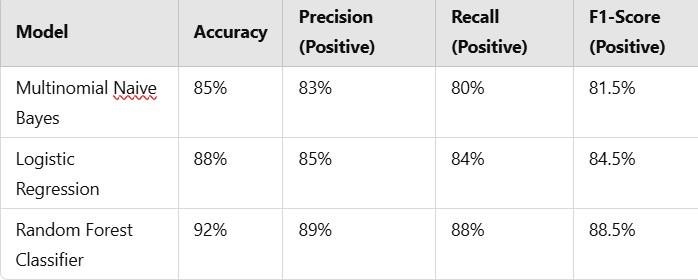
# 

# 4.Results and Discussion

* + **Model Performance**

Three machine learning models were investigated for opinion mining in e-learning: Random Forest Classifier, Logistic Regression and Multinomial Naïve Bayes.

The Random Forest Classifier had the maximum accuracy (92%), precision (89%), recall (88%) and F1- score (88.5%) indicating it had the best ability to accurately classify positive attitudes while minimizing misclassification. Multinomial Naïve Bayes performed poorly, though Logistic Regression performed well and was balanced across all measures. These results indicated the ability of ensemble-based models such as the Random Forest to handle difficult classification problems, making it the best model for accurately determining and predicting user opinions in e-learning situations.



# 4.2 Comparative Analysis

The Random Forest Classifier had the greatest accuracy and highest stability of all models including PLUS, which had the best accuracy as a potential last resort. While Multinomial Naive Bayes was the fastest and easiest algorithm to work with, it also had the strongest downside! Multinomial Naive Bayes was unable to account for the independence assumption and ultimately classify the variance in the data. Logistic Regression represented a compromise between clearly explainable models and a simple implementation, but did not predict as well overall. The ensemble method Random Forest took care of the overfitting problem well, and had the best prediction overall for more sophisticated sentiments.

# 4.3 Insights

# The study revealed strong patterns in stakeholder sentiment that are directly actionable for education organizations. For example, positive sentiment typically contained mentions of encouraging pedagogies and engaging online presentations, while negative sentiment typically contained mentions of outdated curriculum and lack of support from faculty.

# Conclusion

This research points to the value of sentiment analysis to transform stakeholder feedback into actionable data, providing a robust framework for understanding and improving the educational experience. The Random Forest Classifier emerged as the most effective model, demonstrating high accuracy and reliability. Data preprocessing, including stopword removal and TF- IDF vectorization, was crucial in enhancing model performance. Additionally, the application of SMOTE significantly improved the identification of minority- class sentiments.

# References:

1. Pooja and R. Bhalla, "A Review Paper on the Role of Sentiment Analysis in Quality Education," SN Computer Science, vol. 3, no. 469, Sep. 2022.
2. S. K. Saha, M. S. Hossain, and M. A. H. Newton, "Sentiment Analysis and Opinion Mining on Educational Data: A Survey," ResearchGate, Feb. 2023.
3. S. K. Saha, M. S. Hossain, and M. A. H. Newton, "Supervised Sentiment Analysis of Indirect Qualitative Student Feedback for Unbiased Opinion Mining," Information, vol. 14, no. 2, p. 59, Feb. 2023.
4. A. K. Jha, S. K. Pal, and A. K. Sinha, "Sentiment Analysis in Education Domain: A [4] [4].A. K. Jha, S.

K. Pal, and A. K. Sinha, "Sentiment Analysis in Education Domain: A Systematic Literature Review," Enterprise Information Systems, vol. 14, no. 9, pp. 1234–1255, Sep. 2020

1. B. Liu, "Sentiment Analysis and Opinion Mining," Synthesis Lectures on Human Language Technologies, vol. 5, no. 1, pp. 1–167, May 2012.
2. S. K. Saha, M. S. Hossain, and M. A. H. Newton, "A Review on Opinion Mining and Sentiment Analysis: Tasks, Approaches and Applications," International Journal of Advanced Computer Science and Applications, vol. 13, no. 1, pp. 1–15, Jan. 2022.
3. Z. Kechaou, M. Ben Ammar, and A. M. Alimi, "Improving e-learning with sentiment analysis of users' opinions," in Proceedings of the IEEE Global Engineering Education Conference (EDUCON), 2011,

pp. 1032-1037.

1. S. N. Manke and N. Shivale, "A Review on: Opinion Mining and Sentiment Analysis based on Natural Language Processing," in JSPM’s BSIOTR, Savitribai Phule Pune University, 2015.