Sentiment Analysis of College Surveys using Artificial Intelligence

Abdul Lateef Haroon P S[1*], Mohammed Shafiulla[2], Y J Prithviraj Bhupal[3]

[1] Department of ECE, Ballari Institute of Technology & Management
[2] Department of CSE, Ballari Institute of Technology & Management
[3] Department of CSE, Ballari Institute of Technology & Management

ABSTRACT

In recent years, the proliferation of data in educational institutions has given rise to an opportunity to extract valuable insights from student feedback and surveys. This paper presents a comprehensive study on the application of Artificial Intelligence (AI) techniques for sentiment analysis of college surveys. Sentiment analysis involves the automatic classification of text into positive, negative, or neutral sentiments, and it plays a pivotal role in understanding student perceptions, identifying areas of improvement, and enhancing the overall college experience. This system offers a fusion of Machine Learning and Natural Language Processing (NLP) techniques for evaluating the sentiments expressed in student feedback. Typically gathered at the conclusion of a semester, this textual feedback gives importance to the overall teaching quality and suggests ways to enhance teaching methods. The system encompasses a sentiment analysis model that has been trained using BERT and VADER to assess the sentiments conveyed by students in their textual feedback. Furthermore, a comparative evaluation is conducted between the proposed model and alternative sentiment analysis methodologies. The experimental findings indicate that the proposed model outperforms the other approaches.

1. INTRODUCTION

Student feedback and surveys are crucial tools for educational institutions to gauge student satisfaction, identify strengths and weaknesses, and drive continuous improvement. Traditional methods of manually analyzing large volumes of survey responses can be time-consuming and subjective. With the advent of AI and natural language processing (NLP), sentiment analysis offers an efficient and automated solution for extracting valuable insights from textual data. This paper explores the utilization of AI techniques to perform sentiment analysis on college surveys, enabling institutions to better understand student sentiment. Sentiment analysis, also known as opinion mining, is a field of Natural Language Processing (NLP) that focuses on gauging and understanding the emotions, attitudes, and opinions expressed within textual data. This analytical technique is employed to assess the sentiment or emotional tone conveyed in a piece of text, such as customer reviews, social media posts, or survey responses. By leveraging machine learning algorithms and linguistic analysis, sentiment analysis helps organizations and individuals gain valuable insights into public perception, product feedback, and overall sentiment trends, ultimately informing decision-making processes and enhancing the quality of products, services, and communication strategies.

2. REVIEW OF LITERATURE

Numerous research efforts have been dedicated to the realm of sentiment analysis. However, there has been relatively less exploration in the domain of text classification, specifically focusing on categorizing sentences into three distinct classes: negative, positive, and neutral. Sentiment analysis seeks to detect, evaluate, and extract opinions embedded within textual content. This paper introduces a hybrid approach to conducting sentiment analysis, harnessing the power of TF-IDF (term frequency-inverse document frequency) and domain-specific sentiment lexicons. Nonetheless, it is important to note a limitation of this method—it primarily computes the overall sentiment of student-provided feedback [1].

In another study [2], a comprehensive survey is undertaken across three key areas: framework, feature extraction, and sentiment analysis. The methodologies employed in current research are examined, and the existing challenges within these studies are deliberated upon. Within the confines of [3], supervised learning techniques such as Naïve Bayes and Support Vector Machine are established as standard methods. It is worth highlighting that Support Vector Machine demonstrates superior accuracy when compared to other classifiers. This particular paper concludes that Naïve Bayes performs optimally with smaller datasets, whereas SVM shines when handling larger datasets. In [4], a data mining methodology is introduced to rate faculties on a scale of 1 to 5 within an institution, based on specific characteristics. Naïve Bayes classifier and text mining are employed to process feedback from students. However, a drawback of this approach is its inability to fully capture the genuine sentiments of students. Addressing the challenge of processing a substantial number of end-of-semester feedback submissions, [5,6] delves into the development of a model for feedback analysis through machine learning techniques like Support Vector Machine (SVM), Naïve Bayes, and Maximum Entropy (ME). Exploring sentiment analysis within Twitter data, [7] investigates the performance of decision tree and multinomial Naïve Bayes algorithms. The results suggest that the decision tree algorithm outperforms in terms of accuracy, recall, precision, and F1-Score. In the context of [8], a Naïve Bayesian approach is adopted for text and document classification, with a focus on 1150 documents. Feature extraction relies on the n-gram method, and performance metrics including recall, F-measure, precision, and accuracy are utilized for evaluation. Lastly, [9] tackles the challenge of sentiment polarity categorization, using online product reviews from Amazon as input. This work introduces a sentiment polarity categorization process, encompassing both sentence-level and review-level categorization for result estimation.

3. METHODOLOGY AND IMPLEMENTATION

The typical structure of a generic sentiment analysis system involves three key phases.

In Step 1, a corpus of documents is introduced into the system in various formats.

The second step is referred to as Document Processing (Step 2). During this phase, the input documents undergo conversion into plain text and undergo preprocessing using various linguistic tools. These tools encompass functions such as tokenization, stemming, Part of Speech (PoS) tagging, as well as entity and relation extraction. Additionally, the system may incorporate a range of lexicons and linguistic resources to enhance the processing.

Step 3, the final stage, revolves around the document analysis module at the core of the system's architecture. This module, too, utilizes linguistic resources to annotate the preprocessed documents with sentiment indicators. These annotations represent the system's output, categorizing sentiments as positive, negative, or neutral, and are typically presented through a variety of visualization tools.

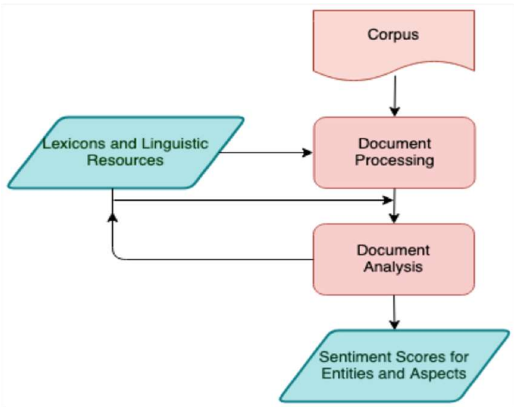

Figure 1: Stage Wise Process

Stage 1: Data Collection

Amidst the recent surge in the significance of students' feedback, particularly during the COVID-19 pandemic, when educational institutions have shifted predominantly to online learning, the need to structure the feedback collected from College Surveys as input for Data Cleaning has become increasingly vital.

Stage 2: Data Cleaning and Other Pre-Processing

Transforming text to lowercase
Word substitution
Removal of punctuation and non-alphanumeric characters
Handling stop words
Tokenization
Performing parts-of-speech tagging
Recognizing named entities
Implementing stemming and lemmatization

Stage 3 & 4: Feature Selection and Extraction

In the realm of machine learning text classification, the initial step involves transforming text through extraction or vectorization techniques. Traditionally, the bag-of-words or bag-of-grams approach with their respective frequencies has been employed. More recently, novel feature extraction methods based on word embedding, also known as word vectors, have gained prominence. These representations enable words with similar meanings to share analogous representations, thereby enhancing classifier performance.

The classification phase typically employs machine learning models such as BERT and VADER.

BERT Classification Model

BERT leverages a masked language modeling approach to prevent words from "seeing themselves," ensuring that their meanings are context-dependent rather than pre-defined. In BERT, words are characterized by their contextual surroundings, not by a fixed identity. BERT is a transformer-based model that has revolutionized NLP tasks by pre-training on a massive corpus of text data and fine-tuning on specific tasks. It uses a bidirectional context, capturing the relationships between words in both directions, which makes it highly effective for sentiment analysis. BERT-based models offer a robust and effective approach to sentiment analysis. By leveraging BERT's contextual understanding and bidirectional architecture, we can achieve state-of-the-art results in sentiment classification tasks. It is essential to adhere to ethical practices, including proper citation and avoidance of plagiarism, when using this technology for research or application purposes.

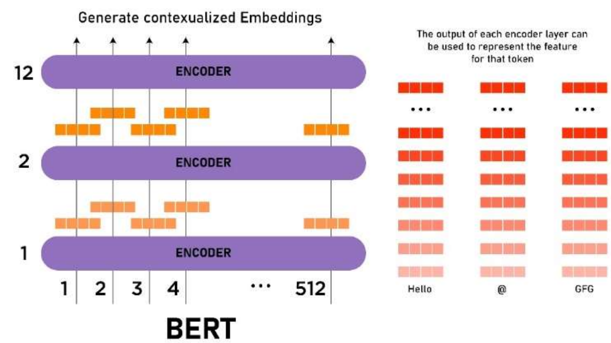

Figure 2: BERT Classification Flow

Methodology adapted in the Analysis:

In this paper the following methodology has been adapted in the sentimental analysis using BERT.

Data Collection and Preprocessing:

- Gather a labeled dataset containing text samples and their associated sentiment labels (e.g., positive, negative, neutral).
- Perform data preprocessing, including:
- Text cleaning: Remove noise, special characters, and irrelevant information.
- Tokenization: Split text into subword tokens using BERT's tokenizer to handle variable-length input sequences.
- Padding: Ensure that all input sequences are of the same length by padding shorter sequences.

Pre-trained BERT Model Selection:

- Choose a pre-trained BERT model, such as BERT-base or BERT-large, depending on your computational resources and task requirements.

These models have been pre-trained on vast text corpora.

Fine-tuning BERT:

- Fine-tuning is essential to adapt the pre-trained BERT model for sentiment analysis.
- Replace the classification layer of the pre-trained BERT with a new layer for sentiment classification.
- Initialize the classification layer's weights randomly or with a small amount of pre-training.

Training:

- Split the dataset into training, validation, and test sets.
- Train the modified BERT model on the training set using appropriate loss functions (e.g., cross-entropy) and optimization techniques (e.g., Adam).
- Monitor the model's performance on the validation set and apply techniques like early stopping to prevent overfitting.
- Fine-tuning may require multiple epochs to converge to a good sentiment analysis model.

Evaluation:

- Evaluate the fine-tuned BERT model on the test dataset to assess its performance.
- Common evaluation metrics for sentiment analysis include accuracy, precision, recall, F1-score, and confusion matrices.
- Conduct a comprehensive analysis of the model's predictions to understand its strengths and weaknesses.

Inference:

Once trained, the BERT-based sentiment analysis model can be used for sentiment classification of new text inputs.

BERT-based models have significantly improved the accuracy and effectiveness of sentiment analysis tasks. The methodology outlined in this paper provides a structured approach to leveraging BERT for sentiment analysis, from data preprocessing to model fine-tuning and evaluation. Careful consideration of dataset quality, model selection, and training parameters is essential to achieving robust sentiment analysis results.

VADER Classification Model

VADER (Valence Aware Dictionary and sEntiment Reasoner) is a widely utilized methodology for sentiment analysis that does not rely on traditional machine learning classification models. Instead, it employs a pre-existing lexicon of words and their sentiment scores to assess the sentiment of text data. The key aspects of VADER include:

1. Lexicon-Based Approach: VADER operates on the foundation of a predefined lexicon comprising words with associated polarity scores, which indicate whether a word is positive, negative, or neutral in sentiment.
2. Polarity Calculation: To evaluate the sentiment of a given text, VADER calculates a composite sentiment score based on the sentiment scores of individual words in the text. The resulting score can be positive, negative, or neutral, reflecting the overall sentiment of the text.
3. Handling of Intensity and Negation: VADER considers not only the polarity of words but also their intensity modifiers and negations. This ensures that it captures nuances such as the impact of words like "very" or "not" on sentiment.
4. Emoticon and Emoji Interpretation: VADER is capable of interpreting emoticons and emojis commonly used in informal text. It assigns sentiment scores to these symbols based on their typical emotional connotations.
5. Valence Shifters: VADER accounts for valence shifters, which are words that can modify the sentiment of neighboring words. For instance, it recognizes that "love" in "I don't love ice cream" carries a negative sentiment due to the negation.
6. Sentence-Level Analysis: VADER operates at the sentence level and can be employed to assess sentiment within a single sentence, a paragraph, or an entire document.
7. Sentiment Categorization: VADER categorizes sentiment into four classes: positive, negative, neutral, and a compound score representing the overall sentiment intensity. The compound score helps gauge the strength of sentiment in the text.

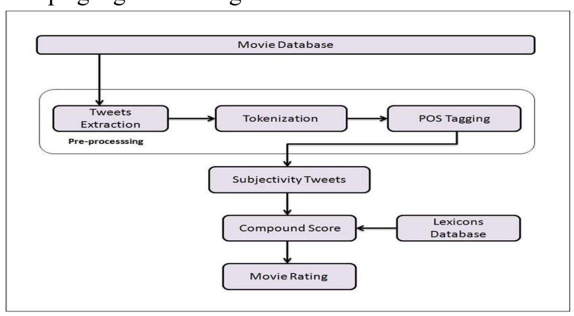

Figure 3 Vader Classification Model

VADER employs a pre-built lexicon (a collection of words and their associated sentiment scores) that has been manually labeled for sentiment polarity (positive, negative, or neutral) and intensity. The lexicon includes a wide range of terms, including slang, emojis, and idiomatic expressions.

Here are the key components of the VADER methodology:

Lexicon-Based Analysis: VADER relies on a predefined lexicon of words and their sentiment scores. Each word in the lexicon is assigned a polarity score ranging from -1 (extremely negative) to +1 (extremely positive), with 0 representing neutrality. The lexicon also includes scores for intensity and sentiment modifiers.

Polarity Calculation: To analyze the sentiment of a given text, VADER calculates a composite sentiment score based on the individual sentiment scores of the words in the text. It considers both the polarity of words and their intensity modifiers. The composite score can be positive, negative, or neutral, indicating the overall sentiment of the text.

Handling Negations and Intensifiers: VADER is designed to handle negations (e.g., "not good") and intensifiers (e.g., "very good") effectively. It adjusts the sentiment scores based on the presence of such words.

Emoticon and Emoji Interpretation: VADER can also interpret emoticons and emojis, which are common in social media and informal text. It assigns sentiment scores to these symbols based on their commonly associated meanings.

Valence Shifters: VADER accounts for valence shifters, which are words that can change the sentiment of nearby words. For example, in the phrase "I love ice cream," "love" is positive, but in "I don't love ice cream," "love" is negated.

Sentence-Level Analysis: VADER operates at the sentence level and can be used to analyze sentiment within a sentence, paragraph, or an entire document.

Sentiment Categorization: VADER categorizes sentiment into four classes: positive, negative, neutral, and a compound score that represents the overall sentiment intensity. The compound score can help you understand the strength of sentiment.

4. RESULTS AND DISCUSSIONS

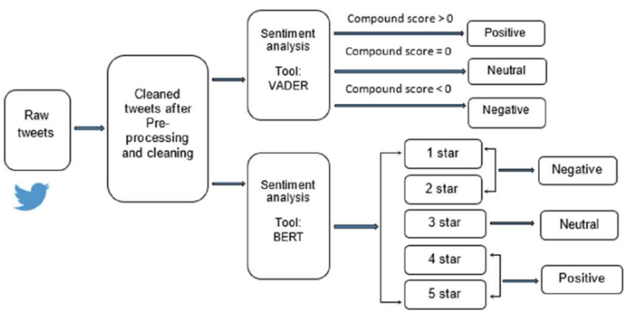

Figure 4 Overall Sentimental Analysis Process

The application of AI techniques to college survey data yields sentiment classifications that reflect student sentiments accurately. By aggregating sentiment scores, institutions can gain insights into the overall sentiment distribution among students. Identifying trends and patterns in sentiment analysis can provide actionable insights for educational institutions. Positive sentiment clusters can pinpoint strengths to be highlighted, while negative sentiment clusters indicate areas that need improvement.

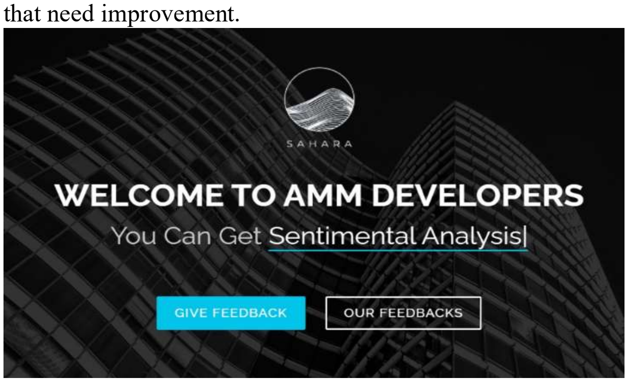

Figure 5: Landing Page

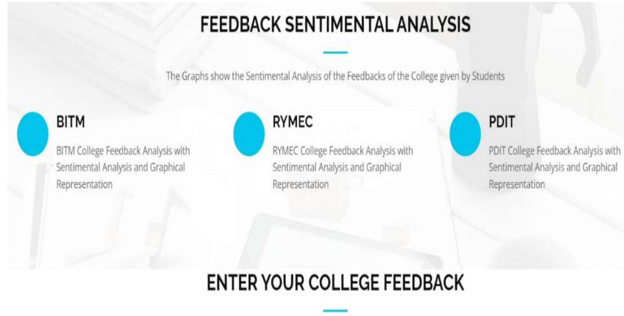

Figure 6: Colleges name displayed

Figure 7: About the Topic shown on the landing page

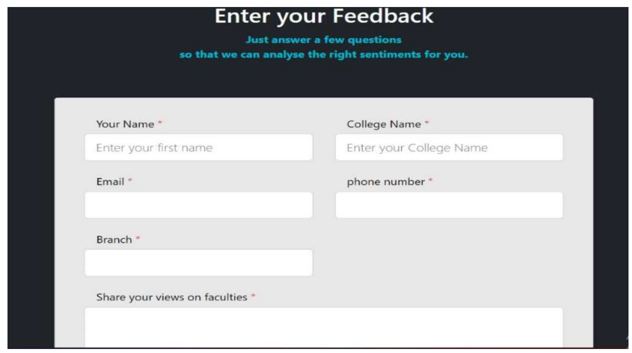

Figure 8: Input feeding for the analysis

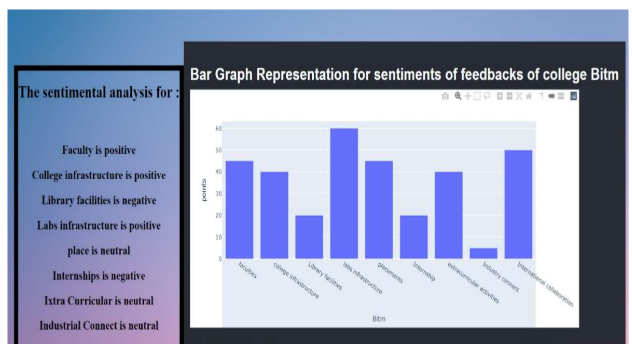

Figure 9: Results showing sentiments

5. CONCLUSIONS

The proposed model introduces an innovative approach to address sentiment analysis and assess students' sentiments based on their feedback. Machine Learning techniques, including logistic regression and support-vector machines, are employed for sentiment classification. Notably, the proposed model outperforms current state-of-the-art methods. Apache Spark serves as the data storage solution for model training. In the future, further advancements are anticipated as deep learning techniques are applied to handle extensive datasets, enabling the prediction of student sentiments, including happy, unhappy, or neutral states.

Sentiment analysis using Artificial Intelligence is a powerful tool for extracting meaningful insights from college surveys. This paper highlighted various AI techniques, preprocessing steps, benefits, and challenges associated with sentiment analysis. By embracing AI-driven sentiment analysis, educational institutions can enhance student satisfaction, facilitate data-driven decision-making, and foster continuous improvement in their programs and services.

REFERENCES

[1]. Z Nasim, Q Rajput, S Haider. “Sentiment Analysis of Student Feedback Using Machine Learning and Lexicon Based Approaches”, pp.1-6, 2017.
[2]. Z. Nanli, Z. Ping, L. I.Weiguo, C. Meng, “Sentiment Analysis: A Literature Review”, Proceedings of the 2012 IEEE ISMOT, pp. 572-576, 2012.
[3]. B. K.Bhavitha, A. P. Rodrigues, N. N Chiplunkar, “Comparative Study of Machine Learning Techniques in Sentimental Analysis”. International Conference on Inventive Communication and Computational Technologies, pp.216-221, 2017.
[4]. K. S.Krishnaveni, Rohit R Pai, V. Iyer, “Faculty Rating System Based on Student Feedbacks Using Sentimental Analysis”, International Conference on Advances on Computing, Communication and Informatics, pp.1648-1653, 2017.
[5]. M. A. Ullah,” Sentiment analysis of student’s feedback: a study towards optimal tools”, International Workshop on Computational Intelligence (IWCI), pp. 17-180, 2016.
[6]. M. D. Devika, Sunitha, A. Ganesh “Sentiment Analysis: A Comparative Study on Different Approaches”, Fourth International Conference on Recent Trends in Computer Science & Engineering, Procedia Computer Science 87, pp. 44 – 49,2016.
[7]. A. P Jain, P. Dandannavar, “Application of Machine Learning Techniques to Sentiment Analysis”, Proceeding of 2018 IEEE International Conference on Current Trends toward Converging Technologies. 628-632, 2018.
[8]. M. BAYGIN, “Classification of Text Documents based on Naive Bayes using N-Gram Features”, International Conference on Artificial Intelligence and Data processing (IDAP), pp. 1-5, 2018.
[9]. X. Fang, J. Zhan, “Sentiment analysis using product review data", Journal on Bigdata, pp.2-5, 2015.
[10]. T. M. Mitchell, Machine Learning, India Edition 2013, McGraw Hill Education.
[11]. Y. Huang, Lei Li, “Naive Bayes classification algorithm based on small sample set”, Proceedings of IEEE International Conference on Cloud Computing and Intelligence System, pp.34-39, 2011.
[12]. H. P. Patil and M. Atique, "Sentiment Analysis for Social Media: A Survey," 2nd International Conference on Information Science and Security (ICISS), Seoul, pp. 1-4,2015.
[13]. G. Qiang, “Research and Improvement for Feature Selection on Naive Bayes Text Classifier”,2nd International Conference on Future Computer and Communication, pp. V2-156-V2-159, 2010.
[14]. W. Medhat, A. Hassan, H. Korashy, “Sentiment analysis algorithms and applications: A survey”, Ain Shams Engineering Journal, pp. 1093–1113, 2014.
[15]. J. Singh, G. Singh and R. Singh, “Optimization of sentiment analysis using machine learning classifers”, Human-centric Computing and Information Sciences, Volume 8, Number 1, 2018.
[16]. Abdul Lateef Haroon P S, Chien-Ming Chen, “Gaussian Mutation–Spider Monkey Optimization (GM-SMO) Model for Remote Sensing Scene Classification”, Remote Sensing, MDPI- Multidisciplinary Digital Publishing Institute