

Enhancing Educational Experiences: A Sentiment Analysis Approach to Student Feedback

Submitted by

Ashwant Manikoth(3741680)

Joel Cherian Jacob (3741215)

Nidhin Joseph (3741627)

Table of Contents

1. PROBLEM STATEMENT	3
A. VALUE OF SOLVING THE PROBLEM	3
I. <i>Enhancing the Depth and Quality of Educational Feedback Analysis</i>	3
II. <i>Empowering Educators with Actionable Insights</i>	3
III. <i>Fostering a More Engaged and Satisfactory Learning Experience</i>	3
IV. <i>Setting a New Standard in Educational Feedback Utilization</i>	4
B. END-USER IMPACT	4
I. <i>Educators and Academic Administrators:</i>	4
II. <i>Students</i>	4
III. <i>Broader Educational Ecosystem</i>	4
C. CONSTRAINTS AND REQUIREMENTS	5
I. <i>Data Quality and Integrity</i>	5
II. <i>Computational Resources and Scalability</i>	5
III. <i>Accuracy and Reliability of the Model</i>	5
IV. <i>Usability</i>	5
D. TYPE OF ML PROBLEM	6
I. <i>Sentiment Analysis with Hierarchical Complexity</i>	6
II. <i>Neural Networks and Complex Data Interpretation</i>	6
III. <i>Balancing Computational Efficiency with Model Accuracy</i>	6
E. SUCCESS METRICS	6
2. SOLUTION DESIGN	7
A. LITERATURE REVIEW	7
I. <i>Has this problem been encountered before?</i>	7
II. <i>How was it solved? What is the state-of-the-art technique?</i>	7
III. <i>What were the limitations to that solution? (Gap in solution)</i>	8
IV. <i>What are you proposing that is “novel”?</i>	8
V. <i>References of previous related work (minimum 5)</i>	9
B. FROM PERSPECTIVE OF ML WORKFLOW	10
I. <i>Data Selection</i>	10
II. <i>Data Pre-processing</i>	11
III. <i>Model Selection</i>	12
IV. <i>Training/Fine-tuning</i>	13
V. <i>Hyperparameter tuning strategy</i>	14
VI. <i>Evaluation Metrics</i>	14
3. IMPLEMENTATION	15
A. CODE	15
B. REPOSITORY – URL	15
C. RESULTS/PLOTS – GRAPHS	15
D. RESULT EXPLANATION –	18
4. CONCLUSION AND DISCUSSIONS	19
5. PROOF OF CONCEPT – WEB DASHBOARD	20
I. <i>Implementation Details</i>	21
II. <i>Functionalities and Features</i>	21

1. Problem Statement

In the changing world of education, student feedback plays a crucial role in shaping teaching methods and course content. Traditionally, feedback has been measured using Likert scales, which provide a broad overview of student satisfaction but fail to capture the nuanced sentiments expressed. This gap in understanding the aspects of feedback poses a significant challenge for educators as it hinders their ability to effectively address students' specific needs and concerns, ultimately limiting the optimization of our learning environment. Our project aims to transform this approach by utilizing advanced neural network models for a more sophisticated analysis of sentiments. By doing so, we seek to convert students' qualitative feedback into actionable insights that give educators a deeper understanding of their experiences beyond simple numerical ratings.

The essence of our project lies in its capacity not only to grasp and categorize overall sentiment from feedback but also to delve into specific aspects of the educational experience. The thorough analysis enables educators to respond effectively and specifically to student feedback. This leads to decision-making that aligns closely with the preferences and needs of students. Such an approach promises to improve educational strategies, foster greater engagement, and increase student satisfaction. Additionally, the project's utilization of models for sentiment analysis represents an exciting convergence of advanced machine-learning techniques and practical applications in education. It introduces innovative technology to address a long-standing challenge, establishing a new benchmark for analyzing and leveraging student feedback.

In the sections, we will explore the distinct benefits of solving this problem, examine its impact on end users, discuss project constraints and requirements, delve into the nature of the machine learning problem being tackled and outline key metrics that will define the success of this endeavour.

A. Value of Solving the Problem

The goal to leverage advanced neural network models include Aspect Based Sentiment Analysis (ABSA) models for sentiment analysis in student feedback is a critical step forward in the realm of educational enhancement. This approach is poised to provide substantial value in several key areas:

I. *Enhancing the Depth and Quality of Educational Feedback Analysis*

The traditional approach of depending on Likert scales imposes constraints on the interpretation of feedback. This project seeks to employ a sophisticated sentiment analysis model to capture the nuances embedded in student feedback. By delving into this analysis, we can understand students' genuine sentiments, frequently overlooked when translating data into numbers. It is not merely about statistics but about connecting with students' authentic experiences, worries and emotions.

II. *Empowering Educators with Actionable Insights*

The real power of student feedback lies in its capacity to assist educators in refining their teaching methods and course structures. At the same time, in a hectic world where time is money, the effort and time required to deduce actionable insights from students' feedback is difficult without a Likert scale-based approach or some categorization. The project's approach guarantees that feedback is gathered, organized, and converted into insights. Educators can use these insights to make informed enhancements that tackle students' needs and concerns directly. As a result, an educational environment that is more responsive is created, where student experiences and feedback influence changes.

III. *Fostering a More Engaged and Satisfactory Learning Experience*

When students realize that their feedback is given consideration, it boosts their involvement and contentment with the learning procedure. The methodology employed by this project to analyze feedback ensures that student opinions are genuinely acknowledged and acted upon regardless of how bulk it is. It can result in enhanced course material, efficient teaching techniques and an enriching educational journey for students.

IV. Setting a New Standard in Educational Feedback Utilization

Using network models for sentiment analysis showcases an exciting fusion of technology and education. It highlights how machine learning can enhance the aspects of education and improve the overall learning environment.

To sum up, solving this problem brings value by enhancing the quality of feedback analysis, empowering educators, and enriching student experiences. The aim is to create an ecosystem that's adaptable, responsive, and deeply tuned to the needs and emotions of students. This project can potentially revolutionize how educational feedback is perceived and converted into actionable and useful insights, making it a potent tool for improving education.

B. End-User Impact

The implementation of neural network-based sentiment analysis for student feedback has a profound impact on its primary end-users: - educators and academic administrators. This multifaceted impact addresses several aspects of the educational process and experience.

I. Educators and Academic Administrators:

The improved sentiment analysis model gives educators a more nuanced understanding of student feedback, surpassing the insights obtained from numerical ratings. This enhanced insight empowers educators to decide on curriculum design, teaching methodologies and strategies for engaging students. The model's ability to identify aspect-specific issues within student feedback is crucial for targeted interventions. For instance, if a consistent pattern of dissatisfaction arises concerning teaching methods or course materials, educators can focus on revising these elements. Additionally, the network model's rapid processing and analytical capabilities enable responses to emerging trends in student feedback. This heightened responsiveness is particularly valuable in promptly addressing issues that may impact students or academic performance, ensuring relevant course adjustments.

II. Students

The detailed analysis of student feedback holds promise for benefiting students significantly. As educators implement changes based on feedback insights, students will likely find themselves in an engaging, supportive and effective learning environment. Furthermore, students can feel empowered knowing that their feedback is thoroughly analyzed and contributes to changes in their experience. This feeling of empowerment, which arises from being valued and heard, can enhance their engagement and satisfaction with the process. It creates an environment where students genuinely feel part of their journey.

III. Broader Educational Ecosystem

Implementing analytics in interpreting feedback signals a cultural shift in the education sector towards data-driven strategies. This shift opens doors to an adaptive and student-centred approach to education, where robust data insights guide informed decisions. Moreover, the successful execution of this project could establish a standard for institutions by showcasing the advantages of integrating advanced technology into feedback analysis and decision-making processes. This model demonstrates how data guides and enhances the educational process, potentially inspiring other institutions to adopt similar approaches for modernizing and improving educational outcomes.

In conclusion, the impact of this project on end users extends beyond improvements in teaching methods and course content. It represents a move towards a data-driven, responsive, and empathetic approach to education where student feedback drives continuous improvement and fosters innovation.

C. Constraints and Requirements

Specific factors need to be considered when deploying a neural network-based sentiment analysis model for student feedback in a setting. These factors play a role in shaping the design, development, and implementation of the project:

I. *Data Quality and Integrity*

The model's effectiveness in sentiment analysis relies on the amount and diversity of feedback data it handles. To truly achieve its potential, the model requires a large amount of varied feedback data that covers a wide range of sentiments and expressions. This diversity ensures that the training dataset encompasses kinds of student feedback accurately.

The model's success relies heavily on high-quality input data, necessitating preprocessing. The first step is cleaning up the text data by removing any characters and standardizing its format. This step is crucial for enhancing data quality. Additionally, any missing or incorrect data entries must be addressed during this stage. Furthermore, the text is broken into tokens to ensure consistency and uniformity during data processing.

II. *Computational Resources and Scalability*

Model's effective deployment and performance in analyzing amounts of text data, such as student feedback, rely heavily on having sufficient computational resources. These complex and deep models require processing power for training and operation. Ensuring the availability of this processing power is crucial during both the training and deployment phases. Additionally, the solution needs to be scalable, meaning it can handle varying volumes of feedback during academic periods when there is a surge in the quantity of feedback. This scalability ensures performance and accuracy in sentiment analysis regardless of the amount of data being processed.

III. *Accuracy and Reliability of the Model*

To be considered reliable and practical, the sentiment analysis model should achieve accuracy in sentiment classification and demonstrate robustness and generalizability. High accuracy guarantees that the model can consistently classify sentiments in student feedback, which is required for practical application. In addition to accuracy, the model must handle and interpret styles and formats of feedback effectively since they can vary significantly among students and academic contexts. Moreover, the model must be able to apply its knowledge across courses and academic settings. This emphasizes its adaptability and usefulness in environments. The combination of performance, resilience and versatility is key to developing a sentiment analysis tool that's reliable and widely applicable in education.

IV. *Usability*

In an educational setting, the effectiveness of the sentiment analysis solution relies on its capabilities and user-friendliness. Educators and administrators should have access to, and a clear understanding of the insights provided by the analysis. The presentation of results should be accessible and intuitive, enabling interpretation and application of data for enhancements. Achieving seamless integration with a user-friendly interface for adoption and effective utilization in environments is essential. Moreover, the model must be able to apply its knowledge across courses and academic settings. This emphasizes its adaptability and usefulness in environments. The combination of performance, resilience and versatility is key to developing a sentiment analysis tool that's reliable and widely applicable in education.

Addressing these limitations and requirements is vital for project success. It involves balancing feasibility and ethical considerations while ensuring scalability and robustness. The goal is to create a user solution that adds value to the educational process.

D. Type of ML Problem

The project at hand addresses a complex and multifaceted machine learning problem, primarily situated within the realm of Natural Language Processing (NLP), but with distinct characteristics that set it apart:

I. *Sentiment Analysis with Hierarchical Complexity*

At its core, this problem involves sentiment analysis, which is a common task in NLP. However, what sets it apart is the need to analyze sentiments through different layers which gives more detailed insights, even at different aspect levels. Unlike typical sentiment analysis that classifies sentiments as positive, negative, or neutral, this problem requires delving into the dimensions and nuances of sentiments. It aims to identify the sentiment while dissecting and understanding layers within it. This complexity makes the approach more advanced and intricate.

II. *Neural Networks and Complex Data Interpretation*

Solving this problem requires utilizing network models known for their effectiveness in handling diverse datasets and interpreting complex patterns within data. The challenge lies in applying networks to process and analyze textual data, often containing intricate patterns, and contextual nuances. Leveraging the capabilities of neural network in these areas is crucial when dealing with the complexities involved in this type of sentiment analysis.

III. *Balancing Computational Efficiency with Model Accuracy*

One of the challenges in this machine learning problem involves finding the balance between efficiently processing large amounts of text data and achieving accurate and nuanced sentiment analysis. The model needs to be designed to focus on processing speed and resource utilization while also providing analysis and interpretation of sentiments. This machine-learning problem combines sentiment analysis, rating classification and aspect-based analysis using network models. It shows how NLP challenges evolve and move beyond surface-level understanding to explore the intricacies and complexities of language as it conveys opinions, emotions, and experiences.

E. Success Metrics

To evaluate the success and effectiveness of the network model, for the project titled "Actionable Insights from Student Feedback: A Sentiment Analysis Approach" we use carefully selected metrics. These metrics are designed to measure how accurately the sentiment analysis performs and assess the model's efficiency and practicality when applied in a setting.

Accuracy and Precision: The main metric we consider is the model's accuracy, which determines how predictions it gets right, including rating-based classifications and more nuanced sentiment and aspect-based analyses. Precision is also crucial as it evaluates how positively the model makes predictions in correctly identifying specific sentiments or aspects within the feedback.

Recall and F1 Score: Recall measures how well the model can correctly identify all category instances, such as a specific sentiment or aspect. The F1 Score provides a metric that considers both precision and recall, making it particularly useful when achieving a balance between these two aspects is important.

Confusion Matrix: Visualizing the model performance across categories becomes easier, with a confusion matrix proving extremely valuable. It helps determine if the model consistently misclassifies sentiments or aspects, which is crucial for making adjustments and enhancements.

Efficiency in Training: Considering the large amount of data involved, how efficiently the model trains itself is an important measure. This includes considering the time required for training and the computational resources needed.

These measures of success together create a framework for evaluating the effectiveness of the network model. They cover performance factors like accuracy and efficiency and practical considerations such as usability and real-world applicability. Through these metrics, our project aims to achieve technical performance and deliver tangible and useful benefits within the education domain.

2. Solution Design

The solution design outlines our approach and methodologies for turning student feedback into actionable insights. The core of this solution lies in the innovative application of neural network models specifically tailored for sentiment analysis, including ABSA in an educational context.

A. Literature Review

I. *Has this problem been encountered before?*

Effectively analyzing student feedback goes beyond sentiment analysis and delves into the nuanced territory addressed by Aspect-Based Sentiment Analysis (ABSA). In the past, understanding student feedback in education has mainly involved interpreting metrics like Likert scales. While this approach provides a perspective on student satisfaction, it overlooks the qualitative insights that often reveal more. Significant efforts have been made in natural language processing and sentiment analysis in response to this limitation. However, these efforts traditionally focused on general sentiment classification without considering aspects or themes within the text.

The emergence of ABSA marked an advancement in sentiment analysis. This approach surpasses general sentiment classification by identifying and evaluating sentiments associated with aspects mentioned in a text.

When it comes to student feedback, it is important not to determine whether it is positive or negative but to understand what aspects of the educational experience are being praised or criticized. These aspects could include teaching quality, course content or learning resources.

ABSA (Aspect-Based Sentiment Analysis) has shown potential in product reviews and customer services, its application in education has been limited. The educational domain presents challenges such as domain language, the sensitivity of feedback content and varying formats of student responses. These challenges have not been extensively addressed through ABSA. Therefore, while sentiment analysis in text data is a known and explored issue, applying ABSA to student feedback in education is a relatively new area of research. It presents an opportunity to improve our understanding and utilization of student feedback by going beyond sentiments and extracting insights about specific aspects. Such insights would be highly valuable for educators and academic institutions as they can take steps based on them.

II. *How was it solved? What is the state-of-the-art technique?*

The evolution of sentiment analysis in recent years has been marked by significant advancements, particularly in Natural Language Processing (NLP). This evolution has shifted from basic text processing methods, like keyword extraction and statistical analysis, to more sophisticated machine learning techniques.

Initially, machine learning algorithms such as Support Vector Machines (SVMs) and Naive Bayes classifiers played a role in advancing sentiment analysis. These algorithms increased sophistication by providing accurate and nuanced text interpretations compared to earlier methods. A pivotal advancement in this field came with the introduction of neural network models, specifically Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Gated Recurrent Units (GRUs). These models are particularly adept at processing sequential data and have been instrumental in enhancing the ability to analyze sentiments in text more deeply and contextually.

Revolution through Transformer Models: The most recent and significant advancement in sentiment analysis has been developing and applying transformer models, such as BERT (Bidirectional Encoder Representations from Transformers). Transformer models represent a leap forward in the field, with their ability to process text bidirectionally and capture a more nuanced understanding of context and subtleties in language.

Aspect-Based Sentiment Analysis (ABSA): Alongside these developments, Aspect-Based Sentiment Analysis (ABSA) has become increasingly relevant. ABSA provides a more granular approach by focusing on specific aspects within

texts and analyzing sentiments associated with each aspect. This technique is particularly useful for dissecting complex feedback into distinct components and understanding the multifaceted nature of sentiments expressed.

In summary, today's state-of-the-art sentiment analysis is characterized by these advanced techniques, each contributing to a more refined and comprehensive understanding of the text.

III. *What were the limitations to that solution? (Gap in solution)*

While current sentiment analysis methods, including advanced techniques like neural networks and BERT, have significantly improved the understanding of textual feedback, there remain key limitations, particularly in the context of educational feedback, which our three-tiered drill-down approach aims to address:

Surface-Level Analysis: Traditional sentiment analysis often provides a surface-level understanding, categorizing feedback into basic positive, negative, or neutral sentiments. This approach lacks the depth to understand students' complexities and specific concerns. Our first tier, employing a Bag of Words combined with neural networks approach for rating-based classification, goes beyond mere positive or negative categorization, offering a nuanced understanding based on a 1-5 rating scale that reflects the varying degrees of student satisfaction or dissatisfaction.

Lack of Detailed Sentiment Dissection: Many existing models do not dissect sentiments to understand underlying emotions or specific aspects of feedback, which is critical in an educational setting. Our second tier, using Bidirectional GRU, delves into the nuanced sentiments, particularly focusing on lower ratings to understand specific negative emotions. This level of analysis is crucial for identifying and addressing the root causes of student discontent.

Absence of Aspect-Specific Insights: Traditional sentiment analysis often overlooks the importance of aspect-based insights, vital in educational feedback for pinpointing specific areas of a course or teaching methodology that need improvement. Our third tier addresses this gap by employing Aspect-Based Sentiment Analysis (ABSA). It allows for a detailed breakdown of feedback into specific educational aspects, providing educators with actionable insights on elements of their course or teaching style.

Integration and Practical Application in Educational Settings: Sentiment analysis tools are often not tailored for or easily integrated into educational environments. Our model is designed with the educational context in mind, ensuring that it can be seamlessly integrated into existing educational platforms and systems. It makes the insights it generates readily accessible and actionable for educators.

In summary, our proposed solution is specifically designed to address the limitations of current sentiment analysis methods in the educational sector. By offering a more detailed rating classification, deeper sentiment dissection, and aspect-specific insights, our solution provides a more comprehensive, actionable, and education-focused analysis of student feedback.

IV. *What are you proposing that is “novel”?*

Our proposal (**Three-Tiered Model**) introduces a novel, hierarchical, drill-down model for sentiment analysis in educational feedback, extending beyond traditional techniques to address the specific nuances of student feedback. This multi-tiered approach, comprising three distinct levels of analysis, represents a unique amalgamation of current state-of-the-art techniques, tailored for the educational context.

- a. Tier 1 - Rating-Based Classification: The first tier aims to classify student feedback which are in the text format to a Likert Scale of 1 to 5. The approach we have followed is a Bag of Words (BoW) combined with neural networks. The Bag of Words technique provides a way to represent the textual information for the neural network to understand. The neural network then learns the patterns over its Dense Layers.
- b. Tier 2 - Nuanced Sentiment Analysis: The second tier delves into a more nuanced analysis of the sentiments. Here, we employ bidirectional GRU layers, which is crucial for learning textual patterns in both forward and

reverse directions, providing a comprehensive view of the context. GRUs are chosen for their ability to efficiently capture dependencies in sequence data, making them ideal for text analysis. This tier is particularly sophisticated in its approach; it can discern and differentiate between negative emotions, such as frustration, disappointment, or confusion, often associated with lower ratings. The insights generated here are more detailed, providing educators with an understanding of the underlying reasons for student dissatisfaction.

- c. *Tier 3 - Aspect-Based Sentiment Analysis (ABSA)*: Our model's third and most detailed tier is the Aspect-Based Sentiment Analysis. This tier focuses on dissecting the feedback into specific aspects of the educational experience, such as teaching methodology, course content, or student support services. It employs an advanced ABSA model that is finely tuned to identify different aspects mentioned in the feedback and analyze the sentiment associated with each of these aspects. This tier's capability to break down feedback into distinct components and evaluate sentiments accordingly offers educators a highly granular view of student feedback. It pinpoints the exact areas within the educational experience that require attention and improvement, enabling targeted actions.

While each tier operates independently, they collectively contribute to a comprehensive understanding of student feedback. The first tier sets the stage with a broad classification, the second tier adds depth by exploring the emotions behind specific ratings, and the third tier offers the most detailed insights by focusing on individual aspects of the educational experience.

By implementing this multi-tiered approach, our model aligns with state-of-the-art sentiment analysis techniques and adapts and extends them to meet the unique challenges of analyzing student feedback in education. The result is a robust, nuanced, and practical tool for educators to understand and respond to student needs effectively.

V. *References of previous related work (minimum 5)*

In exploring sentiment analysis within educational settings, particularly focusing on student feedback, it is crucial to acknowledge the foundational and contemporary works that have shaped this field. The following references have guided our understanding and approach to sentiment analysis using machine learning, NLP, and POS tagging. Each study contributes unique insights into the complexities of interpreting student language, the effective use of various sentiment analysis methodologies, and the practical applications of these techniques in educational contexts. These works collectively provide a comprehensive view of the current state of sentiment analysis in education, highlighting both the challenges and the innovative solutions that have emerged in recent years:

- [1] N. R, P. M. S, P. P. Harithas and V. Hegde, "Sentimental Analysis on Student Feedback using NLP & POS Tagging," 2022 International Conference on Edge Computing and Applications (ICECAA), Tamilnadu, India, 2022, pp. 309-313, doi: 10.1109/ICECAA55415.2022.9936569.
 - This study explores the application of sentiment analysis in education, focusing on analyzing student feedback using NLP and POS (Part-of-Speech) tagging. The paper addresses the challenge of interpreting the language used by students in feedback, a task complicated by the volume and complexity of the data. The proposed method involves an automated analysis of textual feedback to evaluate teaching effectiveness, using ML and NLP techniques to classify sentiments expressed in student comments.
- [2] T. Shaik, X. Tao, C. Dann, C. Quadrelli, Y. Li and S. O'Neill, "Educational Decision Support System Adopting Sentiment Analysis on Student Feedback," 2022 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), Niagara Falls, ON, Canada, 2022, pp. 377-383, doi: 10.1109/WI-IAT55865.2022.00062.
 - This paper proposes a conceptual framework for sentiment analysis on student feedback within educational institutions. The study introduces an innovative approach to processing qualitative feedback from students using tokenization, stemming, and stopword removal. It employs TextBlob for sentiment categorization based on polarity and subjectivity and a Bi-LSTM deep learning model for multi-label feedback classification into 19 aspects of Biggs's model.

- [3] I. Sindhu, S. Muhammad Daudpota, K. Badar, M. Bakhtyar, J. Baber and M. Nurunnabi, "Aspect-Based Opinion Mining on Student's Feedback for Faculty Teaching Performance Evaluation," in IEEE Access, vol. 7, pp. 108729-108741, 2019, doi: 10.1109/ACCESS.2019.2928872.
 - This paper introduces a supervised aspect-based opinion mining system using a two-layered LSTM model to analyze student feedback for faculty performance evaluation. The model predicts aspects of feedback and their sentiment orientation.
- [4] Z. Nasim, Q. Rajput and S. Haider, "Sentiment analysis of student feedback using machine learning and lexicon-based approaches," 2017 International Conference on Research and Innovation in Information Systems (ICRIIS), Langkawi, Malaysia, 2017, pp. 1-6, doi: 10.1109/ICRIIS.2017.8002475.
 - This paper presents a novel approach for sentiment analysis of student feedback, combining machine learning and lexicon-based methods. Focusing on feedback collected at the end of academic semesters, it explores the use of TF-IDF and lexicon-based features to train a sentiment analysis model. The model aims to extract valuable insights about teaching quality from student feedback.
- [5] S. Rani and P. Kumar, "A Sentiment Analysis System to Improve Teaching and Learning," in Computer, vol. 50, no. 5, pp. 36-43, May 2017, doi: 10.1109/MC.2017.133.
 - This paper presents a sentiment analysis system designed to enhance teaching and learning by applying natural language processing and machine learning to student feedback. The system analyzes comments from course surveys and online sources, identifying sentiment polarity, emotions expressed, and levels of satisfaction or dissatisfaction.

Other references include:

- <https://huggingface.co/yangheng/deberta-v3-base-absa-v1.1>
- <https://huggingface.co/datasets/dair-ai/emotion>
- <https://medium.com/@abdulraqibshakir03/sentiment-analysis-on-student-feedback-in-engineering-education-55a913dd7967>
- <https://www.analyticsvidhya.com/blog/2022/01/sentiment-analysis-with-lstm/>
- <https://www.kaggle.com/code/atharvamartiwar/tokenizer-bow-tfidf-using-lstm-nn>.

B. From perspective of ML Workflow

Having established the framework of our innovative Three-Tiered Model for sentiment analysis in the educational context, this section will delve into the specifics of each model. Our approach combines three distinct levels of analysis, each employing state-of-the-art techniques, to provide a comprehensive understanding of student feedback:

1. **Tier 1 - Rating-Based Classification:** This section will explain the Bag of words combined with neural network approach towards classifying feedback to ratings ranging from 1 to 5.
2. **Tier 2 - Nuanced Sentiment Analysis:** We will explore the deployment of bidirectional GRU layers for a deeper, more nuanced sentiment analysis, especially focusing on interpreting emotions and attitudes in student feedback.
3. **Tier 3 - Aspect-Based Sentiment Analysis (ABSA):** The final tier will focus on our advanced ABSA model, which dissects feedback into specific educational aspects and assesses the sentiments tied to each.

While functioning independently, each tier synergistically contributes to a holistic analysis of student feedback. We will now provide detailed insights into each model's functionality, implementation, and the unique value it brings to our comprehensive sentiment analysis solution.

I. Data Selection

In our project, a consistent approach to data selection was adopted for all three models, ensuring uniformity in the testing phase. We crafted a specialized dataset containing feedback from 39 students, specifically tailored to test each model under similar conditions. This approach was designed to provide a coherent and comparable evaluation across all models. For the Aspect-Based Sentiment Analysis (ABSA) model, we took an additional step to

construct a custom small dataset, particularly for evaluation purposes. This decision was necessitated by the lack of readily available student feedback ABSA datasets online. Our objective was to create a dataset that closely mimics real-world scenarios in educational settings, thereby enabling a more accurate assessment of the model's performance in handling specific nuances in student feedback. Below, we provide detailed information on the training datasets used for each model:

Tier 1 – Rating Based Classification

For this model, we have selected the Coursera Course Reviews Dataset : [Kaggle Dataset Link](#). The dataset contains reviews made by students on the MOOC platform Coursera and their corresponding star ratings ranging from 1 to 5. The rating 1 corresponds to Very Negative, 2 - Negative, 3 - Neutral, 4 - Positive and 5- Very Positive. We believe that this dataset best represents the input data for our problem as it is actual reviews given by students on different courses.

Tier 2 – Nuance Sentiment Analysis

In this model, the training datasets employed are a combination of two distinct sources, each bringing unique value to the project:

- Emotion Dataset from Hugging Face: [Hugging Face Dataset Link](#)
This dataset is an extensive collection of texts labelled with emotions, providing a wide range of emotional expressions. Its diversity is crucial for training our model to recognize and interpret various emotional nuances in textual data.
- Twitter Emotion Dataset from Kaggle: [Kaggle Dataset Link](#)
Sourced from Twitter, this dataset includes a broad spectrum of real-world emotional expressions, offering valuable insights into everyday sentiment expression. The inclusion of this dataset aims to enhance the model's applicability in real-life scenarios, particularly in analyzing concise and informal text as often found in social media.

The strategic combination of these datasets is intended to encompass a comprehensive range of emotions, thus enhancing the diversity and robustness of our training data. This approach is designed to improve the model's ability to recognize and understand various emotional expressions, which is central to effective sentiment analysis in text.

Tier 3 – ABSA

The Tier 3 ABSA model has been pre-trained on a comprehensive dataset, making it highly effective for general ABSA tasks, including our specific use case in the realm of student feedback analysis. Given the model's robust pre-training, we opted not to retrain it from scratch. Instead, our focus shifted towards leveraging custom datasets specifically curated for evaluating and testing the model's performance. These datasets are tailored to our unique requirement of analyzing student feedback, encompassing a diverse range of aspects and sentiments that students typically express. This approach allowed us to thoroughly assess the model's efficacy in our specific context, ensuring that it meets our standards for accuracy and reliability in sentiment analysis.

II. Data Pre-processing

Tier 1 – Rating Based Classification

The data for the model is loaded into a DataFrame to be cleaned and transformed suitable for input to the Bag of Words combined with Neural Network model. The preprocessing steps involve removing stopwords, hashtags, mentions and any URLs using regex patterns. We also convert the text to lowercase. These steps ensure that some of the noise from the data is removed and the important data is being fed to the Neural Network.

Tier 2 – Nuanced Sentiment Analysis

Data preprocessing is crucial in preparing the dataset for all models, including the Tier 2 model. This process includes several key tasks: cleaning text data by removing non-essential elements (like non-English characters, numbers, and special symbols), converting text to lowercase, and filtering out stopwords to reduce noise. The emotion labels are simplified and standardized to aid the model training process. Text data is then tokenized and converted into sequences, with consistent length achieved through padding. Finally, the dataset is divided into training and validation sets, setting the stage for the model's training and subsequent evaluation. This thorough

preprocessing ensures the data is optimally formatted for training the emotion analysis model, enhancing its potential performance and accuracy.

Tier 3 – ABSA

The pre-processing stage was crucial for ensuring the model accurately identified and analyzed aspects within student feedback. Natural Language Processing techniques, particularly part-of-speech tagging, were employed to extract relevant nouns and noun pairs from the feedback text. This process was instrumental in identifying significant aspects, such as 'LAB' and 'PROFESSOR'. By filtering out common stop words and redundant nouns, the model could focus on the feedback's most relevant and impactful aspects, which is vital for accurate sentiment analysis.

III. Model Selection

Tier 1 – Rating Based Classification

We have selected the Bag of Networks + Neural Network approach, which was referenced from the notebook : <https://www.kaggle.com/code/atharvamartiwar/tokenizer-bow-tfidf-using-lstm-nn>.

The bag-of-words (BoW) combined with neural network is method described for text classification. The BoW model does not give importance to the sequential order of words in the text and primarily focuses on the word frequency of each word for a text phrase. It builds a sparse matrix in which each row corresponds to a given text phrase and each column represent a unique word that is collected from the entire words (corpus) for a given dataset. This sparse matrix can be then fed as an input into a neural network.

Neural Networks are good for learning complex patterns when given data with high dimensionality. It makes them adept for processing the BoW Sparse Matrix. The networks usually consist of multiple layers of interconnected neurons that learn the relationships between input features and output labels. They do this by applying a mathematical function on each neuron and updating their weights after each iteration with the dataset. With the BoW representation and the Neural Network as a pattern learner we can classify the text based on their corresponding ratings. Therefore, with the above approach we can create a ratings classifier which classifies text phrases to ratings.

We experimented with a few architectures for the Neural Network:

- Neural Network with 2 Dense Layers (10(relu)-->6(softmax))
- Neural Network with 3 Dense Layers (32(relu)-->10(relu)-->6(softmax))
- Neural Network with 4 Dense Layers (64(relu)-->32(relu)-->10(relu)-->6(softmax))
- Neural Network with 5 Dense Layers (128(relu)-->64(relu)-->32(relu)-->10(relu)-->6(softmax))
- Neural Network with 4 Dense Layers (128(relu)-->32(relu)-->16(relu)-->6(softmax))
- Neural Network with 4 Dense Layers (128(relu)-->32(relu)-->6(softmax))
- Neural Network with 4 Dense Layers (128(relu)-->32(relu)-->10(relu)-->6(softmax))

Finalized model structure is as follows:

- Input - Bag of Words vectorization
- Pattern Learner - Neural Network with 4 Dense Layers (128(relu)-->32(relu)-->10(relu)-->6(softmax))

This model architecture had good results for test datasets and worked with most of the edge cases that we tried. It also has good graph of increasing training accuracy - validation accuracy, and comparatively decreasing validation loss graph.

Tier 2 – Nuanced Sentiment Analysis

The model is designed to process textual data and efficiently interpret emotional expressions. Our model's core is an *Embedding layer* essential for natural language processing tasks. It transforms words into 100-dimensional vector representations, capturing their nuanced semantic relationships. This layer forms the foundation for understanding the context embedded in the text. The model then employs *Bidirectional GRU layers*, with 64 and 32 units, respectively. This bidirectional approach is crucial for learning textual patterns in both forward and reverse

directions, providing a comprehensive view of the context. GRUs are chosen for their ability to efficiently capture dependencies in sequence data, making them ideal for text analysis.

To combat overfitting and enhance generalization, we incorporate a *Dense layer with 64 units* followed by *L2 regularization and a Dropout layer*. The Dense layer introduces non-linearity to the model, allowing it to learn complex patterns. L2 regularization penalizes complexity, and the Dropout layer randomly deactivates neurons during training, reducing the model's dependency on specific features. The architecture concludes with a *Dense output layer* with six units, one for each emotion category, utilizing a *softmax* activation function. This setup allows the model to distribute probability over the emotion classes. For compilation, we opt for the *Adam optimizer* with a learning rate 0.0005, balancing speed and stability in training. The model uses *sparse_categorical_crossentropy* as the loss function, which is suitable for multi-class classification. We monitor *accuracy* to track the model's performance. This model combines embedding, bidirectional learning, and regularization techniques to robustly analyze and interpret emotions in text, making it a potent tool for sentiment analysis tasks.

Tier 3 - ABSA

For this tier, the '[yangheng/deberta-v3-base-absa-v1.1](#)' model was selected, a pre-trained model specifically designed for Aspect-Based Sentiment Analysis (ABSA) and available on Hugging Face. This model is adept at understanding sentiments expressed about specific aspects within a text, a capability that aligns perfectly with the nuanced requirements of analyzing student feedback. The model's pre-training on a comprehensive dataset significantly enhances its effectiveness for general ABSA tasks, and its architecture is adept at discerning fine-grained sentiment information about different aspects within texts. This makes it an ideal choice for the project, as it can provide deep insights into the varied sentiments expressed in student feedback.

IV. Training/Fine-tuning

Tier 1 – Rating Based Classification

In this tier, we used the Bag of Words (BoW) technique combined with Neural Networks to predict the ratings from a text phrase in relation to courses. The training was done on multiple architectures and different hyperparameter configuration to find an optimal mode.

Tier 2 – Nuanced Sentiment Analysis

The model is trained over epochs with a batch size of 64. This batch size balances computational efficiency and the ability to update model weights effectively. Training the model for enough epochs is crucial, as it allows the model to iteratively learn from the data, adjusting its weights with each batch. However, to prevent overtraining and ensure efficient use of computational resources, we implement an *EarlyStopping callback*. This mechanism monitors the model's performance on the validation data, particularly looking at the 'val_accuracy'. If there's no improvement in validation accuracy for three consecutive epochs, training is halted, and the model reverts to the weights from its best performance epoch. This approach helps prevent overfitting and helped us identify the optimum number of epochs that we need to use for training the model.

The training and validation accuracy metrics are closely monitored, providing insights into how well the model is learning and generalizing. The goal is to achieve a high validation accuracy, indicating the model's robustness and ability to perform well on new, unseen data.

Tier 3 - ABSA

The tier leveraged the robust pre-training of '[yangheng/deberta-v3-base-absa-v1.1](#)', eliminating the need to retrain the model from scratch. Instead, the focus was on using custom datasets, specifically curated for evaluating and testing the model's performance in the context of student feedback analysis. This approach ensured that the model's performance was thoroughly assessed against the project's specific requirements, thus ensuring its reliability and accuracy in sentiment analysis tasks within the educational domain.

V. *Hyperparameter tuning strategy*

Tier 1 – Rating Based Classification

The following parameters were changed to find a good convergence in the tier 1:

Learning Rate - The learning rate is the rate at which a model learns patterns from the data during the training phase. Too high of a rate can make the model learn too much and it might not accurately learn the pattern. If it is very low, learning will take more time. Finding an optimal learning rate is important to improve the model. We tried the following values mainly - 0.0001, 0.001, 0.005, 0.01. We then finalized on the 0.01 learning rate due it is having less validation error for the same no of epochs and better graph.

Activation Function – It is a mathematical function that triggers a neuron in the Dense Layer. A careful selection of activation function is important for accurate results and faster convergence. We have tried the Relu, Elu and Tanh activation functions in the input and hidden layers. A better convergence was observed with the Relu function, a faster one with Elu, while Tanh resulted in unexpected performance of the model for certain text phrases.

Batch Size - Batch Size is the no of training records for which the weights of a neural network are updated. We have tried with batch size of 256, 128,64 and 32. The model performed most for the batch size of 256, while 128 was also a close contender.

Tier 2 – Nuanced Sentiment Analysis

For tier 2 model, the tuning strategy began with initial explorations, focusing on critical parameters like learning rate and epochs. We recognized early on that the learning rate was a key influencer of the model's training behaviour. A higher learning rate in our initial experiment led to a stagnation of accuracy of around 38%, suggesting that the model was converging too quickly to a suboptimal solution. This was a clear indicator that the learning rate needed to be adjusted. Subsequently, lowering the learning rate to 0.0005 and increasing epochs to 10 showed significant improvement, with the model's validation accuracy rising to around 89%. This adjustment allowed a more gradual and stable learning process, enabling the model to capture the underlying patterns in the data more effectively. In parallel, we adjusted the batch size to find an optimal balance between computational efficiency and the model's ability to learn from the data. Additionally, we explored alterations in the model's architecture, particularly focusing on the activation functions and including dropout layers. These tweaks were aimed at enhancing the model's capability to capture complex patterns and relationships within the sentiment data, which was crucial for the effectiveness of our sentiment analysis. Overall, our hyperparameter tuning strategy was a delicate and thoughtful balancing act. The final model configuration emerged from this rigorous process, demonstrating enhanced performance and efficiency, making it a robust tool for analyzing sentiments in educational feedback.

Tier 3 - ABSA

In this tier model, hyperparameter tuning was not a focal point due to the utilization of the model in its pre-trained state. The model's existing configuration and parameters, honed through extensive pre-training, were deemed sufficient for the project's requirements, allowing for immediate application to the test dataset without further adjustments.

VI. *Evaluation Metrics*

For all models, the following metrics were used. All of them were pivotal for a thorough assessment, given the complexities involved in sentiment analysis, especially in the context of educational feedback.

Accuracy: This fundamental metric was indispensable as it directly indicated the model's overall effectiveness. It measured the proportion of correctly identified sentiments, encompassing positive and negative cases. In student feedback analysis, accuracy is critical as it directly impacts the reliability of the insights drawn from the model.

Precision and Recall: These metrics offered a deeper insight into the model's classification capabilities. Precision, or the model's ability to correctly identify positive sentiments among all positive predictions, was crucial for minimizing false positives. On the other hand, Recall focused on the model's efficiency in capturing all actual positive cases, which was vital in scenarios where missing a positive sentiment could lead to significant oversight. In educational settings, where every piece of feedback can be critical, balancing precision and recall was key to ensuring that the model did not overlook important sentiments while maintaining a low rate of false alarms.

F1 Score: The F1 Score emerged as a critical metric, especially when striking a balance between precision and recall was imperative. As the harmonic mean of precision and recall, it served as a single measure to assess the model's accuracy in classifying sentiments accurately without leaning too heavily towards either precision or recall. This balance is particularly important in educational feedback, where the distinction between different types of sentiments can often be subtle yet significant.

These metrics collectively formed an essential part of our evaluation toolkit. These not only enabled us to gauge the overall accuracy of the models but also allowed us to fine-tune them for a balanced performance. The ability to differentiate between various sentiments was just as important as the model's general accuracy, making these metrics indispensable in our analysis.

3. Implementation

We have organized our source code into three distinct Google Colab notebooks to implement our three-tiered sentiment analysis model. Each notebook corresponds to one tier of the model and is structured to segment the code into manageable sections, with comprehensive explanations provided in accompanying text blocks.

A. Code

Given the large size of Colab notebooks, we are providing direct links to these files committed to GitHub rather than embedding the entire source code within this document.

1. [Tier 1 Model](#)
2. [Tier 2 Model](#)
3. [Tier 3 ABSA Model](#)

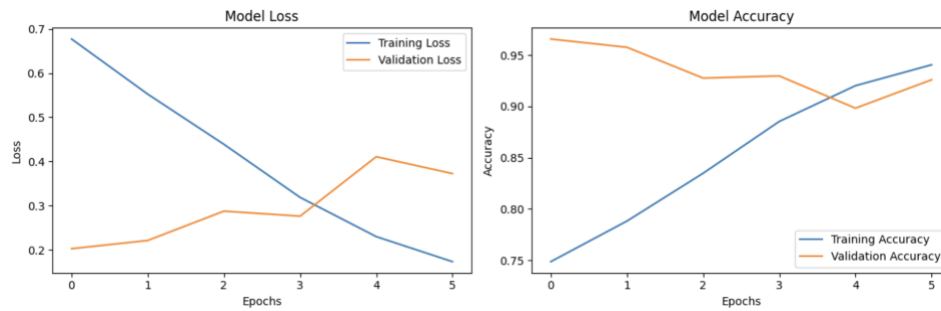
B. Repository – URL

The [GitHub repository](#) contains three primary directories:

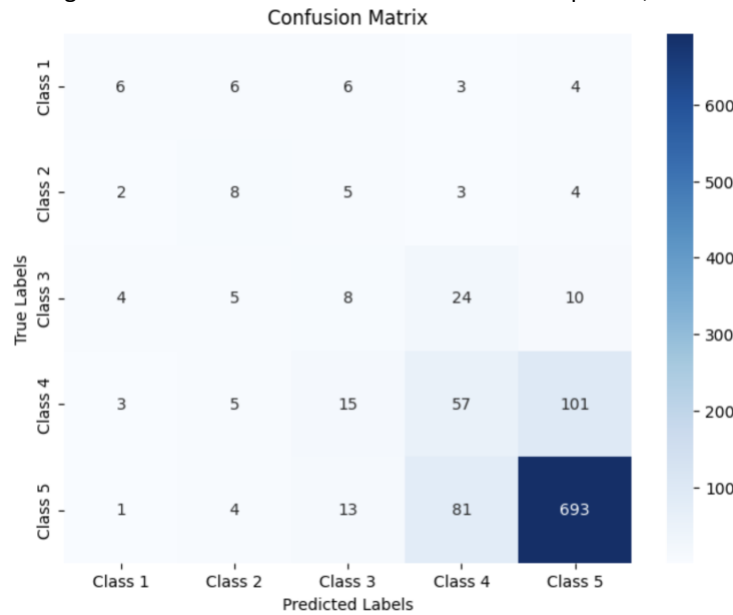
1. Datasets: This directory contains folders dedicated to the training and test datasets utilized in our project. It also includes a separate folder for the output files generated by the models.
2. Models: In this directory, you'll find the Colab files for all three models, meticulously preserved and made accessible.
3. Report: A digital copy of this comprehensive report can be found within this section for easy reference and review.

C. Results/Plots – Graphs

Tier 1



For the tier 1 BoW combined with a 4 Layer Neural Network model, we can see that the Model Accuracy is increasing. Though there is an increase in validation loss after epoch 3, it has a decreasing tendency after 4.

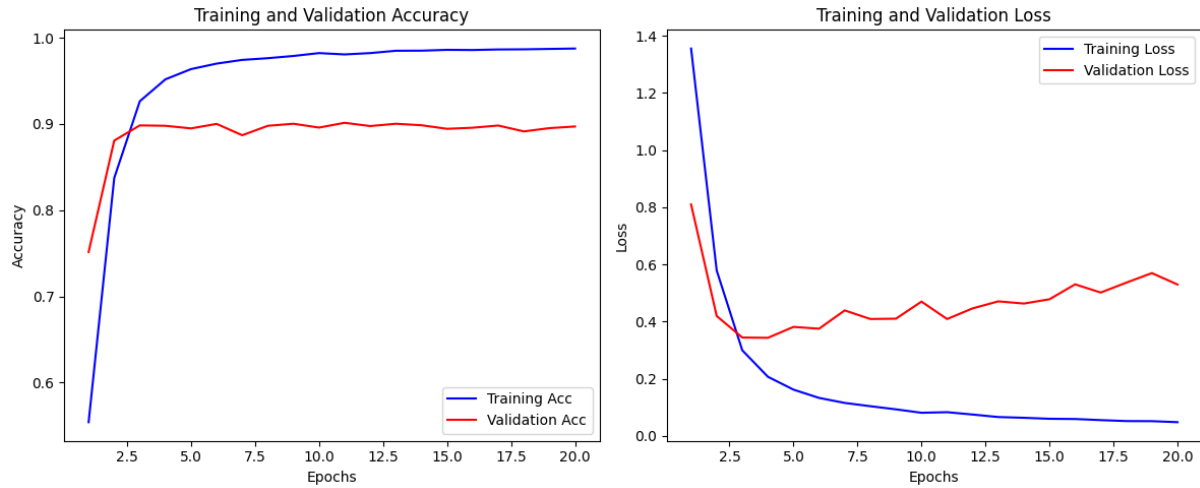


This confusion matrix plot was obtained from the test data for the BoW combined with a 4 Layer Neural Network model in tier 1. From this we confusion matrix we can see that the model has a bias towards the Rating 5, but it can work on most data, and we can see that it is marginally confused between certain classes like 2,3 and 3,4. This scenario is true even in the real world in case of ratings from text phrases, therefore we could use this model for our predictions for the time being. This can be improved further by training on more real datasets with even distribution of data.

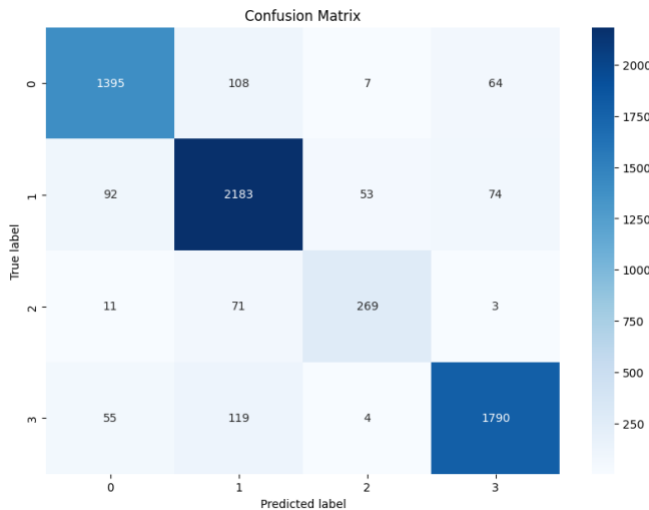
Tier 2:

The model achieved a high overall accuracy of 89.71%. This indicates that the model is quite effective in correctly classifying sentiments in most cases. The model utilizes two levels of Bidirectional GRUs (64 and 32 units, respectively). This design choice allows the model to capture dependencies in both directions (forward and backward) in the text sequence, which is especially beneficial for understanding context and meaning in language.

Following are the Training and validation accuracy charts and Confusion Matrix for the final model:

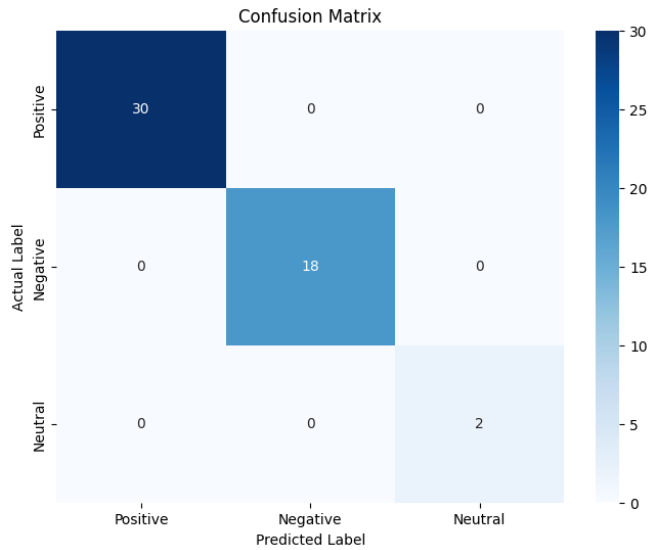


	precision	recall	f1-score	support
0	0.90	0.89	0.89	1574
1	0.88	0.91	0.89	240
2	0.81	0.76	0.78	354
3	0.93	0.91	0.92	1968
accuracy			0.90	6298
macro avg	0.88	0.87	0.87	6298
weighted avg	0.90	0.90	0.89	6298



Tier 3

We chose a pre-trained ABSA model which is already trained on generic ABSA datasets but demonstrated a strong performance with our student feedback test dataset. Given the specific format requirements of ABSA models and the lack of publicly available student feedback datasets tailored for ABSA, we crafted a small dataset using AI content generation technology to evaluate the model effectively. The model's evaluation yielded outstanding results, demonstrating its high precision and effectiveness. It achieved perfect scores across all key metrics: an accuracy of 100%, precision of 100%, recall of 100%, and an F1 score of 100%. These metrics indicate the model's impeccable ability to classify sentiments accurately in the context of student feedback, showcasing its reliability and the potential for practical applications in educational settings. The impressive success metrics of the model underscore the robustness and adaptability of the pre-trained ABSA model to our specific use case.



D. Result Explanation –

Tier 1

For tier 1, as discussed, we used a combination a Bag of Words (BoW) method and Neural Networks for the text to ratings multi – class classification problem. This approached used the BoW representation to transform the texts to a sparse matrix for better learning with the Neural Network. We experimented with different Dense Layer architecture and Hyperparameter optimizations for this model. Each configuration was performance reviewed through the accuracy and loss metrics. Afterwards, we found that the 128-32-10-6 dense layer neural network with the Relu Function, batch size 256 with a learning rate of 0.01 gave comparatively better performance compared to the other combinations. Although the accuracy for the test data is on the lower end due to the bias on rating 5, this model gave good output for the test data we had prepared for prediction. More data records with even distribution of the ratings classes are needed to improve this model.

Tier 2

The main model for sentiment analysis, constructed using TensorFlow's Sequential architecture, presents a sophisticated and strategic approach to text classification. At its core, the model features an Embedding layer designed to transform words into dense vectors, capturing semantic relationships effectively. This is complemented by two levels of *Bidirectional GRU* layers, with 64 and 32 units respectively, enabling the model to grasp dependencies in textual data in both forward and backward directions, a key aspect for understanding context in language processing. The model architecture also includes a Dense layer with 64 units and *ReLU* activation, integrated with L2 regularization to prevent overfitting, and a Dropout layer with a rate of 0.5 to further enhance the model's ability to learn robust features. Optimization is handled by the Adam optimizer, selected for its efficiency with sparse gradients and adaptive learning rate, set at 0.0005 in this case. The choice of *sparse_categorical_crossentropy* as the loss function aligns well with the multi-class nature of the sentiment classification task. The final output layer consists of a Dense layer with 6 units and *softmax* activation, reflecting the number of sentiment classes in the dataset. This model's design not only indicates a deep understanding of the intricacies involved in sentiment analysis but also strikes a balance between capturing complex patterns in text data and ensuring generalizability, making it a solid foundation for effective sentiment analysis and further exploratory tuning.

Experiment Configuration #1:

Utilizing a high learning rate of 0.05, the model displayed poor performance with a stagnating accuracy around 38%, suggesting that the learning rate was too high for effective learning. This configuration underscored the need for a well-calibrated learning rate to avoid premature convergence or instability.

Experiment Configuration #2:

When the learning rate was reduced to 0.0005 and epochs increased to 10, the model's performance improved significantly, achieving a validation accuracy of 89.39%. This demonstrated the importance of a lower learning rate and a higher epoch count for stable learning and better generalization.

Experiment Configuration #3:

Despite increasing the epochs to 15, a higher learning rate of 0.01 did not yield a significant improvement over the second experiment, with a final validation accuracy of 89.60%. This indicated that a higher learning rate, even with more epochs, might still be too aggressive for optimal model training.

Experiment Configuration #4:

Maintaining the learning rate at 0.01 but reducing epochs to 5 and altering the activation functions in the model led to a slightly improved validation accuracy of 90.22%. This change suggested that while architectural adjustments can impact performance, the choice of an appropriate learning rate remains a pivotal factor for model efficacy.

Tier 3:

Our project's third tier, centred around Aspect-Based Sentiment Analysis (ABSA), utilized the 'yangheng/deberta-v3-base-absa-v1.1' model. This model was strategically chosen for its specialized capabilities in extracting and analyzing sentiments associated with specific aspects of text data. Our selection was influenced by the model's proven efficacy in handling complex sentiment analysis tasks, making it an ideal choice for dissecting the multifaceted nature of student feedback.

A crucial part of our approach involved identifying key aspects of student feedback. We employed advanced natural language processing techniques to dynamically extract significant elements like 'LAB' and 'PROFESSOR' from the test dataset, which consisted of genuine student feedback. This aspect identification process was pivotal in tailoring the ABSA to our specific needs and ensured that the analysis was grounded in the actual content of the feedback.

In conclusion, our project's successful implementation and evaluation of the ABSA model showcased its effectiveness in providing deep, actionable insights into student feedback. More details and detailed descriptions on the steps we performed, and the result explanations are added to our Colab source files located [here](#).

4. Conclusion and Discussions

Our project's primary goal was to harness sentiment analysis's transformative power in educational feedback. Developing the Three-Tiered Sentiment Analysis Model, a pioneering approach tailored for educational settings represents a significant stride in understanding and utilizing student feedback effectively.

Conclusion

The Three-Tiered Model, comprising Rating-Based Classification, Nuanced Sentiment Analysis, and Aspect-Based Sentiment Analysis, has proven to be a robust framework for dissecting and interpreting student feedback. Each tier, functioning independently and synergistically, has contributed to a comprehensive understanding of student sentiments.

- *Tier 1 (Rating-Based Classification)* laid the foundation, providing a quick yet insightful glance into the general sentiment trends through numeric ratings. It offered an immediate snapshot of student satisfaction levels, serving as a springboard for deeper analysis.
- *Tier 2 (Nuanced Sentiment Analysis)* brought a more sophisticated layer of interpretation, harnessing the power of BERT to delve into the emotional undercurrents of feedback. This tier was instrumental in distinguishing subtle emotional nuances, from frustration to contentment, offering a deeper understanding of the 'why' behind student ratings.

- The Proof of Concept (POC) Web Dashboard* was a testament to the practical applicability of our model. It visualized data from each tier in an interactive and user-friendly manner, turning complex analyses into accessible insights for educators.

Our model's success lies in its adaptability and depth. It can be tailored to various educational settings, from small classrooms to massive online courses. However, there remain areas for further exploration and enhancement.

- Next Step:*** Besides providing educators with insights into the aspects influencing students' emotions, there is scope for further innovation. A subsequent model could suggest actionable steps based on these insights. This would identify the key areas of concern or strength and offer practical, data-driven recommendations for improvement or reinforcement. Such a model would bridge the gap between understanding student sentiments and implementing effective strategies to enhance the educational experience. It represents the next logical Step in our journey toward a more responsive and student-centric educational environment.

5. Proof Of Concept – Web Dashboard

The figure consists of two parts. On the left is a word cloud where words are sized according to their frequency. The most prominent words are "course", "feedback", "lack", "lab", "The", "Upset", "lacks", "to", "son", "date", "yet", "lecture", "theoretical", "short", "learning", "office", "fine", "subject", "The", "lab", "up", "lab", "pleased", "standard", "looking", "The", "lab", "s", "The", "s", "The". On the right is a bar chart titled "Total number of participants: 39" and "Aspects: LAB, PROFESSOR". The legend indicates three categories: Positive (green), Neutral (yellow), and Negative (red). The Y-axis ranges from 0 to 25.

Aspect	Positive	Neutral	Negative
LAB	8	23	8
PROFESSOR	23	4	12

I. Implementation Details

The POC Web Dashboard offers an intricate portrayal of sentiment analysis, drawing from three JSON files that align with the different tiers of our sentiment analysis model. This integration allows educators to comprehensively understand student feedback, from general sentiment to specific aspects.

Ratings JSON (Tier 1 – Overall Rating Analysis)

This file is the foundation of our analysis, containing student identifiers paired with their overall ratings on a scale of 1 to 5. The dashboard uses this data to construct a pie chart that visualizes the distribution of these ratings. This visualization offers a quick snapshot of the student body's general sentiment and acts as an interactive gateway to deeper insights. For instance, clicking on a segment of the pie chart takes the user to a detailed breakdown of emotions associated with that rating. This tier is crucial as it sets the stage for more nuanced analysis, highlighting whether most of the feedback skews positive, negative, or a neutral score out of 5.

Emotions JSON (Tier 2 – Emotional Sentiment Analysis)

At this level, the dashboard delves into the emotional sentiments associated with each piece of feedback. Each student's feedback is labelled with an emotion such as "Happy", "Frustrated", or "Neutral". The dashboard renders this data in two ways: a summary table listing student IDs and their corresponding emotions and a chart depicting each emotion's prevalence. This dual presentation provides a clear picture of the predominant emotional sentiments and allows educators to correlate specific emotions with certain ratings or aspects of the course. This tier adds depth to the analysis by uncovering the emotional undertones that might influence student perceptions and experiences.

Aspects JSON (Tier 3 – Aspect-Based Sentiment Analysis)

The final tier focuses on dissecting student feedback into specific aspects of the educational experience, such as "Teaching Experience" and "Lab Experience". The file lists these key aspects and links each student's sentiment ratings. In the dashboard, this translates into a detailed view of how each student perceives various aspects of their educational journey. Furthermore, the dashboard aggregates this aspect-specific sentiment data to create a comparative chart, offering educators a bird's-eye view of which areas are faring well and which require attention. This tier provides the most granular insights, pinpointing precise areas for improvement and commendation in the educational experience.

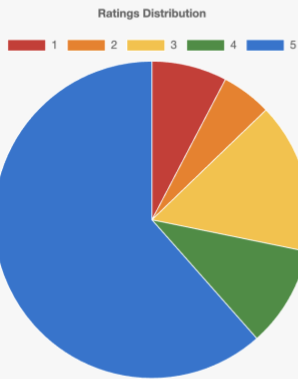
The POC Web Dashboard intricately weaves data from the three tiers to present a multifaceted view of student feedback. Starting from a broad categorization of overall student sentiment, it gradually narrows to emotional nuances and specific aspects of the educational experience. This approach renders a detailed and actionable analysis of student feedback and empowers educators with data-driven insights to enhance teaching methodologies and course content.

II. Functionalities and Features

The POC Web Dashboard encapsulates various functionalities and features designed to transform raw student feedback into actionable insights. This sophisticated interface is engineered to provide a user-friendly experience, enabling educators and administrators to navigate through layers of sentiment analysis effortlessly. The dashboard's design philosophy centers on clarity, interactivity, and comprehensive data representation, ensuring that every element contributes to an intuitive understanding of student sentiment.

Tier 1 – Ratings Overview

The first tier of the dashboard, the focus is on providing an overarching view of student feedback through a Ratings Overview. This segment adeptly captures and visualizes the overall ratings given by students, laying out the foundation for the sentiment analysis process. The Ratings Overview is not just a quantitative display of scores; it is an entry point into deeper analytical layers, setting the stage for more detailed emotional and aspect-based sentiment exploration. By offering a macroscopic view of student ratings, this tier allows educators to gauge the student body's general mood and satisfaction levels briefly, setting the context for further in-depth analysis in subsequent tiers. A pie chart displays the distribution of student ratings (1 - 5) and a table view will be there to represent the data with hyperlinks to display more insights provided by the subsequent models.

Total Feedbacks: **39**

Rating	Percentage
Rating 1	7.69%
Rating 2	5.13%
Rating 3	15.38%
Rating 4	10.26%
Rating 5	61.54%

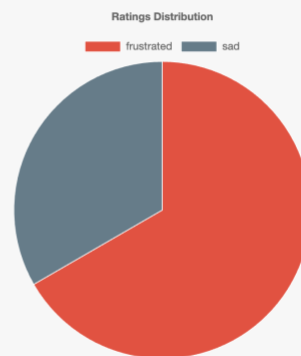
Tier 2 – Emotions Overview

This section features a chart showing the distribution of emotions and a detailed table that can be sorted to display each student's feedback along with their corresponding emotional sentiments. Users can filter the table to view specific emotions associated with chosen ratings. While the initial tier offers insights from our first model, this second tier delves into results from our second sentiment analysis model. In this project, we focused on a range of emotions: Pleased, happy, sad, and Frustrated. This tier's strength lies in its ability to provide educators with deeper insights into the reasons behind a student's 1/5 rating and the significance of their feedback based on their expressed emotions. For instance, a student expressing 'Frustration' may require more attention compared to one who expresses 'Sad' or 'Boring,' highlighting the varying degrees of urgency and concern in student feedback.

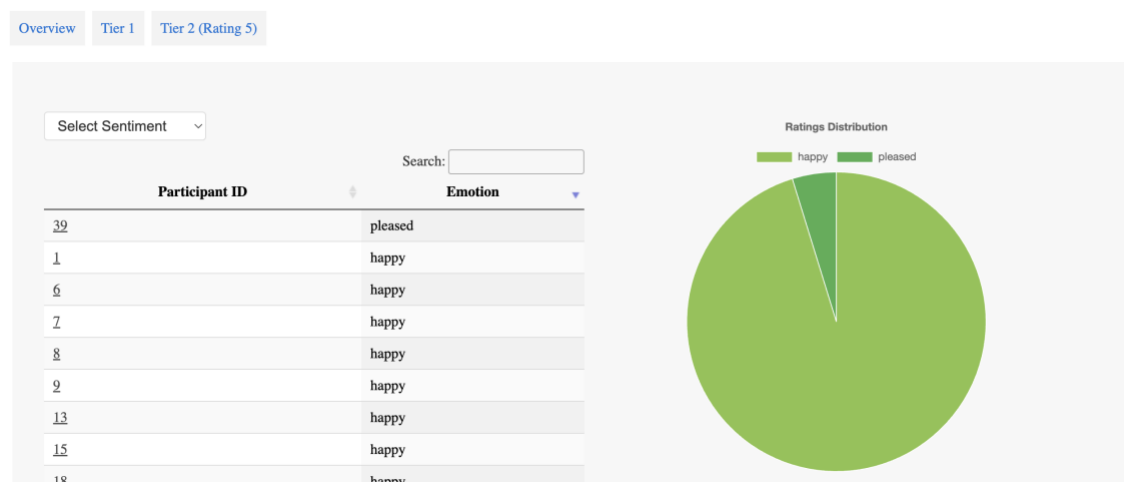
Select Sentiment

Search:

Participant ID	Emotion
10	sad
4	frustrated
19	frustrated

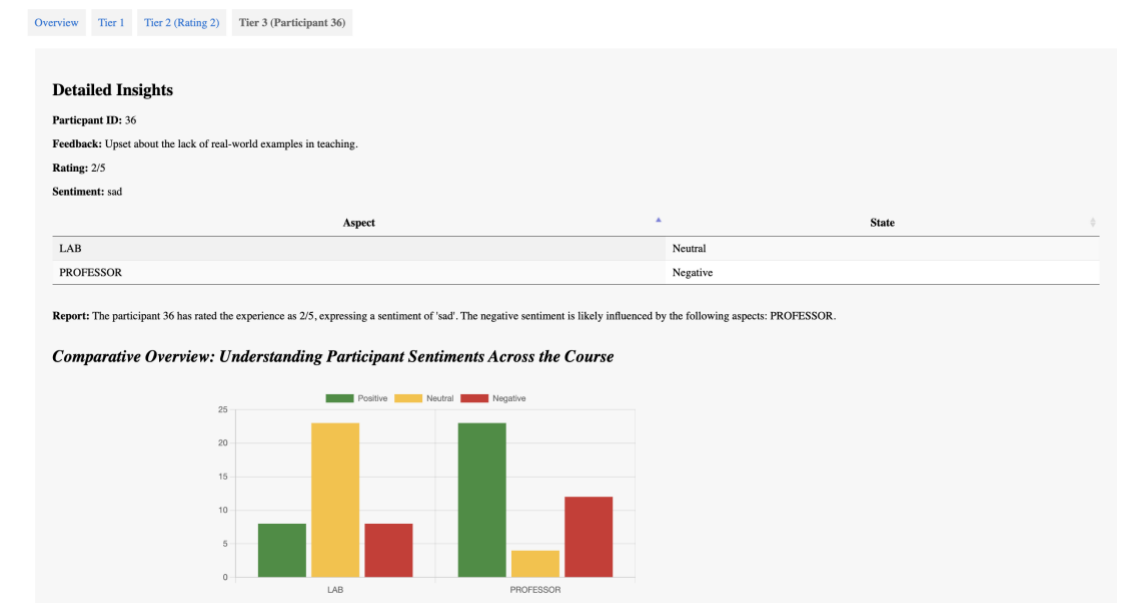


This categorization also applies to positive ratings, and the meaningfulness of this feedback classification is enhanced when we incorporate insights from the third and final model in our three-tiered approach.



Tier 3 – Aspect Analysis

The dashboard highlights which aspects have predominantly positive, negative, or neutral sentiments, offering granular insights into specific areas of the educational experience. The third and most detailed tier of our POC Web Dashboard is dedicated to the Aspect Sentiment Overview. Here, the intricate nuances of student feedback are examined, focusing on specific aspects of the educational experience. This tier utilizes ABSA model output JSON to visualize the aspect-based sentiment overview. The Aspect Sentiment Overview is not just a simple aggregation of opinions; it is a sophisticated breakdown that provides educators with a granular view of student feedback. Each aspect is evaluated for sentiment, giving educators a clear picture of what aspects are excelling and which require attention and improvement.



This tier is instrumental in pinpointing precise areas of the educational experience that impact student satisfaction. By analyzing the sentiment associated with each aspect, educators gain actionable insights, enabling them to tailor their approaches, address student concerns effectively, and enhance the overall educational experience. The Aspect Sentiment Overview transforms student feedback into a roadmap for educational excellence and student satisfaction.