Classifying Sentiments using BERT

Brief Introduction:

Sentiments are the words utter by someone which represents any anger, emotion, or happiness after utilising any service.it is basically a feed-back of a service provided by the user and this feedback is considered an essential equipment for business growth and flourishing, these merely words are then kept into mind for future endeavours of the business.it is in short considered as an opinion about a specific topic.

Methodologies used in sentiment analysis:

There are numerous methodologies adopted by the researchers and analyst to classify sentiments into categories like, positive, neutral, and negative. some of the most widely techniques are different algorithms like **Naïve Bayes** (it uses a probabilistic approach), **Linear Regression** (It is a statistical algorithm which calculates results into two categories only**), Support Vector Machine SVM** (A supervised ML technique like linear regression but more advanced),Deep Learning(It is a hierarchical machine learning technique which involve multiple levels and processes human complex words easily).(Rachel Wolff, 2020)

BERT's introduction and architecture:

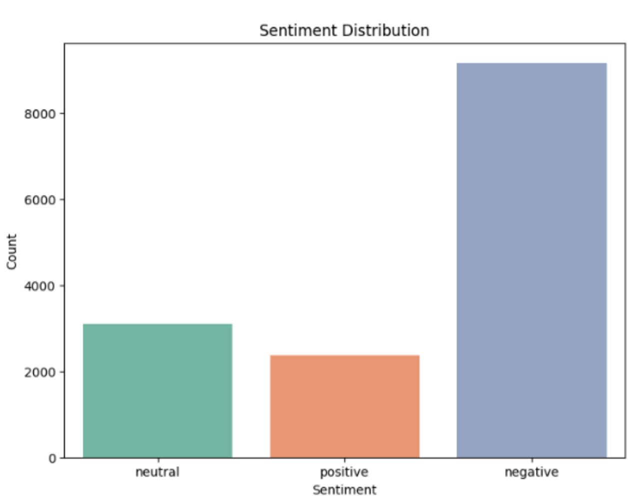
Bidirectional Encoder Representations from Transformers was introduced by researchers at Google AI in 2018. it was developed to help Google understand search queries better. But it turned out to be great at understanding all sorts of language tasks. The foundation of BERT lies in the Transformer architecture, initially devised by Google to handle diverse sequence-to-sequence tasks like translation and chatbots. BERT leverages this Transformer architecture, refining it into a more intelligent tool. The "bidirectional" aspect is pivotal, in most natural language processing endeavours, predicting subsequent words requires knowledge of both preceding and succeeding context. BERT excels here, grasping the entirety of a sentence's context to predict words accurately. Unlike models limited to left or right contexts, BERT's bidirectionality empowers it with holistic context comprehension, enhancing its overall potency in language understanding tasks.(*What Is BERT? An Intro to BERT Models*, 2023)

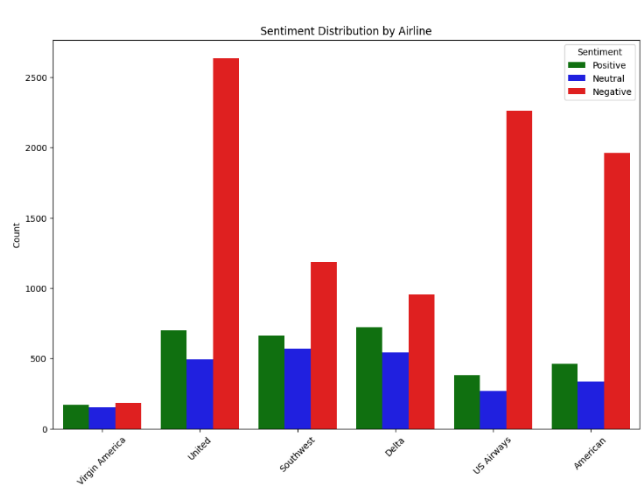
Topic Explanation:

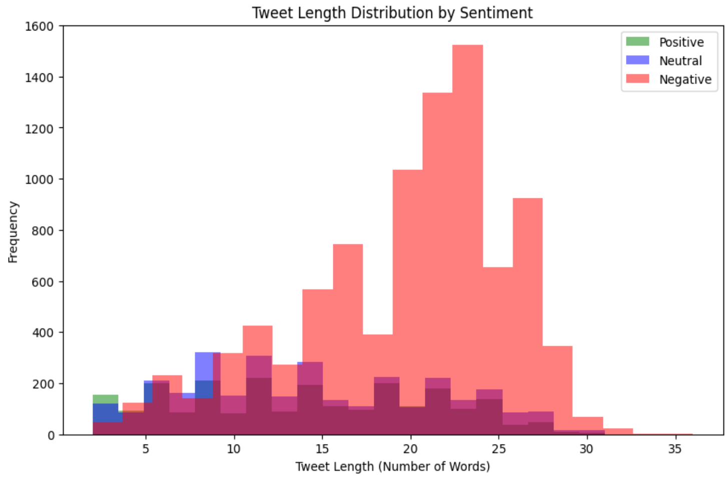
The airline industry has expanded rapidly over the past twenty years, becoming increasingly competitive. Gathering feedback from consumers and analysing it properly has become essential in this environment. One valuable method for this is through effective data collection. This data can then be utilized for sentiment analysis, a technique used to understand the attitudes and emotions conveyed in text or data.(Patel et al., 2023) This report aims to describe an analysis done on airplane reviews to classify sentiments into positive, negative, and neutral using BERT model and fine tuning it according to the need of the data.

Dataset and EDA:

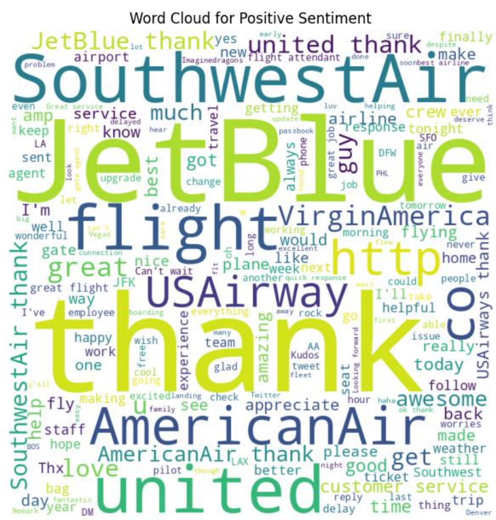
The dataset chosen for this analysis was Airplane reviews and was explored visually to capture complexity, detailed insights, and nature of the data. Dataset shape was viewed and number of sentiments with each category were also printed for better understanding. A bar graph was constructed to visually analyse the distribution of different sentiments within the data and it appears that negative sentiment is the most frequent, followed by neutral sentiment

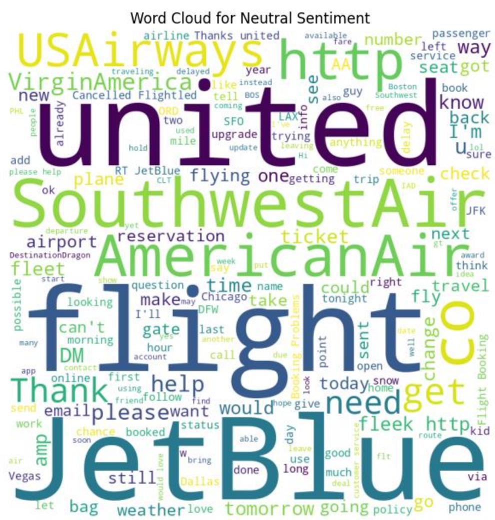






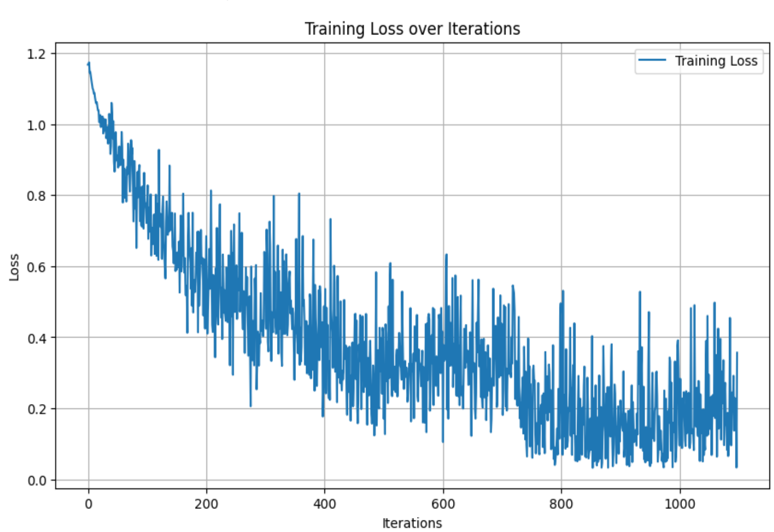
and then positive sentiment. This suggests that most of the text data leans towards negative words or sarcastic tone. After analysing sentiment distribution, the foremost step was to identify the reasons behind most negative sentiments to gain this insight sentiment distribution with respect to airline was illustrated and it came to light that most of the negative comments were received by United airlines followed by US airways on second, American airlines being third highest negative sentiment receiver. To analyse the message or tweet length histogram was plotted, the stacked bars for each tweet length category (positive, neutral, and negative) show how many tweets of that length belong to each sentiment category. This histogram is important in sentiment analysis because it can reveal interesting relationships between tweet length and the emotional tone of the message. The histogram seems to show that a lower proportion of shorter tweets (around 5 words or less) are categorized as positive and neutral sentiment. This could be because people use shorter messages to express quick happiness and satisfaction. Higher proportion of tweets with more words (around 15 words or more) are categorized as negative sentiment. This could be because people use longer messages for revealing their frustration via written communication. Word clouds for all the three sentiments were generated to visualise the most widely used words in all the sentiments which could further help in detecting the biases in the data and could also be fruitful in identifying the balanced or unbalanced dataset.





Model Training:

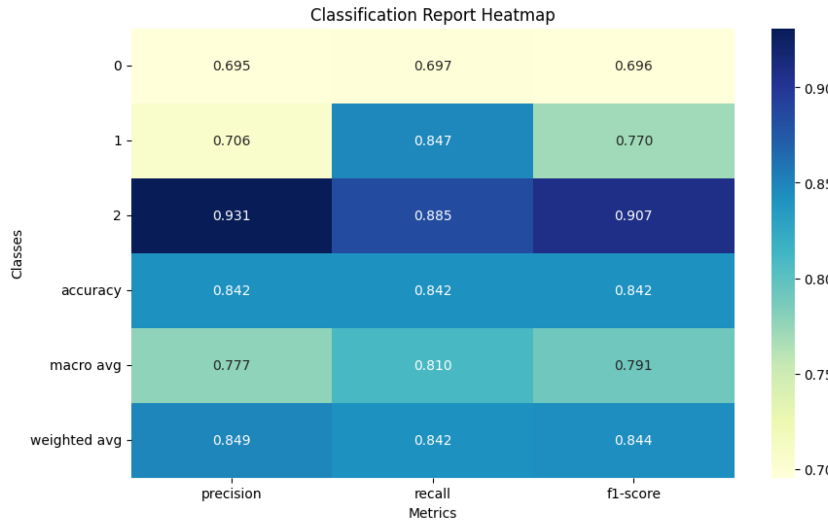
The data was explored and visualised for better understanding and deciding the requirements needed for dataset implementation, BERT as explained above was used to classify airline reviews into positive, negative, or neutral sentiments. Firstly, the sentiment labels were mapped to numerical values, and then the dataset was split into training and testing sets. A custom dataset class was defined to process the text data and tokenize it using the BERT tokenizer. The BERT model was then utilized as a classifier, with a series of fully connected layers added on top for classification. The model was trained using cross-entropy loss and Adam optimizer for a specified number of epochs. During training, the loss values were monitored and printed at regular intervals to track the training progress. Each epoch represents one complete pass through the entire training dataset. Within each epoch, the training process is divided into steps, where a batch of data is processed in each step. The "Loss" value indicates how well the model is performing at each



step, with lower values indicating better performance. In this case, the loss values decrease over the course of training, which suggests that the model is learning to make more accurate predictions as training progresses.

Model testing and Outcomes:

After training the model's performance was evaluated on the test set to finalise that whether model is capable to perform well on unseen data and a classification report was generated. It provides a comprehensive evaluation of the model's performance on a test dataset. And it is visible that the model demonstrates strong performance, with high precision, recall, and F1-score across all sentiment classes, resulting in an overall accuracy of 84%. The macro and weighted average metrics provide insights into the model's performance across all classes, considering both the balanced and class-weighted scenarios, respectively.



Prospects:

This trained and tested model is capable enough to be used in futuristic sentiment analysis projects where core objective is to extract insights and make informed decisions in the airline industry or any field where customer feedback is required. Some of the potential applications are **Customer feedback analysis**, **Operational Decision Making**, **Marketing and Brand Management**, **Competitive Analysis** etc.

Bibliography:

Patel, A., Oza, P., & Agrawal, S. (2023). Sentiment Analysis of Customer Feedback and Reviews for Airline Services using Language Representation Model. *Procedia Computer Science*, *218*, 2459–2467. https://doi.org/10.1016/J.PROCS.2023.01.221

Rachel Wolff. (2020, April 20). *Sentiment Analysis & Machine Learning*. Monkey Learn. https://monkeylearn.com/blog/sentiment-analysis-machine-learning/#:~:text=Sentiment%20analysis%20is%20a%20machine,detect%20sentiment%20without%20human%20input.

What is BERT? An Intro to BERT Models. (2023). https://www.datacamp.com/blog/what-is-bert-an-intro-to-bert-models

Google Collaboratory link:

https://colab.research.google.com/drive/1CmwqCq1wgDt8jLsKWNKm_NLbC86i92-p?usp=sharing