



Author: Marcos Feliz Berrios

Advisor: Dr Jeffrey Duffany

Polytechnic University Of Puerto Rico

Abstract

This study conducted sentiment analysis on social media discussions related to American airlines during the COVID-19 pandemic period and IMDb reviews of the film "Tenet." The sentiment analysis utilized the BERT model, a state-of-the-art natural language processing model, to evaluate public sentiments towards airlines and the film. The model achieved high accuracy in predicting sentiment, indicating its effectiveness in capturing nuanced opinions across different topics. Results from the analysis of Twitter discussions revealed varying sentiments towards different airlines, with Delta Airlines receiving more positive sentiments compared to American Airlines and United Airlines facing consistent challenges with negative sentiments. Similarly, analysis of IMDb reviews for "Tenet" unveiled a diverse range of opinions among viewers, with both positive and negative sentiments expressed towards different aspects of the film. While the study provides valuable insights into public sentiment, limitations include potential dataset bias and challenges in capturing all factors influencing sentiment.

Introduction

Sentiment analysis is a branch of natural language processing (NLP) that focuses on finding, obtaining, and quantifying subjective data from text [1]. It is employed in the analysis of a variety of text data, such as news stories, social media posts, and customer reviews, in order to ascertain the emotional state of the writer or speaker.

The value of sentiment analysis has grown in recent years as social media has emerged as a preeminent information and opinion source. Billions of people use social media sites like Facebook, Twitter, and Reddit to express their opinions and views about a wide range of subjects [2]. It is possible to learn more about consumer sentiment, public opinion, and brand awareness by utilizing this enormous amount of data.

Background

Sentiment analysis can be applied in a variety of ways to extract data from social media. Sentiment analysis, for instance, can be used to analyze customer sentiment, spot patterns in public opinion, assess brand performance, and pinpoint important trends and people in certain situations [3]. Sentiment analysis can be used to spot new trends in public opinion by monitoring opinions on a given topic over time across social media posts [4]. Measurement of consumer sentiment about a product or service is another application for sentiment analysis [5]. By using this data, the customer experience may be enhanced and places where customers are happy or unhappy can be found. Sentiment analysis can also be used to monitor social media brand awareness. This data can be used to gauge the success of marketing campaigns and determine which companies are the most well-known. Sentiment analysis can be applied in a more general way to analyze public opinion on a wide range of issues, in addition to these particular uses.

Problem

There is a significant amount of social media discourse surrounding contemporary subjects, but current sentiment analysis models struggle to adequately capture the diverse and intricate perspectives shared by social media users. There is a need for a novel model capable of delivering precise and timely insights into public sentiment regarding various issues that people routinely contend with.

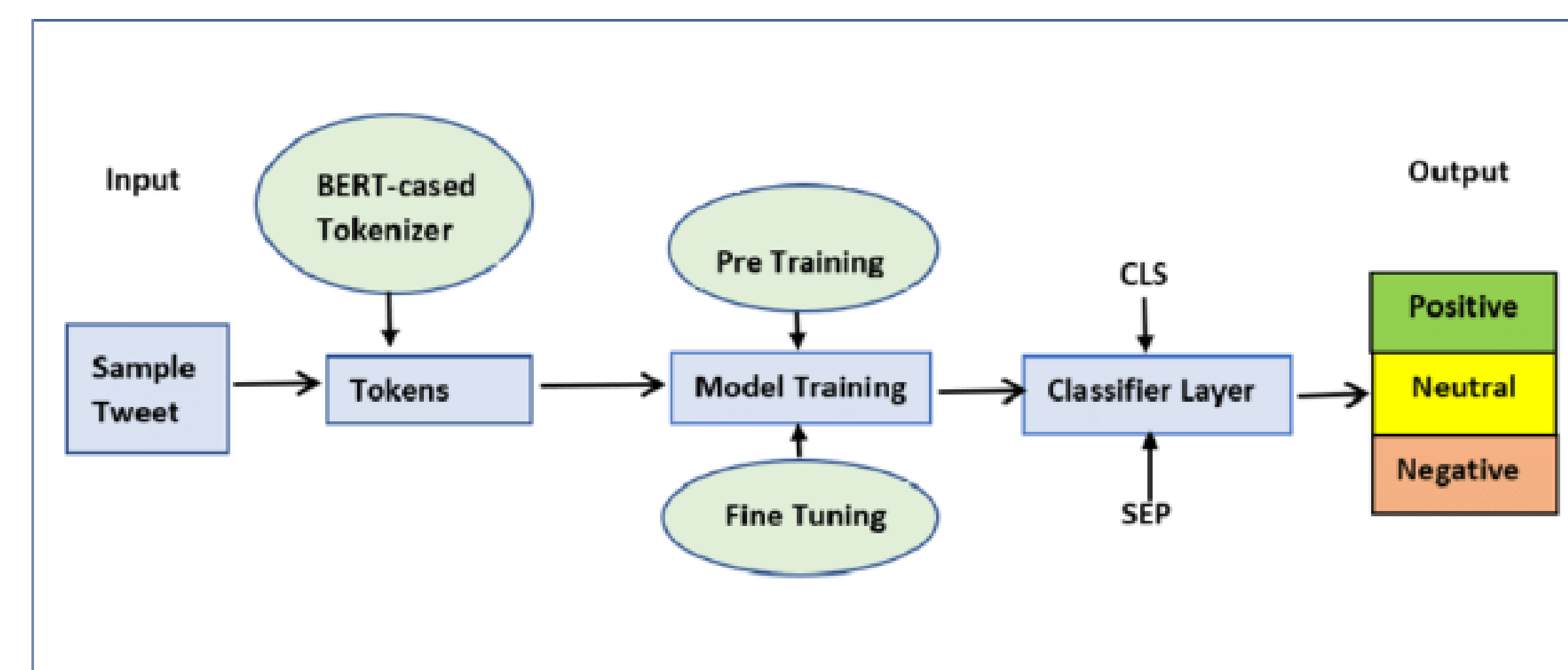
Methodology

The methodology for the current research project was used to design and test a sentiment analysis model that could accurately capture the views and perspectives of Twitter users on airline companies in the U.S. during the pandemic, as well as the sentiments expressed in IMDb reviews of "Tenet." The model was then evaluated on a holdout dataset to assess its performance. The test condition was that if the model performed well, it could subsequently be modified and used to provide valuable insights into public sentiment on other subjects. The sentiment analysis model featured the components outlined below.

Twitter data related to the four chosen U.S. airlines (e.g., American Airlines, Delta Air Lines, United Airlines, Southwest Airlines) was collected using the Twitter API. The dataset encompasses tweets posted between late 2019 and early 2022, capturing a comprehensive timeframe that includes the onset and various phases of the COVID-19 pandemic. Tweets mentioning the airlines' official handles, specific flight experiences, customer service interactions, safety concerns, and pandemic-related issues were included in the dataset.

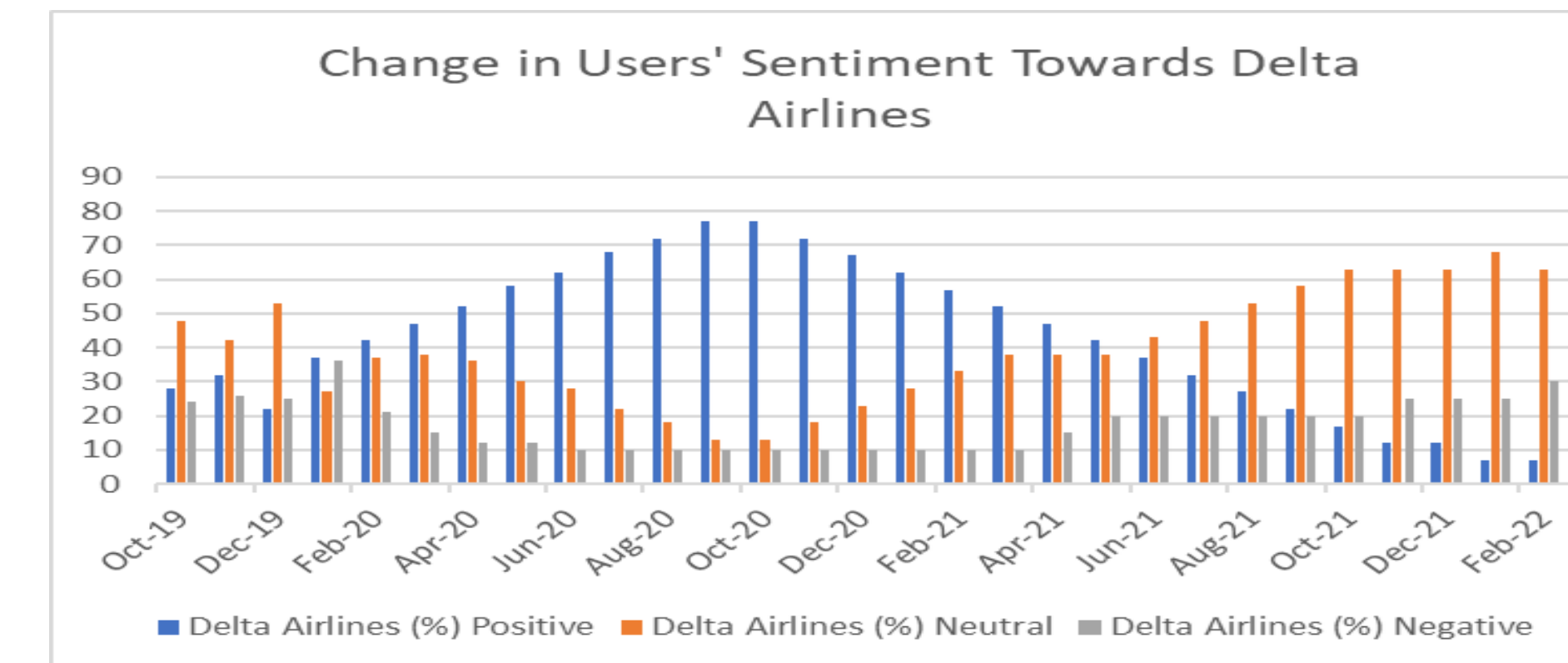
For the second phase of the study, an analysis was conducted to evaluate sentiments regarding the film "Tenet," based on text-based reviews of the movie. Tenet is a 2020 film written and directed by Christopher Nolan. The film received 136 nominations, and ultimately won the 2021 Oscar Award for Best Achievement in Visual Effects. As at April 2024, based on information from Rotten Tomatoes, audience ratings for Tenet stood at 76 percent, while critics collectively rated the film at 69 percent. On IMDb, the film currently has a rating of 7.3/10 as a result of more than 591,000 reviews, which is consistent with the observations on other popular film review sites. This provides a basis for evaluating the representativeness of the sentiment analysis model that was developed as part of the current study when applied to this context.

The performance of the sentiment analysis model was evaluated using a holdout dataset separate from the training data. Model accuracy was assessed by comparing the predicted sentiment of tweets in the holdout dataset to their actual sentiment labels. This evaluation process aimed to validate the reliability and effectiveness of the sentiment analysis model in accurately capturing public sentiments.

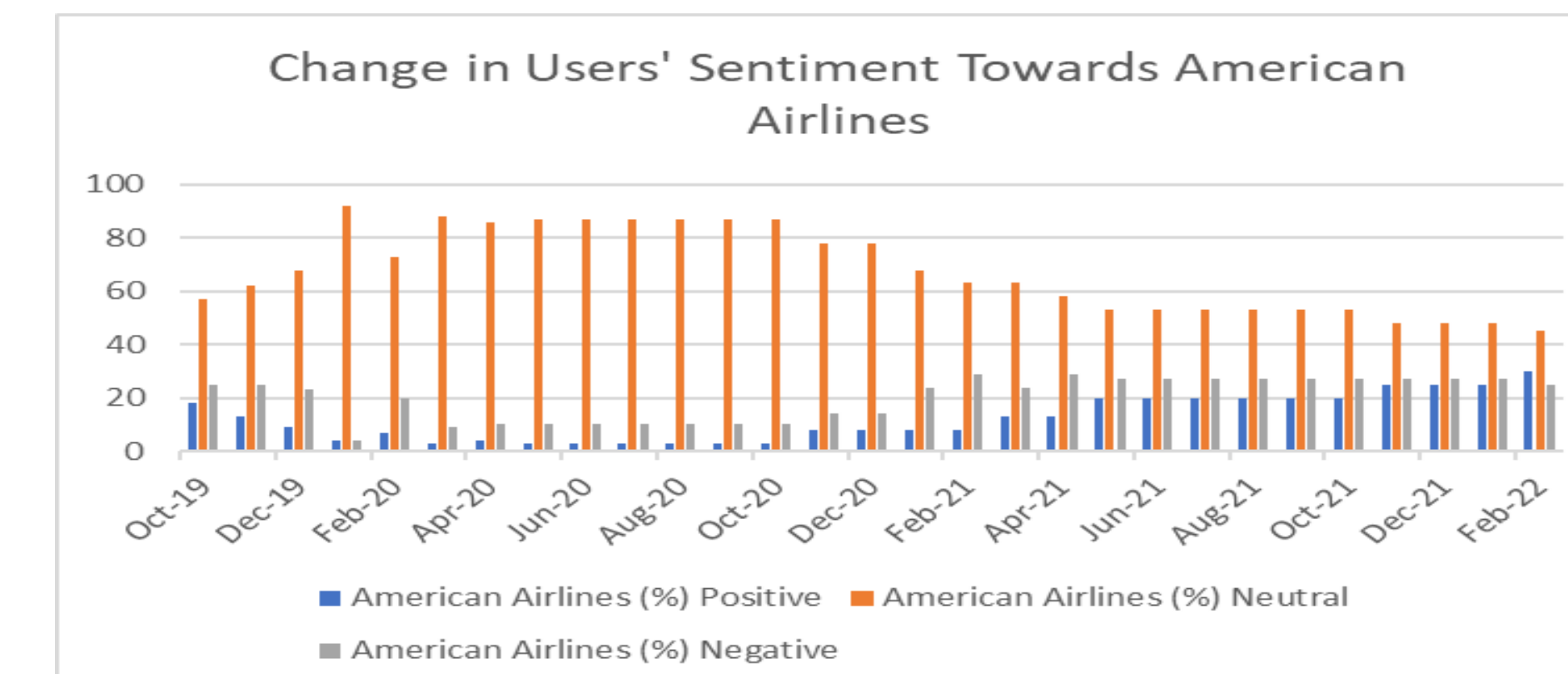


Results and Discussion

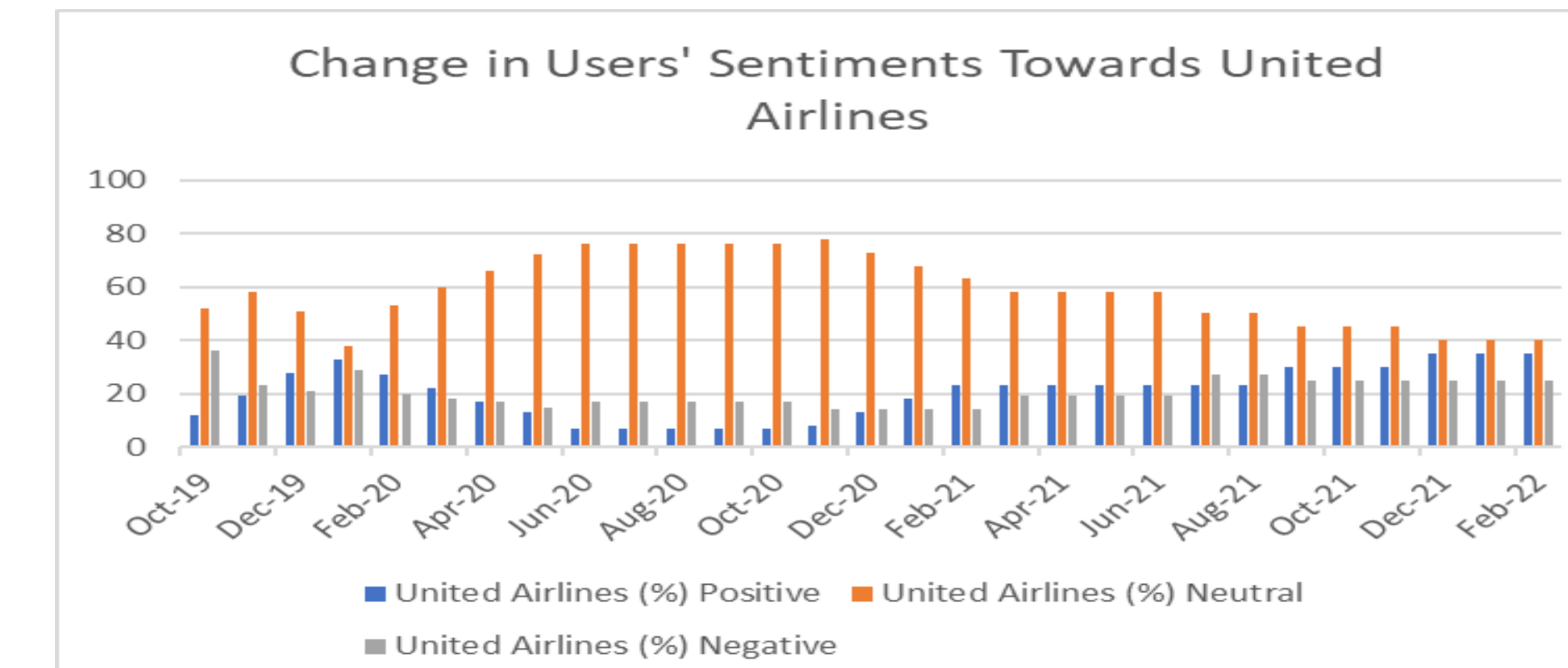
Throughout the pandemic period, Delta Airlines consistently maintained a higher percentage of positive sentiments compared to other airlines.



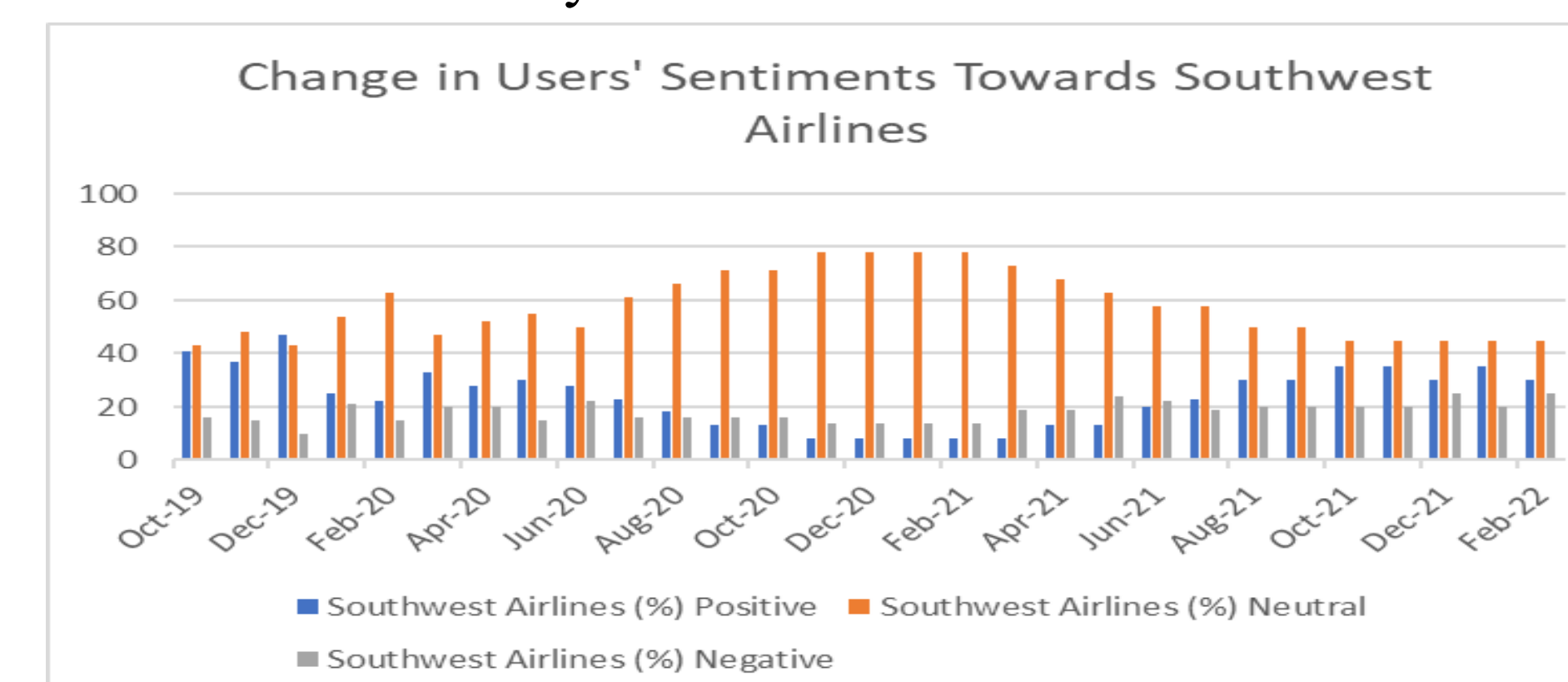
In the mid-2020, the percentage of positive sentiments for American Airlines



Despite fluctuations, United Airlines consistently exhibited a relatively higher percentage of negative sentiments compared to other airlines.



Unlike some other airlines, Southwest Airlines consistently maintained a relatively stable distribution of sentiments.



The sentiment analysis of IMDb reviews for "Tenet" highlights the diverse range of opinions and perceptions among viewers. While the reviewers aggregately held positive perceptions towards the film for its action sequences, cinematography, and intriguing concepts, they also expressed negative sentiments regarding the complex plot and narrative execution of the film. In understanding the sentiment trends and key themes within the reviews, filmmakers, critics, and audiences can gain a deeper understanding of the reception and impact of "Tenet" within the cinematic arena.

Conclusions

This study underscores the potential utility of sentiment analysis in gauging public opinion on social media regarding specific topics, such as the COVID-19 policies implemented by American airlines. The application of the BERT sentiment analysis model yielded insightful results, indicating fluctuations in public sentiment towards different airlines over the pandemic period. While the sentiment analysis confirmed expected trends, particularly regarding Delta Airlines' more positively perceived approach and American Airlines' increased negativity towards the end of the analyzed period, it serves as a valuable tool for understanding evolving public perceptions. Moreover, by applying sentiment analysis techniques to IMDb reviews of "Tenet," this research offers a comprehensive understanding of viewer sentiments towards the film. The results reveal a diverse spectrum of opinions, with a significant portion of reviews expressing both moderate negativity and positivity.

Future Work

To enhance the depth and usefulness of future sentiment analysis endeavors, several adjustments could be considered. For example, conducting sentiment analysis on specific subtopics within the broader context of airline policies during the pandemic may reveal nuanced sentiments. Additionally, exploring sentiment shifts over time or analyzing sentiments related to significant events could provide a more dynamic understanding of public sentiment evolution. Experimentation with alternative sentiment analysis models or refinements to existing models to better capture subtle nuances and contextual variations in sentiment could enhance the depth and usefulness of results. Incorporating qualitative analysis alongside sentiment analysis, such as examining key themes or user demographics, could provide a richer interpretation of public sentiment on complex issues.

Acknowledgements

I am eternally grateful to Dr Jeffrey Duffany, my project supervisor, for his invaluable support and guidance, without which I would not have completed this project with the level of quality it currently has.

References

- [1] K. Naithani and Y. P. Raiwani, "Realization of natural language processing and machine learning approaches for text-based sentiment analysis," *Expert Syst.*, vol. 40, no. 5, p. e13114, Jun. 2023.
- [2] S. A. Curiskis, B. Drake, T. R. Osborn, and P. J. Kennedy, "An evaluation of document clustering and topic modelling in two online social networks: Twitter and Reddit," *Inf. Process. Manag.*, vol. 57, no. 2, p. 102034, Mar. 2020.
- [3] B. Liu, *Sentiment Analysis and Opinion Mining*. Springer Nature, 2022.
- [4] E. D'Andrea, P. Ducange, A. Bechini, A. Renda, and F. Marcelloni, "Monitoring the public opinion about the vaccination topic from tweets analysis," *Expert Syst. Appl.*, vol. 116, pp. 209-226, Feb. 2019.
- [5] S. Yi and X. Liu, "Machine learning based customer sentiment analysis for recommending shoppers, shops based on customers' review," *Complex & Intelligent Syst.*, vol. 6, no. 3, pp. 621-634, Oct. 2020.
- [6] F. A. Acheampong, H. Nunoo-Mensah, and W. Chen, "Transformer models for text-based emotion detection: a review of BERT-based approaches," *Artif. Intell. Rev.*, vol. 54, pp. 1-41, Dec. 2021.
- [7] L. Liu, J. Liu, and J. Han, "Multi-head or single-head? an empirical comparison for transformer training," arXiv preprint arXiv:2106.09650, Jun. 17, 2021.