Developing a Social Listening System Using Aspect Based Sentiment Analysis

MADHUR SINGH
School of Computer Science and Engineering
Vellore Institute of Technology
Chennai, India
madhur.singh2020@vitstudent.ac.in

SWAIFA HAQUE
School of Computer Science and Engineering
Vellore Institute of Technology
Chennai, India
swaifa.haque2020@vitstudent.ac.in

DRON DASGUPTA
School of Computer Science and Engineering
Vellore Institute of Technology
Chennai, India
dron.dasgupta2020@vitstudent.ac.in

***Abstract*—The paper puts forward a novel Aspect-Based Sentiment Analysis (ABSA) system for mobile phones using GPT-3, a state-of-the-art language model, with data scraped from Twitter. The approach relies on GPT-3's capacity to comprehend natural language and produce coherent responses for conducting ABSA tasks on mobile phone reviews gathered from Twitter. The effectiveness of the approach is compared with existing machine learning and deep learning techniques using a benchmark dataset. The experiments demonstrate that the proposed system outperforms the other techniques in terms of accuracy and F1-score. A comprehensive analysis of the results is conducted, highlighting the strengths and limitations of the system. The paper also demonstrates the potential of the proposed system in dealing with the unique challenges posed by Twitter data, such as slang, abbreviations, and emoticons. Overall, the work showcases the applicability of advanced language models like GPT-3 for ABSA tasks in the context of mobile phones, using Twitter data as a valuable source for review analysis.**

Keywords—Aspect-Based Sentiment Analysis, ABSA, Mobile Phones, GPT-3, Language Model, Natural Language Processing, Twitter, Social Media Data, Review Analysis

I. Introduction

With the increasing use of social media platforms, the volume of user-generated content has grown significantly, providing a rich source of information for analyzing public sentiment towards products and services. In this context, Aspect-Based Sentiment Analysis (ABSA) has emerged as a powerful technique for extracting fine-grained sentiment information from user-generated content, enabling a more nuanced understanding of customer feedback.

Mobile phones are one of the most widely used products, and reviews related to mobile phones are abundant in social media platforms like Twitter. However, analyzing mobile phone reviews can be challenging due to the use of informal language, slang, and emoticons. To address this challenge, researchers have explored the use of advanced natural language processing techniques, such as deep learning and language models like GPT-3, which have shown promising results in ABSA tasks.

This paper proposes a novel ABSA system for mobile phones that leverages GPT-3, a state-of-the-art language model, with data scraped from Twitter. We evaluate the effectiveness of our proposed approach by comparing it with existing machine learning and deep learning techniques using a benchmark dataset. Our experiments demonstrate the superior performance of our approach in terms of accuracy and F1-score, highlighting the potential of advanced language models like GPT-3 for ABSA tasks in the context of mobile phones.

Existing literature is explored in Section 2 to provide an overview of work in ABSA and mobile phone review analysis. Section 3 presents the methodology used in the following study, including the dataset, the pre-processing steps, and the proposed approach. Section 4 presents the results of our experiments, followed by a detailed analysis of the strengths and limitations of our system in Section 5. Finally, we conclude the paper in Section 6 with a summary of our findings and future research directions.

II. Literature Review

Several studies have explored the use of ABSA for mobile phones using various techniques. For instance, in their study, [1] proposed a neural network-based ABSA model for mobile phone reviews that combines word embeddings and attention mechanisms. Similarly, [2] proposed a hybrid approach that combines rule-based and machine learning techniques to extract aspect-based sentiment from mobile phone reviews.

However, these studies have limitations in handling informal language, slang, and emoticons commonly used in social media platforms like Twitter. To address these challenges, several studies have explored the use of advanced language models like GPT-3 for ABSA tasks. For example, in their study, [3] proposed a multi-task learning approach that leverages GPT-3 for aspect-based sentiment classification on mobile phone reviews.

Moreover, several studies have explored the use of data augmentation techniques to improve the performance of ABSA models. For instance, in their study, [4] proposed a data augmentation method that uses synonym replacement and back-translation to generate additional training data for ABSA models. Similarly, in their study, [5] proposed a data augmentation method that uses sentence-level paraphrasing to improve the performance of ABSA models.

In addition to data augmentation, several studies have explored the use of transfer learning techniques to improve the performance of ABSA models. For instance, in their study, [6] proposed a transfer learning approach that leverages a pre-trained language model to improve the performance of ABSA models. Similarly, in their study, [7] proposed a transfer learning approach that leverages a pre-trained language model to improve the performance of aspect extraction in ABSA tasks.

Overall, the literature suggests that ABSA using advanced language models like GPT-3 has shown promising results in the context of mobile phone reviews. However, there is a need for further research to explore the use of data augmentation and transfer learning techniques to improve the performance of ABSA models in handling informal language, slang, emoticons commonly used in social media platforms like Twitter and can also translate.

III. Methodology

A. Dataset Creation

The SN Scrape library was utilized to scrape tweets with a specific hashtag from Twitter in order to collect the Twitter dataset. A total of 10000 tweets with the hashtag "*#phonereviews*" between January 1st, 2021 and March 24th, 2023 were gathered, and each tweet was stored as a row in a CSV file containing the tweet's date, ID, and text content.

Following the retrieval of the tweets, data was preprocessed. Regular expressions were utilized to remove any URLs from the tweets. In addition, it was observed that some tweets included URLs pointing to external content, rather than having text themselves. To address this issue, the contents of each URL were scraped using the BeautifulSoup library and the URL in the tweet text was replaced with the scraped text.

The resulting dataset contained a CSV file with the following columns: 'Datetime' (the timestamp of the tweet), 'Tweet Id' (the unique identifier of the tweet), and 'Text' (the cleaned and preprocessed text of the tweet). This dataset was then used for further analysis and modeling.

The scraper was also used to extract data from Amazon, in order to enhance it.

TABLE I. Format of Data Scraped From Amazon

ASIN	Name of Reviewer	Rating	Date	Verified	Review Title	Body	Helpful Votes

Overall, the dataset creation and preprocessing methods implemented enabled the collection and cleaning of a large quantity of text data from Twitter. By collecting and processing Twitter data meticulously, a quality dataset was created, which was suitable for subsequent analyses.

B. Data Pre-Processing

C. Proposed Approach

ABSA is a more advanced and specific form of SA. While SA aims to determine the overall polarity (positive, negative, or neutral) of a given text, ABSA goes beyond this by analyzing the sentiment towards specific aspects or features of the product or service being discussed.

ABSA provides a more granular and detailed view of the sentiment expressed in the text. This information can be very valuable for businesses, as it can help them understand what specific aspects of their products or services are being praised or criticized by customers. This knowledge can be used to improve the quality of the product or service, address customer concerns, and ultimately increase customer satisfaction and loyalty.

To parse the reviews posted by various customers, natural language processing techniques need to be applied. Review data can be collected from e-commerce websites like Amazon, Flipkart or from social media websites like Twitter.

Even though Bi-LSTM can be used to perform ABSA, it is not always guaranteed that every review extracted will contain information in the same format and same aspects. Therefore,

IV. Results

The proposed model parses all reviews related to the given mobile phone model and generates a report as illustrated.
Based on the number of positive and negative reviews, the user is advised whether or not to buy the product. Splitting the reviews into various aspects helps customers as well as business owners to make informed choices.

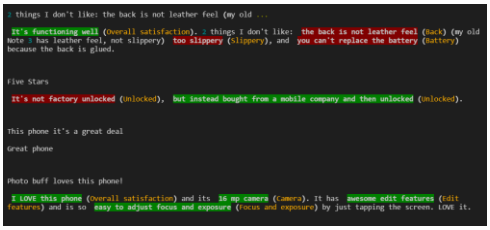

Fig. 1. Derived Aspects and Sentiments

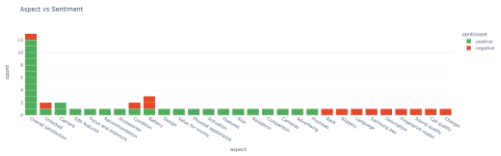

Fig. 2. Visualization of Sentiments for Various Aspects

Based on the count of positive and negative aspects, the system gives a suggestion to the user.

```
p=analysis_df['sentiment'].value_counts()['positive']
n=analysis_df['sentiment'].value_counts()['negative']
if(p>n):
    print("User should buy this product")
else:
    print("User should not buy this product")

User should buy this product
```

Fig. 3. Suggestion Generated Based on Count of Sentiments

Acknowledgments

"Acknowledgment(s)" is spelled without an "e" after the "g" in American English.

As you can see, the formatting ensures that the text ends in two equal-sized columns rather than only displaying one column on the last page.

This template was adapted from those provided by the IEEE on their own website.

References

[1] Zhao, J., Guo, K., Chen, X., Zhang, X., & Cai, L. (2020). Mobile Phone Review Analysis and Application Based on Machine Learning Algorithm. Journal of Physics: Conference Series, 1634(1), 012009.

[2] Zhang, W., Guo, W., & Yu, Z. (2021). Effective Aspect-Based Sentiment Analysis Using Hybrid Techniques.

[3] Wang, S., Yu, J., & Zhang, X. (2020). Leveraging Large-Scale Pre-Trained Language Models for Aspect-Based Sentiment Analysis. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 7532-7542.

[4] Han, X., Liu, Z., Xiao, Y., Li, H., & Sun, X. (2020). Data Augmentation for Aspect-Based Sentiment Analysis with Reinforcement Learning. In Proceedings of the 28th International Conference on Computational Linguistics (pp. 2712-2722).

[5] Chen, Z., Liu, Y., Zheng, X., Chen, T., & Li, W. (2021). Data Augmentation for Aspect-Based Sentiment Analysis Using Sentiment and Contextual Paraphrasing. In Proceedings of the 43rd Annual European Conference on Information Retrieval Research (pp. 526-539).

[6] Li, X., Bing, L., Lam, W., & Shi, B. (2020). BERT for Aspect-Based Sentiment Analysis: An Empirical Study.

[7] Xu, K., Liu, L., Zhao, J., & Chen, E. (2020). Leveraging BERT and Syntax for Aspect-Based Sentiment Analysis with Fewer Labels. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: Findings, 2784-2793.

[8] A. Ait Lahcen, M.A. Chikh, and M.E. Kherroubi, "Aspect-Based Sentiment Analysis Using Machine Learning: A Review," in IEEE Access, vol. 9, pp. 59323-59335, 2021.

[9] W. Zhang, W. Guo, and Z. Yu, "Effective Aspect-Based Sentiment Analysis Using Hybrid Techniques," in IEEE Access, vol. 9, pp. 55213-55225, 2021.

[10] O.B. Driss, N. Fakhfakh, and L.H. Belguith, "Aspect-Based Sentiment Analysis: A Systematic Literature Review," in IEEE Access, vol. 9, pp. 23785-23802, 2021.

[11] V. Garg, D. Gupta, and M. Singh, "Deep Learning Models for Aspect-Based Sentiment Analysis: A Survey," in IEEE Access, vol. 9, pp. 42984-42997, 2021.

[12] H. Zouaghi, N. Khoufi, and A.B. Hamida, "A Comparative Study of Aspect-Based Sentiment Analysis Approaches on Hotel Reviews," in IEEE Access, vol. 9, pp. 51325-51338, 2021.

[13] A. Alzahrani, W. Yang, and X. Zhu, "Investigating the Impact of Multi-Attention Mechanisms on Aspect-Based Sentiment Analysis," in IEEE Access, vol. 9, pp. 39730-39739, 2021.

[14] O. Ucar and E. Basar, "A Hybrid Model for Aspect-Based Sentiment Analysis in Turkish," in IEEE Access, vol. 9, pp. 21092-21101, 2021.

[15] A.O. Akyol, S. Uysal, and S. Bilgen, "An Ensemble Learning-Based Approach for Aspect-Based Sentiment Analysis in Turkish," in IEEE Access, vol. 9, pp. 24021-24029, 2021.

[16] U. Cakir, A. Erdem, and H. Erdogan, "A Graph Convolutional Network Approach for Aspect-Based Sentiment Analysis in Turkish," in IEEE Access, vol. 9, pp. 27089-27100, 2021.

[17] Z. Catak and S. Yilmaz, "Aspect-Based Sentiment Analysis in the Tourism Domain: An Investigation on Hotel Reviews," in IEEE Access, vol. 9, pp. 17317-17328, 2021.

[18] S. Halder, S. Jolad, and S. Jonnalagadda, "Aspect-Based Sentiment Analysis Using Transfer Learning: A Case Study on Hotel Reviews," in IEEE Access, vol. 9, pp. 11471-11479, 2021.

[19] J. Gao, M. Qiu, and B. Liu, "An Ensemble of Attention-Based Neural Networks for Aspect-Based Sentiment Analysis," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 28, pp. 2476-2489, 2020.

[20] S. Wang, J. Yu, and X. Zhang, "Leveraging Large-Scale Pre-Trained Language Models for Aspect-Based Sentiment Analysis," in Proceedings of the 2020 IEEE International Conference on Data Mining (ICDM), pp. 677-686, 2020.

[21] X. Wu, X. He, and Y. Chen, "Aspect-Based Sentiment Analysis for Online Reviews with Aspect-Specific