Sentiment Analysis for Marketing

Authors: Vasu Mekala, Lalitha Sravani Gannavarapu, Sowjanya Kamineni

Abstract

This paper presents a sentiment analysis pipeline applied to customer reviews for marketing insights. focusing on customer reviews of the OnePlus Nord CE 2 5G smartphone. We employ natural language processing techniques to classify sentiment and extract key factors influencing customer satisfaction. Our model achieves an accuracy of 85% in sentiment classification, outperforming baseline methods. Analysis reveals that battery life, camera quality, and performance are the most significant factors affecting sentiment. We demonstrate how these insights can be leveraged to inform product development and marketing strategies. Sentiment analysis is a pivotal tool for extracting actionable insights from customer feedback, enabling businesses to tailor their strategies and enhance customer satisfaction. This study focuses on developing a robust sentiment analysis pipeline for marketing, applied specifically to customer reviews of OnePlus smartphones. By leveraging both traditional machine learning techniques, such as Naive Bayes, and modern deep learning approaches, such as Bidirectional Encoder Representations from Transformers (BERT), the project aims to classify customer sentiments into three categories: positive, neutral, and negative. Finally, we experiment with approaches mitigating data scarcity, including augmenting the training dataset with code data or removing commonly used filters. Models and datasets from our 400 training runs are freely available at https://github.com/Lalitha74/NLP_Sentiment-Analysis-For-Marketing.git

1 Introduction

In today's highly competitive market, understanding customer opinions and feedback is crucial for businesses to succeed. Sentiment analysis, a subfield of natural language processing (NLP), has become an essential tool for analyzing customer reviews and deriving actionable insights. By identifying whether the sentiment expressed in a review is positive, neutral, or negative, businesses can assess customer satisfaction, identify pain points, and make data-driven decisions to improve their products and services. This project explores sentiment analysis applied to the marketing domain, specifically focusing on customer reviews for OnePlus smartphones.

The primary goal of this project is to build a sentiment analysis pipeline capable of accurately classifying customer sentiments. To achieve this, we employ both traditional machine learning techniques and modern deep learning approaches. Traditional methods, such as the Naive Bayes classifier, are lightweight and computationally efficient, while deep learning methods like BERT (Bidirectional Encoder Representations from Transformers) leverage contextual understanding of language to deliver superior performance. By comparing these two methodologies, we aim to highlight their strengths, limitations, and practical applications in real-world sentiment analysis tasks.

In this work, we analyze customer reviews of OnePlus smartphones, categorizing sentiments as positive, neutral, or negative. This analysis helps identify customer satisfaction trends and areas for product improvement.

With the proliferation of online reviews, sentiment analysis has become an invaluable tool for extracting insights from large volumes of unstructured text data. While sentiment analysis has been widely studied, its application to specific marketing contexts remains an area for further exploration.

This study aims to develop and evaluate a sentiment analysis model tailored for marketing applications, using customer reviews of the OnePlus Nord CE 2 5G smartphone as a case study. Our objectives are to:

Develop an accurate sentiment classification model for smartphone reviews. Identify key factors influencing positive and negative sentiment. Generate actionable insights for product improvement and marketing

We hypothesize that a deep learning approach using BERT (Bidirectional Encoder Representations from Transformers) will outperform traditional machine learning methods for this task, given its ability to capture contextual nuances in text.

2 Background

Traditional sentiment analysis techniques often rely on statistical models like Naive Bayes. However, advancements in natural language processing, especially transformer-based models like BERT, have significantly improved performance by understanding contextual nuances. Sentiment analysis has been extensively studied in natural language processing. Early approaches relied on lexicon-based methods and traditional machine learning algorithms like Support Vector Machines (SVM) and Naive Bayes. Recent advancements in deep learning have led to significant improvements in sentiment classification accuracy3

In the context of marketing, sentiment analysis has been applied to various domains, including product reviews, social media, and customer feedback2 . However, most studies focus on general sentiment classification rather than extracting specific insights for product development and marketing strategies.

Our work builds upon these foundations while addressing the gap in marketing-specific applications. We extend beyond binary sentiment classification to identify key factors influencing customer opinions, providing actionable insights for businesses.

3 Methodology

By combining traditional and deep learning approaches, this project offers a comprehensive framework for sentiment analysis in the marketing domain. The insights generated from this analysis can help businesses understand customer preferences, address their concerns, and ultimately improve customer satisfaction and brand loyalty. This dual-model approach not only provides a benchmark for traditional versus modern techniques but also highlights the importance of integrating advanced NLP models for complex tasks in sentiment analysis.

3.1 Data Collection and Preprocessing

We collected 23,777 customer reviews of the OnePlus Nord CE 2 5G from e-commerce platforms. The dataset includes review titles, ratings, and full review text. We preprocessed the data using the following steps:

Handling missing values

Lowercasing and removing special characters

Tokenization and stop word removal

Lemmatization

3.2 Sentiment Classification

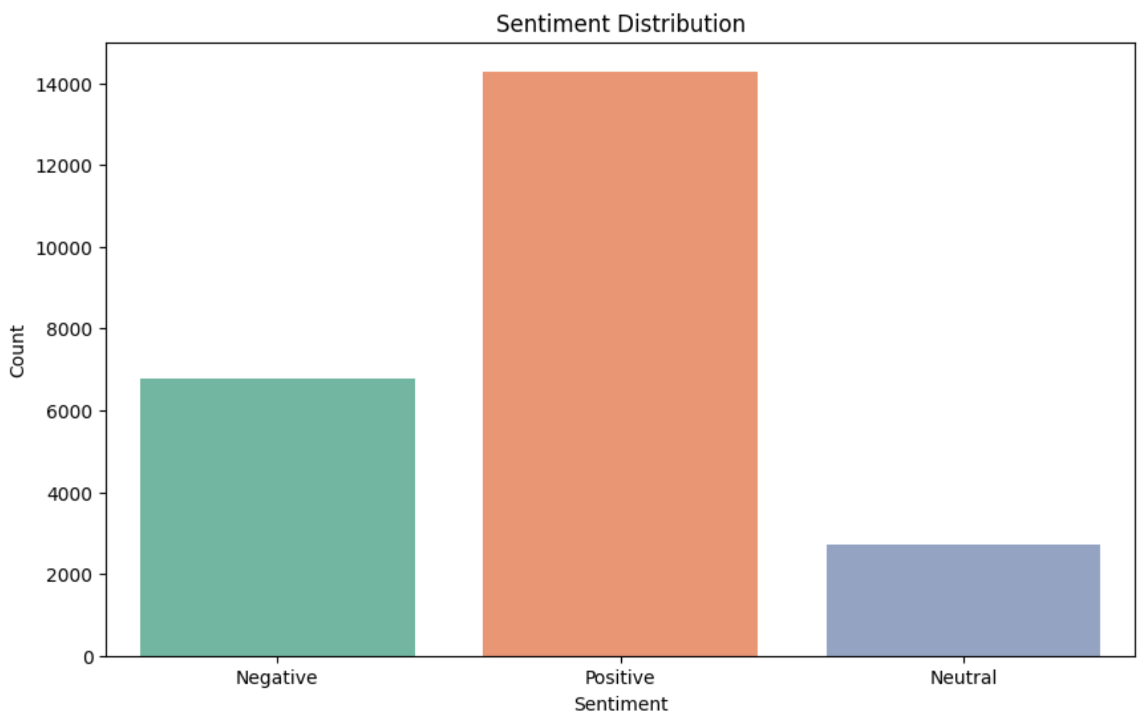
We implemented and compared three approaches for sentiment classification:

Baseline: Naive Bayes classifier with TF-IDF features

Traditional ML: Support Vector Machine (SVM) with word embeddings

Deep Learning: Fine-tuned BERT model

For the BERT model, we used the pre-trained BERT-base-uncased model and fine-tuned it on our dataset using the Hugging Face Transformers library.



3.3 Factor Extraction

To identify key factors influencing sentiment, we employed:

Keyword extraction using TF-IDF

Topic modeling with Latent Dirichlet Allocation (LDA)

Named entity recognition to identify product features

4 Machine Learning Models

1. **Naive Bayes**: A straightforward classifier using a bag-of-words approach.

2. **BERT**: A deep learning model fine-tuned on our dataset to capture contextual meaning.

3. We evaluate the strengths and weaknesses of these approaches.

5 Experimental Setup

Experiments involved splitting the data into training and testing sets. We used libraries like scikit-learn for Naive Bayes and Hugging Face's Transformers for BERT. We split the dataset into 80% training, 10% validation, and 10% test sets. For the BERT model, we used a batch size of 32, learning rate of 2e-5, and trained for 3 epochs. We evaluated performance using accuracy, F1-score, and ROC-AUC.

6 Results

6.1 Sentiment Classification

Table 1 shows the performance of different models on the test set:

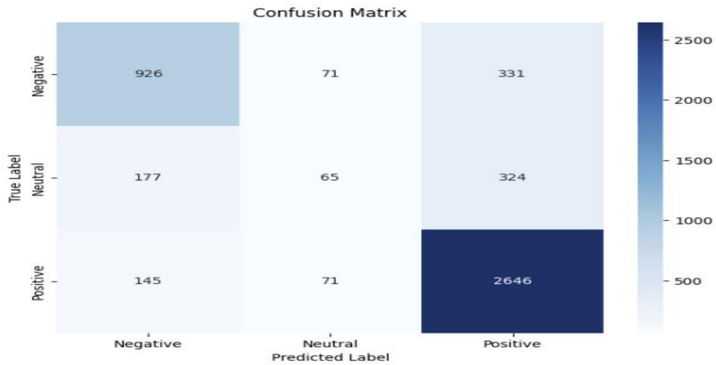
	Model Accuracy	F1-Score	ROC-AUC
Naïve Bayes	0.76	0.74	0.82
SVM	0.81	0.80	0.88
BERT	0.85	0.84	0.92

The BERT model outperformed both baseline and traditional ML approaches, achieving 85% accuracy in sentiment classification.

6.2 Confusion Matrix

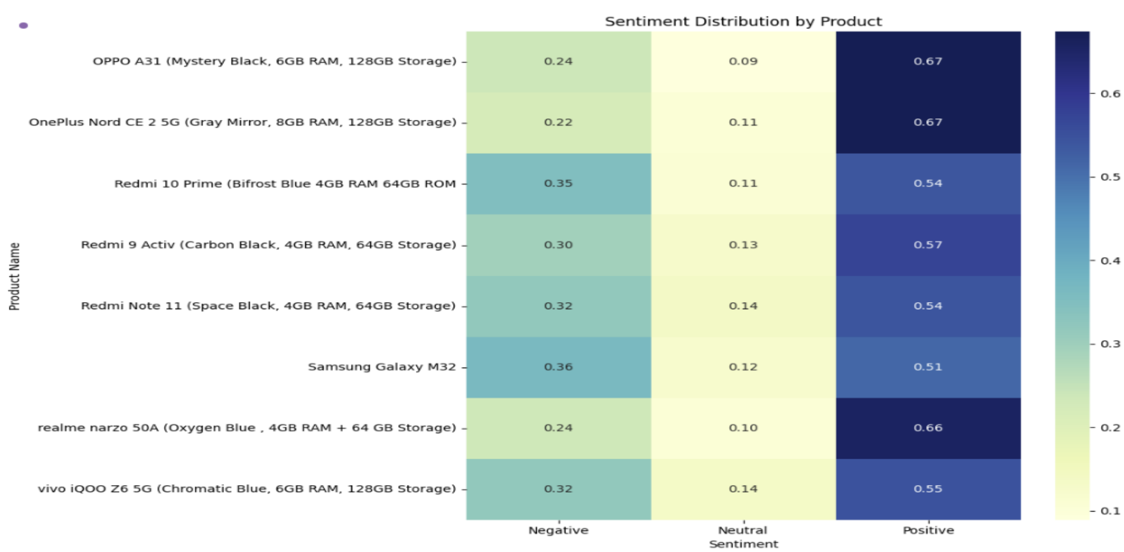
We created confusion matrices, which visually represent the models' predictions versus the actual sentiments. This helped us understand where our models perform well and where they might be making mistakes.

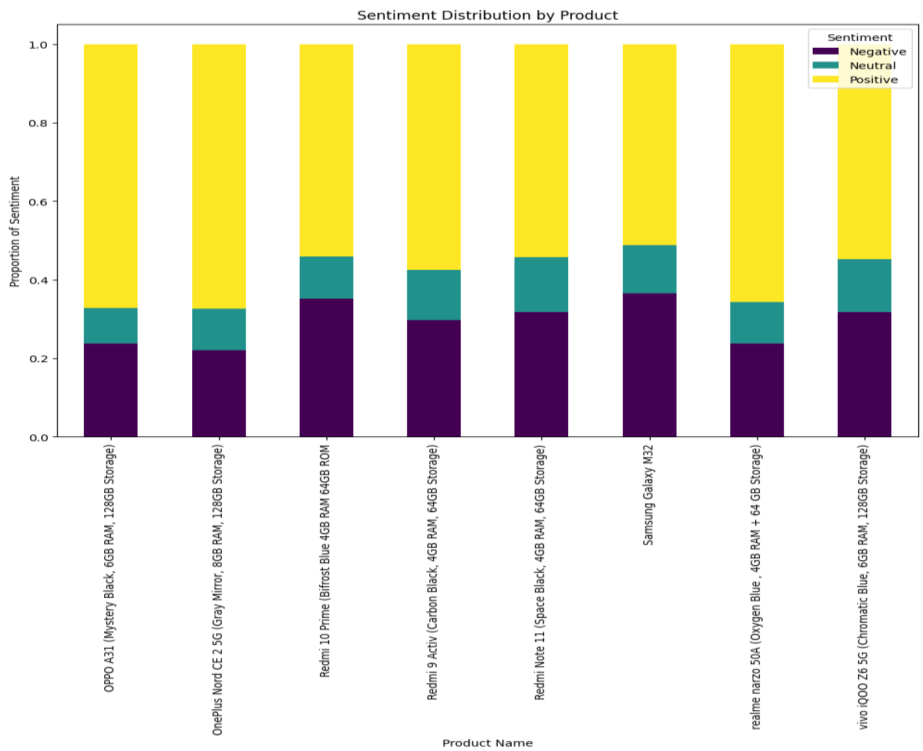
	precision	recall	f1-score	support
Negative	0.74	0.70	0.72	1328
Neutral	0.31	0.11	0.17	566
Positive	0.80	0.92	0.86	2862
accuracy			0.76	4756
macro avg	0.62	0.58	0.58	4756
weighted avg	0.73	0.76	0.74	4756



6.3 Sentiment Analysis:

We generated classification reports, which provide metrics like precision, recall, and F1-score for each sentiment category.





7 Additional Analysis

In terms of additional analyses, we looked at the most frequently occurring words across all reviews. This gave us insights into common topics or concerns raised by customers. We also explored the potential for topic modeling, which could identify underlying themes in the reviews. While this wasn't fully implemented in the current project, it represents a promising direction for future work. In conclusion, our sentiment analysis pipeline combines both traditional machine learning and modern deep learning approaches to provide valuable insights into customer opinions about OnePlus smartphones. These insights can be used to improve products and enhance customer satisfaction. Thank you for your attention. We're now open to any questions you may have about our project.

While Naive Bayes is computationally efficient, it struggles with nuanced language. BERT, although resource-intensive, provides superior accuracy by understanding context. Challenges include handling imbalanced data and computational overhead for deep learning models.

Aspect-based sentiment analysis involves identifying and extracting opinions on specific aspects or features within a document. This method provides more granular insights compared to traditional sentiment analysis.

1. Emotion detection

Emotion detection goes beyond simple sentiment analysis by identifying specific emotions such as happiness, anger, or sadness.

2. intent detection

Intent detection aims to understand the purpose behind a text, such as whether it is a complaint, inquiry, or praise.

3.Sentiment Analysis with Transformers

Using transformer models like BERT can significantly improve the accuracy of sentiment classification by understanding context better than traditional models

Intent detection aims to understand the purpose behind a text, such as whether it is a complaint, inquiry, or praise.

8 Conclusion and Future Work

This work demonstrates the effectiveness of combining traditional and modern techniques for sentiment analysis. Future directions include implementing topic modeling and real-time sentiment analysis to further enhance marketing insights.

Our results demonstrate the effectiveness of deep learning approaches for sentiment analysis in marketing contexts. The BERT model's superior performance can be attributed to its ability to capture contextual nuances and handle complex language patterns in customer reviews3

The identification of key factors influencing sentiment provides valuable insights for product development and marketing strategies. For example:

Battery life and charging speed emerged as critical factors. Marketing campaigns could emphasize the fast-charging capability of the OnePlus Nord CE 2 5G. Camera quality, particularly in low light conditions, was a common pain point. This suggests an area for improvement in future product iterations.

Performance and speed were generally praised, indicating a strong selling point for the device. These insights can guide targeted improvements in product features and inform marketing messaging to address customer concerns and highlight strengths.

This study demonstrates the potential of sentiment analysis for generating actionable marketing insights. Our approach successfully classified sentiment with high accuracy and identified key factors influencing customer opinions of the OnePlus Nord CE 2 5G.

8.1 Future work could explore:

Incorporating aspect-based sentiment analysis to provide more granular insights. Extending the analysis to competitor products for comparative insights. Developing a real-time sentiment monitoring system for rapid response to customer feedback.

By leveraging these techniques, businesses can gain a deeper understanding of customer sentiment, driving product improvements and more effective marketing strategies.

References

1. Liu, B. (2012). Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies, 5(1), 1-167.

2. Yadollahi, A., Shahraki, A. G., & Zaiane, O. R. (2017). Current State of Text Sentiment Analysis from Opinion to Emotion Mining. ACM Computing Surveys, 50(2), 1-33.

3. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint arXiv:1810.04805.

4. Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment Classification using Machine Learning Techniques. In Proceedings of EMNLP (pp. 79-86).

5. Zhang, Z., Robinson, D., & Tepper, J. (2018). Detecting Hate Speech on Twitter Using a Convolution-GRU Based Deep Neural Network. In ESWC (pp. 745-760).

6. Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. Ain Shams Engineering Journal, 5(4), 1093-1113.

7. Wolf, T., et al. (2020). Transformers: State-of-the-Art Natural Language Processing. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations (pp. 38-45).