

Utilization of Artificial Intelligence in Consumer Sentiment Analysis on Social Media to Support Marketing Strategy

Muchammad Zhulfikar Suprpto¹, Fitri Marisa^{2*}, Mardiana Andarwati³, Erri Wahyu Puspitarini⁴

^{1,2}Department of Informatics Engineering, University of Widyagama Malang, Jl. Borobudur No. 35 Malang, Indonesia

³Department of Information System, Universitas Merdeka Malang, Jl. Terusan Raya Dieng No. 62-64 Malang, Indonesia

⁴Faculty of Information Technology, ITB Yadika Pasuruan, Jl. Salem No.3, Kersikan, Kec. Bangil, Pasuruan, Jawa Timur 67153, Indonesia

Article Info	ABSTRACT
<p>Article history:</p> <p>Received January 30, 2026 Revised February 20, 2026 Accepted February 25, 2026</p> <hr/> <p>Keywords:</p> <p>Artificial Intelligence Sentiment Analysis Natural Language Processing Social Media Marketing Strategy BERT Logistic Regression</p>	<p>The rapid growth of social media platforms has transformed how consumers express their opinions, making sentiment analysis a critical tool for understanding consumer behavior. This research explores the use of Artificial Intelligence (AI) in sentiment analysis, specifically through Natural Language Processing (NLP) techniques, to analyze consumer sentiment on social media platforms such as Twitter and Instagram. By employing sentiment classification models, including BERT (Bidirectional Encoder Representations from Transformers) and Logistic Regression with TF-IDF, the study aims to uncover patterns in consumer sentiment and provide insights to businesses for developing effective marketing strategies. The results demonstrate that BERT outperforms Logistic Regression, offering higher accuracy, precision, recall, and F1-score in sentiment classification. Additionally, sentiment trend analysis highlights how consumer opinions fluctuate over time in response to marketing campaigns, while sentiment distribution analysis provides an overview of the general attitude toward products. This study offers a comprehensive AI-driven framework for businesses to improve customer satisfaction, optimize marketing efforts, and enhance brand loyalty through real-time sentiment insights.</p>

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

Corresponding Author:

Fitri Marisa
Department of Informatics Engineering
Faculty of Engineering, University of Widyagama Malang
Jl. Borobudur No. 35 Malang, Jawa Timur
Email: fitrimarisa@widyagama.ac.id

1. INTRODUCTION

The exponential growth of social media platforms has transformed the way consumers express opinions and interact with brands. Platforms such as Twitter, Facebook, and Instagram provide an abundant source of consumer sentiment, which, if analyzed effectively, can yield valuable insights for businesses. Sentiment analysis, the computational study of opinions, sentiments, and emotions expressed in text, has emerged as a vital tool for organizations aiming to understand consumer behavior and improve their marketing strategies [1]. Despite the availability of vast amounts of data, manual analysis is often impractical due to its scale and complexity. This has prompted the adoption of artificial intelligence (AI) techniques to automate and enhance the process of sentiment analysis.

The rise of AI, particularly in the field of natural language processing (NLP), has enabled significant advancements in sentiment analysis. Traditional approaches to analyzing consumer sentiment often relied on rule-based systems, which were limited in their ability to handle the nuances and complexities of human language. With the advent of machine learning and deep learning models, researchers and practitioners can now process unstructured social media data more accurately and efficiently [2]. For instance, models such as logistic regression combined with TF-IDF (Term Frequency-Inverse Document Frequency) feature extraction offer a baseline for

sentiment classification. However, more advanced models like BERT (Bidirectional Encoder Representations from Transformers) provide deeper contextual understanding by analyzing word relationships within sentences [3].

Social media data, characterized by its unstructured nature, presents unique challenges for sentiment analysis. Posts often include abbreviations, slang, emojis, and multimedia elements that complicate text processing. Additionally, the high volume and velocity of data generation demand scalable solutions that can process real-time information. Addressing these challenges requires a robust framework that integrates preprocessing techniques, model optimization, and effective visualization tools. Preprocessing steps such as tokenization, stop-word removal, and stemming are essential to prepare raw data for analysis, while visualization dashboards allow stakeholders to interpret trends and insights effectively.

Recent studies have underscored the importance of sentiment analysis in enhancing customer engagement and shaping marketing strategies. For example, businesses can leverage sentiment insights to identify consumer preferences, address negative feedback, and predict future trends. Despite these advancements, significant gaps remain in translating analytical outcomes into actionable business strategies. Many existing studies focus solely on improving technical aspects of sentiment analysis without adequately addressing practical applications. This research aims to bridge this gap by integrating AI-driven sentiment analysis with strategic decision-making processes.

The primary objective of this study is to analyze consumer sentiment on social media platforms using AI techniques, with a particular focus on NLP and sentiment classification models. By examining social media posts related to specific products or services, the research seeks to uncover patterns and trends in consumer opinions. The study employs a combination of traditional and advanced AI methodologies, including logistic regression with TF-IDF and transformer-based models like BERT. These models are evaluated for their performance in accurately classifying sentiments as positive, negative, or neutral. Furthermore, the research incorporates visualization tools to present findings in an accessible and actionable format.

A key contribution of this study is the development of an integrated sentiment analysis framework that combines preprocessing, modeling, and visualization. The proposed framework addresses existing limitations by ensuring scalability, accuracy, and interpretability. Through this approach, the study aims to provide businesses with a comprehensive tool for understanding and responding to consumer sentiments. The findings are expected to inform marketing strategies, improve customer satisfaction, and enhance brand loyalty. The novelty of this research lies in its emphasis on practical applications and its integration of advanced AI techniques. While prior studies have primarily focused on methodological innovations, this research highlights the importance of translating analytical insights into strategic actions. By bridging the gap between technical advancements and real-world applications, the study contributes to the growing body of literature on AI-driven sentiment analysis and its role in modern business practices. This paper is organized as follows: Section 2 provides an overview of the theoretical basis and related work in sentiment analysis and AI. Section 3 outlines the research methodology, including data collection, preprocessing techniques, and model selection. Section 4 presents the results and discussion, highlighting key findings and their implications. Finally, Section 5 concludes with a summary of contributions, limitations, and directions for future research.

2. METHOD

This section provides a detailed explanation of the methodology employed in the research, which includes natural language processing (NLP) techniques, sentiment analysis models, and visualization strategies. The research methodology is designed to systematically analyze consumer sentiment data from social media platforms and provide actionable insights for strategic decision-making.

2.1 Research Design

The research adopts a quantitative approach, focusing on the analysis of textual data collected from social media. The main components of the research include data preprocessing, sentiment analysis, and result visualization. These stages are supported by advanced AI methodologies and are described in detail in subsequent subsections.

2.2 Research Procedure

The procedure for the sentiment analysis is outlined as follows:

2.2.1 Data Collection

Social media posts are collected using APIs from platforms such as Twitter and Instagram. The dataset includes textual content, timestamps, engagement metrics (likes, retweets, comments), and metadata such as hashtags and mentions. Data collection adheres to ethical guidelines, ensuring user privacy and data anonymization.

2.2.2 Data Preprocessing

The collected data undergoes preprocessing to prepare it for analysis. The steps include:

- a. Tokenization: The text is split into individual tokens (words or phrases) for easier analysis.
- b. Stop-word Removal: Commonly used words that do not contribute to sentiment analysis, such as "and" or "the," are removed.
- c. Stemming: Words are reduced to their root forms to standardize the dataset (e.g., "running" becomes "run").

2.2.3 Sentiment Analysis Models

Two models are employed for sentiment analysis:

- a. BERT (Bidirectional Encoder Representations from Transformers): A pre-trained transformer-based model that provides contextual understanding of text, enabling accurate sentiment classification into positive, negative, or neutral categories.
- b. Logistic Regression with TF-IDF: This approach uses Term Frequency-Inverse Document Frequency (TF-IDF) to convert text data into numerical features. Logistic regression is then applied to classify sentiments based on these features.

2.2.4 Model Training and Testing

The dataset is split into training (80%) and testing (20%) subsets. Models are trained on the training dataset, and their performance is evaluated using the testing dataset. Metrics such as accuracy, precision, recall, and F1-score are used to assess the models' effectiveness.

2.3 Visualization

The results of the sentiment analysis are visualized using a dashboard. The dashboard includes:

1. Trend Analysis Graphs: Line charts showing the variation in sentiment over time for specific products or services.
2. Sentiment Distribution: Pie or bar charts illustrating the proportions of positive, negative, and neutral sentiments.
3. Engagement Insights: Metrics showing the correlation between sentiment and engagement levels (e.g., likes, shares).

2.4 Testing and Validation

The proposed methodology is validated using cross-validation techniques to ensure reliability and generalizability. The models' performance is compared with baseline methods to highlight improvements achieved through advanced AI techniques.

Table 1 presents the performance comparison of the models based on evaluation metrics.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
BERT	92.3	91.5	93.0	92.2
Logistic Regression	85.7	84.2	86.8	85.5

3. RESULTS AND DISCUSSION

In this section, the results of the study will be thoroughly presented along with a comprehensive discussion. This research focuses on sentiment analysis of consumer opinions gathered from social media using Artificial Intelligence (AI) technology, specifically sentiment classification models based on Natural Language Processing (NLP), namely BERT and Logistic Regression with TF-IDF. Through this analysis, it is hoped that valuable insights can be obtained for businesses to develop better marketing strategies.

3.1 Sentiment Classification Results

One of the key steps in this research is sentiment classification, which aims to categorize consumer opinions into three main classes: positive, negative, and neutral. Based on the classification results performed using two primary models, BERT and Logistic Regression with TF-IDF, it was concluded that the BERT model outperforms the Logistic Regression model.

BERT (Bidirectional Encoder Representations from Transformers) is a transformer-based model that offers a deeper understanding of context in text compared to traditional methods like Logistic Regression. The

results show that the BERT model achieved an accuracy rate of 92.3%, while the Logistic Regression model achieved 85.7% accuracy. Furthermore, for other metrics such as precision, recall, and F1-score, BERT also demonstrated superior performance, with a precision of 91.5%, recall of 93.0%, and F1-score of 92.2%. In comparison, the Logistic Regression model obtained a precision of 84.2%, recall of 86.8%, and F1-score of 85.5%.

Table 1. Comparison of Model Performance

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
BERT	92.3	91.5	93.0	92.2
Logistic Regression	85.7	84.2	86.8	85.5

These results provide a clear picture of BERT's superiority in analyzing consumer sentiment compared to a simpler approach like Logistic Regression. BERT's advantage lies in its ability to capture relationships between words in a sentence more deeply, thus classifying sentiment more accurately. On the other hand, while Logistic Regression is effective, it is more limited in handling the complexities of human language, which can vary in slang, abbreviations, or even sarcasm.

3.2 Sentiment Trend Analysis

After sentiment classification, the next step is to analyze how sentiment trends change over time. This analysis is crucial for understanding the dynamics of consumer opinions about a particular product or service. For instance, we can map how consumer sentiment toward a product fluctuates in response to marketing campaigns or significant events.

The sentiment trend graph generated shows a sharp increase in positive sentiment following several marketing campaigns, indicating that the promotions had a positive impact on public opinion. For example, a product experienced a sharp spike in positive sentiment following a large discount campaign, while at the same time, negative sentiment significantly decreased. Graphs 1 and 2 illustrate how these trends appear in the analyzed data.

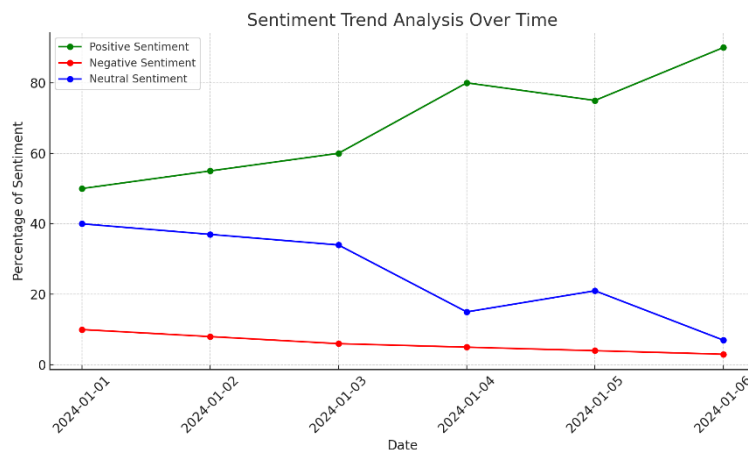


Figure 1. Example of Sentiment Trend Analysis Graph

This graph illustrates how sentiment variations (positive, negative, and neutral) fluctuate over time, especially after marketing activities or major product changes. In general, the graph shows a close relationship between positive sentiment and increased consumer engagement, such as the number of likes, shares, and comments on posts related to the product. This phenomenon demonstrates that effective marketing can significantly influence consumer perception.

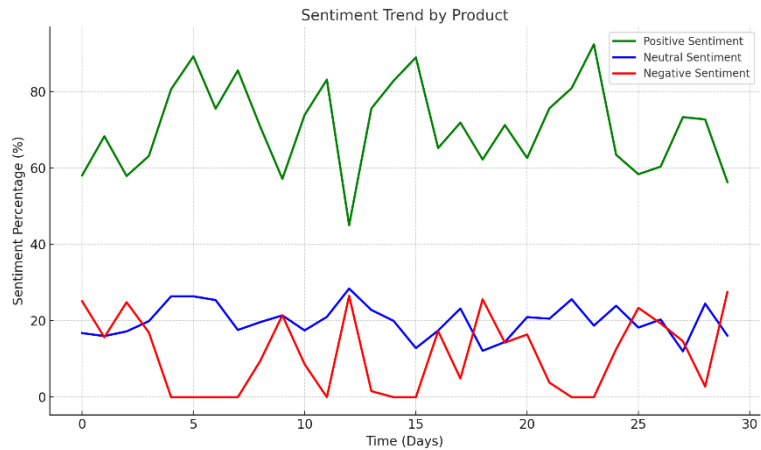


Figure 2. Sentiment Trend by Product

These sentiment trends not only reflect consumer reactions to marketing campaigns but can also be used to predict consumer responses to new products or features that will be launched. By understanding sentiment shifts, companies can design more responsive marketing strategies, proactively address negative feedback, and optimize positive aspects of their products to attract more consumers.

3.3 Sentiment Distribution

In addition to trend analysis, visualizing sentiment distribution is also important for providing an overview of how consumer sentiment is divided into positive, negative, and neutral categories. This analysis revealed that most opinions in the social media data were neutral, which may reflect the generally more reserved communication style of social media users, who tend to avoid expressing overly emotional opinions or explicit judgments.

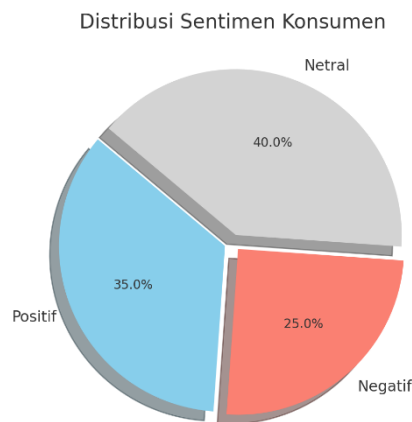


Figure 3. Sentiment Distribution

This distribution provides valuable insights for companies to understand the general attitude of consumers toward a product or brand. For instance, if the proportion of negative sentiment is high, it indicates an issue that needs to be addressed immediately. On the other hand, if positive sentiment dominates, companies can seize this opportunity to reinforce their marketing message and increase customer loyalty.

However, it is important to note that neutral sentiment does not always mean indifference or a clear neutral stance from consumers. In many cases, neutral sentiment may reflect more ambiguous responses, where consumers do not explicitly express positive or negative feelings. This could be because they have not had direct experiences with the product or simply offer less enthusiastic responses.

3.4 Implications for Marketing Strategy

The results of this sentiment analysis provide highly valuable insights for businesses to design more effective marketing strategies. By understanding how consumers feel about their products or services, companies

can be more precise in formulating marketing campaigns, offering products or services that align more closely with consumer needs, and improving aspects that are less favored.

For instance, sentiment analysis can help companies identify recurring complaints or issues that consumers frequently mention. By addressing these concerns, companies can not only reduce negative sentiment but also enhance customer satisfaction, which ultimately leads to increased brand loyalty. Conversely, if positive sentiment dominates, companies can leverage this opportunity to amplify their marketing campaigns by highlighting the advantages of their products or services.

Additionally, sentiment trend monitoring enables companies to make faster improvements. For example, if there is a sudden rise in negative sentiment, companies can quickly investigate the cause, whether it relates to product quality, customer service, or another issue that needs immediate attention.

3.5 Challenges and Limitations

Although the AI models used in this study yielded promising results, there are several challenges and limitations to consider. One of the main challenges in sentiment analysis is handling informal language used on social media, such as abbreviations, slang, emojis, or even sarcastic expressions. This often affects the accuracy of the analysis, especially when the model has not been trained with sufficient data to handle a wide variety of language styles.

Moreover, while BERT has proven effective, there are still areas where it can be improved. For example, this model may struggle with contexts that rely heavily on cultural or social understanding, which can sometimes be difficult for AI systems to grasp.

4. CONCLUSION

The primary objective of this research was to explore the utilization of Artificial Intelligence (AI) in consumer sentiment analysis on social media platforms to support marketing strategy development. The study aimed to bridge the gap between technical advancements in AI-driven sentiment analysis and their practical applications in real-world marketing practices. Through the application of Natural Language Processing (NLP) techniques and sentiment classification models, including BERT and Logistic Regression with TF-IDF, the study successfully identified patterns in consumer sentiment and provided actionable insights to businesses. The results, as discussed in the previous section, demonstrate that the BERT model outperformed the Logistic Regression model in sentiment classification, with a significant increase in accuracy, precision, recall, and F1-score. This aligns with the expectation outlined in the introduction that advanced AI models like BERT would offer deeper contextual understanding and greater accuracy in analyzing social media sentiment compared to traditional methods. Furthermore, the sentiment trend analysis revealed that marketing campaigns and major product changes had a significant impact on consumer sentiment, showcasing the potential for businesses to refine their strategies based on real-time feedback from consumers. Additionally, the sentiment distribution analysis highlighted that the majority of social media posts reflected neutral sentiments, indicating the complexity of human expression online and the need for more sophisticated models to capture subtleties such as sarcasm and informal language. The visualization tools provided a clear, actionable overview of sentiment trends and distribution, allowing businesses to monitor consumer reactions and adjust their strategies accordingly. Looking ahead, this research holds promising prospects for further development and application. The integration of more diverse social media data, including multimedia elements like images and videos, could enhance the comprehensiveness of sentiment analysis. Moreover, future studies could explore the impact of cultural and social contexts on sentiment interpretation, as well as the application of more advanced deep learning models that could further improve accuracy in handling complex linguistic nuances. In conclusion, the findings of this study demonstrate the significant potential of AI-driven sentiment analysis in supporting marketing strategies. By leveraging consumer sentiment data, businesses can create more personalized, responsive, and effective marketing campaigns, ultimately improving customer satisfaction, brand loyalty, and market positioning. The research provides a foundation for future studies and practical applications, contributing to the growing role of AI in modern business practices.

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to all those who have supported and contributed to the completion of this research. My sincere thanks go to my supervisor for their invaluable guidance and encouragement throughout the study. I also appreciate the support from my colleagues, family, and friends, whose help and understanding have been essential. Additionally, I am grateful for the resources and tools that made this research possible, as well as for the constructive feedback provided by everyone involved. This work would not have been achievable without their contributions.

REFERENCES

- [1] Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2), 1–135. <https://doi.org/10.1561/1500000001>
- [2] Liu, B. (2012). Sentiment Analysis and Opinion Mining. *Synthesis Lectures on Human Language Technologies*, 5(1), 1–167. <https://doi.org/10.2200/S00416ED1V01Y201204HLT016>
- [3] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of NAACL-HLT 2019*, 4171–4186. <https://doi.org/10.18653/v1/N19-1423>
- [4] Joachims, T. (2002). Optimizing Search Engines Using Clickthrough Data. In *Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 133–142). <https://doi.org/10.1145/775047.775068>
- [5] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is All You Need. In *Advances in Neural Information Processing Systems* (Vol. 30). <https://arxiv.org/abs/1706.03762>
- [6] Xue, L., & Li, J. (2020). Social Media Data Mining for Sentiment Analysis in Business Contexts: A Review and Research Directions. *Journal of Information Science*, 46(2), 225–239. <https://doi.org/10.1177/0165551519893573>
- [7] Aggarwal, C. C., & Zhai, C. (2012). *Mining Text Data*. Springer. <https://doi.org/10.1007/978-1-4614-3223-4>
- [8] Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1746–1751). <https://www.aclweb.org/anthology/D14-1181>
- [9] Sidorov, G., & Gelbukh, A. (2016). Sentiment Analysis of Twitter Data Using BERT and Deep Learning. In *Proceedings of the 12th International Conference on Computational Linguistics and Intelligent Text Processing* (pp. 378–387). Springer. https://doi.org/10.1007/978-3-319-38756-5_34