International Journal of Research Publication and Reviews, Vol 6, Issue 5, pp 17920-17922 May 2025

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Consumer Sentiment Analysis

Mayank, Vidushi, Vaibhav Pandey & Prabhat Jha

Information Technology, Inderprastha Engineering College, Ghaziabad

ABSTRACT:

The Consumer Sentiment Analysis project aims to bridge the gap between consumers and brands by leveraging advanced AI technologies to analyze consumer feedback and behavior. The platform allows consumers to review products purchased from various sources while earning benefits for their participation. Simultaneously, it empowers brands to gain valuable insights into consumer sentiments and preferences, enabling data-driven decision-making.

At the core of the system is an AI-driven sentiment analysis model, powered by GPT-based Generative AI, integrated using Flask with Python. This model processes and interprets customer reviews, providing a deep understanding of consumer emotions, trends, and preferences. The platform also incorporates the collection of Net Promoter Score (NPS) data through consumer product sharing, offering additional metrics to evaluate customer loyalty and satisfaction.

Future enhancements include real-time sentiment analysis for live feedback, AI-driven recommendations for brands based on sentiment trends, and the expansion to social media sentiment tracking for broader data acquisition. By offering these features, the project aspires to create a comprehensive solution that improves brand strategies, enhances consumer experiences, and fosters stronger brand-consumer relationships.

This innovative approach not only highlights the transformative potential of AI in sentiment analysis but also underscores its significance in modern, consumer-centric industries.

1. INTRODUCTION

Consumer sentiment analysis is a technique used to determine the opinions, emotions, and attitudes of customers regarding products, services, or brands. Using Natural Language Processing (NLP) and machine learning techniques, consumer sentiment analysis processes large volumes of consumer data from social media, reviews, and other text sources to provide valuable insights into consumer behavior. This project will focus on developing a model to analyze consumer sentiment with the goal of supporting better decision-making in marketing and customer service.

The modern consumer landscape is highly dynamic, with customers continuously sharing their experiences and opinions online. The ability to quickly assess these sentiments can provide businesses with a competitive edge by identifying customer satisfaction levels, market trends, and potential areas for improvement. Traditional methods of collecting consumer feedback, such as surveys and questionnaires, are often time-consuming and may not accurately relect the latest customer sentiments. Sentiment analysis, on the other hand, leverages real- time data from diverse sources, enabling companies to respond swiftly to customer needs.

2. LITERATURE REVIEW

Markdown has become an indispensable tool for documenting open-source projects, enabling clear and concise project descriptions. This literature review examines the significance of Markdown in software documentation and its impact on project engagement.

2.1 Development of system software Documentation: The proposed project aims to leverage state-of-the-art NLP and deep learning techniques to build an efficient sentiment analysis model that addresses some of the existing challenges, ultimately providing actionable insights for businesses.

2.2 Lexicon-Based Approaches: Lexicon-based methods use predefined sentiment dictionaries to assign polarity scores to words. Though simple and interpretable, they struggle with context, sarcasm, and negation, often leading to inaccurate sentiment detection in complex or informal

2.3 Machine Learning-Based Methods: Supervised machine learning models like SVM and Naive Bayes use labeled data for sentiment classification. They improve accuracy over rule-based systems but lack deep contextual understanding and require extensive feature engineering and large annotated datasets.

2.4 Deep Learning Techniques: Deep learning models such as CNNs and LSTMs learn semantic patterns directly from raw text. They outperform traditional methods by capturing long-range dependencies but require significant computational resources and large training datasets for optimal performance.

2.5 Transformer-Based Models: Transformers like BERT and RoBERTa revolutionized sentiment analysis with context-aware, bidirectional understanding. These pre-trained models achieve high accuracy and adaptability through fine-tuning, outperforming earlier models on sentiment classification benchmarks.

2.6 Aspect-Based Sentiment Analysis (ABSA): ABSA targets specific aspects within text, identifying sentiments linked to attributes like price or service. It offers more detailed insights than general sentiment analysis, helping businesses pinpoint strengths and weaknesses in products or services.

2.7 Multilingual and Cross-Lingual Sentiment Analysis: Multilingual sentiment analysis addresses non-English feedback using translation, embeddings, or multilingual models like mBERT. It enables broader sentiment understanding across languages but faces challenges with idiomatic expressions and translation accuracy.

2.8 Sentiment Analysis on Social Media: Social media sentiment analysis handles informal, short, and emoji- rich content from platforms like Twitter. It supports real-time opinion tracking and event-based analysis but requires specialized preprocessing and robust models to handle noisy, dynamic language.

2.9 Emerging Trends and Future Directions: Trends include zero-shot learning, multimodal sentiment analysis, and emotion detection. These aim to improve performance with less data, broader input types, and deeper emotional understanding, making sentiment analysis more practical and comprehensive.

2.10 Drawback - Domain Dependency: Many models perform well only within the domain they were trained on (e.g., product reviews), but fail to generalize across different domains like healthcare or finance, limiting the portability of sentiment analysis solutions.

2.11 Real-Time Sentiment Monitoring: Real-time sentiment monitoring is increasingly applied in brand reputation and crisis management. It combines streaming data analytics with NLP models to detect and respond to consumer sentiment trends as they emerge in online platforms.

3. METHODOLOGY

3.1 Sentiment Classification and Evaluation: Trained models classify text into sentiment categories such as positive, negative, or neutral. Performance is measured using metrics like accuracy, precision, recall, and F1- score. Cross-validation and confusion matrix analysis ensure the robustness of the classification

3.2 Feedback Loop and Continuous Improvement: User feedback is integrated into the system to refine models and preprocessing rules. Model retraining and adjustments are periodically conducted to address domain drift and new data patterns, ensuring the analysis remains up-to-date.

3.3 Deployment and Maintenance: The final system is deployed with APIs or web interfaces for real-time or batch sentiment analysis. Regular maintenance is scheduled to update models, improve accuracy, and ensure compatibility with new data formats and industry standards.

3.4 Data Collection and Requirement Analysis: The project begins by identifying data sources such as product reviews, social media posts, and customer feedback. Both structured and unstructured datasets are analyzed to understand domain-specific sentiment expressions. This phase also involves defining project goals and evaluating existing sentiment analysis solutions to establish a baseline.

3.5 Data Preprocessing and Cleaning: Collected data undergoes preprocessing to remove noise, including stop words, special characters, and irrelevant information. Techniques like tokenization, lemmatization, and handling negations or emojis are applied to prepare the data for sentiment modeling. Domain-specific preprocessing rules are formulated for better accuracy.

3.6 Feature Extraction and Representation: Relevant features are extracted using methods such as TF-IDF, word embeddings (Word2Vec, GloVe), or contextual embeddings (BERT). These representations capture both syntactic and semantic characteristics of text, enabling models to better understand consumer opinions.

3.7 Model Selection and Training: Various models are evaluated, including traditional classifiers (Naive Bayes, SVM), deep learning models (LSTM, CNN), and transformer-based architectures (BERT, RoBERTa). Models are trained using annotated datasets, and hyperparameters are fine-tuned to achieve optimal performance.

3.8 User Feedback and Iterative Refinement: Continuous improvements are made based on user feedback, ensuring that documentation remains relevant and user-friendly.

3.9 Ongoing Maintenance: Regular updates are applied to ensure documentation stays aligned with evolving project requirements and industry standards.

4. DISCUSSION

The tool developed in this project aims to simplify the consumer trends and help them understand the brand performance and its presence. This also helps brands to identify the areas of improvements and brand marketing over the platforms.

5. CONCLUSION

This project successfully demonstrates the potential of AI-driven technologies in bridging the gap between consumers and brands. By integrating GPT-based Generative AI with Flask and Python, the platform provides accurate and actionable insights into consumer sentiment and behaviour. The incorporation of features such as real-time sentiment tracking, Net Promoter Score (NPS) collection, and incentivized consumer engagement ensures a comprehensive approach to understanding and leveraging customer feedback.

This system not only empowers brands to make data-driven decisions but also enhances the consumer experience by rewarding their participation and creating a user-centric platform. The innovative combination of AI models, scalable infrastructure, and intuitive design positions the project as a significant step forward in sentiment analysis and market intelligence

6. ACKNOWLEDGMENT

The authors would like to thank the faculty, mentors, and peers for their valuable insights and support during the course of this research. We also acknowledge the availability of open-source datasets and tools that significantly contributed to the successful development and analysis presented in this study.

REFERENCES

1. Esuli, A., & Sebastiani, F. (2006). *SentiWordNet: A publicly available lexical resource for opinion mining*. Proceedings of LREC.
2. *Nielsen, F. Å. (2011). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. arXiv preprint arXiv:1103.2903.*
3. Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). *Linguistic Inquiry and Word Count (LIWC): A text analysis program*.
4. Pang, B., Lee, L., & Vaithyanathan, S. (2002). *Thumbs up? Sentiment classification using machine learning techniques*. Proceedings of EMNLP.
5. Sebastiani, F. (2002). *Machine learning in automated text categorization*. ACM Computing Surveys (CSUR), 34(1), 1–47.
6. Kim, Y. (2014). *Convolutional neural networks for sentence classification*. Proceedings of EMNLP.
7. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of deep bidirectional transformers for language understanding*. Proceedings of NAACL-HLT.
8. Zhang, L., Wang, S., & Liu, B. (2018). *Deep learning for sentiment analysis: A survey*. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4), e1253.
9. Akhtar, M. S., Ekbal, A., & Bhattacharyya, P. (2017). *Aspect based sentiment analysis: A survey*. ACM Computing Surveys (CSUR), 50(3), 45.
10. Rosenthal, S., Farra, N., & Nakov, P. (2017). *SemEval-2017 Task 4: Sentiment Analysis in Twitter*. Proceedings of SemEval.