International Journal of Interdisciplinary Innovative Research & Development (IJIIRD)
ISSN: 2456-236X
Vol. **11** Special Issue **02** | **2026**

CNN–Genetic Algorithm Approach for Customer Review Sentiment Classification

Snehal Dilip Deshmukh[1], Prof. A.S.Narkhede[2]
[1] *Student, Department of Computer Engineering ,Padm.Dr.V.B.Kolte college of engineering, malkapur, Maharashtra, India*
[2] *Assistant Professor, Department of Computer Engineering,Padm. Dr. V. B.Kolte college of engineering, malkapur, Maharashtra ,India*

DOI: 10.5281/zenodo.19288051

ABSTRACT

The rapid growth of e-commerce platforms has led to an enormous volume of user-generated reviews, making it challenging for consumers and businesses to extract meaningful insights manually. Sentiment analysis provides an effective approach to automatically determine opinions and emotional polarity expressed in textual reviews. This study focuses on sentiment analysis of product reviews using natural language processing and machine learning techniques. The proposed system pre-processes review data through tokenization, stop-word removal, and feature extraction to convert unstructured text into meaningful representations. Various classification algorithms are employed to categorize reviews into positive, negative, and neutral sentiments. The performance of the system is evaluated using metrics such as accuracy, precision, recall, and F1-score. Experimental results demonstrate that the proposed approach effectively captures customer sentiment and provides valuable insights into product quality and user satisfaction. The outcome of this work can assist consumers in making informed purchase decisions and help sellers improve products and services based on customer feedback.

Keywords-: Sentiment Analysis, Product Reviews, Natural Language Processing, Machine Learning, Text Classification, Opinion Mining, Customer Feedback Analysis

1. INTRODUCTION

Sentiment Analysis on Product Reviews has emerged as a crucial area of research within the broader field of Natural Language Processing (NLP) and opinion mining, as it focuses on automatically identifying, extracting, and interpreting the emotional tone and subjective information expressed by customers in their written feedback. With the rapid expansion of e-commerce platforms, millions of product reviews are generated daily, providing a rich source of unstructured textual data that reflects customer satisfaction, preferences, expectations, and dissatisfaction. These reviews significantly influence purchasing decisions, brand reputation, and product development strategies, making it essential for businesses to accurately understand consumer sentiment. However, due to the informal nature of user-generated content, which often includes slang, abbreviations, sarcasm, mixed emotions, and grammatical inconsistencies, manual analysis becomes impractical and inefficient. Therefore, automated sentiment analysis systems play a vital role in transforming this vast volume of textual data into meaningful insights by classifying opinions into positive, negative, or neutral categories. By leveraging advanced machine learning and deep learning techniques such as Support Vector Machines, Naïve Bayes, Recurrent Neural Networks, and Transformer-based models, researchers can enhance the accuracy and reliability of sentiment prediction [1]. This analytical approach not only aids companies in monitoring product performance and customer satisfaction in real time but also supports strategic decision-making, competitive analysis, and personalized recommendation systems. Consequently, sentiment analysis on product reviews stands as a powerful tool for extracting actionable intelligence, improving user experience, and driving data-driven business growth in an increasingly digital and customer-centric marketplace [2].

Sentiment analysis, a subfield of natural language processing (NLP), focuses on identifying and classifying opinions expressed in textual data into sentiment categories such as positive, negative, or neutral [2]. By applying sentiment analysis techniques to product reviews, it becomes possible to automatically extract meaningful insights from customer feedback. This enables consumers to make informed purchasing decisions and helps sellers and manufacturers understand market trends, product strengths, and areas requiring improvement [3].

This study focuses on developing a sentiment analysis model for product reviews using NLP and machine learning approaches [4]. The proposed system preprocesses review text, extracts relevant features, and classifies sentiments effectively. By analyzing large-scale review datasets, the system aims to deliver accurate sentiment predictions that support data-driven decision-making for both consumers and businesses [7].

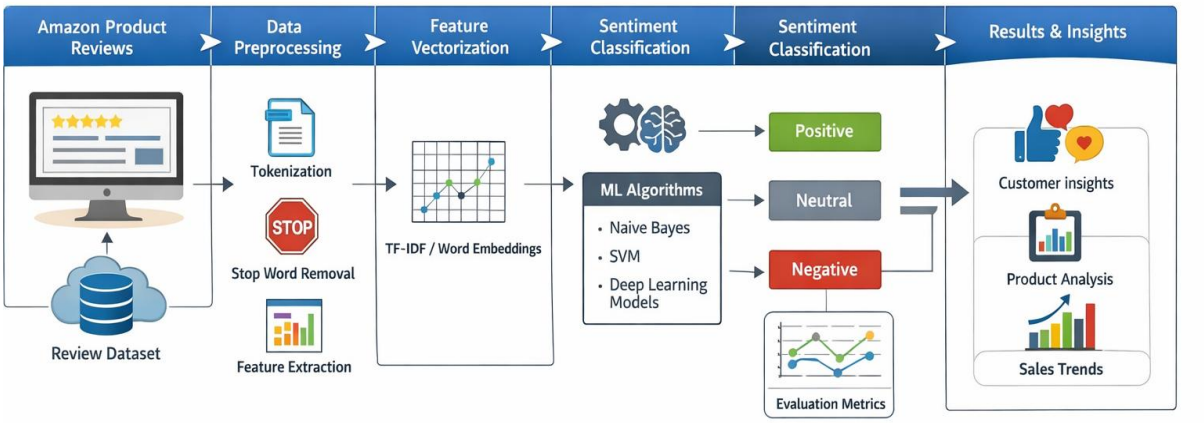

Fig 1: Architecture Design

The architecture for Sentiment Analysis on Product Reviews is designed to automatically extract and classify customer opinions from large-scale textual review data. The process begins with the collection of product reviews from a structured dataset, which serves as the primary input to the system. These raw reviews are first passed through a data preprocessing module where noise is removed by performing tokenization, stop-word removal, normalization, and feature extraction to convert unstructured text into a clean and meaningful format [8]. The preprocessed text is then transformed into numerical representations using feature factorization techniques such as Term Frequency Inverse Document Frequency (TF-IDF) or word embeddings, enabling effective machine learning processing. The resulting feature vectors are fed into the sentiment classification module, where machine learning or deep learning algorithms classify each review into sentiment categories such as positive, negative, or neutral. Model performance is evaluated using standard metrics including accuracy, precision, recall, and F1-score to ensure reliable sentiment prediction. Finally, the classified sentiments are aggregated and analyzed to generate results and insights, providing valuable information for customer feedback analysis, product evaluation, and business decision-making [14].

1.1 Problem Statement

The rapid growth of e-commerce platforms such as Amazon has resulted in an overwhelming volume of customer product reviews, making it difficult for consumers and businesses to analyse and interpret feedback manually [2]. While star ratings provide a quick indication of customer satisfaction, they often fail to capture the true sentiment and detailed opinions expressed in textual reviews. The unstructured nature of review data, presence of noise, and variability in language usage further complicate accurate sentiment interpretation [6]. Existing manual and traditional analysis methods are time-consuming, subjective, and inefficient when dealing with large-scale data. Therefore, there is a need for an automated and intelligent sentiment analysis system that can efficiently process product reviews, accurately classify sentiments, and extract meaningful insights to support informed decision-making for both customers and sellers [9].

1.2. Research Objectives

- To collect and pre-process product review data by removing noise, irrelevant symbols, and inconsistencies to ensure data quality and reliability.
- To design and implement sentiment classification models capable of identifying the emotional polarity of customer reviews.
- To compare the performance of different machine learning and deep learning algorithms for sentiment prediction accuracy.
- To analyse customer opinion trends across different product categories based on review data.
- To identify common positive and negative factors influencing customer satisfaction through sentiment mining.
- To visualize sentiment distribution and patterns for better interpretation of results.
- To evaluate the effectiveness of the proposed model using standard performance metrics such as accuracy, precision, recall, and F1-score.
- To reduce manual effort in analysing large volumes of customer feedback by automating the sentiment analysis process.
- To enhance business strategies by providing actionable insights derived from customer sentiments.
- To contribute to the improvement of product quality and customer experience by identifying key areas of concern expressed in reviews.

1.3 Motivation

The motivation for conducting Sentiment Analysis on Product Reviews arises from the exponential growth of online shopping and the overwhelming volume of customer feedback generated daily, which contains invaluable insights into consumer experiences, expectations, and satisfaction levels. As hosts millions of products across diverse categories, customers rely heavily on reviews to evaluate product quality and make purchasing decisions, while businesses depend on this feedback to monitor performance and improve their offerings. However, manually analyzing such an extensive collection of reviews is highly time-consuming, inefficient, and prone to human bias, making it difficult to extract consistent and reliable conclusions [2]. This creates a strong need for automated systems capable of accurately interpreting customer emotions and opinions at scale. By implementing sentiment analysis techniques, organizations can efficiently identify patterns of customer satisfaction and dissatisfaction, detect emerging trends, understand recurring complaints, and evaluate overall brand perception. This approach not only enhances decision-making for consumers but also enables sellers and manufacturers to refine product designs, enhance service standards, and respond proactively to customer concerns [4]. Furthermore, sentiment analysis supports the development of personalized recommendations and strategic marketing initiatives by aligning products more closely with user preferences. Therefore, the primary motivation behind this study is to transform raw and unstructured textual data into meaningful insights that improve customer experience, strengthen business strategies, and contribute to the advancement of intelligent data-driven solutions in the evolving e-commerce landscape [6].

1.4 Research Gap

- Most existing approaches rely on traditional machine learning models that fail to capture contextual and semantic meaning accurately.
- Limited ability to handle sarcasm, implicit opinions, and complex linguistic expressions in customer reviews.
- Majority of studies focus only on overall sentiment polarity, ignoring aspect-level and sentiment intensity analysis.
- Insufficient handling of noisy, unstructured, and imbalanced review datasets.
- Limited consideration of fake, biased, or spam reviews affecting sentiment accuracy.
- Lack of domain adaptability across different product categories on product.
- Many models are trained on static datasets and do not adapt to evolving language patterns over time.
- Evaluation is often restricted to limited metrics, without real-world deployment analysis.

2.LITERATURE REVIEW

The literature review on Sentiment Analysis of Product Reviews highlights the extensive research carried out in the domains of Natural Language Processing, opinion mining, and machine learning to effectively analyse and interpret customer feedback in e-commerce environments [1]. Previous studies emphasize the growing importance of understanding consumer sentiment as online reviews significantly influence purchasing behavior and brand perception. Researchers have explored a wide range of approaches, including traditional machine learning algorithms such as Naïve Bayes, Support Vector Machines, and Logistic Regression, as well as advanced deep learning models like Convolutional Neural Networks, Long Short-Term Memory networks, and Transformer-based architectures to improve classification accuracy [15]. The reviewed literature also addresses key challenges such as handling noisy textual data, detecting sarcasm, managing class imbalance, and dealing with domain-specific vocabulary commonly found in reviews. Various pre-processing techniques, including tokenization, stemming, lemmatization, stop-word removal, and feature extraction methods like Term Frequency–Inverse Document Frequency (TF-IDF) and word embedding's, have been widely implemented to enhance model performance [13]. Furthermore, existing studies compare the effectiveness of rule-based, lexicon-based, and hybrid approaches, highlighting the evolution from simple polarity detection to more complex emotion and aspect-based sentiment analysis. Overall, the literature demonstrates continuous progress toward building more accurate and scalable sentiment analysis systems, while also identifying limitations such as contextual ambiguity and the need for real-time adaptability, thereby establishing a strong foundation for further exploration and improvement in analyzing product reviews [14].

1. **Sepideh Paknejad *"Sentiment classification on reviews using machine learning approaches"*** The author applies supervised machine-learning methods (Naïve Bayes and SVM) on beauty-product reviews from product. Both models obtain promising accuracies (≈ 80%+), with SVM outperforming Naïve Bayes for larger datasets. This demonstrates the viability of classical ML techniques for product review sentiment classification.
2. **Ravi Kumar Singh & Dr. Kamalraj Ramalingam "Product Review Sentiment Analysis with Machine Learning"** This study processes a large product-review dataset and applies supervised learning (e.g. Logistic Regression, SVM, Decision Trees) to classify sentiment (positive/negative). The work underscores the usefulness of review-based sentiment classification to help prospective buyers.
3. **Roobaea Alroobaea *"Sentiment Analysis of Amazon Product Reviews by"*** The authors employ machine-

learning algorithms with feature-extraction techniques to analyze product reviews. They contribute to empirical knowledge about how different methods perform on real-world review data, helping one understand trade-offs between accuracy, complexity, and practicality.

4. **MF bin Harunasir et al. "Sentiment Analysis of Product Reviews by"** The authors employ machine-learning algorithms with feature-extraction techniques to analyze Amazon reviews. They contribute to empirical knowledge about how different methods perform on real-world review data, helping one understand trade-offs between accuracy, complexity, and practicality.
5. **H. Ali et al.** ***"Analyzing Products Sentiment: A Comparative Study"*** MDPI Electronics, This recent study compares multiple classifiers including transformer-based models and finds that a BERT-based approach yields ~ 89% accuracy, outperforming many classical and earlier deep methods. This suggests that modern transformer models may offer improved performance for Amazon-review sentiment tasks.
6. **Xuechun Li, Xueyao Sun, Zewei Xu, Yifan Zhou** ***"Explainable Sentence-Level Sentiment Analysis for Product Reviews"*** This paper performs sentence-level sentiment classification using a BiLSTM with attention mechanism, and emphasizes model interpretability. The study shows accuracy up to 0.96, and demonstrates that attention weights often highlight aspect terms as much (or more) than generic sentiment words, indicating the importance of aspect-based reasoning.

3. PROPOSDED METHODOLOGY

The proposed methodology for sentiment analysis on product reviews begins with the collection of a large dataset of customer reviews containing review text and associated metadata. Initially, the raw reviews undergo a pre-processing stage to remove noise and inconsistencies through tokenization, lowercasing, stop-word removal, punctuation elimination, and lemmatization [4]. This step ensures that the textual data is clean and suitable for analysis. Next, the pre-processed reviews are transformed into numerical representations using feature extraction techniques such as Term Frequency–Inverse Document Frequency (TF-IDF) or word embedding's to capture important textual patterns [5]. These features are then fed into sentiment classification models based on machine learning or deep learning algorithms to categorize each review into positive, negative, or neutral sentiment classes. The performance of the proposed model is evaluated using standard metrics such as accuracy, precision, recall, and F1-score. Finally, the classified sentiments are aggregated and analysed to generate meaningful insights regarding customer satisfaction, product performance, and user preferences, thereby supporting informed decision-making for both consumers and sellers [7].

4. LIMITATIONS

- Difficulty in detecting sarcasm and implicit sentiments in textual reviews.
- Ambiguity in natural language can lead to misclassification of sentiments.
- Presence of noisy, irrelevant, or incomplete review data affects accuracy.
- Imbalanced datasets with more positive than negative reviews can bias results.
- Fake, spam, or biased reviews may reduce sentiment reliability.
- Domain dependency limits model performance across different product categories.
- Contextual meaning may be lost in traditional feature extraction methods.
- Limited handling of multilingual or mixed-language reviews.
- Sentiment polarity classification may ignore sentiment intensity and aspects.
- Model performance depends heavily on quality of preprocessing and training data.

5. CONCLUSION

The sentiment analysis of product reviews demonstrates the effectiveness of natural language processing and machine learning techniques in extracting meaningful insights from large volumes of unstructured customer feedback [1]. By systematically pre-processing review text, extracting relevant features, and applying sentiment classification models, the proposed approach successfully categorizes reviews into positive, negative, and neutral sentiments. The results highlight how automated sentiment analysis can assist consumers in making informed purchasing decisions and help sellers understand customer preferences, product strengths, and areas for improvement [3]. Although challenges such as language ambiguity and noisy data remain, the study confirms that sentiment analysis provides a scalable and efficient solution for analyzing online reviews, contributing to enhanced customer satisfaction and data-driven decision-making in e-commerce platforms [5].

6. REFERENCES

[1] Sepideh Paknejad *"Sentiment classification on Amazon reviews using machine learning approaches"* — Degree Project, KTH Royal Institute of Technology, Stockholm (2018)

[2] Ravi Kumar Singh & Dr. Kamalraj Ramalingam *"Amazon Product Review Sentiment Analysis with Machine Learning"* — International Journal of Trend in Scientific Research and Development (IJTSRD), Vol. 5, Issue 4, June 2021
[3] Roobaea Alroobaea *"Sentiment Analysis on Amazon Product Reviews using the Recurrent Neural Network (RNN)"* — International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 13, Issue 4, 2022
[4] MF bin Harunasir et al. *"Sentiment Analysis of Amazon Product Reviews by"* 2023
[5] H. Ali et al. *"Analyzing Amazon Products Sentiment: A Comparative Study"* — MDPI Electronics, 2024
[6] Xuechun Li, Xueyao Sun, Zewei Xu, Yifan Zhou *"Explainable Sentence-Level Sentiment Analysis for Amazon Product Reviews"* arXiv preprint, Nov 2021
[7]Pang, B., Lee, L., & Vaithyanathan, S. "Sentiment classification using machine learning techniques" *Proceedings of the ACL Conference*, (2002).
[8]Pang, B., & Lee, L. "Opinion mining and sentiment analysis" *Foundations and Trends® in Information Retrieval*, (2008).
[9]Liu, B. (2012). "Sentiment analysis and opinion mining" Morgan & Claypool Publishers, (2008).
[10]McAuley, J., Pandey, R., & Leskovec, J. "Inferring networks of substitutable and complementary products" *Proceedings of the ACM SIGKDD Conference*, (2015).
[11]Medhat, W., Hassan, A., & Korashy, H. "Sentiment analysis algorithms and applications: A survey" *Ain Shams Engineering Journal*, (2014).
[12]Turney, P. D. "Semantic orientation applied to unsupervised classification of reviews. *Proceedings of ACL*, (2002).
[13]Feldman, R. "Techniques and applications for sentiment analysis" *Communications of the ACM*, (2013).
[14]Maas, A. L., Daly, R. E., Pham, P. T., et al. learning word vectors for sentiment analysis. *Proceedings of the ACL Conference*, (2011).
[15]Kim, Y. Convolutional neural networks for sentence classification. *Proceedings of EMNLP*, (2014).
[16]Ishaq, A., Asghar, S., & Gillani, S. A. "Aspect-Based Sentiment Analysis Using a Hybridized Approach Based on CNN and GA" *IEEE Access,(2020)*
*[17]*George, S. C. G., & Sumathi, B. "Genetic Algorithm Based Hybrid Model of Convolutional Neural Network and Random Forest Classifier for Sentiment Classification" *Turkish Journal of Computer and Mathematics Education, (2021)*
[18] Sravya, G., & Sreedevi, M. (2020). "Genetic Optimization in Hybrid Level Sentiment Analysis for Opinion Classification" *International Journal of Advanced Trends in Computer Science and Engineering (2020).*
[19] Fitria, M. A., & Setiawan, E. B. (2024). "Hybrid Deep Learning with GloVe and Genetic Algorithm for Sentiment Analysis on X: 2024 Election" *Scientific Journal of Informatics,(2024).*
[20] Al-Harbi, O., Hamed, A., & Alzoubi, M."A Deep Neural Network Optimized by a Genetic Algorithm to Improve Arabic Sentiment Classification" *Ingénierie des Systèmes d'Information*, (2023)
[21] Dangi, D., Bhagat, A., & Dixit, D. K. "Sentiment Analysis on Social Media Using Genetic Algorithm with CNN". *Computers, Materials & Continua (CMC)*. (2022)
[22] Fliss, I. (2022). "Toward a New Hybrid Intelligent Sentiment Analysis using CNN-LSTM and Cultural Algorithms" *ICAART,(2022).*
[23] Li, C. (2021). "Genetic Algorithm based hyper-parameters optimization for transfer convolutional neural network" *arXiv preprint. (GA optimization for CNN; useful if you describe GA-based CNN tuning in your methodology)* (2021).
[24] Saidi, W. "Optimal architecture for a sentiment analysis transformer with multihead attention and genetic crossover" *Scientific Reports. (GA crossover used for sentiment-model optimization; good "GA in sentiment" reference),(2025)*
[25] A. Ishaq, S. Asghar, and S. A. Gillani, "Aspect-based sentiment analysis using a hybridized approach based on convolutional neural network and genetic algorithm," *IEEE Access*, vol. 8, pp. 135499–135512, 2020.