

Machine Learning-Based Sentiment Analysis and Predictive Forecasting for E-Commerce Product Reviews

Neeraj Sharma¹, Rahul Misra², Ravi Joshi³

¹School of Computer Application, JECRC University, Jaipur

²Department of Engineering & Technology, Jagannath University, Jaipur

³Department of Computer Science, Apex University, Jaipur

rahul.misra@jagannathuniversity.org, neeraj.sharma@jecrcu.edu.in, ravijoshi93@gmail.com

Abstract: The rapid growth of e-commerce platforms has generated large volumes of customer reviews that provide valuable insights into consumer preferences and market trends. This paper presents a machine learning-based approach for sentiment analysis and forecasting using Amazon product reviews. Various machine learning and deep learning models, including Support Vector Machine (SVM), Naïve Bayes, K-Nearest Neighbors (KNN), Convolutional Neural Networks (CNN), Artificial Neural Networks (ANN), and Bidirectional Long Short-Term Memory (Bi-LSTM), are utilized to classify customer sentiments and identify purchasing patterns. Forecasting techniques are further employed to predict future market trends and customer behavior, supporting informed decision-making in inventory management, marketing, and business strategy. The study demonstrates that integrating sentiment analysis with predictive forecasting enhances business intelligence, improves customer satisfaction, and

enables more effective data-driven decision-making in modern e-commerce environments.

Keywords — Sentiment Analysis, Data Forecasting, Machine Learning, Support Vector Machine, Naïve Bayes, KNN, CNN, Artificial Neural Network, Bi-LSTM.

1 Introduction

In today's digital age, the massive volume of user-generated content on e-commerce platforms has become an invaluable resource for understanding customer preferences, behavior, and satisfaction levels. Among these platforms, Amazon stands out due to its extensive repository of customer reviews, where users express their opinions about products and services. This wealth of data has paved the way for advanced analytical techniques to decipher consumer sentiment and provide actionable insights [1-3].

Sentiment analysis also referred to as opinion mining, plays a crucial role in interpreting the emotional tone of customer reviews. By

analyzing whether reviews convey positive, negative, or neutral sentiments, businesses gain a deeper understanding of how their offerings are perceived. This understanding enables companies to refine their products, tailor their services, and enhance the overall customer experience. On platforms like Amazon, where reviews reflect a diverse range of customer experiences and expectations, sentiment analysis serves as a powerful tool for gauging public opinion [4-5].

By delving into customer reviews, businesses can uncover specific aspects that matter most to consumers, identify areas needing improvement, and detect emerging trends in customer preferences. For instance, a pattern of negative reviews about a product's durability might indicate the need for better materials, while positive feedback about a feature can highlight unique selling points to emphasize in marketing [6-7].

When combined with forecasting techniques, sentiment analysis becomes even more impactful. Forecasting uses historical data to predict future trends in areas such as product demand, sales performance, and customer satisfaction. By integrating sentiment analysis with forecasting, businesses can anticipate market developments and proactively respond to changing consumer needs. This integration allows companies to manage inventory more effectively, optimize

supply chains, and align their strategies with anticipated customer behaviors [8-9].

The application of advanced technologies like machine learning and natural language processing (NLP) has further revolutionized sentiment analysis and forecasting. These tools can process large datasets from platforms like Amazon with remarkable speed and accuracy, identifying nuanced patterns and insights that were previously inaccessible. Machine learning algorithms, for example, can automatically classify reviews by sentiment, while NLP techniques can analyze specific keywords and phrases to detect recurring themes in customer feedback.

The synergy of sentiment analysis and forecasting has made these tools indispensable in the e-commerce sector. By leveraging these technologies, businesses can fine-tune their strategies, enhance decision-making, and maintain a competitive edge in today's data-driven marketplace. Ultimately, this approach empowers companies to stay ahead of market trends, deliver superior customer experiences, and foster long-term loyalty in an ever-evolving digital landscape.

2 Sentiment Analysis

Sentiment analysis, also known as opinion mining, is a natural language processing (NLP) technique used to identify and classify emotions,

opinions, or attitudes expressed in textual data. By categorizing content into sentiment types positive, negative, or neutral this method provides valuable insights into user feelings, preferences, and experiences. It has become increasingly significant in the digital era, as individuals frequently share their thoughts and opinions through various online platforms, including blogs, forums, social media, and product reviews [11-12].

For e-commerce businesses, sentiment analysis serves as a powerful tool to gain a deeper understanding of customer opinions, assess product quality, and make data-driven decisions to enhance their offerings. Customers, too, benefit from sentiment analysis, as it allows them to quickly gauge others' opinions and make informed purchasing decisions. Beyond the realm of e-commerce, sentiment analysis has applications in a wide range of industries, including politics, healthcare, and entertainment, where understanding public opinion is critical for shaping strategies and policies [13].

The field of sentiment analysis has undergone significant advancements, driven by the proliferation of large-scale text data and the emergence of cutting-edge technologies like machine learning and deep learning. Techniques such as Long Short-Term Memory (LSTM) networks, Naive Bayes classifiers, and Support Vector Machines (SVM) have become integral to

processing and analyzing unstructured text data. These methods enhance the accuracy of predictions by leveraging sophisticated approaches like feature selection and vectorization, which extract relevant information and represent text data in structured formats.

Recent innovations, including transfer learning and hybrid models, have further expanded the scope and precision of sentiment analysis. Transfer learning enables models trained on one dataset to be adapted to other related datasets, improving their performance in new contexts. Hybrid models, which combine multiple techniques, offer more nuanced and context-aware sentiment classification by addressing limitations inherent in individual approaches.

Overall, sentiment analysis has evolved into a highly effective tool for interpreting textual data, enabling businesses and individuals to make more informed decisions, anticipate trends, and respond proactively to user sentiments in a variety of fields.

Sentiment analysis, also known as opinion mining, identifies the emotional tone of text and is classified into various types based on its scope and focus. Fine-grained analysis provides detailed sentiment levels, such as very positive, positive, neutral, negative, or very negative, while binary analysis simplifies this by categorizing text as positive or negative. Aspect-based analysis (ABSA) targets specific features

or aspects of products, helping businesses improve. Emotion detection identifies emotions like joy, anger, or sadness, aiding customer service and mental health evaluations. Multilingual analysis processes text in multiple languages, supporting global businesses, and intent-based analysis identifies the purpose behind text, such as complaints or suggestions, enhancing customer support systems. Contextual analysis considers the context to interpret complex sentiments, while neutral analysis focuses on balanced statements often found in news or formal content. Comparative analysis evaluates sentiments in comparisons, such as "Product A is better than Product B," providing insights into competition. Visual and multimodal analysis examines emotions expressed in text, images, or videos, offering deeper insights into user attitudes, particularly in social media and marketing. Each type addresses specific needs, chosen based on the analysis goals and context.

3 Data Forecasting

Data forecasting is a vital area of study with applications across various industries, such as healthcare, finance, and business. At its core, forecasting involves predicting future values or outcomes based on historical and current data. These predictions, while not always perfectly accurate, provide an informed estimate of what could happen in the future. As a result, forecasts

serve as valuable tools for planning and decision-making in many fields.

One common example of data forecasting is weather prediction. By analyzing past weather patterns and trends, meteorologists can estimate future weather conditions. Similarly, in healthcare, forecasting can help identify patients who might be at risk of developing certain diseases. Early detection through forecasting enables timely interventions, improving patient outcomes and reducing healthcare costs.

Forecasting also plays a critical role in business operations. Companies analyze past sales data, customer demand, and potential risks to predict future revenue trends and market behavior. This helps businesses make strategic decisions, allocate resources efficiently, and plan budgets effectively. For instance, a company expecting higher sales in the coming months can use this information to increase production, hire additional staff, or adjust inventory levels.

Investors and stock market analysts also rely heavily on forecasting. Businesses with expected revenue growth may see a rise in stock prices, influencing investment decisions. On a broader scale, forecasting techniques are used to predict economic indicators such as GDP growth, unemployment rates, or inflation. These predictions help analysts understand how these trends might impact financial markets and long-term economic stability.

In research and statistical analysis, forecasting is an essential tool for making predictions about future events. It helps decision-makers across various fields to identify patterns, understand risks, and develop strategic plans. Whether it's guiding a company's growth, analyzing stock market trends, or improving healthcare outcomes, forecasting empowers individuals and organizations to make well-informed, data-driven decisions. By leveraging insights from historical data, forecasting reduces uncertainty and enhances precision, ensuring better preparedness for the future.

4 Sentiment Analysis and Forecasting Using Machine Learning

Chauhan et al. (2024) analyzed Amazon customer reviews using Support Vector Machine (SVM) and Naïve Bayes classifiers to categorize sentiments into positive, negative, and neutral classes. Their models achieved accuracies of 76% and 80%, respectively, demonstrating the usefulness of machine learning in customer feedback analysis and business decision-making.

Sridhar et al. (2024) proposed a GRCA-SBFNN-based sentiment analysis framework for reviews collected from Facebook, Twitter, and Amazon. Their model achieved 98% accuracy with high precision and recall, highlighting the effectiveness of advanced neural network architectures for opinion mining.

Chhabra et al. (2024) employed a Bidirectional Long Short-Term Memory (Bi-LSTM) network for Amazon review classification. Using text tokenization and encoding techniques, the proposed model achieved an accuracy of 93.92%, outperforming several conventional machine learning approaches.

Byali et al. (2024) presented a hybrid sentiment analysis framework that combined machine learning and rule-based techniques for text mining. The study demonstrated that Decision Tree models and Hadoop-based processing effectively handle large-scale Amazon review datasets with improved computational efficiency.

Liashenko et al. (2024) integrated sentiment analysis with machine learning for Tesla stock price forecasting. Their Multivariate Regression model significantly reduced prediction errors by incorporating investor sentiment, demonstrating the importance of sentiment-aware forecasting.

Deepa et al. (2023) developed a Novel Convolutional Neural Network (N-CNN) for emotion detection in Amazon product reviews. The proposed model achieved an accuracy of 98.3%, illustrating the effectiveness of deep learning for customer sentiment classification.

Kumar et al. (2023) analyzed more than 80,000 Amazon product reviews using machine learning algorithms to classify customer sentiments and support data-driven business decisions through accurate opinion prediction.

Kausar et al. (2023) compared Decision Tree and Logistic Regression models for Amazon review classification. The Decision Tree classifier achieved 99% accuracy, while word cloud visualization provided additional insights into customer preferences and frequently used terms.

Sangeetha et al. (2023) introduced a Pearson Correlation-based Harris Hawks Optimization with an RNN-LSTM model for sentiment analysis. The proposed framework achieved 95.8% accuracy, demonstrating strong performance on large-scale review datasets.

Santhosh et al. (2023) investigated machine learning-based sentiment analysis for public health monitoring using social media data. Their study showed that classifiers such as Naïve Bayes can effectively identify public sentiment, enabling faster and more cost-effective epidemic surveillance.

5 Conclusions

This study demonstrates the significant potential of combining sentiment analysis with forecasting techniques, particularly in the context of e-commerce platforms like Amazon. By leveraging machine learning models such as SVM, Naïve Bayes, and deep learning architectures like LSTM networks, businesses can gain valuable insights into customer sentiment, leading to improved product offerings

and customer satisfaction. Forecasting models further enhance this analysis by predicting future trends and customer behavior, which is essential for strategic planning and resource optimization. The research also highlights the importance of using hybrid models and advanced processing frameworks like Hadoop for efficient analysis of large datasets. As machine learning and natural language processing continue to evolve, the integration of sentiment analysis and forecasting will remain a key tool for businesses seeking to stay ahead in the competitive digital marketplace. Ultimately, this work illustrates how data-driven insights derived from customer feedback can empower companies to refine their strategies, foster customer loyalty, and achieve long-term success.

References

- [1] M. S. Meghana, D. Abhijith, S. Aysha and P. K. Kollu, "Sentiment Analysis on Amazon Product Reviews using LSTM and Naive Bayes," IEEE 7th International Conference on Computing Methodologies and Communication (ICCMC), pp. 626-631, 2023.
- [2] H. Kumar, A. Kumar, A. Anand, A. Sabu and T. Jain, "Sentiment Analysis on Amazon Electronics Product Reviews using Machine Learning Techniques," IEEE 4th IEEE

- Global Conference for Advancement in Technology (GCAT), pp. 1-6, 2023.
- [3] P. Juyal and A. Kundaliya, "A Comparative Study of Hybrid Deep Sentimental Analysis Learning Techniques with CNN and SVM," IEEE World Conference on Applied Intelligence and Computing (AIC), pp. 596-600, 2023.
- [4] P. Gaur, S. Vashistha and P. Jha, "Twitter Sentiment Analysis Using Naive Bayes-Based Machine Learning Technique", Sentiment Analysis and Deep Learning. Advances in Intelligent Systems and Computing, pp. 367-376, 2023.
- [5] V. Joshi, S. Patel, R. Agarwal and H. Arora, "Sentiments Analysis using Machine Learning Algorithms," 2023 Second International Conference on Electronics and Renewable Systems (ICEARS), pp. 1425-1429, 2023.
- [6] H. Arora, M. Kumar, T. Rasool and P. Panchal, "Facial and Emotional Identification using Artificial Intelligence", IEEE 6th International Conference on Trends in Electronics and Informatics (ICOEI), pp. 1025-1030, 2022.
- [7] H. Arora, G. K. Soni, R. K. Kushwaha and P. Prasoon, "Digital Image Security Based on the Hybrid Model of Image Hiding and Encryption", 2021 6th International Conference on Communication and Electronics Systems (ICCES), pp. 1153-1157, 2021.
- [8] P. Shi, S. Song, Q. Lang and X. Liu, "Futures Price Forecasting Based on LSTM Model Using Public Opinion Sentiment Analysis," 2023 42nd Chinese Control Conference (CCC), pp. 8683-8689, 2023.
- [9] Z. Zhou, M. Gao, Q. Liu and H. Xiao, "Forecasting stock price movements with multiple data sources: Evidence from stock market in China." Physica A-statistical Mechanics and Its Applications 542, 2020.
- [10] G.K. Soni, A. Rawat, S. Jain and S.K. Sharma, "A Pixel-Based Digital Medical Images Protection Using Genetic Algorithm with LSB Watermark Technique". Smart Systems and IoT: Innovations in Computing" in Smart Innovation Systems and Technologies, Singapore:Springer, vol. 141, 2020.
- [11] J. C. Gope, T. Tabassum, M. M. Mabrur, K. Yu and M. Arifuzzaman, "Sentiment Analysis of Amazon Product Reviews Using Machine Learning and Deep Learning Models," IEEE International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE), pp. 1-6, 2022.
- [12] S. Mishra, N. K. Tiwari, K. Kumari and V. Kumawat, "Prediction of Heart Disease using Machine Learning," 2023 2nd International Conference on Applied

- Artificial Intelligence and Computing (ICAAIC), Salem, India, 2023, pp. 1617-1621
- [13] H. Arora, P. Kumar Sharma, K. Mitanshi and A. Choursia, "Enhanced Security of Digital Picture and Text Information with Hybrid Model of Hiding and Encryption Techniques," 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), Erode, India, 2022, pp. 1238-1241, doi: 10.1109/ICSCDS53736.2022.9760822.
- [14] H. Arora, R. Agarwal, P. Sharma, G. Shankar and D. Arora, "Image Security Utilizing Hybrid Model of Steganography and Asymmetric Cryptography Methods," 2023 International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT), Bengaluru, India, 2023, pp. 277-280,
- [15] Shilpi Mishra, Naveen Kumar Tiwari, Kashish Adlakha, Harsh Jain and Mehul Vardiya, "Machine Learning in Banking Products", *Annals of the Romanian Society for Cell Biology*, vol. 25, no. 6, pp. 11750-11755, 2021.
- [16] Vipin Singh, Manish Choubisa and Gaurav Kumar Soni, "Enhanced Image Steganography Technique for Hiding Multiple Images in an Image Using LSB Technique", *TEST Engineering & Management*, vol. 83, pp. 30561-30565, May-June 2020.
- [17] G. K. Soni, D. Yadav, A. Kumar and L. Sharma, "Flexible Antenna Design for Wearable IoT Devices," 2023 3rd International Conference on Technological Advancements in Computational Sciences (ICTACS), Tashkent, Uzbekistan, 2023, pp. 863-867,
- [18] G. Chauhan, A. Sharma and N. Dwivedi, "Amazon Product Reviews Sentimental Analysis using Machine Learning," 2024 IEEE International Conference on Computing, Power and Communication Technologies (IC2PCT), pp. 1202-1206, 2024.
- [19] A. S. Sridhar and S. Nagasundaram, "Product Reviews Based Sentimental Analysis by Feature Extraction and Classification Using Machine Learning Algorithm," 2024 5th International Conference on Electronics and Sustainable Communication Systems (ICESC), pp. 891-897, 2024.
- [20] D. Chhabra, S. Sah, R. Rani, N. Bisla, A. Tomar and A. Sharma, "Deep Learning-based Sentiment Analysis of Amazon Product Reviews," 2024 1st International Conference on Trends in Engineering Systems and Technologies (ICTEST), pp. 1-7, 2024.

- [21] M. S. Byali, S. Kanakaraddi, T. Badagi and A. Chikaraddi, "Sentimental Analysis of Amazon Reviews using Map Reducer Technique in Hadoop," 2024 11th International Conference on Computing for Sustainable Global Development (INDIACom), pp. 961-967, 2024.
- [22] O. Liashenko, T. Kravets and V. Plushchov, "Stock Price Forecasting using Sentiment Analysis of Stock Tweets," IEEE 14th International Conference on Advanced Computer Information Technologies (ACIT), pp. 351-354, 2024.
- [23] N. Deepa, J. S. Priya and T. Devi, "Sentimental Analysis Recognition in Customer review using Novel-CNN," IEEE International Conference on Computer Communication and Informatics (ICCCI), pp. 1-4, 2023.
- [24] K. Kumar, S. K. Singh, A. Kumar and A. S. Singh, "Sentimental Analysis of Customer's Reviews Using Machine Learning," IEEE 5th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), pp. 326-330, 2023.
- [25] M. A. Kausar, S. O. Fageeri, A. Soosaimanickam, "Sentiment Classification based on Machine Learning Approaches in Amazon Product Reviews", Engineering, Technology & Applied Science Research, Vol. 13, No. 3, pp. 10849-10855, 2023.
- [26] Sangeetha J, Kumaran U, "Sentiment analysis of amazon user reviews using a hybrid approach", Measurement: Sensors, Vol. 27, pp. 1-11, 2023.
- [27] R. Santhosh, S. V. Vignesh, E. Rithish and P. D. Mahendhiran, "Sentimental Analysis on Amazon Camera Reviews using Naive Bayes Algorithm," 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS), pp. 1-8, 2023.
- [28] A. Upadhyay, R. Misra, S. K. Henge, Y. Bhardwaj, "Protection of Digital Image and Text Information Security Using LSB and Crossover Techniques", Computational Vision and Bio-Inspired Computing. Advances in Intelligent Systems and Computing, vol 1439. 2023.
- [29] A. Kumar, N. Gupta, R. Misra, S. Sharma, D. Chaudhary, G. Sharma, "Deep learning based highway vehicles detection and counting system using computer vision", Journal of Information & Optimization Sciences, Vol. 44, Issue. 5, pp. 997-1008, 2023.
- [30] S. Kumari, S. Sharma, A. Johari, G. K. Soni, H. Kaushik and S. Jain, "Twitter Sentiment Analysis using Machine Learning Techniques," 2026 6th International Conference on Expert Clouds and Applications (ICOECA), pp. 1711-1717, 2026.

- [31] Aayushi, L. Yadav, P. Paliwal, A. Johari, R. Ajmera and G. K. Soni, "A Simulation-based Evaluation of Machine Learning Models for Algorithmic Trading in Equity Markets," 2026 6th International Conference on Expert Clouds and Applications (ICOECA), pp. 1049-1054, 2026.
- [32] K. Gautam, G. K. Soni, R. Ajmera, N. Hemrajani, J. Ahuja and M. K. Jha, "Deep Reinforcement Learning for Stock Market Portfolio Optimization," 2026 5th International Conference on Communication, Computing and Electronics Systems (ICCCES), pp. 1835-1839, 2026.
- [33] A. Jangir, A. Agrawal, C. Sharma, G. K. Soni, R. Ajmera and A. Johari, "Comparative Performance Analysis of Deep Learning and Traditional Algorithms for Facial Recognition and Image Classification," 2025 4th International Conference on Automation, Computing and Renewable Systems (ICACRS), pp. 1172-1175, 2025.
- [34] D. Saxena, J. Sharma, G. K. Soni, Y. Rao, S. Sharma and S. Lavania, "Sentimental Analysis and Forecasting using Machine Learning Algorithms," 2025 4th International Conference on Automation, Computing and Renewable Systems (ICACRS), pp. 917-921, 2025.
- [35] S. A. Saiyed, N. Sharma, H. Kaushik, P. Jain, G. K. Soni and R. Joshi, "Transforming portfolio management with AI and ML: shaping investor perceptions and the future of the Indian investment sector," Parul University International Conference on Engineering and Technology 2025 (PiCET 2025), pp. 1108-1114, 2025.