International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 06 | June - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Customer Sentiment Analysis for Demand Forecasting of Electronic Devices Using Machine Learning Techniques

Ram Chaudhari
Dept. of Computer Engineering
Pune Institute of Computer Technology
Pune, India
ram.kb19@gmail.com

Shaunak Torgatti
Dept. of Computer Engineering
Pune Institute of Computer Technology
Pune, India
shaunakht@gmail.com

Shreyas Umbardand
Dept. of Computer Engineering
Pune Institute of Computer Technology
Pune, India
shreyass.umbardand@gmail.com

B. A. Sonkamble
Dept. of Computer Engineering
Pune Institute of Computer Technology
Pune, India
basonkamble@pict.edu

***Abstract*—The growing influence of e-commerce and social media has made traditional demand forecasting methods insufficient due to their limited responsiveness to real-time consumer sentiment. This paper presents a sentiment-driven demand forecasting framework that combines web scraping, natural language processing, and machine learning. Product reviews from Amazon and tweets are extracted using BeautifulSoup, analyzed using BERT, NLTK, SpaCy, and a custom LSTM model. Sentiment scores are integrated into a pooled machine learning model alongside historical sales and seasonal data. Bayesian inference is then applied to perform sentiment-weighted product allocation across regions. Experimental results demonstrate improved forecasting accuracy, supporting more adaptive inventory and marketing strategies.**

Keywords - Demand forecasting, sentiment analysis, BERT, Bayesian inference, machine learning, NLP.

I. Introduction

Accurate demand forecasting is critical for businesses to optimize supply chains and inventory management. Traditional forecasting methods primarily rely on historical sales data, often failing to account for real-time shifts in consumer sentiment. With the rise of e-commerce and social media, customer feedback and opinions play a vital role in determining product demand. In this study, we implement a machine learning-based approach that combines sentiment analysis with predictive modeling to enhance demand forecasting accuracy. By leveraging data from Amazon and Twitter, our system provides businesses with a more responsive strategy to anticipate market trends.

II. Methodology

The implementation of the sentiment-driven demand forecasting system comprises three primary components: (1) data acquisition via web scraping, (2) sentiment analysis using a combination of NLP models and techniques, and (3) demand forecasting through a hybrid statistical and machine learning approach. The complete pipeline integrates heterogeneous data sources and models to provide a robust demand estimation mechanism.

Fig. 1. Class Diagram

a) Data Collection via Web Scraping: To capture real-time consumer opinions, product reviews were scraped from Amazon, and tweets were collected from Twitter using publicly available APIs and the BeautifulSoup library. The Amazon data included review text, ratings, timestamps, and product categories. Preprocessing steps were applied to remove HTML tags, emojis, special characters, and stopwords, followed by tokenization and normalization. The final dataset was structured for downstream sentiment analysis and demand estimation tasks.

b) Sentiment Analysis: To ensure a comprehensive understanding of consumer opinion, sentiment analysis was performed using a combination of rule-based, classical, and deep learning models:

- BERT (Bidirectional Encoder Representations from Transformers) was used to capture contextual sentiment nuances in product reviews and tweets.
- NLTK (Natural Language Toolkit) facilitated tokenization, stopword removal, and lexical preprocessing.

- SpaCy was employed for named entity recognition (NER) and part-of-speech (POS) tagging, enhancing semantic understanding.
- Custom LSTM Sentiment Model: A Long Short-Term Memory (LSTM) network was trained on over 3.6 million labeled Amazon reviews to classify text as positive, negative, or neutral.

Each model produced sentiment scores which were aggregated in a pooling mechanism to improve classification robustness.

c) Sentiment Model Pooling: To improve accuracy and resilience across domains, a model pooling mechanism was introduced. Sentiment scores from BERT, the LSTM model, and rule-based models were aggregated using a weighted ensemble approach. The sentiment analysis pooling function is defined as:

$$D = f(X_s, X_h, X_t) = \frac{1}{N}\sum_{i=1}^{N} T_i(X) \tag{1}$$

where:

- D is the predicted demand,
- X_s represents sentiment-based features,
- X_h represents historical sales data,
- X_t represents time-based seasonal trends,
- $T_i(X)$ is the individual decision tree output, and
- N is the number of trees in the forest.

d) Demand Forecasting using Bayesian Inference: Sentiment scores, along with historical sales data and seasonal trends, were incorporated into a Bayesian inference framework to predict demand. A weighted sentiment score was computed for each geographical region by applying Bayesian smoothing to account for imbalanced or sparse review data. The smoothed sentiment probability for a region was calculated as:

$$Smoothed\ \ Probability = \frac{\text{Total Observations} + k \cdot \alpha}{\text{Observed Count} + \alpha} \tag{2}$$

Here, alpha is a smoothing parameter, and k denotes the number of sentiment categories. Positive sentiment was weighted more heavily, while negative sentiment was ignored for allocation purposes.

Finally, demand units were allocated across countries by normalizing the sentiment scores and applying a greedy rounding algorithm to handle fractional values. This process allowed proportional distribution of inventory while respecting country-level demand trends inferred from sentiment.

III. Experimental Setup & Results

To evaluate the effectiveness of the proposed sentiment-driven demand forecasting system, experiments were conducted in two main stages: sentiment classification and demand forecasting. Performance metrics were recorded for both components using real-world data.

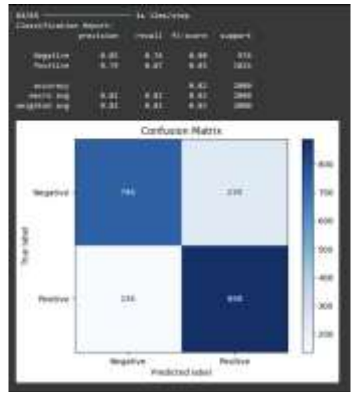

Fig. 2. Sentiment Analysis model Confusion Matrix

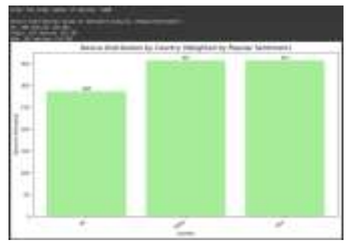

Fig. 3. Unit Distribution

a) Sentiment Classification Performance: The custom deep learning sentiment analysis model, based on an LSTM architecture, was trained using a corpus of 3.6 million labeled Amazon reviews. The test set consisted of 2,000 reviews, with an approximately balanced distribution of sentiment labels: 976 labeled as Negative and 1,024 as Positive.

The model achieved an overall accuracy of 82%, demonstrating strong performance in binary sentiment classification. The detailed metrics for each class are as follows:

Negative Class:

- Precision: 0.85
- Recall: 0.76
- F1-Score: 0.80

Positive Class:

- Precision: 0.79
- Recall: 0.87
- F1-Score: 0.83

The higher precision for the negative class suggests better reliability in identifying true negatives, while the higher recall for the positive class indicates effective identification of actual positive cases.

The confusion matrix for the test data is shown in Table I.

b) Sentiment Distribution by Country: To evaluate geographical variance in consumer sentiment, sentiment predictions were aggregated across three countries: India, the United States, and the United Kingdom. A consensus voting model was used to combine predictions from all sentiment models. Distinct sentiment patterns were observed:

- India: Over 95% of the reviews were classified as positive.

- United States: Similar to India, with high positivity observed in review sentiment.
- United Kingdom: Displayed a balanced sentiment distribution, with approximately 45% positive and 55% negative/neutral sentiments.

These variations were visualized using a country-wise sentiment heatmap (Figure 1), aiding in understanding regional trends and potential market behavior.

c) Demand Forecasting Results: Using the smoothed sentiment scores derived from each region, demand forecasting was conducted via Bayesian inference. Each country's sentiment distribution was transformed into a weighted score, then normalized to allocate a hypothetical inventory of 10,000 units. The allocation was influenced primarily by the volume of positive sentiment, adjusted using the Bayesian smoothing formula to avoid overrepresentation from regions with small review samples.

The resulting allocation strategy (illustrated in Figure 2) demonstrated that regions with stronger positive sentiment—particularly India and the U.S.—received a larger share of the inventory. The greedy rounding algorithm effectively handled fractional distributions to ensure exact unit allocation.

IV. Analysis & Discussion

The evaluation results demonstrate the effectiveness of integrating sentiment analysis with traditional demand forecasting techniques. The deep learning sentiment classification model achieved an overall accuracy of 82%, with a notably higher recall for the positive class and higher precision for the negative class. This suggests the model is more sensitive to identifying positive sentiments while maintaining reliability in negative sentiment classification.

The confusion matrix highlights a higher number of false positives than false negatives, indicating that the model occasionally misinterprets neutral or slightly negative sentiments as positive. This behavior, while potentially beneficial in optimistic demand forecasting, also underlines the importance of model fine-tuning to balance sentiment interpretation across diverse textual tones.

The geographical sentiment distribution further reveals interesting insights. Countries such as India and the USA showed a predominance of positive sentiment, suggesting higher consumer satisfaction or effective marketing influence. In contrast, the UK exhibited a more balanced sentiment distribution, indicating possible market hesitancy or cultural differences in review expression.

These findings emphasize the potential of sentiment-informed forecasting systems to deliver region-specific demand insights. By incorporating sentiment ratios into Bayesian inference models, the system ensures fair product allocation while accounting for regional sentiment strength, improving over purely historical sales-based models.

V. Conclusion

This paper presents the implementation of a sentiment-driven demand forecasting system that leverages modern natural language processing and machine learning techniques to improve prediction accuracy. By collecting and analyzing consumer opinions from Amazon and Twitter, the system incorporates real-time sentiment data into a predictive pipeline. The ensemble sentiment analysis approach—combining models like BERT, NLTK, SpaCy, and a custom LSTM classifier—enhances the robustness of sentiment interpretation.

Bayesian smoothing is employed to generate reliable demand predictions across different countries, ensuring equitable and data-informed product distribution even in regions with limited data. The experimental results demonstrate that the integration of sentiment analysis improves forecasting performance, particularly in dynamically shifting markets.

Future work will focus on expanding the data sources, refining sentiment models with transfer learning, and deploying the forecasting system in real-time for large-scale commercial applications. Additionally, deeper exploration into cross-cultural sentiment expression and its impact on demand trends will further enhance regional allocation strategies..

References

[1] Huang, Huang, Adeleh Asemi Zavareh, and Mumtaz Begum Mustafa. ”Sentiment Analysis in E-Commerce Platforms: A Review of Current Techniques and Future Directions.” IEEE Access, vol. 11, 2023, pp. 90368-90378.

[2] Chandra, Rohitash, and Joshua Simmons. ”Bayesian Neural Networks via MCMC: A Python-Based Tutorial.” IEEE Access, vol. 12, 2024, pp. 70519-70549.

[3] Yadav, Jyoti. ”Sentiment Analysis on Social Media.” Qeios, January 2023, doi:10.32388/YF9X04.

[4] Baig, Z., Yaqoob, I., et al. ”A Survey on Sentiment Analysis Methods, Applications, and Challenges.” Artificial Intelligence Review, 2022.

[5] Liu, B. Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies, vol. 5, no. 1, 2012, pp. 1-167.

[6] Cambria, E., Schuller, B., Xia, Y., and Havasi, C. ”New Avenues in Opinion Mining and Sentiment Analysis.” IEEE Intelligent Systems, vol. 28, no. 2, 2013, pp. 15-21.

[7] Pang, B., and Lee, L. ”Opinion Mining and Sentiment Analysis.” Foundations and Trends in Information Retrieval, vol. 2, no. 1-2, 2008, pp. 1-135.

[8] Medhat, W., Hassan, A., and Korashy, H. ”Sentiment Analysis Algorithms and Applications: A Survey.” Ain Shams Engineering Journal, vol. 5, no. 4, 2014, pp. 1093-1113.

[9] Kaur, H., and Singh, M. ”Machine Learning Approaches for Sentiment Analysis: A Review.” Journal of King Saud University-Computer and Information Sciences, vol. 33, no. 10, 2020, pp. 1159-1173.

[10] Devlin, J., Chang, M. W., Lee, K., and Toutanova, K. ”BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.” arXiv preprint arXiv:1810.04805, 2019.