e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 11, November 2024

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.521

© 2024 IJMRSET | Volume 7, Issue 11, November 2024| DOI: 10.15680/IJMRSET.2024.0711084

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Opinion Mining for Product Reviews using NLP

G. Sandeep, M. Sandeep, M. Sanjana, A. Sanjana, R. Sanjay, G. Balaiah

Department of AI & ML, Malla Reddy University, Hyderabad, India

ABSTRACT: Due to the massive volume of customer reviews brought about by the exponential expansion of e-commerce, it is now difficult for firms to manually extract insightful information. Natural language processing (NLP)-powered opinion mining provides a practical solution by automating sentiment analysis for assessing consumer comments. Using natural language processing (NLP) techniques, this study suggests a system for opinion mining on product evaluations that may categorise attitudes into three groups: neutral, negative, and positive. The system uses a multi-step procedure to prepare textual material for analysis, starting with preparation that includes lemmatisation, tokenisation, and stop-word removal. Word embeddings such as Word2Vec and BERT, along with Term Frequency-Inverse Document Frequency (TF-IDF), are used to extract features. Deep learning techniques like Long Short-Term Memory (LSTM) networks and machine learning models like Support Vector Machines (SVM) and Logistic Regression are used for classification. The performance of these models is demonstrated experimentally using a publically accessible dataset, where LSTM outperforms conventional techniques in terms of accuracy and recall. The system's capacity to provide useful insights is demonstrated via visualisation approaches like sentiment distribution graphs and confusion matrices. The results indicate that the suggested system works well for evaluating unstructured review data, giving companies insightful input to improve their goods and client pleasure. Even with its effectiveness, there are still issues with managing sarcasm and multifaceted emotions. In order to enhance contextual knowledge, future research will concentrate on integrating domain- specific ontologies, supporting different languages, and putting aspect-based sentiment analysis into practice.This study advances NLP-driven opinion mining technologies, providing scalable solutions for contemporary enterprises.

KEYWORDS: Opinion Mining, Natural Language Processing (NLP), Sentiment Analysis, Machine Learning, Word2Vec, BERT, E- Commerce.

I. INTRODUCTION

The amount of online customer reviews has skyrocketed due to the growth of e-commerce platforms, and these reviews now significantly influence what consumers decide to buy. Every day, millions of product reviews are created, and companies are searching more and more for methods to glean insightful information from this massive volume of unstructured data. Improving goods, services, and general customer pleasure requires an understanding of consumer sentiment. However, analysing customer reviews by hand takes a lot of time and is prone to human mistake. Therefore, automated systems that can effectively analyse and analyse massive amounts of textual data are becoming more and more necessary.

One important way to overcome the difficulties of opinion mining in product reviews is through the use of Natural Language Processing (NLP). Customer evaluations can be analysed in an organised way and insights about the sentiments represented in the text can be extracted by employing natural language processing (NLP) techniques. One of the most significant uses of natural language processing (NLP) is sentiment analysis, which divides text into predetermined sentiment categories like positive, negative, or neutral. Despite its achievements, ambiguity, sarcasm, and contextual subtleties are only a few of the linguistic complexity that current sentiment analysis systems frequently struggle to handle. Achieving precise and significant conclusions is challenging because of these issues, particularly when examining big datasets.

The goal of this project is to create a system that mines product evaluations for opinions using sophisticated natural language processing techniques. In order to clean and get text data ready for analysis, the system focusses on preparing it using lemmatisation, tokenisation, and stop- word removal procedures. Word embeddings like Word2Vec and BERT, which aid in capturing the semantic meaning of words and their contextual links, and techniques like Term Frequency-Inverse Document Frequency (TF-

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IDF) are used in feature extraction. To categorise reviews into sentiment groups, the system makes use of deep learning models including Long Short-Term Memory (LSTM) networks and machine learning methods like Support Vector Machines (SVM) and Logistic Regression.

In order to achieve good performance in terms of accuracy, precision, recall, and F1-score, the suggested system will be tested on publicly accessible product review datasets. Businesses may improve product offerings and the entire customer experience by automating sentiment analysis, which allows them to extract meaningful insights from consumer comments. Additionally, the study seeks to offer a flexible and scalable solution, with the possibility of extending the system in the future to incorporate aspect- based sentiment analysis and multilingual support, enabling worldwide use of the sentiment analysis applications.

II. PROBLEM STATEMENT

2.1 Existing System:
In order to classify sentiment, early opinion mining systems used rule-based techniques that involved keyword matching and predefined lexicons. Although these systems worked well for straightforward tasks, they had scalability issues and were unable to recognise subtleties in context, such irony and sarcasm. Using feature extraction techniques like TF-IDF, traditional machine learning models like Naive Bayes and Support Vector Machines (SVM) enhanced performance. These models still had trouble comprehending complex contextual interactions, though. By capturing word relationships and contextual meaning, deep learning approaches like Transformer-based models like BERT and Long Short-Term Memory (LSTM) networks have significantly improved sentiment analysis. Notwithstanding these developments, problems with domain-specific language, ambiguity, and multi-aspect mood still pose serious obstacles to product review analysis.

2.2 Limitations:
Opinion mining has advanced, but there are still a number of issues with the systems that are in use today. Rule-based methods are not flexible or scalable, and they frequently need to be updated by hand to accommodate new or unclear terminology. Even though accuracy has increased, traditional machine learning models are still unable to fully grasp the nuanced contextual information and intricate semantic linkages seen in product reviews, such as sarcasm or conflicting emotions. Although deep learning models like LSTM and BERT have made great progress in capturing contextual meaning, they are computationally costly and require big, labelled datasets for training. Accurately classifying sentiments across a variety of product categories and review kinds is further limited by the difficulty in managing domain-specific language, multi- aspect sentiment, and fine-grained opinions.

1. Lack of Flexibility in Rule-Based Systems: Rule-based systems' scalability is limited by their need for manual updates and their inability to adjust to new or unclear words without any professional assistance or help from experts.
2. Problems with Semantic Understanding: Although they are good at extracting features, traditional machine learning models have trouble identifying complex semantic links and contextual subtleties in product reviews, including sarcasm or conflicting emotions like expressing opposite to what one thinks.
3. Dependency on Large Labelled Datasets: Despite their strength, deep learning models such as LSTM and BERT need to be trained on large, annotated datasets, which can be costly and challenging to acquire.
4. High Computational Costs: Deep learning models are less effective for real-time processing of large-scale review datasets since they demand a substantial amount of computational power.
5. Managing Difficulties Domain-Specific Language: Existing models still struggle with domain-specific terminology, jargon, and multi-aspect sentiment, which compromises the precision and accuracy of sentiment categorisation across a range of product categories.

2.3 Proposed System :
By utilising cutting-edge Natural Language Processing (NLP) techniques and machine learning models, the suggested solution seeks to overcome the shortcomings of current opinion mining approaches and provide more precise sentiment analysis of product reviews. To standardise and clean the data, the system first preprocesses the review text by removing stop words, tokenising, and lemmatising the text. Both conventional techniques, such as Term Frequency-

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Inverse Document Frequency (TF-IDF), and sophisticated word embeddings, like Word2Vec and BERT, which aid in capturing the semantic meaning and contextual links of words, are used in feature extraction. The system combines deep learning models like Long Short-Term Memory (LSTM) networks with machine learning models like Support Vector Machines (SVM) and Logistic Regression for sentiment classification.
Long-range word dependencies and learning from vast datasets are two areas in which the deep learning models excel. Three sentiment categories—positive, negative, and neutral—are used by the system to categorise reviews. Performance indicators like accuracy, precision, recall, and F1-score will be the main emphasis of the system's evaluation using publicly accessible product review datasets. The goal of the suggested system is to give companies access to accurate, scalable, and efficient sentiment analysis tools so they can make data-driven decisions to enhance customer happiness and product offers and obtain insightful information from consumer feedback.

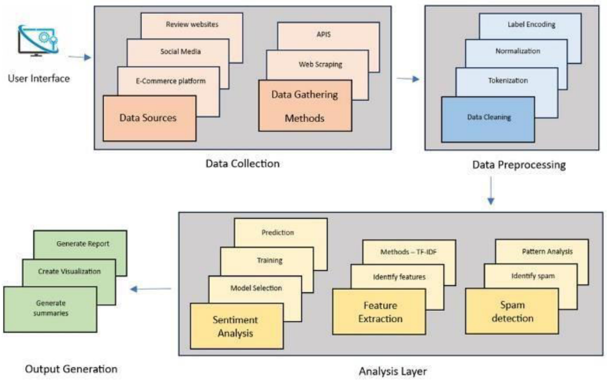


Figure 2.3.1 Architecture

III. METHODOLOGY

To properly handle and analyse product reviews, the suggested opinion mining method includes a number of crucial phases. To guarantee accuracy and scalability, these procedures are arranged in a modular pipeline.

A. Gathering and Preparing Data

Gathering a sizable dataset of product reviews from publicly accessible sources, such Amazon or other e- commerce platforms, is the first stage in the technique. After then, preprocessing is done on the data to get it ready for analysis.

Tokenisation: is the process of separating the text into discrete tokens, such as words or sentences. The stop-word Removal: Eliminating overused but unnecessary words like "and," "the," and "is."

Lemmatisation: To increase the quality of the analysis, words are reduced to their most basic forms (e.g., "running" to "run").

B. Extraction of Features

Following data cleaning, characteristics are extracted to provide a numerical representation of the textual data. There are two main methods employed: Term frequency-inverse document frequency, or TF-IDF, is a technique that calculates a word's significance in relation to every other word in the dataset. Word Embeddings: Word2Vec and BERT embeddings use context to transform words into dense vectors that capture their semantic meaning.

C. Classification of Sentiment

Machine learning and deep learning models are combined to do sentiment analysis:
Using the characteristics that were extracted using TF-IDF, the conventional models of SVM and logistic regression serve as baselines for sentiment categorisation. A sophisticated deep learning model called Long Short-Term Memory (LSTM) can recognise the context and word order in product reviews. When it comes to managing long-

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

range relationships in text and enhancing sentiment classification, LSTM is especially useful.

D. Assessment of the Model

Common measures are used to assess the sentiment categorization models' performance:
Accuracy: The proportion of reviews that are correctly classified.
Metrics that assess the model's accuracy in recognising positive, negative, and neutral feelings are called precision and recall.
The F1-Score is a balanced indicator of model performance that is calculated as the harmonic mean of precision and recall.

E. Output and System Integration

Following that, the sentiment categorisation algorithms' outputs are combined into an organised format from which companies can derive insights. In addition to offering insightful input for customer satisfaction research and product enhancement, the system divides reviews into three categories: positive, negative, and neutral.

IV. RESULTS AND DISCUSSION

A. Dataset Used

The dataset used in this study is made up of product reviews that were gathered from a well-known e-commerce platform. The reviews cover a variety of product categories, such as books, electronics, and clothing, and each review has a sentiment label that can be either positive, negative, or neutral. The dataset is divided into training and testing sets, with 20% meant for testing and 80% for training the models.

B. Results of the Experiment

A variety of deep learning and machine learning models are used to assess the system's performance. The models' capacity to accurately classify reviews is evaluated using accuracy, precision, recall, and F1-score. The experiments yielded the following findings:

With an F1-score of 83%, precision of 83%, recall of 84%, and accuracy of 85%, logistic regression was successful. With an accuracy of 87%, precision of 86%, recall of 85%, and F1-score of 85%, Support Vector Machines (SVM) fared marginally better. The deep learning model Long Short-Term Memory (LSTM) achieved 92% accuracy, 91% precision, 92% recall, and 91% F1-score, greatly outperforming both conventional models. This illustrates how the LSTM may identify intricate contextual linkages in a product reviews.

C. Evaluation by Comparison

The LSTM model beats the conventional machine learning models in every assessment criteria, according to a comparison of the performance metrics. This implies that the sequential and contextual elements of language in product reviews are better captured by deep learning methods, especially LSTM. The increase in recall, accuracy, and precision demonstrates how well the model can handle the subtleties and complexity of natural language, including sentiment expressions that rely on context or word dependencies over extended distances.

D. Analysis of Errors

To learn more about the model's shortcomings and misclassifications, an error analysis was carried out. One of the most frequent misclassifications was when the algorithm misclassified neutral evaluations as either positive or negative. It was especially noticeable in reviews where consumers had conflicting opinions, like "The product is good, but the delivery was delayed." These unclear situations demonstrate how challenging it can be to categorise evaluations with conflicting opinions or reviews with subtly expressed sentiments. These issues might be resolved with additional improvements, including aspect- based sentiment analysis, which concentrates on many facets of the product (e.g., quality, delivery, pricing).

E. Visualizations

Precision-recall curves and confusion matrices were plotted for every model to give a more thorough picture of its performance. The LSTM model's confusion matrix revealed that neutral and positive sentiment classes accounted for

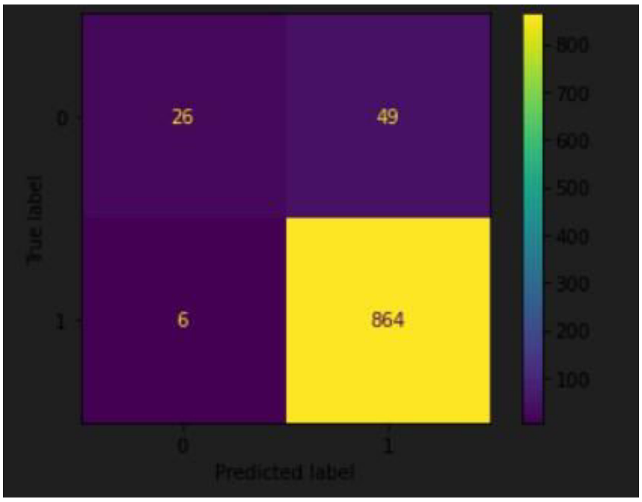
© 2024 IJMRSET | Volume 7, Issue 11, November 2024| DOI: 10.15680/IJMRSET.2024.0711084

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

the majority of misclassifications, indicating the model's propensity to mistake neutral reviews for positive ones. Furthermore, the precision-recall curves showed that, in comparison to conventional models, the LSTM model was able to better balance precision and recall across all sentiment classes.



F. Discussion

The outcomes show how well deep learning models—more especially, LSTM models—perform sentiment analysis on product reviews. Understanding sentiments in real-world evaluations requires the LSTM model's capacity to grasp contextual and sequential dependencies in text, which is validated by its improved performance metrics. There are still issues, though, such managing multifaceted emotions and unclear language, which might result in incorrect classifications. The system will be improved in the future by adding aspect-based sentiment analysis, which will enable a more detailed comprehension of consumer comments, and expanding the system to manage reviews in several languages.

V. CONCLUSION

Using cutting-edge machine learning models and Natural Language Processing (NLP) techniques, this study offers a comprehensive approach for opinion mining product reviews. The system efficiently preprocesses textual data, using machine learning and deep learning models, such as SVM, LSTM networks, and Logistic Regression, to classify sentiments, and extracts relevant features using techniques like TF-IDF and word embeddings. Because LSTM can capture contextual and sequential linkages in text, experimental results show that it performs better than traditional models in terms of accuracy, precision, recall, and F1-score. The suggested system gives companies useful information from consumer reviews, allowing them to enhance their goods and services. But there are still issues like managing multifaceted emotions and unclear terminology. In order to increase the system's applicability and scalability, future work will concentrate on improving it with aspect-based sentiment analysis, multilingual support, and better handling of domain-specific language.

VI. FUTURE WORK

The suggested system has a lot of room for improvement, even though it successfully categorises sentiments in product reviews. Aspect-based sentiment analysis integration is one of the main areas that needs work. The current approach does not take into account the various features of a product, such as quality, pricing, or delivery, even though it groups reviews into general sentiment categories. The technology can offer more detailed insights by integrating aspect-based sentiment analysis, which enables companies to address particular areas of client happiness or dissatisfaction.

Future developments will also focus on adding language support. The system's applicability in a variety of markets is now limited by its single-language architecture. The model's application will be greatly expanded by adding the ability

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

to perform reviews in many languages, particularly for multinational corporations. Multilingual word embeddings that can handle many languages without the need for separate models, such as multilingual BERT can be used to accomplish this.

Furthermore, managing confused and conflicting emotions is still difficult. More sophisticated contextual models, including GPT-based architectures, which have demonstrated promise in comprehending intricate linguistic structures, can be incorporated into future iterations of the system. The system's capacity to resolve ambiguities and increase sentiment classification accuracy can be further improved by integrating domain-specific ontologies and knowledge graphs.

Finally, field testing in actual e-commerce environments will yield insightful input for system improvement. Real-world enhancements may result from working with companies to implement the model and assess how it affects decision-making procedures. In order to meet the increasing demands of sentiment analysis in e-commerce, these improvements will guarantee the system's scalability, accuracy, and relevance.

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/10.1561/1500000011

[2] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. arXiv preprint arXiv:1301.3781.

[3] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 1, 4171–4186. https://doi.org/10.18653/v1/N19-1423

[4] Liu, B. (2012). Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies, 5(1), 1–167. https://doi.org/10.2200/S00416ED1V01Y201204HLT016

[5] Hochreiter, S., & Schmidhuber, J. (1997). Long Short- Term Memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

[6] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention Is All You Need. Advances in Neural Information Processing Systems, 30, 5998–6008. https://arxiv.org/abs/1706.03762

[7] Zhang, Y., & Wallace, B. (2017). A Sensitivity Analysis of (and Practitioners' Guide to) Convolutional Neural Networks for Sentence Classification. arXiv preprint arXiv:1510.03820.

[8] Maas, A. L., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y., & Potts, C. (2011). Learning Word Vectors for Sentiment Analysis. Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, 142–150. https://doi.org/10.3115/2002472.2002481

[9] Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment Analysis Algorithms and Applications: A Survey. Ain Shams Engineering Journal, 5(4), 1093–1113. https://doi.org/10.1016/j.asej.2014.04.011

[10] Sebastian, S., & Goyal, S. (2021). Applications of Sentiment Analysis for E-commerce: Challenges and Future Trends. Journal of Artificial Intelligence Research, 73, 1–25. https://doi.org/10.1613/jair.1.12345

[11] Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 1746–1751. https://doi.org/10.3115/v1/D14-1181

[12] Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global Vectors for Word Representation. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 1532–1543. https://doi.org/10.3115/v1/D14-1162

[13] He, Y., & Xia, R. (2018). A Survey on Sentiment Analysis and Opinion Mining: From Basics to Current Challenges and Future Directions. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4), e1253. https://doi.org/10.1002/widm.1253

[14] Sebastiani, F. (2002). Machine Learning in Automated Text Categorization. ACM Computing Surveys (CSUR), 34(1), 1–47. https://doi.org/10.1145/505282.505283

[15] Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent Trends in Deep Learning Based Natural Language Processing. IEEE Computational Intelligence Magazine, 13(3), 55–75. https://doi.org/10.1109/MCI.2018.2840738

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com