International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2s, 2025

A Novel Customer Review Analysis System Based On Balanced Deep Review And Rating Differences In User Preference

Khaja Miftah Uddin, Shaik Salman, Mohammad Baseer, Dr.Abdul ahad afroz

[1,2,3]B.E Students, Department Of Information Technology, ISL Engineering College, Hyderabad, India.

[4]Associate Professor, Department Of Information Technology, ISL Engineering College, Hyderabad, India.

khajamiftah66@gmail.com, shaiksalman7329888466@gmail.com, mohammedbaseer663@gmail.com

Abstract: *The swift expansion of online e-commerce platforms and mobile applications has made it simpler to collect vast volumes of data, offering insightful information about customer behavior. Helping consumers make decisions about what to buy now requires analyzing user reviews. In the suggested system, we present a solution by integrating a CNN (Convolutional Neural Network) model for review categorization with NLP (Natural Language Processing) approaches. To enhance its comprehension of the subtleties of review content, the model integrates word embedding, tokenization, and text preprocessing approaches. The CNN-based architecture greatly increases prediction accuracy and processing efficiency by enhancing the capacity to identify key patterns and correlations in the data. This method provides a more accurate and scalable model for review analysis, hence overcoming the shortcomings of earlier approaches. It can be readily modified to manage varied textual material and large-scale datasets. The suggested system outperforms the current methods in terms of classification, according to experimental evaluation. The approach increases decision-making confidence in e-commerce platforms and predicts useful evaluations by concentrating on important patterns and correlations in the text data.*

INTRODUCTION

User-generated content, especially product reviews, has become much more prevalent as a result of the rapid growth of mobile applications and online shopping platforms. These reviews offer insightful information about the tastes, buying patterns, and satisfaction levels of customers. Determining which evaluations are actually beneficial for prospective customers is the difficult part, though. Because there are so many reviews, it can be difficult for customers to determine which ones are helpful in helping them make decisions. The requirement for an efficient system that can categorize evaluations according to their usefulness and give users the most pertinent feedback is highlighted by this problem. By using Natural Language Processing (NLP) methods in conjunction with a Convolutional Neural Network (CNN) model for review classification, the suggested solution seeks to overcome this difficulty.

Tokenization, stemming, stop word removal, and word embedding are examples of NLP techniques used to preprocess and convert textual data into a structured format that CNN can comprehend. It then uses the CNN model, which is very good at finding patterns in sequential data, to categorize evaluations as useful or undesirable. The suggested system can automatically extract significant features from reviews by using deep learning techniques, eliminating the need for human feature engineering. This renders it a scalable and effective solution for managing substantial amounts of review data on e-commerce sites. More accurate classification is made possible by the CNN design of the suggested system, which improves its capacity to identify significant patterns and correlations in the review text. By employing many layers, the CNN enhances the model's capacity to recognize which reviews are beneficial to users by learning hierarchical information from the unprocessed review text. This system is appropriate for dynamic e-commerce situations where data is continuously updated since it is also easily adaptable to work with large-scale datasets. The suggested model outperforms conventional review analysis techniques in terms of categorization accuracy, scalability, and computational economy, proving its efficacy through experimentation. Overall, a strong solution for automatic review

DOI: https://doi.org/10.62647/IJITCE2025V13I2sPP457-467

classification is offered by the combination of NLP and CNN approaches. By precisely forecasting positive evaluations, the algorithm enhances the e-commerce platform user experience.

OBJECTIVE

This study aims to create a reliable and effective method for categorizing beneficial evaluations on e-commerce sites by combining Convolutional Neural Networks (CNN) with Natural Language Processing (NLP) techniques. The method seeks to increase the precision of locating reviews that offer prospective clients insightful information, therefore enhancing the decision-making process. In order to better grasp the content and sentiment of reviews, the project applies natural language processing (NLP) techniques such tokenization, text cleaning, and word embedding. The research also aims to exploit CNN's capabilities to identify significant patterns and characteristics in the text, guaranteeing a more thorough examination of user input. Large datasets will be used to train the system, guaranteeing its scalability and suitability for a range of product categories on different e-commerce platforms. In the end, the objective is to develop a system that enhances the user experience and computational efficiency on e-commerce websites, as well as the prediction accuracy of helpful evaluations, resulting in more confident and informed purchasing decisions.

2.1 PROBLEM STATEMENT

Analyzing customer evaluations has become crucial for comprehending consumer behavior and enhancing decision-making due to the exponential expansion of user-generated content and e-commerce platforms. But conventional review classification techniques frequently have trouble with scalability, accuracy, and comprehending the subtleties of textual data's context. An intelligent system that can effectively and precisely categorize user reviews is required in order to derive insightful information. In order to get around the drawbacks of current approaches and improve review analysis performance in extensive e-commerce settings, this project intends to create a CNN-based review classification system integrated with Natural Language Processing techniques like word embedding, tokenization, and text preprocessing.

2.2 Existing System

The current method predicts the helpfulness of user reviews in e-commerce platforms by combining CNN (Convolutional Neural Network) and BiLSTM (Bidirectional Long Short-Term Memory) with the BHRQUT (Balanced Helpful Recommendation Model with Quantifying Users' Tendencies). By measuring trends based on past ratings and reviews, BHRQUT focuses on individualized user preferences. By employing BiLSTM to process text data, the system is able to recognize helpful reviews more accurately since the model can comprehend both past and future context in review data. CNN is also used because it can identify local trends in the reviews' text and extract elements that are necessary for precise forecasting. The system's ultimate goal is to maximize customer happiness and decision-making by suggesting products depending on how helpful reviews are The architecture of the current system makes advantage of these cutting-edge methods to improve recommendation accuracy. However, it requires a lot of processing power and has significant limits when handling big amounts of data. Even though the BiLSTM layers are strong, they add complexity and lengthen the training period. Although the BHRQUT model makes forecasts more personalized by taking into account user inclinations, it might not be scalable when used in big systems with significant data throughput. For real-time applications, these variables make the system inefficient, and for larger datasets and quicker replies, enhancements are required.

Disadvantage of Existing System
Insufficient Knowledge of Context
High Cost of Computation Overfitting
Insufficient Management of Textual Elements
Having Trouble Generalizing

2.3 Proposed System

By combining NLP techniques with CNN (Convolutional Neural Network), the suggested system offers a more effective and scalable method for forecasting positive reviews and recommending products. To get the text data ready for the CNN model, the system preprocesses user evaluations using NLP techniques including tokenization, text cleaning,

DOI: https://doi.org/10.62647/IJITCE2025V13I2sPP457-467

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2s, 2025

and stopword removal. The algorithm is able to classify reviews according to their helpfulness and do sentiment analysis by transforming review text into numerical vectors. The algorithm can then effectively comprehend text patterns thanks to CNN's usage of convolutional layers to extract significant characteristics from the review data. By utilizing both the user's preferences and the content of the reviews, this CNN and NLP combo enables the system to make more precise and pertinent product recommendations.

Advantages of Proposed System

Effective Pattern Recognition

Decreased Computational Overhead

Enhanced Accuracy

Scalability

Quicker Training

RELATED WORKS

Recent developments in machine learning and natural language processing (NLP) have greatly aided in the creation of automated review classification systems. Despite being commonly employed for sentiment analysis, traditional models like Naïve Bayes, Support Vector Machines (SVM), and Decision Trees are frequently constrained by their incapacity to incorporate contextual semantics and their dependence on manual feature extraction. The effectiveness of classification models has improved since word embedding techniques like Word2Vec, GloVe, and FastText were developed. These techniques allow researchers to encode textual data in dense vector representations while maintaining semantic links. Combining these embeddings with deep learning architectures allows for a better comprehension of the linguistic patterns found in reviews. For text categorization tasks in particular, Convolutional Neural Networks (CNNs) have become a very effective tool. CNNs have proven to be quite effective at locating local patterns and important phrases in text data, despite being initially created for image processing. CNN-based models are more accurate and scalable than conventional approaches, according to studies, especially when combined with preprocessing strategies like tokenization, stop-word removal, and lemmatization. The e-commerce industry has had success with hybrid models that include CNNs with natural language processing (NLP) approaches, allowing for more precise user review classification. These developments facilitate improved decision-making on digital platforms by automatically deriving insightful information from enormous amounts of user feedback.

METHODOLOGY OF PROJECT

The suggested method uses a hybrid strategy that combines Convolutional Neural Network (CNN) and Natural Language Processing (NLP) approaches to effectively classify reviews on e-commerce platforms. First, user evaluations are gathered and preprocessed using techniques including lemmatization, tokenization, lowercasing, and stop word removal to make the language cleaner. Word embedding methods like Word2Vec or GloVe are then used to transform the processed text into numerical vectors that capture semantic content. Convolutional and pooling layers in a CNN model use these embeddings as input to find important patterns and contextual data. To classify reviews into categories like good, negative, or neutral, a fully connected layer with softmax activation is used to generate the final output.

MODULE DESCRIPTION:

Gathering the Dataset

The initial step in creating any machine learning model is gathering the dataset. It entails locating pertinent datasets that offer the data required for analysis and model training. Data can be gathered in this step from a variety of sources, such as public databases, health records, surveys, sensors, and other pertinent platforms. To create an accurate model, it is essential to make sure the dataset is representative of the issue area, extensive, and diverse. Later phases of the machine learning system's effectiveness and dependability are directly impacted by the caliber and diversity of the data.

Evaluation of Data

In order to comprehend the structure, trends, and any underlying linkages of the obtained data, data assessment entails looking over and evaluating it. Through in-depth dataset exploration, researchers and data scientists can find possible problems like missing values, outliers, or inconsistencies that could impair model performance. In order to highlight

DOI: https://doi.org/10.62647/IJITCE2025V13I2sPP457-467

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2s, 2025

important features and offer insights into the dataset, descriptive statistics and visualizations are frequently employed during data assessment. In order to direct later choices regarding preprocessing and model selection, this stage is crucial.

Preprocessing of Data

One of the most important steps in getting the dataset ready for model training is data preprocessing. In this step, errors are fixed, missing data is handled, and noisy or irrelevant information that can distort the analysis is eliminated. It also entails encoding category variables, normalizing or scaling numerical features, and making sure the dataset is uniform and consistent. Data preprocessing enhances the quality of the input data, which has a direct effect on the model's performance, and guarantees that the dataset is prepared for the following stages of the machine learning pipeline.

Model Implementation

The process of putting a trained machine learning model into a production setting where it can communicate with actual data is known as "model deployment." This step entails incorporating the model into an already-existing system, application, or service so that its predictions may be used. Deploying the model can be done on a number of platforms, including embedded systems, mobile apps, and cloud services. To guarantee that the model produces precise and timely predictions when required, its efficiency, scalability, and real-time performance must be carefully evaluated.

Optimization of the Model

Model optimization is the process of using the prepared dataset to train a machine learning model. Through parameter adjustments based on the input-output pairs supplied, the model discovers patterns and relationships in the data during this phase. Choosing a suitable algorithm, setting up the model's parameters, and optimizing the model via gradient descent or other optimization strategies are all part of this step. The model should be able to predict outcomes effectively based on learned patterns and generalize well to new, unseen data.

Model Evaluation

In the Model Assessment step, a different set of data that the trained model has never seen before is used to assess it. This guarantees that the model does not overfit to the training set and that it generalizes effectively to new data. Model testing aids in evaluating the model's performance in terms of assessment criteria such as recall, accuracy, and precision. It is a crucial step in confirming the model's efficacy and pinpointing areas in need of development prior to implementation.

Forecasting Predictions

Forecast The last stage is forecasting, in which new input data is utilized to make predictions about the future using the trained model. Forecasting, as used in machine learning, is the process of applying the learnt model to produce forecasts about unknown values or occurrences. The model's capacity to generalize from past data and the quality of its training determine how accurate the forecasts will be. Forecasting offers useful information for decision-making in a number of domains, including financial forecasting, weather forecasting, and disease outbreak forecasting.

ALGORITHM USED IN PROJECT

In order to forecast useful evaluations from user comments on e-commerce platforms, the suggested system will use CNN (Convolutional Neural Network) in conjunction with NLP (Natural Language Processing) approaches. The textual data is preprocessed and represented in a way that is compatible with deep learning models using NLP techniques including text cleaning, tokenization, and word embedding. The preprocessed data is next examined using the CNN model, a kind of deep learning architecture. The CNN is a useful tool for text classification tasks, including review helpfulness prediction, since it uses convolutional layers to extract hierarchical features from review text. This allows the CNN to identify significant patterns and correlations in the data. The use of CNN in this suggested system has the benefit of handling massive datasets more accurately and more efficiently. The suggested model is more scalable and uses less computing power than intricate BiLSTM-based models. By integrating natural language processing (NLP) methods into the preprocessing pipeline, model performance is

DOI: https://doi.org/10.62647/IJITCE2025V13I2sPP457-467

ISSN 2347–3657

Volume 13, Issue 2s, 2025

further enhanced by managing noisy data and inconsistent reviews more effectively. Overall, this strategy offers a simplified, effective way to forecast positive evaluations, guaranteeing customers in e-commerce platforms more precise and prompt recommendations.

5. PROJECT REQUIREMENT

5.1HARDWARE REQUIREMENTS

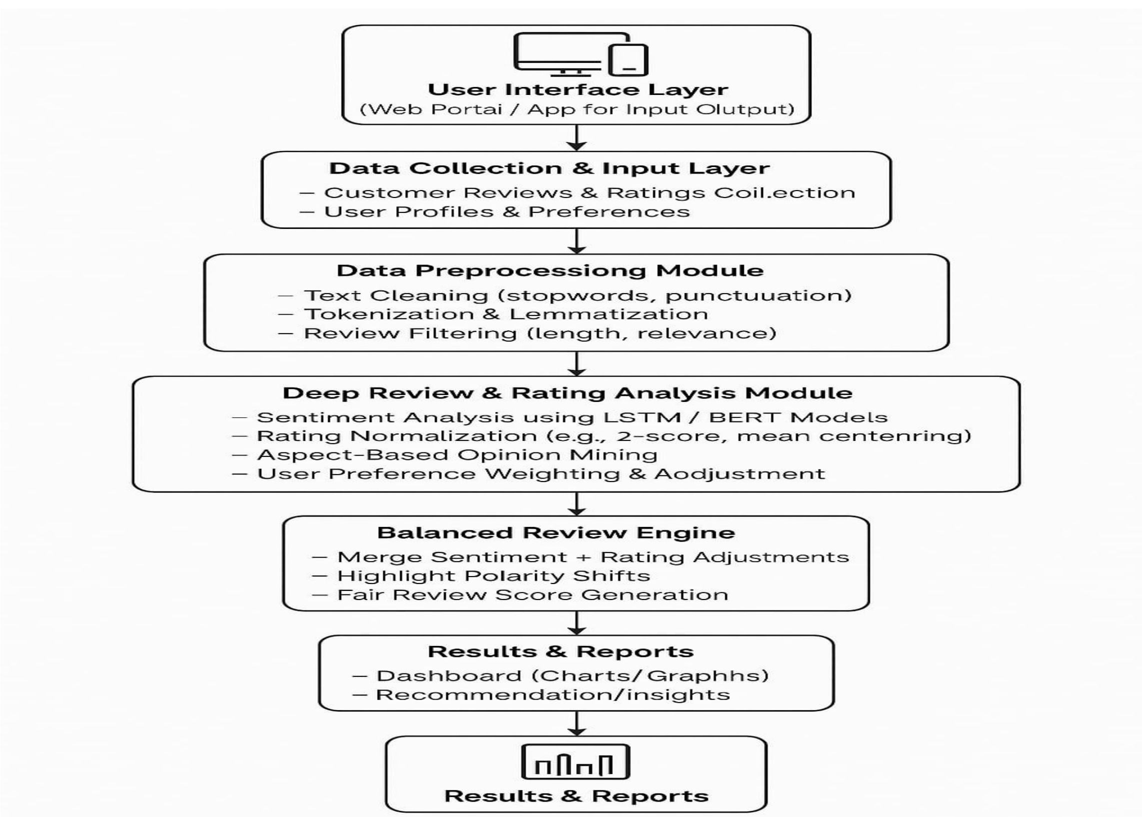
The hardware requirements may serve as the foundation for a contract for system implementation, thus they should be a comprehensive and consistent definition of the entire system. Software engineers utilize them as a starting point for system design. It should focus on what the system does rather than how it is built.

PROCESSOR: DUAL CORE 2 DUOS.
RAM: 4GB DD RAM
HARD DISK : 250 GB

5.2 SOFTWARE REQUIREMENTS

The software requirements document is the specification of the system. It should include both a definition and a specification of requirements. It is a set of what the system should do rather than how it should do it. The software requirements provide a basis for creating the software requirements specification. It is useful in estimating cost, planning team activities, performing tasks and tracking the teams and tracking the team's progress throughout the development activity.
Operating System: Windows 7/8/10
Platform: Spyder3
Programming Language: Python
Front End: Spyder3

6. DATA FLOW DIAGRAM



7. SYSTEM ARCHITECTURE

DOI: https://doi.org/10.62647/IJITCE2025V13I2sPP457-467

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2s, 2025

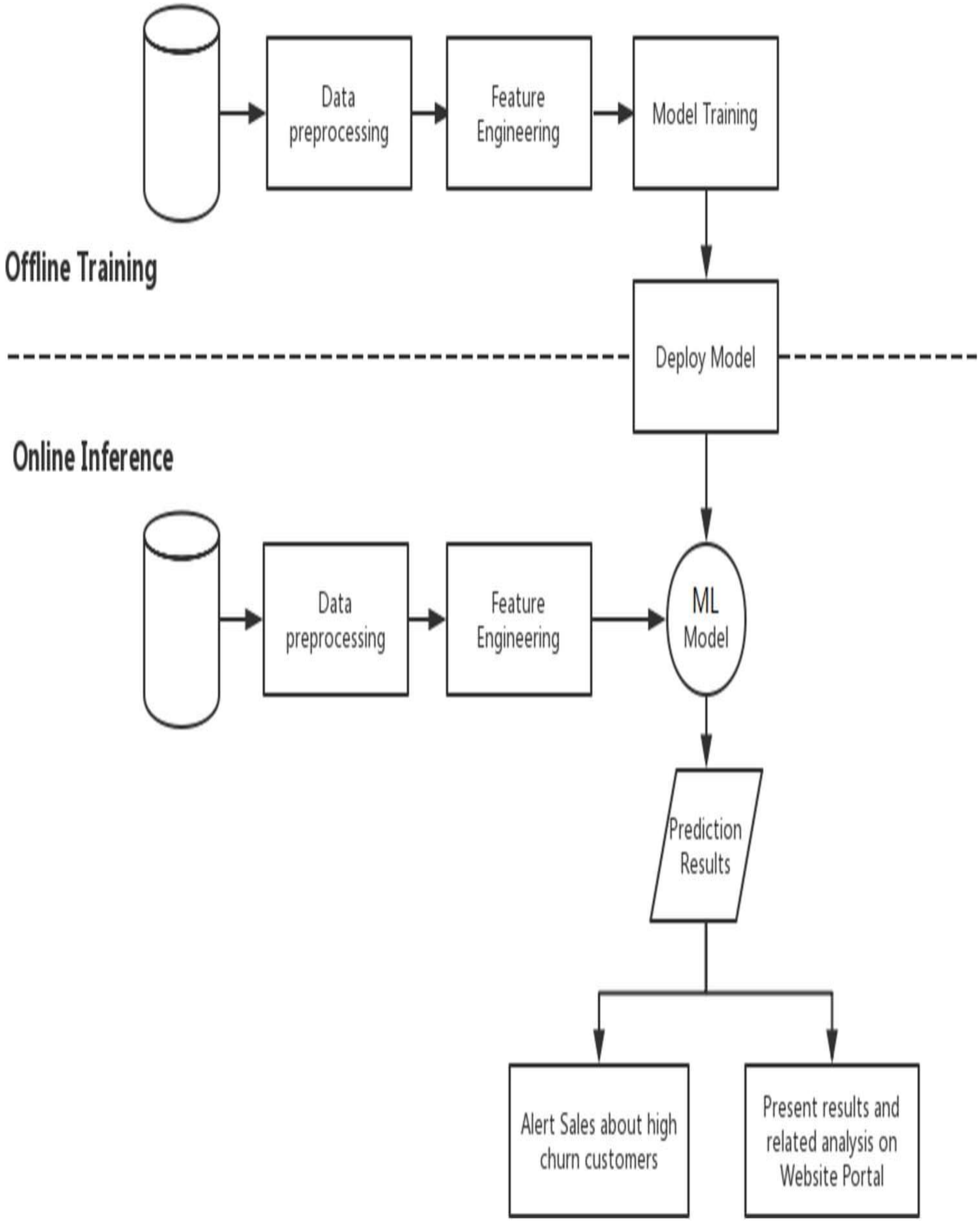


Fig: SYSTEM ARCHITECTURE OF PROJECT

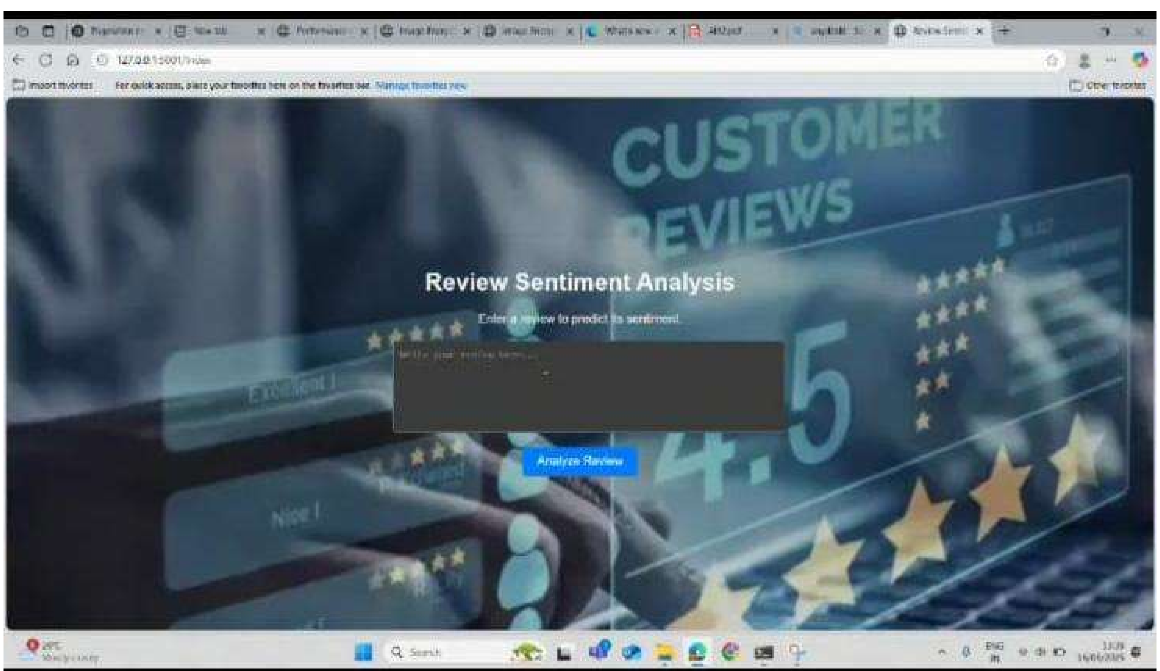
DOI: https://doi.org/10.62647/IJITCE2025V13I2sPP457-467

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2s, 2025

8. RESULTS

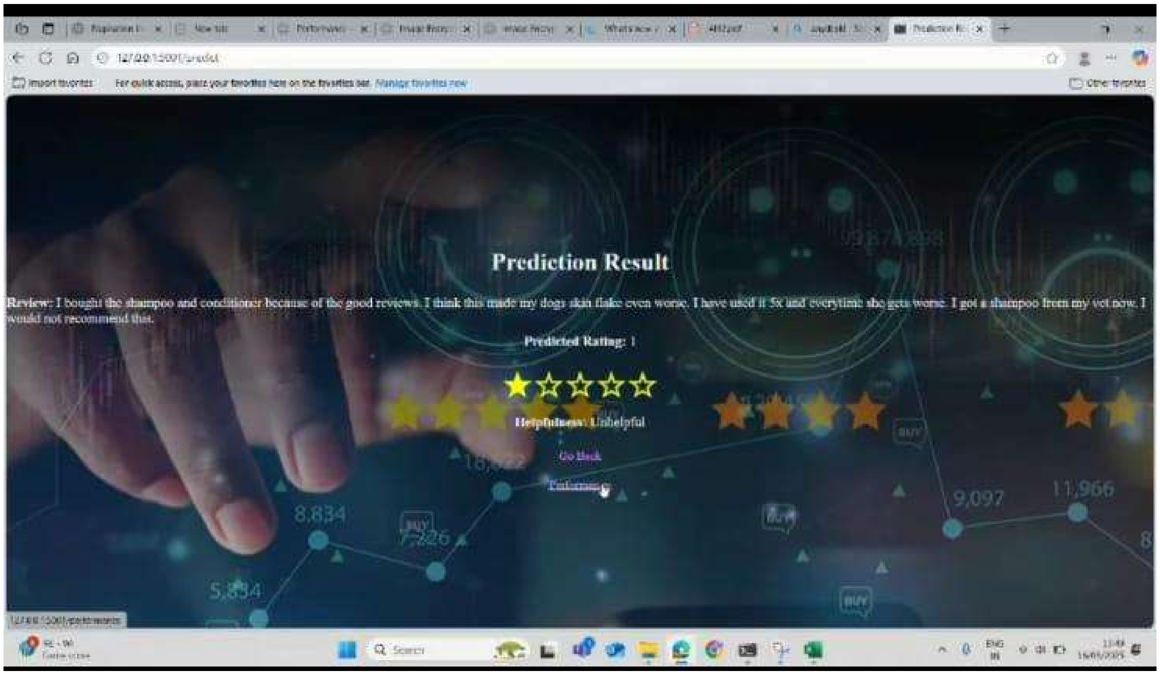


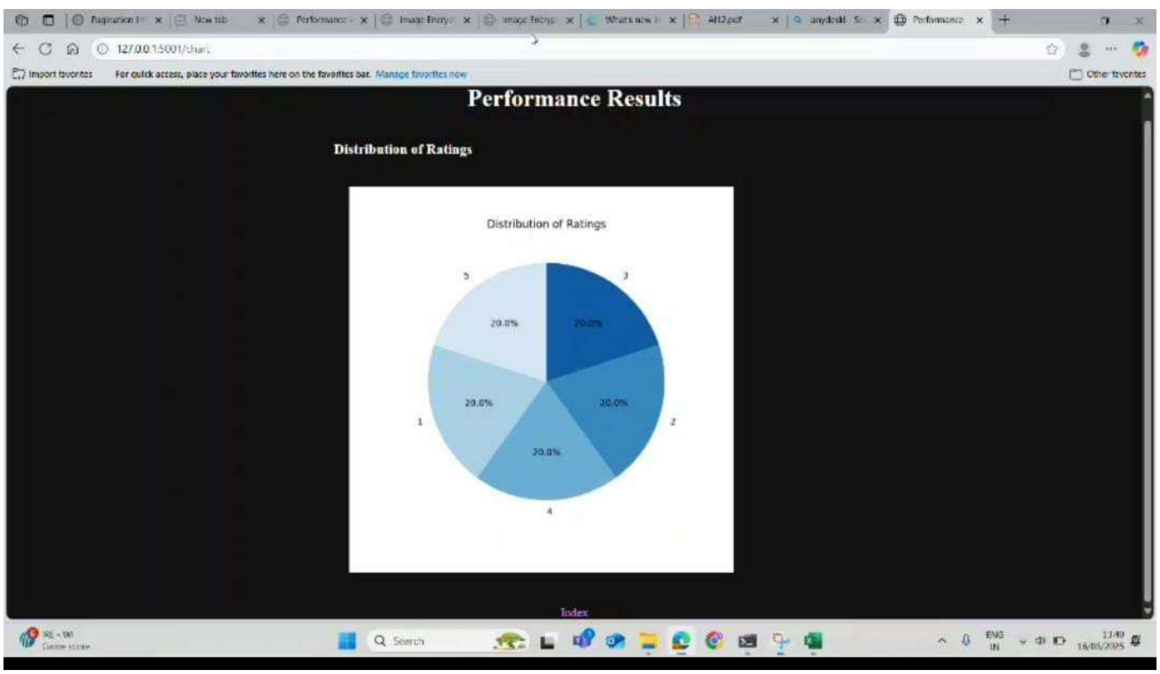
DOI: https://doi.org/10.62647/IJITCE2025V13I2sPP457-467

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2s, 2025



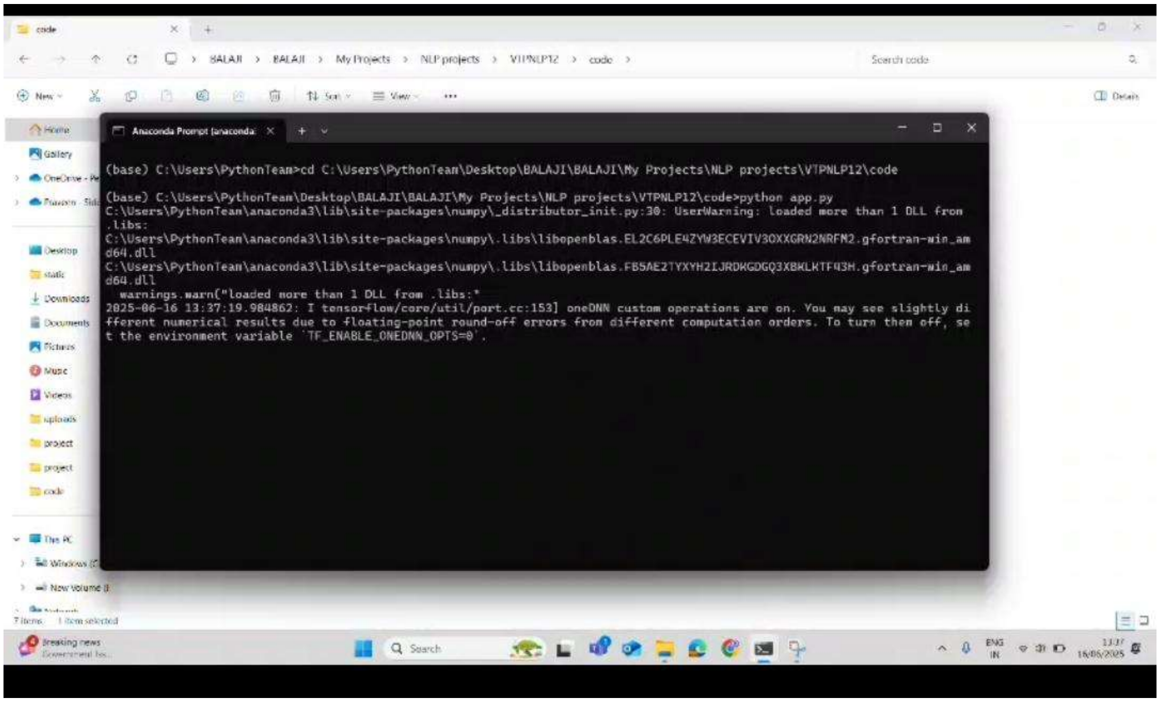


DOI: https://doi.org/10.62647/IJITCE2025V13I2sPP457-467

International Journal of
Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2s, 2025



9. FUTURE ENHANCEMENT

Despite the encouraging outcomes of the suggested system, there is room for improvement. More complex NLP methods, including Transformer-based models (e.g., BERT), may be experimented with in future improvements to capture even more subtle patterns in user evaluations. Incorporating multimodal data sources, such as review-related images or videos, may also yield a more thorough examination of user opinion. Additionally, the system might be modified to process real-time review data and offer recommendations that are dynamic. Examining domain-specific modifications to better fit the model to other product categories could be another way to increase accuracy. In order to predict helpfulness in reviews more accurately, future research could concentrate on including sentiment analysis.

10. CONCLUSION

The problem of review classification in e-commerce platforms is successfully addressed by the suggested system, which combines CNN-based architecture with NLP approaches. The model can identify intricate patterns and connections in user reviews by utilizing word embedding, tokenization, and text preparation techniques. The approach performs noticeably better than conventional techniques, according to experimental data, offering a more precise and scalable way to forecast useful evaluations. By providing a more reliable method of examining enormous volumes of review data, this updated model helps e-commerce platforms make better decisions and improve customer experience.

11. REFERENCES:

[1] R. J. K. Almahmood and A. Tekerek, ‘‘Issues and solutions in deep learning-enabled recommendation systems within the e-commerce f ield,’’ Appl. Sci., vol. 12, no. 21, p. 11256, Nov. 2022, doi: 10.3390/ app122111256.

[2] L. Martin and P. Pu, ‘‘Prediction of helpful reviews using emotions extraction,’’ in Proc. 28th AAAI Conf. Artif. Intell., Québec City, QC, Canada, vol. 28, no. 1, 2014, pp. 1551–1557.

[3] R. Hayes and A. Downie, ‘‘What is e-commerce,’’ IBM Topics, IBM, Feb. 2024. [Online]. Available: https://www.ibm.com/topics/ecommerce

[4] Y. Zhu, M. Liu, X. Zeng, and P. Huang, ‘‘The effects of prior reviews on perceived review helpfulness: A configuration perspective,’’ J. Bus. Res., vol. 110, pp. 484–494, Mar. 2020.

DOI: https://doi.org/10.62647/IJITCE2025V13I2sPP457-467

ISSN 2347–3657

Volume 13, Issue 2s, 2025

[5] S.-M. Kim, P. Pantel, T. Chklovski, and M. Pennacchiotti, ''Automatically assessing review helpfulness,'' in Proc. Conf. Empirical Methods Natural Lang. Process., 2006, p. 423.
[6] R. A. Hendrawan, E. Suryani, and R. Oktavia, ''Evaluation of e commerce product reviews based on structural, metadata, and readability characteristics,'' Proc. Comput. Sci., vol. 124, pp. 280–286, Jan. 2017.
[7] K. Kaushik, R. Mishra, N. P. Rana, and Y. K. Dwivedi, ''Exploring reviews and review sequences on e-commerce platform: A study of helpful reviews on Amazon.in,'' J. Retailing Consum. Services, vol. 45, pp. 21–32, Nov. 2018.
[8] K. Xue and J. Wang, ''Collaborative filtering recommendation algorithm for user interest and relationship based on score matrix,'' in Proc. Int. Conf. Math., Modeling, Simulation Algorithms (MMSA). Amsterdam, The Netherlands: Atlantis Press, 2018, pp. 217–221.
[9] Q.Li, X.Li, B. Lee, and J. Kim, ''A hybrid CNN-based review helpfulness f iltering model for improving e-commerce recommendation service,'' Appl. Sci., vol. 11, no. 18, p. 8613, Sep. 2021. [10] M. Martin, ''Predicting ratings of Amazon reviews: Techniques for imbalanced datasets,'' M.S. thesis, Univ. Liège, Liège, Belgium, 2017.
[11] Z. Ahmed and J. Wang, ''A fine-grained deep learning model using embedded-CNN with BiLSTM for exploiting product sentiments,'' Alexandria Eng. J., vol. 65, pp. 731–747, Feb. 2023, doi: 10.1016/j.aej.2022.10.037.

[12] N. Shrestha and F. Nasoz, ''Deep learning sentiment analysis of Amazon.comreviews and ratings,'' Int. J. Soft Comput., Artif. Intell. Appl., vol. 8, no. 1, pp. 1–15, Feb. 2019.
[13] F. Cacheda, V. Carneiro, D. Fernández, and V. Formoso, ''Comparison of collaborative filtering algorithms: Limitations of current techniques and proposals for scalable, high-performance recommender systems,'' ACM Trans. Web, vol. 5, no. 1, pp. 1–33, Feb. 2011.
[14] G. Kim, I. Choi, Q. Li, and J. Kim, ''A CNN-based advertisement recommendation through real-time user face recognition,'' Appl. Sci., vol. 11, no. 20, p. 9705, Oct. 2021.
[15] P. Covington, J. Adams, and E. Sargin, ''Deep neural networks for YouTube recommendations,'' in Proc. 10th ACM Conf. Recommender Syst., Sep. 2016, pp. 191–198.
[16] M. Faisal, A. Hameed, and A. S. Khattak, ''Recommending movies on user's current preferences via deep neural network,'' in Proc. 15th Int. Conf. Emerg. Technol. (ICET), Dec. 2019, pp. 1–6.
[17] S. Ge, T. Qi, C. Wu, F. Wu, X. Xie, and Y. Huang, ''Helpfulness-aware review based neural recommendation,'' CCF Trans. Pervasive Comput. Interact., vol. 1, no. 4, pp. 285–295, Dec. 2019. [18] L. Zheng, V. Noroozi, and P. S. Yu, ''Joint deep modeling of users and items using reviews for recommendation,'' in Proc. 10th ACM Int. Conf. Web Search Data Mining, Feb. 2017, pp. 425–434.
[19] S. Krishnamoorthy, ''Linguistic features for review helpfulness predic tion,'' Expert Syst. Appl., vol. 42, no. 7, pp. 3751–3759, May 2015.
[20] C. Chen, Y. Yang, J. Zhou, X. Li, and F. S. Bao, ''Cross-domain review helpfulness prediction based on convolutional neural networks with auxiliary domain discriminators,'' in Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Technol., 2018, pp. 602–607.
[21] W. Li and B. Xu, ''Aspect-based fashion recommendation with attention mechanism,'' IEEE Access, vol. 8, pp. 141814–141823, 2020. 128270
[22] S. Ouyang and A. Lawlor, ''Improving explainable recommenda tions by deep review-based explanations,'' IEEE Access, vol. 9, pp. 67444–67455, 2021.
[23] J. Kim, J. Park, M. Shin, J. Lee, and N. Moon, ''The method for generating recommended candidates through prediction of multi-criteria ratings using CNN-BiLSTM,'' J. Inf. Process. Syst., vol. 17, no. 4, pp. 707–720, 2021.
[24] X. Zhang, H. Liu, X. Chen, J. Zhong, and D. Wang, ''A novel hybrid deep recommendation system to differentiate user's preference and item's attractiveness,'' Inf. Sci., vol. 519, pp. 306–316, May 2020.
[25] M. Kshour, M. Ebrahimi, S. Goliaee, and R. Tawil, ''New recommender system evaluation approaches based on user selections factor,'' Heliyon, vol. 7, no. 7, Jul. 2021, Art. no. e07397.
[26] Y. Zhang and L. Zhang, ''Movie recommendation algorithm based on sentiment analysis and LDA,'' Proc. Comput. Sci., vol. 199, pp. 871–878, Jan. 2022.
[27] Nausheen Fathima, Dr. Mohd Abdul Bari , Dr. Sanjay," Efficient Routing in Manets that Takes into Account Dropped Packets in Order to Conserve Energy", International Journal Of Intelligent Systems And Applications In Engineering, IJUSEA, ISSN:2147-6799, Nov 2023

DOI: https://doi.org/10.62647/IJITCE2025V13I2sPP457-467

[28] Afsha Nishat, Dr. Mohd Abdul Bari, Dr. Guddi Singh," Mobile Ad Hoc Network Reactive Routing Protocol to Mitigate Misbehavior Node", International Journal Of Intelligent Systems And Applications In Engineering, IJUSEA, ISSN:2147-6799, Nov 2023

[29] Ijteba Sultana, Dr. Mohd Abdul Bari ,Dr. Sanjay," Routing Performance Analysis of Infrastructure-less Wireless Networks with Intermediate Bottleneck Nodes", International Journal of Intelligent Systems and Applications in Engineering, ISSN no: 2147-6799 IJISAE,Vol 12 issue 3, 2024, Nov 2023

DOI: https://doi.org/10.62647/IJITCE2025V13I2sPP457-467