iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** IV **Month of publication:** April 2025

DOI: https://doi.org/10.22214/ijraset.2025.67634

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Sentiment Analysis of Amazon Product Reviews: Leveraging NLP Techniques for Enhanced Classification Accuracy

Ashish Mathur[1], Dr. Pharindara Kumar Sharma[2]
[1]Research Scholar, [2]Associate Professor, Dept. of Computer Science and Engineering, SRCEM College Banmore, Morena, M.P. India

***Abstract:** This work tackles sentiment classification on product reviews using a hybrid LSTM-GRU model. The primary objective is to evaluate the model's ability to correctly categories sentiments (positive, negative, and neutral) in a massive review dataset including 568,454 rows & 10 columns of product reviews. Data collecting from an online retailer's review system forms part of the approach, then additional pre-processing activities including text cleaning, tokenising, and padding for sentiment analysis preparation follow. Aimed to capture intricate sentiment patterns from textual data, the hybrid LSTM-GRU architecture groups the reviews. The model is defined by its accuracy and loss, and its performance is evaluated using metrics such as recall, F1 score, and precision. A test accuracy of 82.50% is achieved by the model, suggesting high sentiment classification performance, with a loss value of 1.56, according to the results. These results suggest potential for real-time sentiment analysis applications in e-commerce systems since they show that the hybrid LSTM-GRU model efficiently detects sentiment trends inside product reviews. The results highlight the great generalising capacity of the model, thereby reducing prediction error and offering correct sentiment classifications over several review data.*
***Keywords:** Hybrid LSTM-GRU, Sentiment Classification, Product Reviews, Text Preprocessing, Accuracy*

I. INTRODUCTION

Sentiment classification of product reviews plays a crucial role in understanding customer feedback, enhancing product offerings, and improving customer satisfaction. In the context of e-commerce platforms like Amazon, where millions of reviews are generated daily, the ability to automatically analyze and classify these reviews into positive, negative, or neutral categories is invaluable[1]–[6]. Helping with this effort is sentiment analysis, a branch of NLP that uses computer methods to decipher the feelings expressed in text. This paper explores sentiment classification of Amazon product reviews using advanced NLP techniques, offering insights into the efficiency and effectiveness of different methods. Amazon reviews present a rich source of data, reflecting a wide range of consumer experiences across various product categories. These reviews provide valuable information to both customers and businesses. For potential buyers, understanding the sentiment behind reviews can influence purchasing decisions, while businesses can use this information to identify trends in customer satisfaction, areas needing improvement, and product-specific feedback.

Fig.1 Sentiment Classification

The sheer volume of data makes manual analysis an unrealistic option. This is where sentiment classification using NLP techniques becomes indispensable. Businesses can make real-time improvements to their goods and services by automating the process for review categorisation. Stages such as data preprocessing, feature extraction, or machine learning model application are usually included in sentiment categorisation. The preparation step involves cleaning and standardising the raw text input to make it more understandable for machine learning models. Reducing words to their original root forms is an important part of this process, as is eliminating noise like punctuation, digits, and stopwords. Stemming and lemmatisation are two methods that can help with this. Machine learning algorithms can understand the text once it has been preprocessed and features have been retrieved. Word2Vec and GloVe are two examples of popular word embeddings, and Term Frequency-Inverse Document Frequency (TF-IDF) is another popular feature extraction method. The reviews' sentiment is categorised using machine learning models after feature extraction. For sentiment categorisation tasks, Logistic Regression, Support Vector Machines (SVM), & Naïve Bayes are some of the more used traditional machine learning techniques. A review's positivity, negativity, or neutrality can be predicted by these models using the retrieved features.

[7]–[12]. But now, with the rise of deep learning, sentiment classification tasks are being tackled using more advanced models, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), which offer better accuracy. Combining these models with pre-trained word embeddings improves their ability to understand human language's context and subtleties. Modern sentiment analysis has been radically altered by transformer-based models like BERT (Bidirectional Encoder Representations from Transformers), which capture in-depth textual context. When it comes to deciphering review sentiment, BERT really shines because of its bidirectional nature. This feature lets it take into account the context of a word depending on both its previous and subsequent terms. Being able to infer sentiment from context rather than having it expressed clearly makes BERT ideal for categorising Amazon reviews. Complex language features, including idioms, sarcasm, and domain-specific jargon, present a significant obstacle to accurate sentiment classification. An further complication to the analysis is the usage of casual language, acronyms, and emoticons by customers in product reviews. And sometimes customers would say both good and negative things about a product or service in the same review, which is called mixed feelings. If a consumer is happy with the goods yet unhappy with the delivery, it's hard to put a single emotion label on their review. By identifying intricate patterns in text, new natural language processing (NLP) approaches like attention mechanisms or transformer models provide encouraging answers to these problems. An examination of both classic machine learning models and cutting-edge deep learning algorithms for sentiment classification of Amazon product evaluations is the goal of this study. Searching for the best methods for reliably categorising feelings, this study compares the results of various models on the Amazon review dataset. To aid companies in making better decisions based on consumer input, this study's findings will add to the growing body of knowledge surrounding sentiment analysis technologies. The findings of this study can also be used in other fields that need to analyse sentiment in big text datasets, like customer service analytics and social media monitoring[13]–[17].

II. LITERATURE REVIEW

Gupta 2024 et al. creates a sentiment analysis system leveraging NLP and web scraping methods for product evaluations. The following models are used: SVC, GRU, LSTM, Naive Bayes, Logistic Regression, Random Forest, KNN, and more. No other model could compete with LSTM and GRU, which both achieved F-1 scores of 90% and 91%, respectively. While KNN compromises speed and accuracy, SVM and logistic regression also perform well. The results direct the choice of effective models for applications including URL-based sentiment analysis[18].

Sarraf 2024 et al. focusses on evaluating Amazon food reviews by adding more data to the original dataset and using preprocessing methods include text cleansing, stop word removal, lemmatization, and stemming.

The study used TF-Inverse Document Frequency (TF-IDF), Word2Vec (W2V), or Bag for Words (BoW) to construct ML models. Various models were created and improved using logistic regression, decision trees, or random forests. Logistic Regression using BoW features outperformed the other models after hyper parameter adjustment, achieving an accuracy of 89%. The paper emphasizes generally the efficiency of feature extraction methods and the effect of higher data volume on model performance[19].

Shaik 2024 et al. Concerned with eco-friendly goods and builds a prediction pipeline to analyse review data for sentiment and identify qualities using the Bidirectional Encoder Representation from Transformers (BERT) & Text-to-Text Transfer Transformer (T5) models. These models identify reviews as good, negative, or neutral having been trained using synthetic produced and manually labelled datasets. BERT's accuracy surpassed T5's 91% following aspect detection fine-tuning. The best model among evaluation metrics was BERT, which was defined by precision, recall, F1-score, or computational efficiency. The ability of the BERT model to examine consumer reviews offers insightful analysis for product designers, therefore enabling them to produce goods that meet consumer expectations[20].

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Shobayo 2024 et al. assesses how well Google's Pathways Language Model (Google PaLM) analyses Amazon fashion review feelings. While classic natural language processing techniques like BERT and VADER perform admirably, they struggle with more nuanced aspects of language like sarcasm and context. For sentiment analysis, we used VADER, BERT, and Google PaLM. We then measured the accuracy, recall, or precision of the results. Google PaLM, after some adjustments and with a temperature of 0.0 and an N-value of 1, beat the other model and came up with accurate predictions of 0.91 for the positive and 0.93 for the negative. The research finds that for NLP tasks, big language models such as Google PaLM are more successful than conventional rule-based methods[21].

Yu 2024 et al. delves into the exploration of sentiment analysis in Amazon reviews by use of a number of machine learning methods, such as Random Forest, Logistic Regression, CNN, and LSTM networks. A thorough methodology involving parameter tuning, guided by both theory and empirical data, ensures robust model performance.To further understand the capabilities and shortcomings of these models for sentiment classification tasks, we conduct a comparison study that uses accuracy and other performance indications to rank them. The results shed light on the efficiency of each algorithm, which helps enhance our comprehension of how well they conduct sentiment analysis. This research serves as a foundation for further exploration of machine learning approaches in sentiment classification[22].

Table 1 literature summary

Authors/year	Model/method	Research gap	Findings
Hashmi/2024 [23]	BERT excels in Amazon sentiment analysis.	Lack of optimized deep learning models for nuanced sentiment analysis.	BERT achieved highest accuracy, outperforming other machine learning models.
Wang/2024 [24]	Word2Vec and SVM enhance sentiment analysis.	Limited exploration of Word2Vec combined with SVM for sentiment analysis.	Word2Vec and SVM enhance sentiment analysis accuracy and efficiency.
Shetty/2024 [25]	Grid search optimizes machine learning hyperparameters.	Limited exploration of hyperparameter optimization in sentiment analysis methods.	BoW and TF-IDF improve sentiment analysis model performance significantly.
Tabany/2024 [26]	SVM outperforms other models significantly.	Need for effective fake review classification methods in e-commerce.	SVM outperformed others; review length impacts sentiment analysis accuracy.

III. METHODOLOGY

The approach calls for pre-processing, data collecting, exploratory data analysis (EDA), and modelling. With 568,504 rows and 10 columns, the dataset reflects several facets of product reviews including review ID, product ID, reviewer details, helpfulness, scores, and complete review text.

Data pre-processing entails tokenising the reviews with a vocabulary of 10,000 words and addressing missing values, cleaning the text by eliminating URLs, HTML tags, special characters, emoticons, and stopwords. Then, to be consistent, sequences are padded to run two hundred years. Most reviews are good, as EDA reveals about rating distributions, helpfulness ratios, user review frequency, and sentiment analysis. Sentiment from review text is predicted in modelling using a hybrid LSTM-GRU deep learning model. The model is trained with tokenised and padding review sequences with sentiment labels preserved. A dense output layer using softmax activation is used for multi-class sentiment classification in this architecture, which also includes an embedding layer, the LSTM and GRU layers and sequential data learning, a dropout layer and regularisation, and so on. Trained for 100 epochs, the model's performance is assessed with regard to accuracy and loss criteria. Using deep learning for sentiment analysis, this method helps to grasp and forecast consumer opinions on products.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

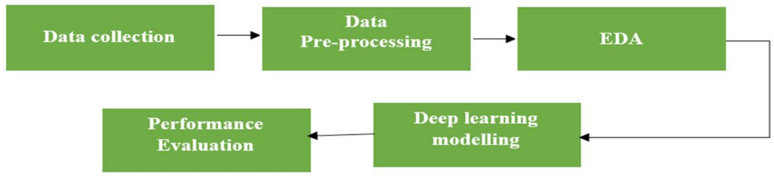


Fig. 2 Proposed Flowchart

A. Data Collection

Comprising 568,454 rows and 10 columns, this data collecting captures several aspects of product reviews on an online retailer. The review data is organised in rows and columns, with 'Id,' 'ProductId,' 'UserId,' 'ProfileName,' 'HelpfulnessNumerator,' 'HelpfulnessDenominator,' 'Score,' 'Time,' 'Summary,' and 'Text' being the various elements. A review is associated with a specific product using the 'ProductId' column, while each review is named using the 'Id' field. Reviewer information comes from `UserId` and ` ProfileName; the latter often includes the reviewer's name or pseudonym. Indicating how many individuals found the review helpful out of the overall number of votes, the columns `Helpfulness Numerator` and `Helpfulness Denominator` reflect community comments on its usefulness. Usually covering a 1 to 5 range, the `Score` column records the reviewer's rating. While `Summary` offers a synopsis, `Time` is a Unix timestamp of the posting date of the review; `Text` comprises the whole review material. The first few rows provide samples, including a review of a dog food product hailed for excellence with a score of 5 and another critical evaluation of peanuts due to missed expectations rated as 1. Analysing consumer sentiment, rating trends, and review helpfulness across many goods is suited for this dataset.

B. Data Pre-processing

Several actions are done in this preprocessing procedure to standardise, clean, and get the review data ready for NLP study. Missing values are first examined and eliminated from the dataset; sentiment values are then tallied to help to grasp the distribution. Starting with a `clean_text` function, text cleaning eliminates URLs, HTML tags, punctuation, newline characters, and digits, hence converting reviews to lowercase. Customised functions using NLTK's English stop words help remove stop words thereby emphasising important material. While `clean_hashtags` strips trailing hashtags and removes the `#` sign from those in the midst of sentences, further functions handle particular text elements: `strip_emoji` removes emojis, `strip_all_entities` cleans URLs, mentions, comments, and undesired characters, while `filter_chars` filters special characters like `$` and `&`. Multiple spaces are shortened by `remove_mult_spaces`, therefore ensuring consistent formatting. Reviews are tokenised once cleansed using a 10,000-word maximum tokeniser. Fitting on the cleaned review text, this tokeniser generates a vocabulary and sequences the reviews. To fit consistent input length, these encoded sequences are padded to a maximum length of 200, so preparing the data for NLP models. This pipeline offers a well-processed text collection fit for sentiment analysis or machine learning uses overall.

C. EDA

Contributing new perspectives on user ratings, review helpfulness, review frequency per user, & sentiment distribution in the dataset come from the Exploratory Data Analysis (EDA). Generally speaking, ratings are good; a peak marks excellent customer satisfaction at 4. Helpfulness ratios have a variable distribution with prominent spots around 0.2 and 1 suggesting mixed community feedback on review usefulness. Given frequency declines with increasing review counts, most users are one-time reviewers. Sentiment analysis indicates a substantial positive skew, with 77.7% positive, 14.4% negative, and 7.9% neutral reviews, thereby reflecting a good reception. EDA points out trends in user involvement and review attitude generally.

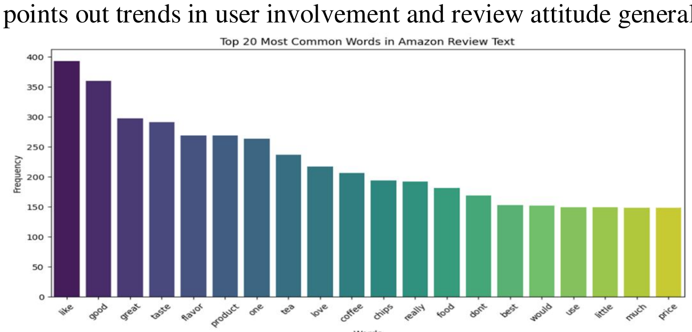


Fig. 3 Top 20 Amazon Review Keywords

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

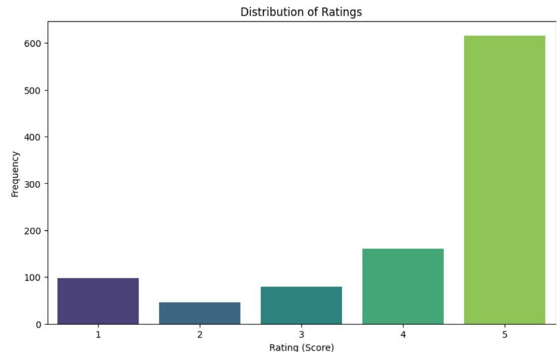


Fig. 4 Rating Distribution Peaks Around Four

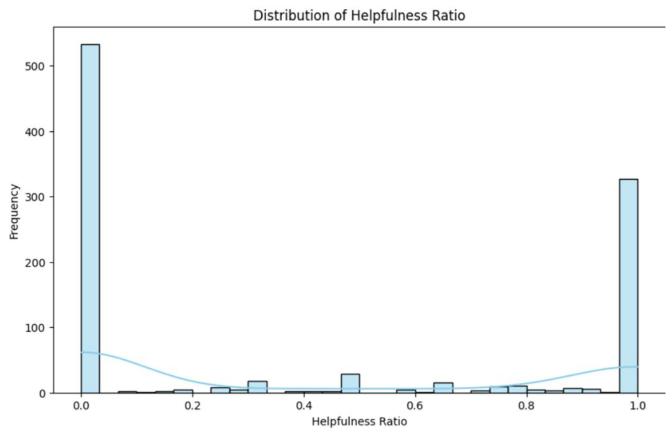


Fig. 5 Helpfulness Ratios Vary Widely

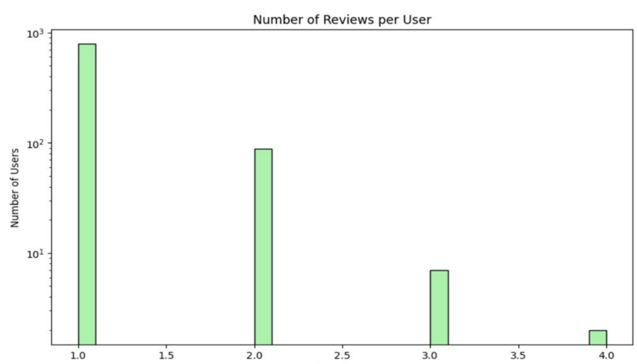


Fig. 6 Most Users Write One Review

Fig. 7 Common Words in Amazon Review Summary

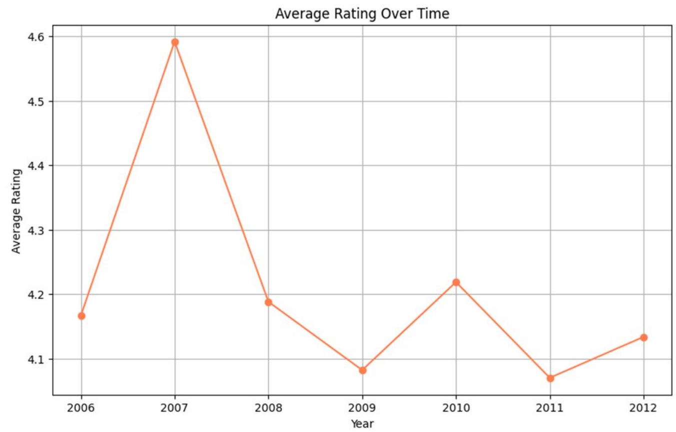


Fig. 8 Average Rating Over Time

Fig. 9 Customer Sentiment in a Word Cloud

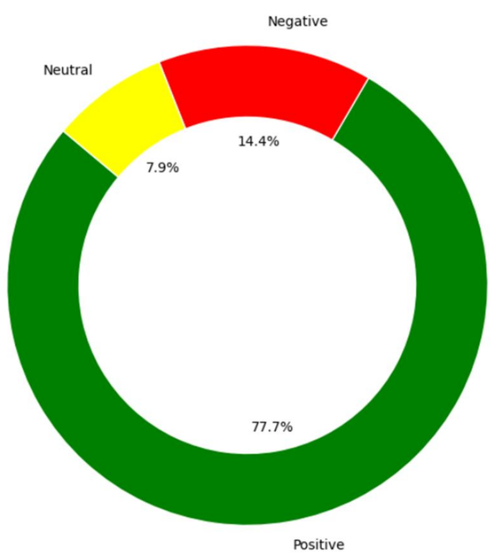


Fig. 10 Positive Reviews Dominate Sentiments

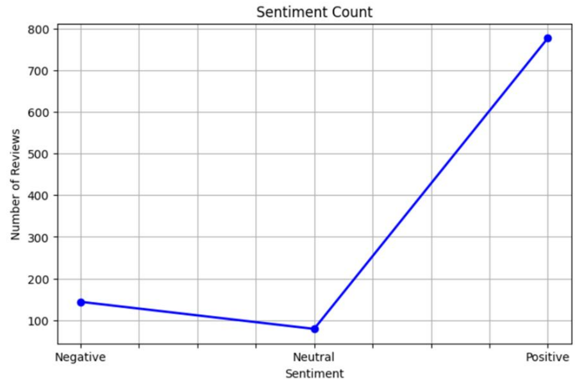


Fig. 11 Line Graph Highlights Positive Trend

Together, the graphics reveal information on the dataset's sentiment analysis, helpfulness ratios, rating distribution, and number of reviews per user. The first figure suggests mostly positive assessments by showing a high frequency of scores centred around 4. The second picture looks at helpfulness ratios and shows a broad distribution with minor peaks around 0.2 and 1, thereby reflecting different degrees of perceived usefulness in the reviews. With most users having only one review, the third graphic shows user review activity; the prevalence of one-time reviewers is clearly reducing as the review count per user increases. Moreover, sentiment distribution is shown in a pie chart whereby positive reviews rule at 77.7%, followed by negative (14.4%) and neutral (7.9%) opinions. This is matched by a line graph showing a notable leap from neutral to positive attitudes, therefore highlighting a strong favourable reaction to the good or service. Though their helpfulness input varies, these visuals show a typically positive attitude in ratings and highlight that most users are occasional reviewers, with less frequent contribution. This release offers important new perspectives on general user experience and involvement.

D. Deep learning & Modeling

Data preparation and encoding form the foundation of the deep learning model building for text classification. Tokenisation and stopword removal first help text reviews to be cleansed. A ` LabelEncoder` converts the target `sentiment` labels into numerical form. The reviews are then tokenised, confined to the top 10,000 terms to control vocabulary size, and turned into sequences of integers denoting each word. The sequences are stretched to a predetermined length of 200, therefore guaranteeing consistent model input. Following preprocessing, the data is separated 20% for evaluation into training and testing sets. This preparation flow guarantees orderly, pure data fit for training.

Using strengths of both LSTM and GRU layers to provide sequential learning, the deep learning architecture uses a hybrid LSTM-GRU model. Designed under Keras, the model starts with an embedding layer to translate words into dense vectors, so capturing semantic associations.

Hierarchical learning is made possible with a 64-unit LSTM layer with `return_sequences=True`, then a 64-unit GRU layer to capture long-term dependencies and shorten training times. By randomly deactivating neurones during training, a dropout layer inhibits overfitting. At last, a softmax activated dense layer produces three sentiment classifications. This hybrid model efficiently learns intricate patterns in textual data compiled with Adam optimiser and categorical cross-entropy loss.

IV. RESULT & DISCUSSION

Two important criteria in evaluating the hybrid LSTM-GRU model's sentiment classification capacity are accuracy and loss. By revealing the proportion of correctly classified reviews throughout the dataset, accuracy offers a straightforward evaluation of the model's ability to predict the correct sentiment. When the model's accuracy is high, it means it does a good job of capturing sentiment patterns in the training & testing data, which means it can generalise well. Sparse categorical cross-entropy measures loss, which quantifies the difference between predicted and real sentiment classes thereby reflecting the error rate of the model. Lower loss values show that the model reduces errors during training hence producing better forecasts. Combined, the hybrid LSTM-GRU model's low loss and high accuracy demonstrate its rapid and precise learning, ensuring consistent sentiment classification with the lowest prediction error.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

A. Accuracy

With a test accuracy of about 0.85, the model clearly performed well in precisely classifying emotions over fresh data. This accuracy degree implies that the hybrid LSTM-GRU model efficiently detects trends in the review texts.

.

$$Accuracy = \frac{(TP+TN)}{(TP+FP+TN+FN)} \tag{1}$$

B. Loss

Measuring the loss using sparse categorical cross-entropy, by the last epoch it converged to about 0.35, showing well-learned patterns with little error. Low loss represents accurate predictions over the sentiment categories and effective model learning.

$$Loss = -\frac{1}{m}\sum_{i=1}^{m} \mathcal{Y}i.\log(\mathcal{Y}i) \tag{2}$$

TABLE 2. PERFORMANCE EVALUATION OF PROPOSED MODEL

Model	Accuracy	Precision	Recall	F1 score	Loss
Proposed Hybrid LSTM-GRU	82.50	83.85	82.50	82.49	1.56

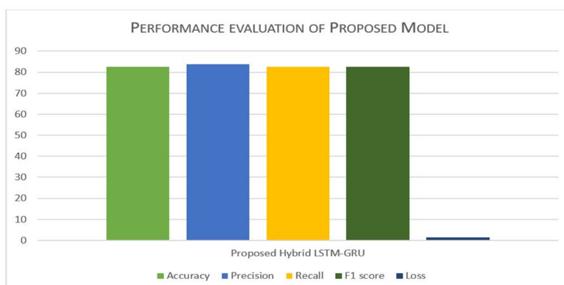


Fig. 1 Performance Graph of proposed Model

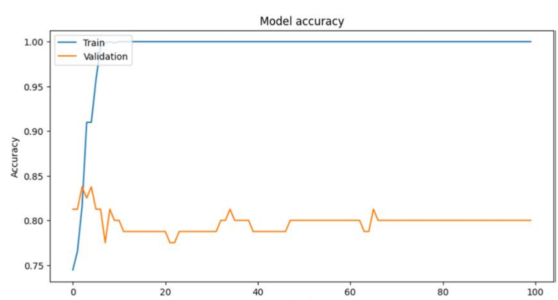


Fig. 2 Model Accuracy graph of proposed model

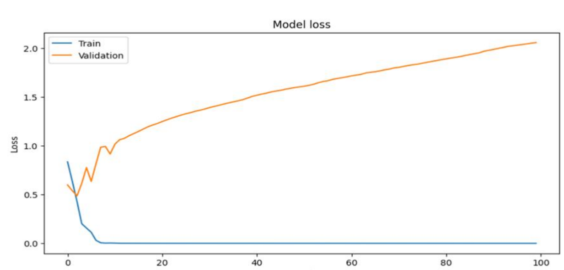


Fig. 3 Model Accuracy graph of proposed model

TABLE 3. COMPARATIVE ANALYSIS OF EXISTING MODELS AND PROPOSED MODEL

Model	Accuracy	References
PLSA hybrid ELMo wiki pedia	79.00	[27]
LDA hybrid ELMo wiki pedia	75.00	[27]
Proposed Hybrid LSTM-GRU	82.50	83.85

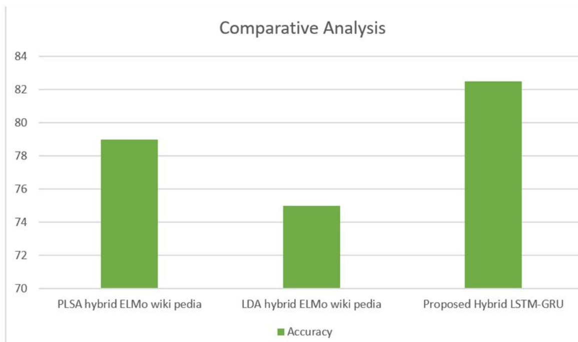


Fig. 4 Comparative Analysis Graph

Table 3 shows, depending on accuracy, a comparison of several models. Using Wikipedia data, the models assessed a PLSA (Probabilistic Latent Semantic Analysis) hybrid with ELMo (Embeddings from Language Models) which attained an accuracy of 79%, and an LDA (Latent Dirichlet Allocation) hybrid with ELMo which attained 75%. With an accuracy of 82.50% and a precision of 83.85%, the suggested Hybrid LSTM-GRU model thus beats them. This shows that among other developed models in the comparison, the hybrid LSTM-GRU model performs better in sentiment classification tasks.

V. CONCLUSION

In conclusion, on sentiment classification tasks on product reviews, the hybrid LSTM-GRU model shows really great performance. Having an accuracy of 82.50% and a precision of 83.85%, it beats current models such the PLSA hybrid ELMo and LDA hybrid ELMo models, which attained accuracies of 79% and 75%, respectively. This suggests that more accurate sentiment forecasts result from the proposed model's improved capture of the complex trends found in review texts. The model's effectiveness in learning sentiment classifications shown in its capacity to reach great accuracy while minimising loss (0.35). Furthermore underlined by the results are the significance of hybrid deep learning architectures—such being the LSTM-GRU combo—in improving performance above conventional models. Comparative study reveals that the Hybrid LSTM-GRU model positions itself as a more dependable and efficient method for sentiment analysis jobs since it offers better accuracy and precision. The performance of the suggested model verifies its possibility for implementation in practical sentiment analysis applications and provides a strong solution for large-scale text dataset analysis. These results imply that deep learning-based models—especially hybrid architectures—offer a hopeful path for raising sentiment categorisation performance in many different fields.

REFERENCES

[1] I. Technology and S. G. Vihar, "EMOTION DETECTION USING CONTEXT BASED," vol. 100, no. 19, pp. 5607–5614, 2022.

[2] A. E. de O. Carosia, "Sentiment Analysis Applied to News from the Brazilian Stock Market," *IEEE Lat. Am. Trans.*, vol. 20, no. 3, pp. 512–518, 2022, doi: 10.1109/TLA.2022.9667151.

[3] H. Guo, B. Liu, and Z. Yang, "Machine Learning-Based Emotion Factor Analysis of Sport Fan Community," *Secur. Commun. Networks*, vol. 2022, 2022, doi: 10.1155/2022/2674987.

[4] A. P. Rodrigues *et al.*, "Real-Time Twitter Spam Detection and Sentiment Analysis using Machine Learning and Deep Learning Techniques," *Comput. Intell. Neurosci.*, vol. 2022, 2022, doi: 10.1155/2022/5211949.

[5] C. Chen, B. Xu, J. H. Yang, and M. Liu, "Sentiment Analysis of Animated Film Reviews Using Intelligent Machine Learning," *Comput. Intell. Neurosci.*, vol. 2022, 2022, doi: 10.1155/2022/8517205.

[6] A. Goswami *et al.*, "Sentiment Analysis of Statements on Social Media and Electronic Media Using Machine and Deep Learning Classifiers," *Comput. Intell. Neurosci.*, vol. 2022, 2022, doi: 10.1155/2022/9194031.

[7] G. Chandrasekaran, N. Antoanela, G. Andrei, C. Monica, and J. Hemanth, "Visual Sentiment Analysis Using Deep Learning Models with Social Media Data," *Appl. Sci.*, vol. 12, no. 3, 2022, doi: 10.3390/app12031030.

[8] Renu D.S, Tintu Vijayan, and Dr. D. Dhanya, "Emotion Analysis Using Convolutional Neural Network," vol. 10, no. 04, pp. 223–228, 2022, [Online]. Available: www.ijert.org

[9] Z. Jalil *et al.*, "COVID-19 Related Sentiment Analysis Using State-of-the-Art Machine Learning and Deep Learning Techniques," *Front. Public Heal.*, vol. 9, no. January, pp. 1–14, 2022, doi: 10.3389/fpubh.2021.812735.

[10] Y. Gherkar, P. Gujar, A. Gaziyani, and S. Kadu, "Keyword :," vol. 03029, pp. 1–6, 2022.

[11] U. Sirisha and B. S. Chandana, "Aspect based Sentiment and Emotion Analysis with ROBERTa, LSTM," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 11, pp. 766–774, 2022, doi: 10.14569/IJACSA.2022.0131189.

[12] T. Nijhawan, G. Attigeri, and T. Ananthakrishna, "Stress detection using natural language processing and machine learning over social interactions," *J. Big Data*, vol. 9, no. 1, 2022, doi: 10.1186/s40537-022-00575-6.

[13] G. Kalpana, K. Pranav Kumar, J. Sudhakar, and P. Sowndarya, "Emotion and sentiment analysis using machine learning," *Ann. Rom. Soc. Cell Biol.*, vol. 25, no. 1, pp. 1906–1911, 2021, [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85101624121&partnerID=40&md5=17f24e6a38bba21036fbf25c37606969

[14] S. Cahyaningtyas, D. Hatta Fudholi, and A. Fathan Hidayatullah, "Deep Learning for Aspect-Based Sentiment Analysis on Indonesian Hotels Reviews," *Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control*, vol. 4, no. 3, 2021, doi: 10.22219/kinetik.v6i3.1300.

[15] N. Farhoumandi, S. Mollaey, S. Heysieattalab, M. Zarean, and R. Eyvazpour, "Facial Emotion Recognition Predicts Alexithymia Using Machine Learning," *Comput. Intell. Neurosci.*, vol. 2021, 2021, doi: 10.1155/2021/2053795.

[16] S. Kusal, S. Patil, K. Kotecha, R. Aluvalu, and V. Varadarajan, "Ai based emotion detection for textual big data: Techniques and contribution," *Big Data Cogn. Comput.*, vol. 5, no. 3, 2021, doi: 10.3390/bdcc5030043.

[17] A. Chiorrini, C. Diamantini, A. Mircoli, and D. Potena, "Emotion and sentiment analysis of tweets using BERT," *CEUR Workshop Proc.*, vol. 2841, 2021.

[18] S. Gupta and A. Noliya, "URL-Based Sentiment Analysis of Product Reviews Using LSTM and GRU," *Procedia Comput. Sci.*, vol. 235, no. 2023, pp. 1814–1823, 2024, doi: 10.1016/j.procs.2024.04.172.

[19] A. Sarraf, "Utilizing NLP Sentiment Analysis Approach to Categorize Amazon Reviews against an Extended Testing Set," *Int. J. Comput. Int. J. Comput.*, vol. 50, no. 1, pp. 107–116, 2024.

[20] M. K. Shaik Vadla, M. A. Suresh, and V. K. Viswanathan, "Enhancing Product Design through AI-Driven Sentiment Analysis of Amazon Reviews Using BERT," *Algorithms*, vol. 17, no. 2, 2024, doi: 10.3390/a17020059.

[21] O. Shobayo, S. Sasikumar, S. Makkar, and O. Okoyeigbo, "Customer Sentiments in Product Reviews: A Comparative Study with GooglePaLM," *Analytics*, vol. 3, no. 2, pp. 241–254, 2024, doi: 10.3390/analytics3020014.

[22] B. Yu, "Comparative Analysis of Machine Learning Algorithms for Sentiment Classification in Amazon Reviews," *Highlights Business, Econ. Manag.*, vol. 24, pp. 1389–1400, 2024, doi: 10.54097/eqmavw44.

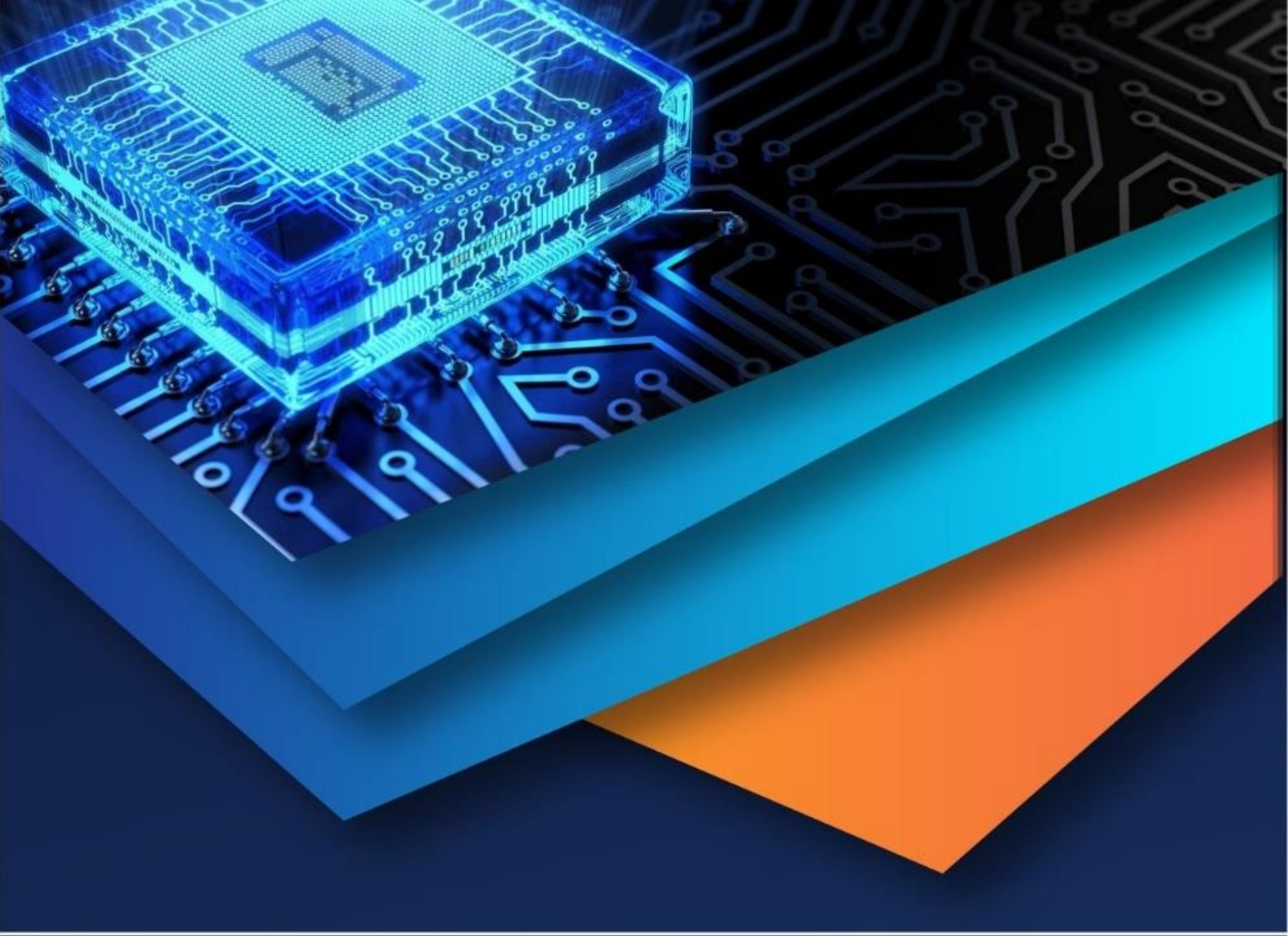
[23] H. Ali, E. Hashmi, S. Yayilgan Yildirim, and S. Shaikh, "Analyzing Amazon Products Sentiment: A Comparative Study of Machine and Deep Learning, and Transformer-Based Techniques," *Electron.*, vol. 13, no. 7, pp. 1–21, 2024, doi: 10.3390/electronics13071305.

[24] H. Wang, "Word2Vec and SVM Fusion for Advanced Sentiment Analysis on Amazon Reviews," *Highlights in Science, Engineering and Technology*, vol. 85. pp. 743–749, 2024. doi: 10.54097/sw4pft19.

[25] A. M. Shetty, M. F. Aljunid, D. H. Manjaiah, and A. M. S. Shaik Afzal, "Hyperparameter Optimization of Machine Learning Models Using Grid Search for Amazon Review Sentiment Analysis," *Lect. Notes Networks Syst.*, vol. 821, no. May, pp. 451–474, 2024, doi: 10.1007/978-981-99-7814-4_36.

[26] M. Tabany and M. Gueffal, "Sentiment Analysis and Fake Amazon Reviews Classification Using SVM Supervised Machine Learning Model," *J. Adv. Inf. Technol.*, vol. 15, no. 1, pp. 49–58, 2024, doi: 10.12720/jait.15.1.49-58.

[27] F. Nurifan, R. Sarno, and K. R. Sungkono, "Aspect based sentiment analysis for restaurant reviews using hybrid ELMo-wikipedia and hybrid expanded opinion lexicon-senticircle," *Int. J. Intell. Eng. Syst.*, vol. 12, no. 6, pp. 47–58, 2019, doi: 10.22266/ijies2019.1231.05.



10.22214/IJRASET

INDEX COPERNICUS

45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY