IJARETY

International Journal of Advanced Research in Education and TechnologY (IJARETY)

Volume 11, Issue 3, May-June 2024

Impact Factor: 7.394

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

crossref

निस्केयर NISCAIR

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 7.394| A Bi-Monthly, Double-Blind Peer Reviewed & Referred Journal |

|| Volume 11, Issue 3, May-June 2024 ||

DOI:10.15680/IJARETY.2024.1103038

Sentiment-Based Rating Forecasting with Bert and LSTM

Mr. Kolisetty Balaji, Mr. Arjun Kuruva, Dr. D. William Albert

M. Tech Student, Department of Computer Science and Engineering, Bheema Institute of Technology and Science, AP, India

Associate Professor, Department of Computer Science and Engineering, Bheema Institute of Technology and Science, AP, India

Associate Professor & Head of the Department, Department of Computer Science and Engineering, Bheema Institute of Technology and Science, AP, India

ABSTRACT: Online reviews have become an increasingly important factor of the decision-making process is online buying has become more and more popular in recent years. Reviews not only give prospective consumers useful information, but they also foster trust and give businesses insightful feedback. We thoroughly analysed the Amazon reviews dataset, which spans many product categories, for our study. Our main goal was to use deep learning, machine learning, ensemble learning, and natural language processing to reliably categorise feelings. Our study process included a number of important phases. We examine the approaches for gathering data, preprocessing stages such as tokenization and normalization, and feature extraction using the TF-IDF and Bag-of-Words techniques. A range of machine learning methods, such as Multinomial Naive Bayes, Random Forest, Decision Tree, and Logistic Regression, were utilized in our research. Furthermore, we used Bagging as a method for group learning. Additionally, we investigated transformer-based models like XLNet and BERT as well as deep learning-based algorithms like CNNs and Bidirectional LSTM. Using criteria like accuracy, precision, recall, and F1 score, we conducted extensive tests and found that the BERT algorithm performed better than the others, obtaining an astounding accuracy rate of 99%..

KEYWORDS: Convolution Neural Network, Long Short Term Memory, Deep Learning, Sentiment Analysis

I.INTRODUCTION

Online shopping is becoming more and more popular; according to a recent estimate, 2.64 billion people will be shopping online by 2023 [1]. User-generated content has increased dramatically since the advent of the internet, which has made it possible for people to voice their thoughts and participate in conversations on a variety of platforms, including blogs, forums, e-commerce websites, and online social networks. There is an enormous amount of user-generated data as a result of this development [2,3]. From these data, people, organisations, and governments must extract and apply pertinent information. The increasing amount of these data emphasises the difficulty of quickly and effectively gathering pertinent information, underscoring the necessity of computational linguistic methods in data analysis [4].

Online shoppers frequently base their judgements on the opinions of previous customers. Reviews offer insightful information about the calibre of goods and services as well as the consumer experience. Good evaluations assist companies in growing their clientele and gaining credibility. Another benefit of negative reviews is that they may provide businesses input on how to make their goods and services better. In general, companies and customers may benefit greatly from the use of customer reviews [5]. They assist companies in enhancing their goods and services and assist customers in making wise purchases. Sentiment analysis (SA) involves the technique of figuring out what a text's emotions or ideas are. This may be done for social network postings, reviews from customers, and other content [6, 7]. SA makes it possible for you to learn from product evaluations by highlighting neutral, positive, and negative feelings in the comments. You may utilise this knowledge to better your product, pinpoint areas that need work, and learn what your target market wants. SA has a difficult time identifying the various emotions expressed in customer feedback. This is due to the fact that various individuals express themselves in different ways and that terms such as "great" or "bad" can have several meanings depending on the context. This might make it challenging to get precise information about client satisfaction, which makes it challenging to make product improvements. Being able to swiftly evaluate a large number of user evaluations is essential in the digital world of today, where they play a significant role. Emotions that

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 7.394| A Bi-Monthly, Double-Blind Peer Reviewed & Referred Journal |

|| Volume 11, Issue 3, May-June 2024 ||

DOI:10.15680/IJARETY.2024.1103038

lie halfway between positive and negative are particularly difficult for current technologies to accurately capture. For this reason, in order for a system to provide accurate forecasts and continue to function well over time, it must be able to comprehend emotions in context.

II.LITERATURE REVIEW

Y. Abbas et al. [8] used a dataset of reviews from Amazon spanning four product categories to create a supervised machine learning algorithm to identify faulty items in online reviews. They used statistical methods for feature selection, concentrating on characteristics like emotional tone and emotions, such as correlation [9] and information gain [10]. With an accuracy score of 0.84, the Random Forest classifier produced its best results. In the meanwhile, a hybrid machine learning approach was put out by C. Ahmed et al. [11] to study customer-service provider interactions with the goal of forecasting sentiment shifts. They looked at five thousand Twitter exchanges. D. Suhartono et al. [12] used weighted word embeddings and Deep Neural Networks (DNNs) to investigate a three-category classification problem in SA of medicinal product reviews. They combined the two word embedding techniques, Word2Vec [14] and GLOVE [13], then processed the embeddings using several convolutional neural network (CNN) layers. When K. S. Kumar et al. [15] used machine learning models to mine Amazon reviews for three products—the Redmi Note 3, the Samsung J7, and the Apple iPhone 5S—they found that Naïve Bayes (NB) performed better than LR in classifying reviews either positive or negative based on recall, accuracy, and F-measure. By combining CNNs with word embeddings, M. Qorich et al. [16] addressed text sentiment analysis on reviews from Amazon. They used three distinct datasets in their study to test two-word embedding methods, Word2Vec and FastText [17]. In every dataset, the CNN model they created outperformed the pre-established baselines and performed better than conventional ML- and DL-based techniques. Stanford University's Xu Yun et al. [18] predicted Yelp review scores using supervised learning methods, including the Support Vector Machine (SVM), NB, and Perceptron algorithm. They employed a 70–30% training–testing divide for cross-validation and tested a number of classifiers to get the values of accuracy and recall. From a pool of 21,500 reviews on Amazon, 3000 English reviews were chosen at random by A. S. Rathor et al. [19]. Preprocessing was done on these evaluations, and repetitive letters were eliminated. In order to create a business model, M. S. Elli et al. [20] extracted sentiment from reviews and used MNB and SVM as their main classifiers, resulting in high accuracy. In their analysis of 300 Amazon reviews of electronic devices, A. Cernian et al. [21] used SentiWordNet for phrase-to-word vector conversion and evaluated CNNs, Word2Vec, LR, SGD, and NB using different feature extraction techniques; Lime was utilised to explain review classifications. CNNs and Word2Vec achieved 91% accuracy. In an attempt to improve readability, M. Nasret al. [22] chose less complex algorithms; yet, they performed very well when paired with SVM, and they could have problems with bigger datasets. W. Tan et al. [23] used a variety of machine learning algorithms, including as NB analysis, SVM, K-Nearest Neighbour (KNN) techniques, and DNNs like RNN, to investigate the relationship between Amazon product reviews and customer ratings. Ultimately, in [24], the ICTCLAS 4 system was used to scan more than 100,000 Chinese clothes product evaluations from Amazon. Using the Weibo tweet texts dataset, Huang et al. [25] focused exclusively on long-range contexts and achieved a 64.1% accuracy rate in sentiment categorization. They used a Hierarchical LSTM network to do this. Unigram features were used in combination of the Naive Bayes method by Hasan et al. [26]. They used SentiWordNet, Word Sense Disambiguation, TextBlob, and other sentiment analyzers to apply this method to translated Urdu twitter data. Their efforts paid off, as they were able to classify sentiment binary with 79% accuracy. A CNN model with two convolution layers and two sequential pairs of pooling layers is included by Deriu et al. [27]. With an F1-score of 67.79%, this CNN model was used to categorise multilingual sentiment datasets made out of tweet data. Jin et al.'s [28] integration of BERT embeddings with a modified TF-IDF model improved the BERT-based model in a multi-label classification challenge. They used a dataset of restaurant patron ratings to conduct an evaluation, and they obtained a 64% accuracy rate. A system with three sets of convolutional and pooling layers that integrates Word2Vec with a CNN model was presented by Ouyang et al. [29]. Using this system, they classified five labels from the corpus of the MR dataset with an extremely fine sentiment level of 45.4%.

III.METHODOLOGY

Our approach comprised a number of crucial phases. To begin, we imported the BERT model and divided the evaluations into manageable tokens using its tokenize. Then, in order to handle the data effectively, we set up a PyTorch DataLoader that was set up for batch processing with a maximum number of batches of 32. We used the AdamW optimizer, which was created especially to deal with weight decay problems and improve neural network training, to optimize our model. The training parameter changes were controlled by setting the learning rate at 2×10^{-5}. We then used our training dataset to train the model for three epochs, and the result was an accuracy of 89%

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 7.394| A Bi-Monthly, Double-Blind Peer Reviewed & Referred Journal |

|| Volume 11, Issue 3, May-June 2024 ||

DOI:10.15680/IJARETY.2024.1103038

An essential stage in NLP activities is data preparation, which maximizes the effectiveness of knowledge discovery. It includes methods like as reduction, integration, cleansing, and transformation of data with the goal of enhancing and preparing the dataset for more efficient analysis [33]. All of the preprocessing stages that were done for this study are shown in Figure 1

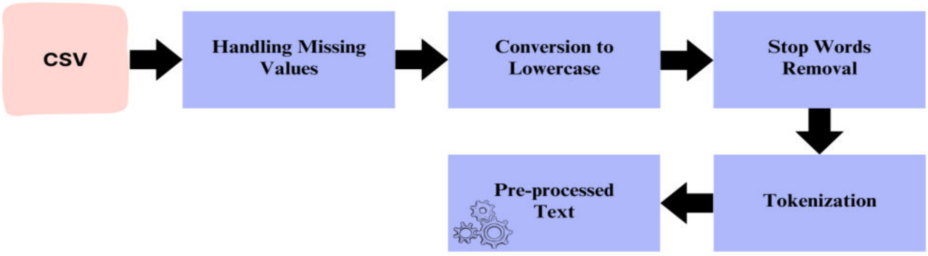


Figure 1: Data Pre processing.

Handling Missing Values:
Our main concern while dealing with the missing values in the dataset is how to handle the missing entries in the "review body" and "star rating" features. The importance of these characteristics in connection to feelings and the results that follow is the reason for this emphasis.

Lowercase Conversion
We change all review words to lowercase in this phase. For instance, "amazing" and "great" become "great" and "amazing." Lowercasing treats words case-insensitively, which helps to standardise the text and decreases the dimensionality of the data?

Removal of Stop Words
Stop words are now words or phrases that are considered irrelevant in all text mining domains. All HTML elements, punctuation, and stop words have been removed from the reviews in our corpus. This preparation stage enhances computing performance and lowers noise in the data.

Tokenization
Tokenization of both sentences and words was used in our study. Tokenization is the process of dividing a text sequence into discrete elements called tokens. Singular words, phrases, or even whole sentences may be included in these tokens [34, 35]. After that, these tokens are used as inputs for a number of procedures, including text mining and parsing. It assists models in concentrating on the significance of discrete pieces instead of interpreting the text as a whole.

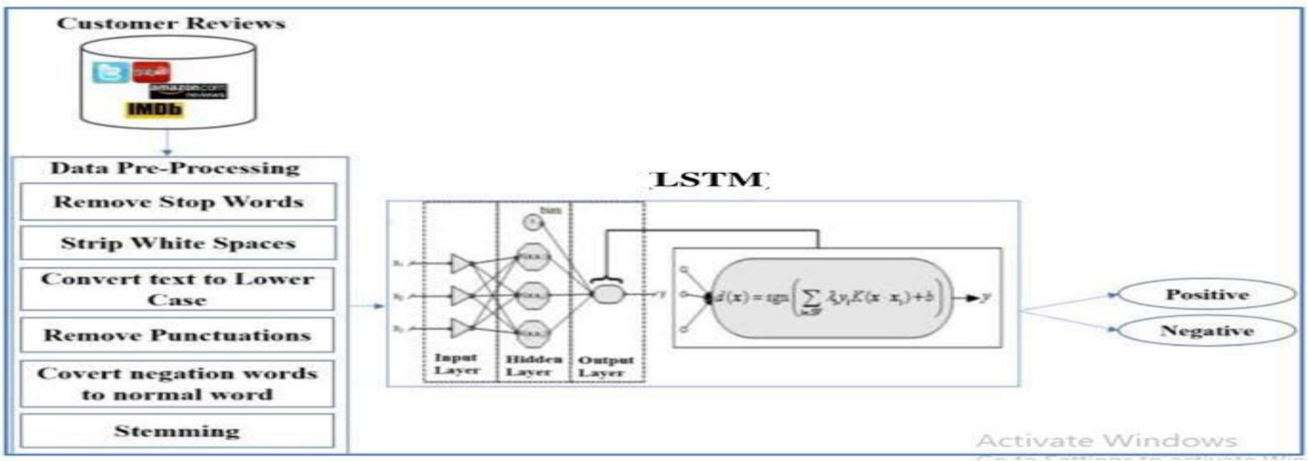


Figure 2: System Architecture

International Journal of Advanced Research in Education and TechnologY(IJARETY)

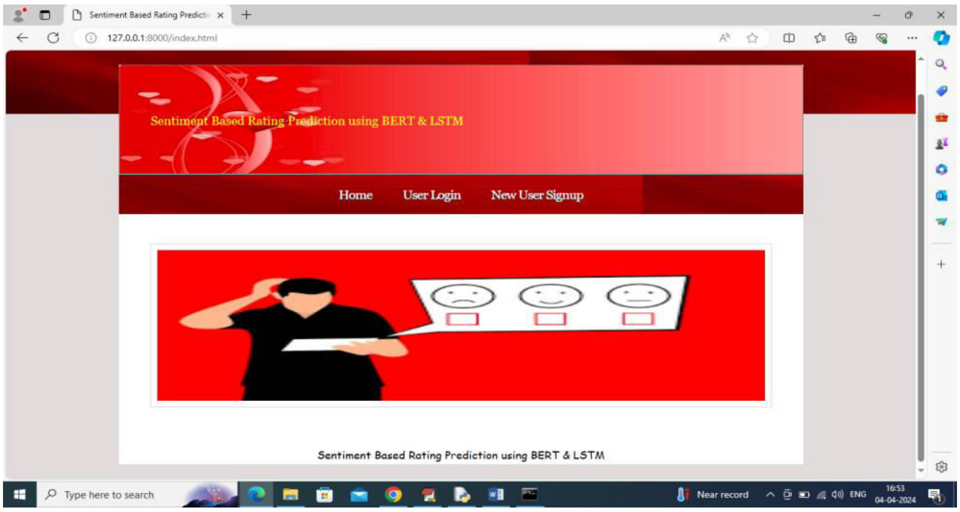
| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 7.394| A Bi-Monthly, Double-Blind Peer Reviewed & Referred Journal |

|| Volume 11, Issue 3, May-June 2024 ||

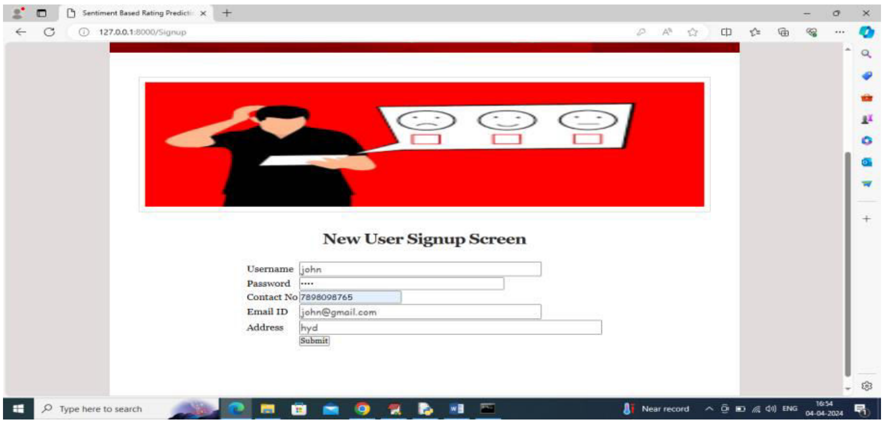
DOI:10.15680/IJARETY.2024.1103038

IV. RESULTS

We used the same fundamental BERT-based neural network architecture as our classifier in all of the trials. The architecture of the classifier used in our tests is shown in Figure 4.4. Textual data is received by the input layer, which is the first layer, and is then sent to the subsequent layer. The pre-processing layer comes next, where the input is converted to numeric token ids and organised into several Tensors that BERT uses as input. The BERT_encoder layer comes next, which generates the embeddings for the whole input review and contains the pre-trained BERT model. Next comes the dropout layer, which helps keep the model from overfitting by arbitrarily choosing which neurons to ignore.



In above screen click on 'New User Signup' link to get below page

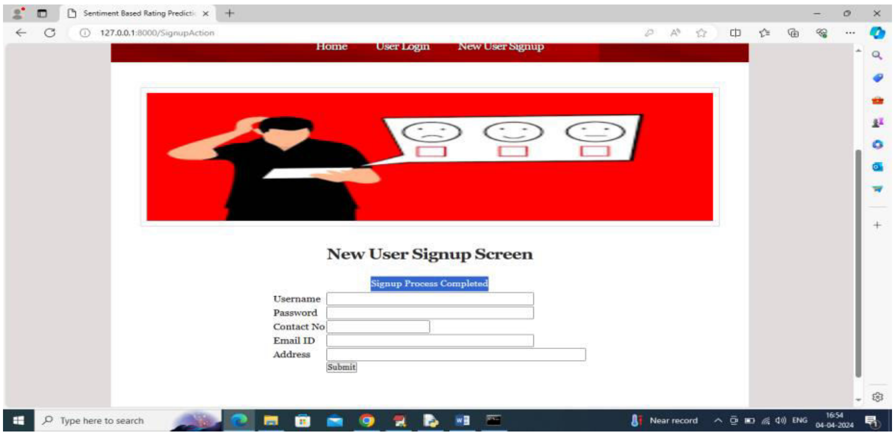


In above screen user is entering sign up details and then press button to get below page

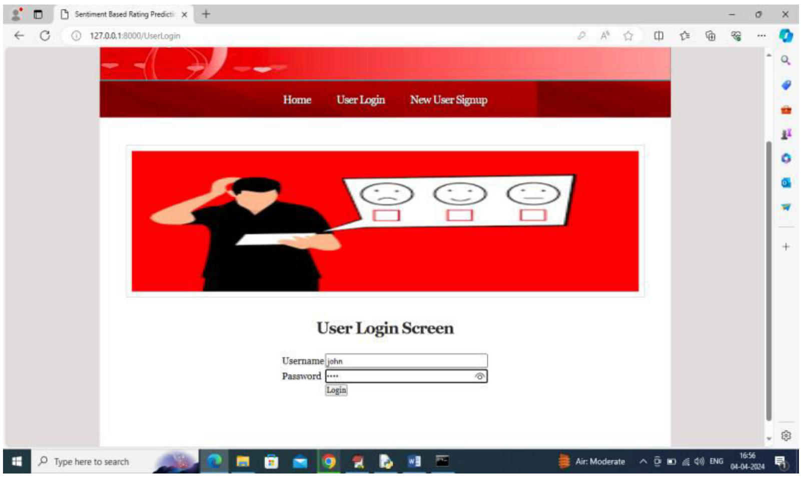
| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 7.394| A Bi-Monthly, Double-Blind Peer Reviewed & Referred Journal |

|| Volume 11, Issue 3, May-June 2024 ||

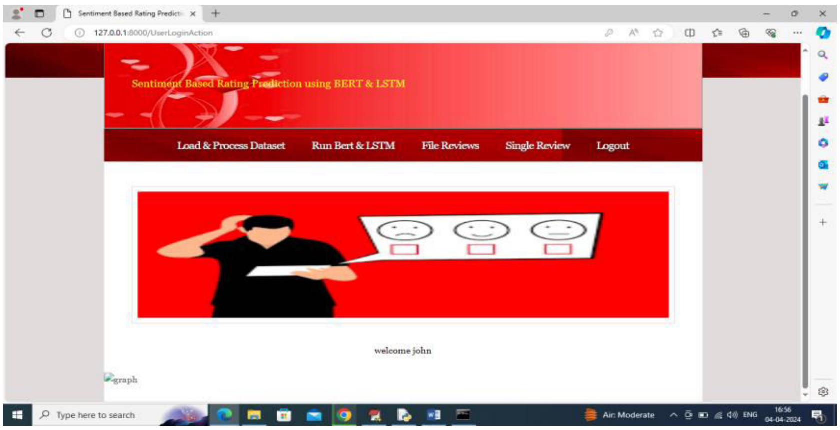
DOI:10.15680/IJARETY.2024.1103038



In above screen sign up completed and now click on 'User Login' link to get below page



In above screen user is login and after login will get below page



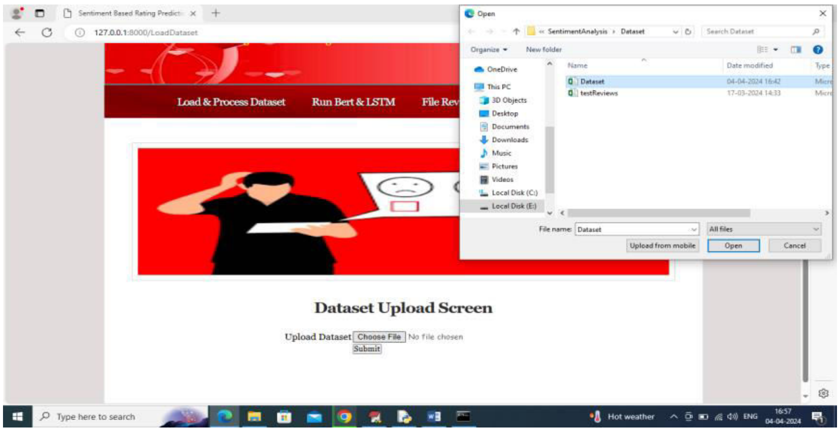
In above screen user can click on 'Load & Process Dataset' link to upload and process data

International Journal of Advanced Research in Education and TechnologY(IJARETY)

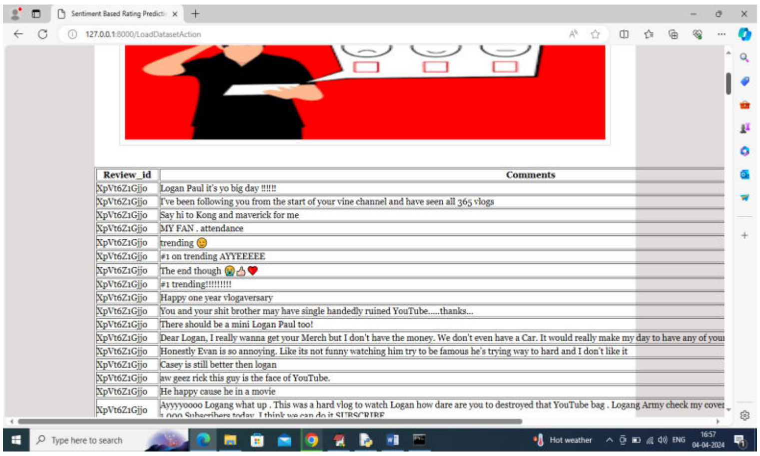
| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 7.394| A Bi-Monthly, Double-Blind Peer Reviewed & Referred Journal |

|| Volume 11, Issue 3, May-June 2024 ||

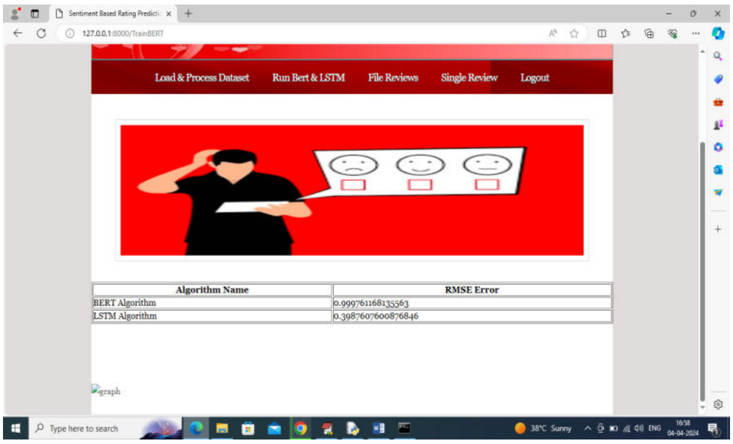
DOI:10.15680/IJARETY.2024.1103038



In above screen selecting and uploading ‘Reviews dataset’ file and then click on ’Open and Submit’ button to load dataset and get below page



In above screen dataset loaded and now click on ‘Train BERT & LSTM Algorithms’ link to train models and get below page

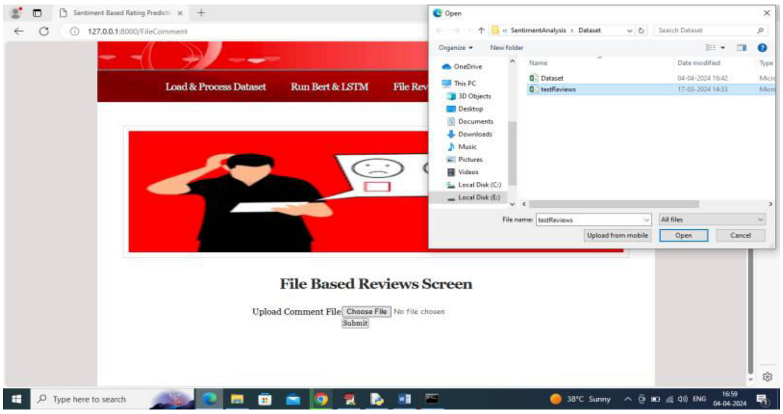


In above screen BERT got 99% and LSTM got 0.39% accuracy and now click on ‘File Reviews’ link to upload test reviews and get below page

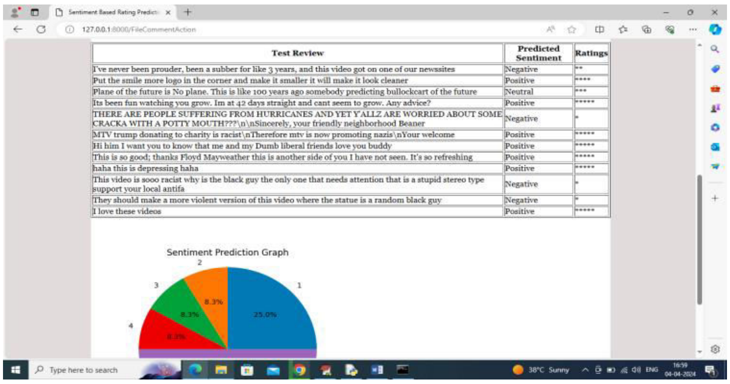
| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 7.394| A Bi-Monthly, Double-Blind Peer Reviewed & Referred Journal |

|| Volume 11, Issue 3, May-June 2024 ||

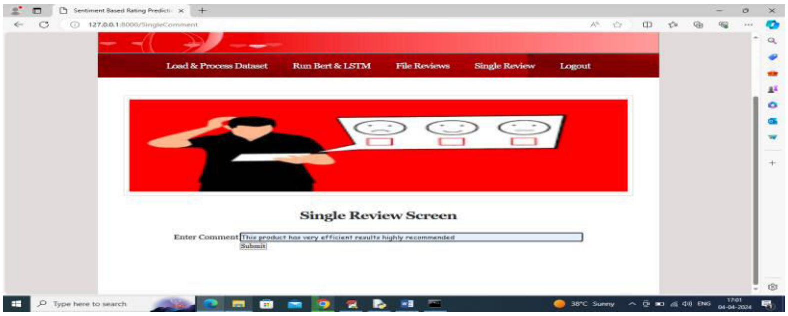
DOI:10.15680/IJARETY.2024.1103038



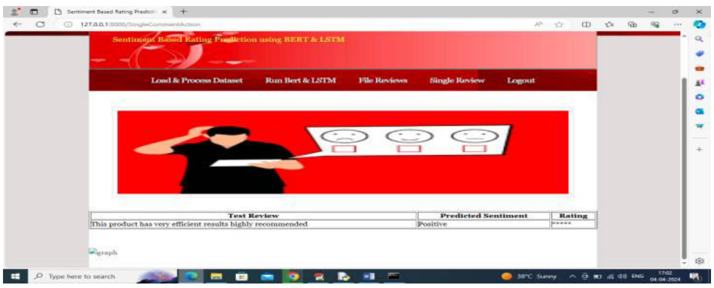
In above screen selecting and uploading ‘Test Reviews’ and then click on ‘Submit’ button to get below output



In above screen in first column can see TEST Reviews and in second column can see sentiment value and in 3rd column can see predicted ratings as number of stars and in pie chart can see number of reviews predicted in different ratings from 1 to 5 and now click on ‘Single Review’ link to type review manually and predict sentiment



In above screen I entered some review and then press button to get below output



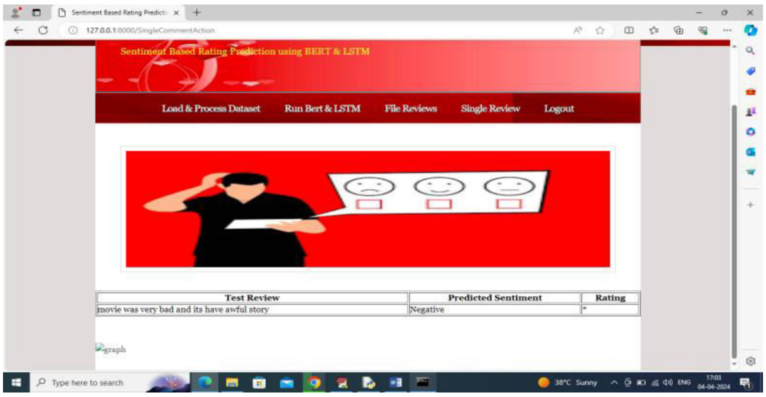
In above screen review is predicted as ‘Positive’ sentiment with 5 start ratings

International Journal of Advanced Research in Education and TechnologY(IJARETY)

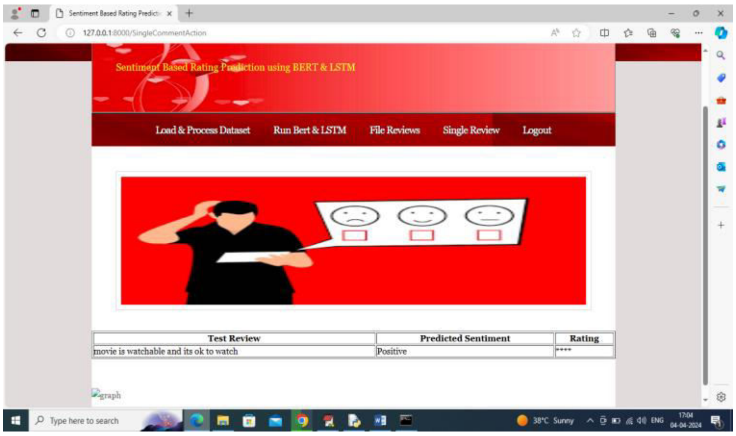
| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 7.394| A Bi-Monthly, Double-Blind Peer Reviewed & Referred Journal |

|| Volume 11, Issue 3, May-June 2024 ||

DOI:10.15680/IJARETY.2024.1103038



In above screen reviews predicted as 'Negative' with 1 rating



In above screen predicted as Positive with 4 star ratings

V.CONCLUSION AND FUTURE WORK

In summary, a promising method with a number of important discoveries and consequences is the use of BERT with LSTM for sentiment-based rating prediction. First off, a more sophisticated comprehension of textual reviews is made possible by the combination of LSTM's sequential modeling capabilities and BERT's contextual embeddings, which capture both the temporal relationships and semantic context inside the data. As a result, user rating predictions based on sentiment analysis of textual reviews are more robust and accurate. Second, the suggested approach outperforms conventional techniques, taking use of deep learning architectures' advantages to manage ambiguity, sarcasm, complicated linguistic structures, and domain-specific terminology. The model improves the user experience in general in rating predictions systems by more accurately and consistently predicting user ratings by capturing the finer points of user emotion. The suggested approach may also be scaled and generalized, which makes it flexible enough to fit a variety of datasets and practical uses. The model may be used in a variety of contexts, such as e-commerce, social media, and consumer feedback research, because of its interpretable representations and strong generalization to new or unknown data.

REFERENCES

1. Do, H.H.; Prasad, P.W.; Maag, A.; Alsadoon, A. Deep learning for aspect-based sentiment analysis: A comparative review. Expert Syst. Appl. **2019**, 118, 272–299. [**Google Scholar**] [**CrossRef**]
2. Wang, J.; Xu, B.; Zu, Y. Deep learning for aspect-based sentiment analysis. In Proceedings of the 2021 International Conference on Machine Learning and Intelligent Systems Engineering (MLISE), Chongqing, China, 9–12 July 2021; IEEE: New York, NY, USA, 2021; pp. 267–271. [**Google Scholar**]
3. Rahman, M.M.; Islam, M.N. Exploring the performance of ensemble machine learning classifiers for sentiment analysis of COVID-19 tweets. In Sentimental Analysis and Deep Learning: Proceedings of ICSADL 2021; Springer: Berlin/Heidelberg, Germany, 2021; pp. 383–396. [**Google Scholar**]

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 7.394| A Bi-Monthly, Double-Blind Peer Reviewed & Referred Journal |

|| Volume 11, Issue 3, May-June 2024 ||

DOI:10.15680/IJARETY.2024.1103038

4. Wankhade, M.; Rao, A.C.S.; Kulkarni, C. A survey on sentiment analysis methods, applications, and challenges. Artif. Intell. Rev. **2022**, 55, 5731–5780. [**Google Scholar**] [**CrossRef**]
5. Xu, Q.A.; Chang, V.; Jayne, C. A systematic review of social media-based sentiment analysis: Emerging trends and challenges. Decis. Anal. J. **2022**, 3, 100073. [**Google Scholar**] [**CrossRef**]
6. Medhat, W.; Hassan, A.; Korashy, H. Sentiment analysis algorithms and applications: A survey. Ain Shams Eng. J. **2014**, 5, 1093–1113. [**Google Scholar**] [**CrossRef**]
7. Abbas, Y.; Malik, M. Defective products identification framework using online reviews. Electron. Commer. Res. **2023**, 23, 899–920. [**Google Scholar**] [**CrossRef**]
8. Crnovrsanin, T.; Di Bartolomeo, S.; Wilson, C.; Dunne, C. Indy Survey Tool: A framework to unearth correlations in survey data. In Proceedings of the 2023 IEEE Visualization and Visual Analytics (VIS), Melbourne, Australia, 21–27 October 2023; Volume 2023. [**Google Scholar**]
9. Dogra, V.; Verma, S.; Chatterjee, P.; Shafi, J.; Choi, J.; Ijaz, M.F. A complete process of text classification system using state-of-the-art NLP models. Comput. Intell. Neurosci. **2022**, 2022, 1883698. [**Google Scholar**] [**CrossRef**]
10. Ahmed, C.; ElKorany, A.; ElSayed, E. Prediction of customer's perception in social networks by integrating sentiment analysis and machine learning. J. Intell. Inf. Syst. **2023**, 60, 829–851. [**Google Scholar**] [**CrossRef**]
11. Suhartono, D.; Purwandari, K.; Jeremy, N.H.; Philip, S.; Arisaputra, P.; Parmonangan, I.H. Deep neural networks and weighted word embeddings for sentiment analysis of drug product reviews. Procedia Comput. Sci. **2023**, 216, 664–671. [**Google Scholar**] [**CrossRef**]
12. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint arXiv:1810.04805.
13. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.
14. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30, 5998-6008.
15. Zhang, Y., Sun, M., & Hu, Y. (2019). Bert4Rec: Sequential recommendation with bidirectional encoder representations from transformer. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management (pp. 1441-1450).
16. Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692.
17. Wang, A., Singh, A., Michael, J., Hill, F., Levy, O., & Bowman, S. R. (2018). GLUE: A multi-task benchmark and analysis platform for natural language understanding. arXiv preprint arXiv:1804.07461.
18. Pennington, J., Socher, R., & Manning, C. (2014). GloVe: Global vectors for word representation. In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP) (pp. 1532-1543).
19. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. arXiv preprint arXiv:2005.14165.
20. Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. R., & Le, Q. V. (2019). XLNet: Generalized autoregressive pretraining for language understanding. arXiv preprint arXiv:1906.08237.
21. Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep contextualized word representations. arXiv preprint arXiv:1802.05365.
22. Nasr, M.M.; Shaaban, E.M.; Hafez, A.M. Building sentiment analysis model using Graphlab. Int. J. Sci. Eng. Res. 2017, 8,11551160.
23. 23. Tan, W.; Wang, X.; Xu, X. Sentiment Analysis for Amazon Reviews. 2018; pp. 1–5. Available online: https://cs229.stanford.edu/proj2018/report/122.pdf (accessed on 9 March 2024).
24. 24. Zhang, D.; Xu, H.; Su, Z.; Xu, Y. Chinese comments sentiment classification based on word2vec and SVMperf. Expert Syst. Appl.2015, 42, 1857–1863. [CrossRef]
25. 25. Huang, M.; Cao, Y.; Dong, C. Modeling rich contexts for sentiment classification with lstm. arXiv 2016, arXiv:1605.01478.
26. Hasan, A.; Moin, S.; Karim, A.; Shamshirband, S. Machine learning-based sentiment analysis for twitter accounts. Math. Comput. Appl. 2018, 23, 11. [CrossRef]
27. Deriu, J.; Lucchi, A.; De Luca, V.; Severyn, A.; Muller, S.; Cieliebak, M.; Hofmann, T.; Jaggi, M. Leveraging large amounts of weakly supervised data for multi-language sentiment classification. In Proceedings of the 26th International Conference on World WideWeb, Perth, Australia, 3–7 April 2017; pp. 1045–1052.

IJARETY

International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975 Impact Factor: 7.394

www.ijarety.in editor.ijarety@gmail.com