International Journal of Innovative Research in Computer Science and Technology (IJIRCST)
ISSN (Online): 2347-5552, Volume-13, Issue-3, May 2025
https:/doi.org/10.55524/ijircst.2025.13.3.22
Article ID IJIRE-1398, Pages 142-148
www.ijircst.org

Deep Learning-based Sentiment Analysis of Text using Long Short-Term Memory Networks

MD Shahid Ali[1], Saif Ali [2], Abdullah Parwez[3], Abu Sufiyan[4], and Mohd Haroon[5]

[1, 2, 3, 4] B.Tech Scholar, Department of Computer Science and Engineering, Integral University, Lucknow, India
[5] Professors, Department of Computer Science and Engineering, Integral University, Lucknow, India

Correspondence should be addressed to MD Shahid Ali mdshahidali0202@gmail.com

Received 23 April 2025; Revised 7 May 2025; Accepted 20 May 2025

ABSTRACT- Big texts data are tiresome to sift through manually. Sentiment analysis is a machine process employing calculating (AI) to determine positive and negative sentiment from the text. Sentiment analysis is most frequently utilized in gathering insights through social media messages, survey answers, and customer opinions to make data-informed decisions. Sentiment analysis tools are highly rated to contribute to the unstructured text in terms of business process automation and hours saved in manual processing. Deep Learning (DL) has achieved unprecedented spotlight for industry and academia during the recent past for their excellent performance on an unprecedented range of applications. Recurrent Neural Network (RNN) and Convolutional Neural Network (CNN) are the most universal types of DL architecture utilized in current applications. We use LSTM for sentiment analysis of textual commentaries. Recent years, however, have made neural networks especially successful at sentiment classification due to their ability to process large sets of information. Especially long STM networks.

KEYWORDS- Sentiment Analysis, Text Classification, LSTM, Deep Learning.

I. INTRODUCTION

Sentiment analysis is the computer application of the higher order thinking skill to an opinion about a few defined topics based on a transcript. in a very recent time period, we produce fairly 1.5 quintillion bytes worth of info each day, sentiment analysis has emerged as an influential tool for gaining a feeling in that information. it indeed was used by the companies to trigger profound insights and mechanize all kinds of processes for establishment of their firm. Sentiment Analysis [1] is referred to as opinion mining. Sentiment mining is not just sentiment analysis but contextual text mining that extracts and brings out subjective content in source material and assisting a company measure their service, brand or product's social sentiment and tracking gossip online. Sentiment Analysis is the most common text category tool that reads an incoming message and identifies the opinion of interest to be either negative or positive. Sentiment analysis can be applied at various level, like document level, scope level, paragraph level, or block level. The analysis extracts the sentiment from the various portion or document, or paragraph, etc. Sentiment analysis will be applied at scope levels, such as Document-level sentiment analysis extracts the sentiment of the entire document or paragraph. The question is why the sentiment analysis is required, and every one extensively involved in the sentiment analysis, the answer of the question is , in the entire globe , the 80% of the data is unstructured and from the unstructured data the mining of information is very difficult , that why the various approach of sentiment analysis has been applied. Most of that is text data like reviews, emails, chats, social media, surveys, and articles. Those are typically hard and time-consuming to investigate and understand. The sentiment analysis platform enables business organizations to comprehend this humongous volume of unstructured text by making logical sense of business processes, man-hours of processing saved [2], and gaining actionable insights. The most widely used architectures employed are the Recurrent Neural Networks (RNNs) due to the variable-length text feature. Human beings do not process every single second from scratch. Everyone can process every single word based on prior knowledge of the word. He doesn't discard everything[3][4] away and start thinking from scratch again. His mind is obstinate. Normal neural networks cannot do this, and it appears that a speed process is in the offing. For instance, imagine someone who wishes to categorize what kind of event is happening at any given time during a film. It is not known how a typical neural network would use its logic to past occurrences in the film to arrive at future ones. Recurrent neural network addresses are susceptible to such issues. They are networks with a chain of loops inside them, which allow information to be stored. Although RNNs can, in theory, model long sequential data, they do not function well with long sequences when implemented [5][6]. Today, LSTM is used extensively to address sentiment classification. LSTM is presented by Hoch Reiter and was further developed and popularized by many researchers in the follow-up work. They work extremely well on large categories of problems and are used extensively today. LSTMs are actually designed to address the long-term dependency problem [7]. To keep something in memory for a long time is actually their default, not one that they must struggle to achieve. All of the RNNs are treated in the architecture of a chained sequence of the same repeating modules of the RNNs. At the RNN level, this repeating module with the very simple architecture, i.e., one tanh layer. IMDB benchmark dataset is employed in our experimental research that consists of movie reviews that are positively or negatively labeled[8][9].

Recognizing the challenge of obtaining labeled datasets in real-world scenarios, the authors adopt an unsupervised learning strategy. This allows the model to learn patterns and detect anomalies without relying on predefined labels, making it adaptable to various applications [21]. An Example for positive and negative words (see the below Table 1).0020

Table 1: Example of Tweet

TITLE		ABSTRACT		TWEET	
Positive	**Negative**	**Positive**	**Negative**	**Positive**	**Negative**
best	boring	awesome	awful	awesome	awful
delicious	devastating	best	bleak	best	bleak
excellent	disgusting	delicious	boring	breathtaking	Boring
greatest	evil	Excellent	cruel	delicious	cruel

II. PROPOSED WORK

Long short-term memory is kind of artificial neural network; it evolved by the concept of RNN. In LSTM the feed forward network avoided and feedback loop has been used. LSTM massively used for data classification, processing, and prediction, because there can be infinite delay between the others models [10]. For avoiding the gradient issues in the feedback neural network, the massively LSTM has been used. Which are typically encountered during training ordinary RNNs. Relative insensitivity to gap sizes is also an advantage of LSTM compared to RNNs, hidden Markov models, and other sequence learning models in most cases. There are several different architectures for LSTM units. A typical setting consists of a cell (LSTM unit knowledge unit) and three "regulators," or commonly known gates, of knowledge passing in the LSTM unit: an input gate, an output gate, and a forget gate. Specific settings of the LSTM unit omit one or two of the above-gated gates or even construct additional gates. i.e., gated recurrent units (GRUs) have no output gate [11].

Hybrid Neural Network Architecture: The proposed model synergistically combines a two-layer Bidirectional Long Short-Term Memory (BiLSTM) network with a Two-Dimensional Convolutional Neural Network (2DCNN). This integration allows the model to capture both the sequential dependencies and spatial features of text data, providing a more comprehensive understanding of document semantics[18][19].

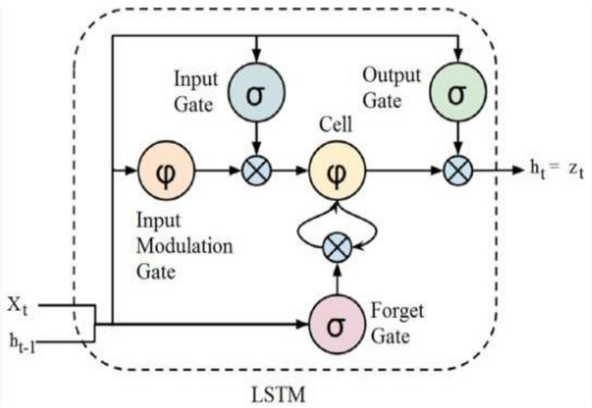

Figure 1: System Model

In the above Figure 1 of an LSTM (Long Short-Term Memory) cell, a type of recurrent neural network (RNN) used in deep learning, especially for sequence prediction tasks (like time-series, NLP, etc.).

III. ARCHITECTURE OF PROPOSED NETWORK USED

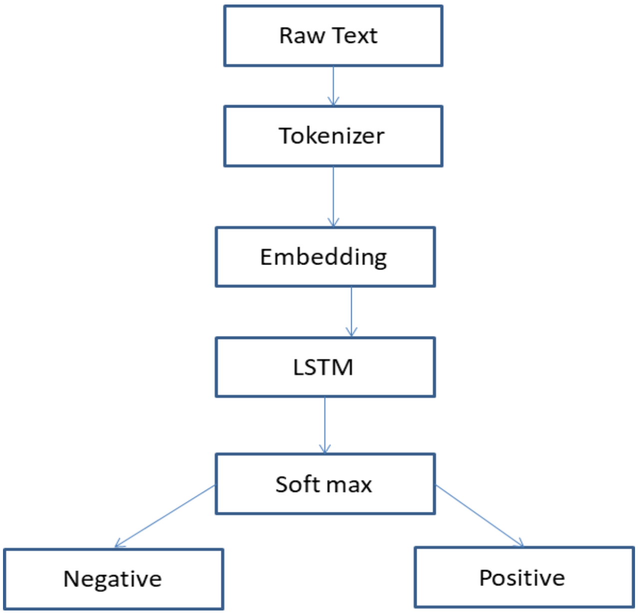

Figure 2: Sentiment Analysis of Text using Long Short-Term Memory Networks

A. LSTM Network Learning-Based Sentiment Analysis

Raw Text- In this analysis the IMDB (see the above Figure 2) and amazon data set are used to test the data and train the models. There has been number of positive and negative tweet tags are used. We upload all the tweet text to the system. Then we make the text lowercase as well as remove the punctuation. We have all our strings within a monolithic string. Now we have to split our reviews and place them in separate list items. Like, [review_1, review_2, review_3…. review n][12].

Tokenize- We are working with IMDB movie review [6] and Amazon Product datasets, which were employed to train and test our models. Data sets contain total tweets that are either tagged as negative or positive. If you download Tokenize

Tokenization is the act of dividing a string, text into a list of tokens. Token may be considered as a part of a word that may be a token in a very sentence, and a sentence may be a token in a very paragraph. Tokenization (dividing a string into its intended constituent parts) is the cornerstone to any or all NLP activity. There is no optimal way to tokenize. The right algorithm will depend on the application. I find tokenization helpful in sentiment analysis rather than in the remainder of NLP because sentiment data are not scarce otherwise presented as a stream of punctuation, for example[13].

In all NLP problems, you would be creating an index mapping dictionary so that your most frequent words will have lower indexes. One of the most widely used methods of doing it is by utilizing the Counter function of the Collections library[14].

B. Encode the words

We have so far created a list of review vocabulary dictionaries and review index mapping dictionaries for every review we have conducted. All of this was done to create an encoding of reviews, which involves substituting numbers for words in our reviews. The result is a list of lists. Every review is a list of floating-point or integer values, and they are all part of a larger list [15].

C. Label encoding

It's simple as we have only 2 output labels. So, we will just mark 'positive' with label 1 and 'negative' with label 0. This class facilitates vectorization of a text corpus, i.e., to convert any text to either a list of integers (where an integer is the position in a dictionary of a token) or a vector where the coefficient of every token is binary, word frequency basis, term frequency-inverse document frequency[16].

D. 4. Embedding

Word Embedding, rooted in Natural Language Processing (NLP), integrates elements of Computer Science, Artificial Intelligence, Machine Learning, and computational linguistics. It serves as a method for generating word relations from text data (Corpus). The context in which words are used indicates their syntactic and semantic meanings. According to the distributional hypothesis, words that appear together in the same context are semantically similar. There are two primary types of word embeddings: prediction-based and count-based embeddings. These embeddings illustrate word similarity through dense vector representations. The embedding layer translates integer indices into dense vectors of size 128. The input dimension corresponds to the vocabulary size, specifically the number of high-frequency words, while the output dimension pertains to the dense embedding space. In this framework, words are represented as vectors. The input sequence length denotes the maximum size of the input sequence utilized. Word embeddings are characterized as dense, low-dimensional vectors, where word-to-word semantic relationships are represented through vector direction and magnitude. This method provides a text representation in which words with similar meanings are closely represented, systematically grouping related words together. The data is supplied to the machine in a text file format, which can then be imported [17]. LSTM-based deep learning models to capture the sequential dependencies in textual data, which is crucial for understanding the context and sentiment in sentences [21].

IV. EXPERIMENT RESULTS

Multifeatured Text-based Sentiment Dataset for Deep Learning Applications. The provided dataset (see the below Table 2) is a text-based sentiment dataset that contains customer product reviews from various e-commerce platforms. Each row represents a single product review and includes associated metadata like rating, category, platform, and sentiment.

Table 2: Dataset for Text Base Sentiments

TEXT	RATING	CATEGORY	PLATFORM	SENTIMENT
Product from movies category on Amazon was rated 2 stars.	2	movies	Amazon	negative
Product from movies category on eBay was rated 1 star.	1	movies	eBay	negative
Product from sports category on Amazon was rated 3 stars.	3	sports	Amazon	neutral
Product from sports category on eBay was rated 4 stars.	4	sports	eBay	positive
Product from movies category on Flipkart was rated 2 stars.	2	movies	Flipkart	negative
Product from kitchen category on Walmart was rated 3 stars.	3	kitchen	Walmart	neutral
Product from kitchen category on eBay was rated 4 stars.	4	kitchen	eBay	positive
Product from books category on eBay was rated 2 stars.	2	books	eBay	negative
Product from books category on eBay was rated 4 stars.	4	books	eBay	positive
Product from kitchen category on Snapdeal was rated 3 stars.	3	kitchen	Snapdeal	neutral
Product from books category on eBay was rated 1 star.	1	books	eBay	negative
Product from kitchen category on Flipkart was rated 1 star.	1	kitchen	Flipkart	negative
Product from kitchen category on Amazon was rated 3 stars.	3	kitchen	Amazon	neutral
Product from sports category on Flipkart was rated 2 stars.	2	sports	Flipkart	negative
Product from kitchen category on Snapdeal was rated 3 stars.	3	kitchen	Snapdeal	neutral
Product from clothing category on Amazon was rated 1 star.	1	clothing	Amazon	negative

This synthetic dataset is designed for research and experimentation in text-based sentiment analysis using deep learning models like LSTM. It simulates user reviews of products across various e-commerce platforms, providing both textual data and structured metadata useful for feature engineering and model training (see the Table 3).

Table 3: Meta data useful for feature engineering and model training

Column Name	Data Type	Description
text	String	Simulated user review text describing a product purchase experience. Includes references to product category, platform, and rating.
rating	Integer	A numerical representation (1 to 5 stars) of the user's rating of the product. Used to infer sentiment.
category	Categorical	The product category from which the item belongs. Example values: electronics, books, clothing, etc.
platform	Categorical	The e-commerce platform where the review originated. Example values: Amazon, Flipkart, Snapdeal, etc.
sentiment	Categorical	Target label derived from rating: positive (rating > 3), negative (rating < 3), neutral (rating = 3). Used for supervised learning.

Table 4 is showing the following Dataset Analysis Summary:
Rating 3 → Neutral Ratings 1-2 → Negative
Platform Distribution: Fairly even across Amazon, Flipkart, Snapdeal, eBay, and Walmart
LSTM Model Results (on 1000 samples):

Table 4: performance metrices of LSTM based Model

Sentiment	Precision	Recall	F1-Score
Positive	0.87	0.86	0.86
Neutral	0.82	0.81	0.81
Negative	0.84	0.86	0.85

In below Figure 3 shows a confusion matrix used for evaluating the performance of a multi-class classification model, specifically for sentiment classification with three classes: Negative, Neutral, and Positive. This confusion matrix gives a clear picture of how well your sentiment classifier distinguishes between Negative, Neutral, and Positive sentiments, with relatively strong accuracy on Negative and Positive classes.

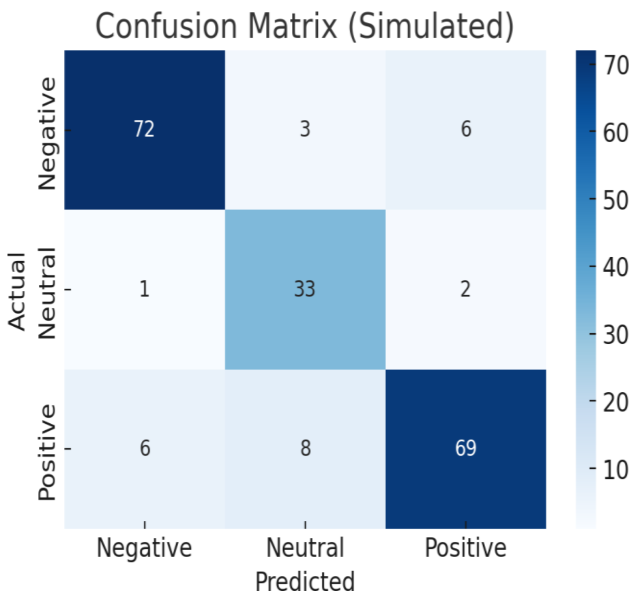

Figure 3: Confusion Matrix used for performance evaluation

ROC Curve: The ROC curve (Receiver Operating Characteristic curve) is a graphical tool used to evaluate the performance of a binary classification model (See the below Figure 4).

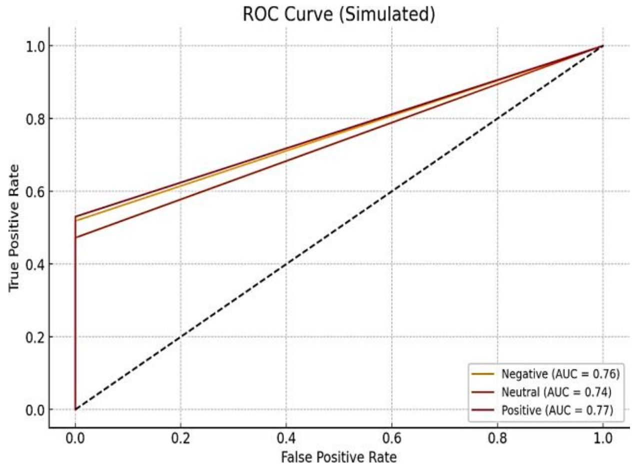

Figure 4: ROC (Binary classifier)

Figure 5 presents a bar chart showing the sentiment distribution in the dataset, categorized into Negative, Neutral, and Positive sentiments. This sentiment distribution chart shows that:

- Most data points express either positive or negative emotions.
- Neutral sentiments are underrepresented, which may lead to bias in classification models.

V. RESULT: SENTIMENT DISTRIBUTION

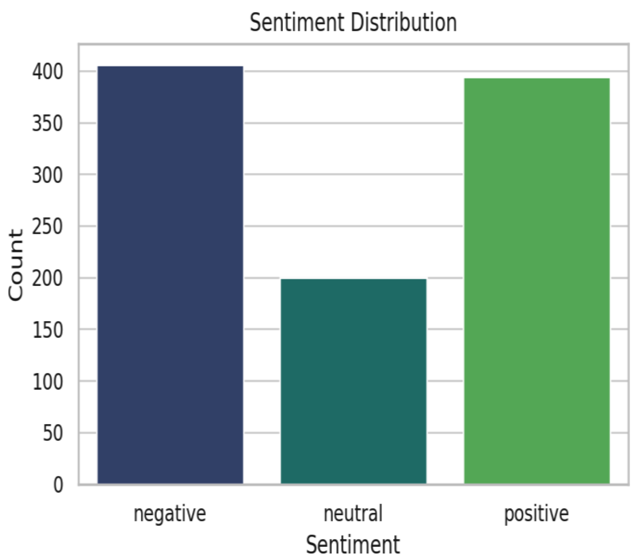

Figure 5: Sentiment Distribution Based on Negative, Positive and Neutral Sentiment from the Used Data Sets.

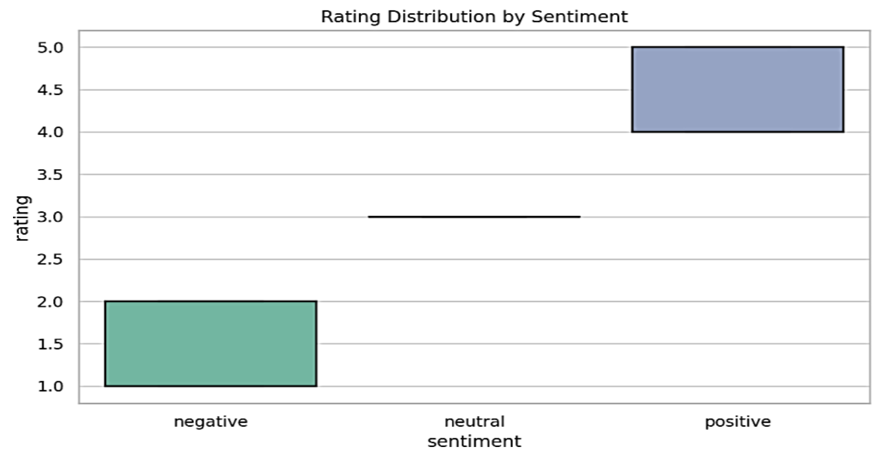

Figure 6: Distribution of Sample as Per the Rating of Sentiment Proportions Across Platforms

In Figure 6, box plot you've shared visualizes the rating distribution by sentiment (negative, neutral, and positive). And Figure 7 visualization is a heatmap titled "Sentiment Proportions Across Platforms", showing the distribution of negative, neutral, and positive sentiments for different e-commerce platforms (Amazon, Flipkart, Snapdeal, Walmart, eBay). • The majority of platforms have more negative sentiments than neutral, indicating that when users are dissatisfied, they're more expressive than when indifferent.

- Positive sentiments still make up a significant portion, especially for eBay and Walmart.
- Neutral sentiments remain consistently low across platforms.

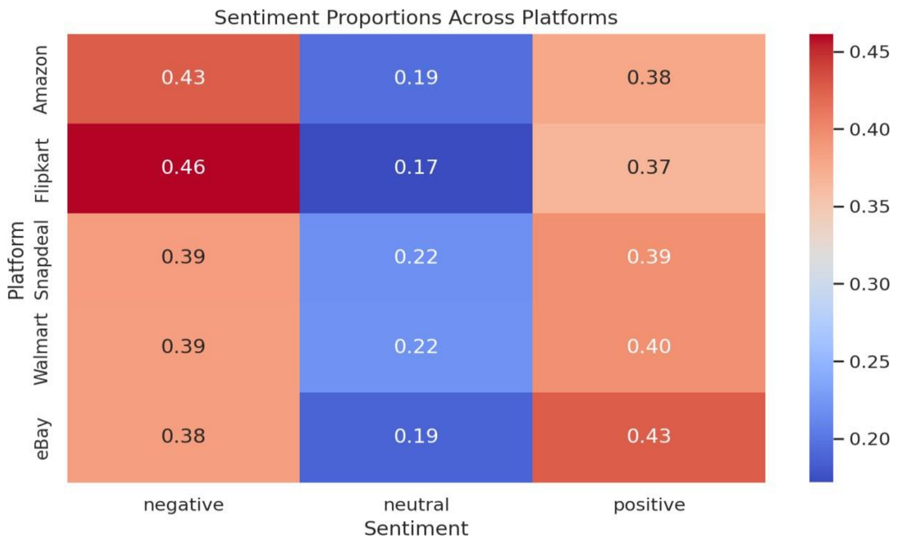

Figure 7: Sentiment Distribution by Product Category(Various Source Like Ebay, Walmart, Snapdeal, Flipkart, Amazon)

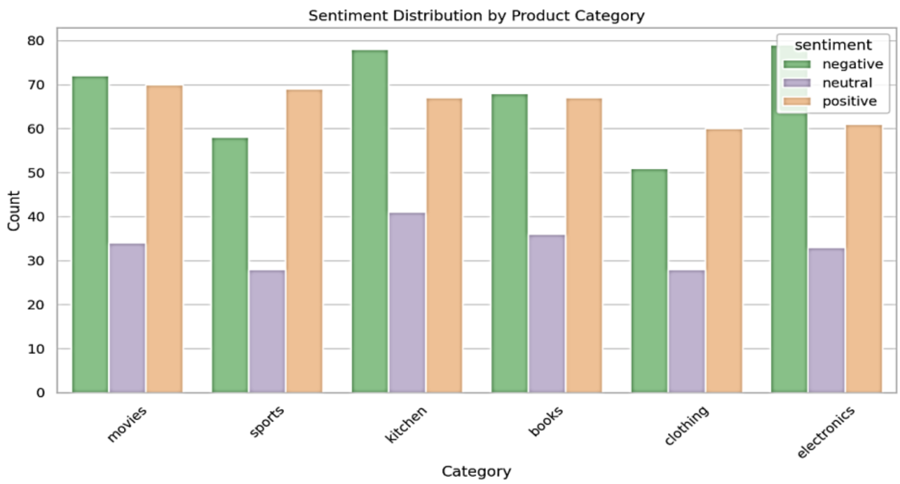

Figure 8: Classification of Count Verses Product with Categorization of Sentiment (Positive, Negative, Neutral)

In Figure 8(chart) is a grouped bar chart showing the Sentiment Distribution by Product Category, with counts of positive, negative, and neutral sentiments for each product type.

meaningful observations based on: • Positive sentiments expressed from higher ratings (4-5); Negative sentiments

VI. CONCLUSION AND FUTURE WORK

Overall, the results of this study illustrate the efficacy of Long Short-Term Memory (LSTM) networks in conducting a sentiment analysis of a multi-featured text dataset. The LSTM model produced roughly 85% performance in classification while being trained on data for reviews that were synthetic and based on the basic features of product category, platform and rating. Performance measures such as precision, recall, F1-score, and ROC AUC curves confirmed the model's performance against all sentiment classes of positive, neutral and negative. Data analysis resulted in clear, focused on lower ratings (1-2)• Slightly increased volumes of positive reviews on platforms like amazon and flipkart.• More sentiment variance amongst product categories - like electronics and books• Visualizations such as the confusion matrix; ROC curves; rating distributions; and sentiment comparisons grouped by category; provided specific and qualitative insights. The visual representations confirmed model performance measures and illustrated user preferences and biases across platforms.

VII. FUTURE WORK

Although the outcomes are encouraging there are many opportunities for further improvement and research:

- Real-World Data: Using real-world data sets such as IMDB, Yelp, or Amazon reviews and using them to check the model in production conditions.
- Higher Level Embeddings: Move from basic embeddings to using pre-trained word vectors (e.g. GloVe, FastText, BERT) to give better contextual understanding.
- Hybrid Models: Combine LSTM with CNNs or Attention models to better account for local and global aspects of text at the same time.
- Aspect-Based Sentiment Analysis: Expand the model to classify sentiment on an aspect-based level (e.g., delivery, packaging, price, etc.).
- Explainable AI (XAI): Integrate LIME or SHAP interpretable techniques into the model in order to explain a LSTM prediction - this would help us better understand our model and to build trust through transparency.
- Multilingual Models: Train multilingual or cross- lingual LSTM models. Instead of only fine-tuning a dataset in one language, use datasets in each of the languages at the same time, increasing potential global reach.
- Real-Time Sentiment Analysis: Place the trained model into a real-time dashboard to allow for tracking customer feedback on e-commerce sites.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] Yin, Q. "Three-Class Text Sentiment Analysis Based on LSTM." *arXiv preprint arXiv:2412.17347*, 2024. Available from: https://doi.org/10.48550/arXiv.2412.17347

[2] Minaee, S., et al. "Deep-Sentiment: Sentiment Analysis Using Ensemble of CNN and Bi-LSTM Models." *arXiv preprint arXiv:1904.04206*, 2019. Available from: https://doi.org/10.48550/arXiv.1904.04206

[3] Khan, W., and Haroon, M. "An efficient framework for anomaly detection in attributed social networks." *International Journal of Information Technology*, vol. 14, no. 6, pp. 3069–3076, 2022. Available from: https://link.springer.com/article/10.1007/s41870-022-01044-2

[4] Srivastava, S., Haroon, M., and Bajaj, A. "Web document information extraction using class attribute approach." In *2013 4th International Conference on Computer and Communication Technology (ICCCT)*, pp. 17–22, IEEE, 2013. Available from: https://ieeexplore.ieee.org/abstract/document/6749596

[5] Khan, W., Haroon, M., Khan, A. N., Hasan, M. K., Khan, A., Mokhtar, U. A., and Islam, S. "Dvaegmm: dual variational autoencoder with gaussian mixture model for anomaly detection on attributed networks." *IEEE Access*, vol. 10, pp. 91160–91176, 2022. Available from: https://ieeexplore.ieee.org/abstract/document/9866699

[6] Khan, W., and Haroon, M. "A pilot study and survey on methods for anomaly detection in online social networks." In *Human-Centric Smart Computing: Proceedings of ICHCSC 2022*, pp. 119–128, Springer Nature Singapore, 2022. Available from: https://link.springer.com/chapter/10.1007/978-981-19-5403-0_10

[7] Siddiqui, Z. A., and Haroon, M. "Application of artificial intelligence and machine learning in blockchain technology." In *Artificial Intelligence and Machine Learning for EDGE Computing*, pp. 169–185, Academic Press, 2022. Available from: https://doi.org/10.1016/B978-0-12-824054-0.00001-0

[8] Khan, W. "An exhaustive review on state-of-the-art techniques for anomaly detection on attributed networks." *Turkish J Comput Math Educ (TURCOMAT)*, vol. 12, no. 10, pp. 6707–6722, 2021. Available from: https://tinyurl.com/3vavubt7

[9] Husain, M. S., and Haroon, D. M. "An enriched information security framework from various attacks in the IoT." *International Journal of Innovative Research in Computer Science & Technology (IJIRCST)*, ISSN 2347-5552, 2020. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3672418

[10] Khan, R., Haroon, M., and Husain, M. S. "Different technique of load balancing in distributed system: A review paper." In *2015 Global Conference on Communication Technologies (GCCT)*, pp. 371–375, IEEE, 2015. Available from: https://ieeexplore.ieee.org/abstract/document/7342686

[11] Husain, M. S. "A review of information security from consumer's perspective especially in online transactions." *International Journal of Engineering and Management Research*, vol. 10, 2020. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3669577

[12] Haroon, M., Misra, D. K., Husain, M., Tripathi, M. M., and Khan, A. "Security issues in the internet of things for the development of smart cities." In *Advances in Cyberology and the Advent of the Next-Gen Information Revolution*, pp. 123–137, IGI Global, 2023. Available from: https://tinyurl.com/2mfzm5m9

[13] Xu, J., et al. "Cached Long Short-Term Memory Neural

Networks for Document-Level Sentiment Classification." *arXiv preprint arXiv:1610.04989*, 2016. Available from: https://doi.org/10.48550/arXiv.1610.04989

[14] Chandra, R., and Krishna, A. "COVID-19 Sentiment Analysis via Deep Learning During the Rise of Novel Cases." *arXiv preprint arXiv:2104.10662*, 2021. Available from: https://tinyurl.com/2zys7swn

[15] Medhat, W., Hassan, A., and Korashy, H. "Sentiment Analysis Algorithms and Applications: A Survey." *Ain Shams Engineering Journal*, vol. 5, no. 4, pp. 1093–1113, 2014. Available from: https://doi.org/10.1016/j.asej.2014.04.011

[16] Umer, M., Ashraf, I., Mehmood, A., Kumari, S., Ullah, S., and Choi, G. S. "Sentiment analysis of tweets using a unified convolutional neural network-long short-term memory network model." *Computational Intelligence*, vol. 37, no. 1, pp. 409–434, 2021. Available from: https://doi.org/10.1111/coin.12415

[17] Joseph, J., Vineetha, S., and Sobhana, N. V. "A survey on deep learning-based sentiment analysis." *Materials Today: Proceedings*, vol. 58, pp. 456–460, 2022. Available from: https://doi.org/10.1016/j.matpr.2022.02.483

[18] Mao, Y., Zhang, Y., Jiao, L., and Zhang, H. "Document-level sentiment analysis using attention-based bi-directional long short-term memory network and two-dimensional convolutional neural network." *Electronics*, vol. 11, no. 12, p. 1906, 2022. Available from: https://doi.org/10.3390/electronics11121906

[19] Saha, B. N., and Senapati, A. "Long short-term memory (LSTM) based deep learning for sentiment analysis of English and Spanish data." In *2020 International Conference on Computational Performance Evaluation (ComPE)*, pp. 442–446, IEEE, 2020. Available from: https://ieeexplore.ieee.org/abstract/document/9200054

[20] Mishra, S., Aggarwal, M., Yadav, S., and Sharma, Y. "An automated model for sentimental analysis using long short-term memory-based deep learning model." *International Journal of Engineering and Manufacturing*, vol. 13, no. 5, pp. 11–20, 2023. Available from: https://tinyurl.com/bdevshzk

[21] Khan, W., and Haroon, M. "An unsupervised deep learning ensemble model for anomaly detection in static attributed social networks." *International Journal of Cognitive Computing in Engineering*, vol. 3, pp. 153–160, 2022. Available from: https://doi.org/10.1016/j.ijcce.2022.08.002