Bi-LSTM model for Sentimental Analysis for Recommendation and Classification on X Dataset

Rainita R[1], Rakshitha K G[2], Rashi Ramesh[3],Sumera Banu[4] and Ms.Suchitra H L[5]
Department of Computer Science and Engineering
PES Institute of Technology & Management
Shivamogga, Karnataka, India
{email rainita.r020620031, kgrakshitha2003,rashiramesh262and sumera2banu}@gmail.com,
suchitrahl@pestrust.edu.in

Abstract- **In the era of rapid digital communication, social media platforms such as X offer valuable insights into public sentiment on various topics, including products and services. This project leverages a Bidirectional Long Short-Term Memory (Bi-LSTM) model integrated with sentiment analysis to classify and recommend products based on user sentiments from X datasets. By analyzing tweets and categorizing them as positive, negative, or neutral, the system enhances the accuracy of sentiment-driven recommendations. This model addresses challenges like informal language and contextual nuances, which traditional models often overlook. Through rigorous data preprocessing, Bi-LSTM's ability to capture bidirectional dependencies, and performance evaluation using metrics like accuracy, precision, recall, and F1 score, our approach demonstrates significant improvements in sentiment classification accuracy. The developed system is not only capable of predicting sentiments but also providing actionable recommendations, making it a valuable tool for businesses and researchers.**

Keywords—Sentiment Analysis, Bi-LSTM, X Dataset, Product Recommendation, Text Classification, Deep Learning

I. INTRODUCTION

Social media platforms have reshaped global communication, with X serving as a major source of real-time public feedback on products, services, and events. This vast repository of data offers companies and researchers an opportunity to analyze public sentiment, enabling them to understand consumer opinions, predict market trends, and make data-driven decisions. Sentiment analysis, a branch of Natural Language Processing (NLP), involves extracting and categorizing emotional tones from text to determine whether the sentiment expressed is positive, negative, or neutral. However, analyzing tweets presents unique challenges due to their informal nature, use of abbreviations, hashtags, emojis, and limited character count, making traditional machine learning models inadequate for understanding complex linguistic patterns. To address these challenges, this project employs a Bidirectional Long Short-Term Memory (Bi-LSTM) model for sentiment analysis and recommendation based on tweets. Unlike traditional unidirectional models, Bi-LSTM captures context from both past and future words, making it highly effective in understanding the nuances of human language. This bidirectional ability enables the model to classify tweets with greater accuracy, even when dealing with complex sentence structures or ambiguous language. By categorizing tweets into positive, negative, or neutral sentiments, the model further enhances decision-making processes by offering sentiment-based product recommendations.

Figure 1: Sentimental Analysis Overview

The project workflow begins with comprehensive data preprocessing, involving text cleaning, tokenization, and word embedding. These steps are critical for transforming noisy and unstructured tweet data into a format suitable for deep learning models. The Bi-LSTM model then processes this data, leveraging its sequential learning capabilities to identify and classify sentiment patterns. Performance is evaluating using key metrics such as accuracy, precision, recall, and F1-score, ensuring a robust and reliable classification system. This approach not only enhances sentiment detection accuracy but also supports real-time implementation, providing businesses with timely insights into customer feedback. Furthermore, this project goes beyond basic sentiment analysis by offering actionable insights that can be applied in various domains, including marketing, customer support, and social monitoring. By analyzing real-time tweets, businesses can anticipate consumer needs, adjust marketing strategies, and improve customer engagement. Future enhancements, may include integrating the model with multilingual datasets and expanding its application to detect more complex emotions, such as sarcasm or irony, which are often challenging for conventional models. In conclusion, this Bi-LSTM-based sentiment analysis system is a significant advancement in handling social media data. Its ability to process and interpret complex language structures makes it a powerful tool for businesses seeking to leverage social media insights for strategic decision-making. As social media continues to evolve, such models will become increasingly essential in understanding and responding to the ever-changing landscape of consumer sentiment.

II. LITERATURE REVIEW

Recent developments in hybrid deep learning models have made big strides in enhancing sentiment analysis, emotion detection, and depression prediction using social media data. One such study by Kour and Gupta [1] proposed a hybrid approach combining Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (biLSTM) networks to detect depression from X data. The CNN component of the model excels in extracting spatial features by identifying local dependencies and patterns within textual data. Meanwhile, biLSTM captures sequential dependencies and context in both forward and backward directions, providing a more comprehensive understanding of the tweet's sentiment. This combination of CNN and biLSTM helps to achieve higher accuracy in classifying tweets as depressive or non-depressive. The authors further emphasized the importance of feature-rich text preprocessing to enhance the representation of the textual data and reduce noise, which is often prevalent in social media content. However, the model's performance heavily depends on the quality and representativeness of the dataset, as noisy or biased data can adversely affect accuracy.

In a similar vein, Vattikundala and Prasad [2] explored a hybrid model for sentiment and emotion detection, particularly in recognizing depression-related behaviors on X. Their model combines Particle Swarm Optimization (PSO) and Ant Colony Optimization (ACO) for feature selection, along with a Support Vector Machine (SVM) classifier for the classification task. The use of PSO and ACO enhances the feature selection process by efficiently reducing dimensionality while maintaining the most relevant features, improving the overall accuracy of the model. The inclusion of swarm intelligence techniques, inspired by nature, adds robustness to the feature selection and classification process, thus enabling the model to handle complex data more effectively. However, the integration of these optimization techniques and the SVM classifier makes the overall model more complex, requiring a deep understanding of each component for the successful implementation.

Pandey et al. [3] focused on sarcasm detection in multilingual social media posts, a task complicated by an informal language, mixed languages, and grammatical errors commonly found in platforms like X. Their proposed hybrid CNN model is designed to handle these complexities by using both character-level and word-level embeddings. This approach allows the model to effectively capture intricate patterns in text, such as slang or mixed-language usage, which are often missed by traditional models. Tested across English, Hindi, and Hinglish datasets,

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 01 | Jan - 2025 SJIF Rating: 8.448 ISSN: 2582-3930

the model achieved high F1-scores and outperformed other state-of-the-art sentiment analysis models in sarcasm detection. Despite its superior performance, the hybrid CNN model requires significant computational resources, especially when trained on large multilingual datasets, posing a challenge for scalability in real-world applications.

Koshy and Elango [4] introduced a multimodal tweet classification model for disaster response systems that processes both textual and image data. By leveraging transformer-based models such as RoBERTa for text and Vision Transformer (ViT) for images, the authors demonstrated the potential of combining text and visual information to improve classification accuracy in disaster scenarios. Their early fusion strategy integrates both types of data at the beginning of the pipeline, which enhances the model's ability to understand cross-modal relationships and provides a comprehensive understanding of each tweet's context. This multimodal approach improves performance in classifying informative posts, which is crucial in emergency response systems. However, one of the challenges faced by the model is the risk of overfitting when trained on a limited or imbalanced dataset, a common issue in disaster-related data that is often sparse and contextually specific.

Samuel et al. [5] focused on analyzing public sentiment during the COVID-19 pandemic using machine learning techniques. The authors employed the Naive Bayes and logistic regression models to classify tweets related to the coronavirus pandemic, with Naive Bayes achieving a high classification accuracy of 91% for short tweets. This result underscores the effectiveness of Naive Bayes in handling concise textual data. However, both models performed poorly on longer tweets, highlighting the limitations of traditional machine learning techniques in capturing the nuances of longer texts. The study provided valuable insights into the public's fear and sentiment during the pandemic, emphasizing the need for more advanced models that can better handle varying tweet lengths and complex sentiments. Roy et al. [6] proposed an LSTM-based model for real-time hate speech detection on X, comparing its performance against other models like Support Vector Machine (SVM), Naïve Bayes (NB), and BERT. By leveraging feature extraction techniques such as Term Frequency-Inverse Document Frequency (TF-IDF) and n-grams, the LSTM model achieved an impressive 95% accuracy, outperforming its counterparts. The model's ability to capture sequential data makes it particularly effective at detecting hate speech, where context plays a crucial role in determining the sentiment of a tweet. However, the LSTM model is resource-intensive, requiring significant computational power, especially when processing real-time data streams. Furthermore, the dataset's potential biases, arising from the X API's data collection method, could skew the model's performance and limit its generalizability.

In conclusion, while these hybrid deep learning models demonstrate significant improvements in the accuracy and efficiency of sentiment analysis, emotion detection, and behavior prediction on social media platforms, challenges such as data quality, model complexity, and resource requirements remain. Future research should aim to address these limitations, ensuring that models are both efficient and scalable for real-time applications.

III. SYSTEM OVERVIEW

The proposed sentiment analysis system is designing to efficiently classify textual data, such as tweets, reviews, or social media posts, by utilizing both traditional machine learning techniques and advanced deep learning models. The system processes raw, unstructured data, transforms it into structured features, and uses these features to predict sentiment classes like positive, negative, or neutral. The workflow begins with data input, where raw textual data is collected. This data can contain unwanted elements such as URLs, hashtags, mentions, or special characters, which may introduce noise and impact the accuracy of sentiment classification.

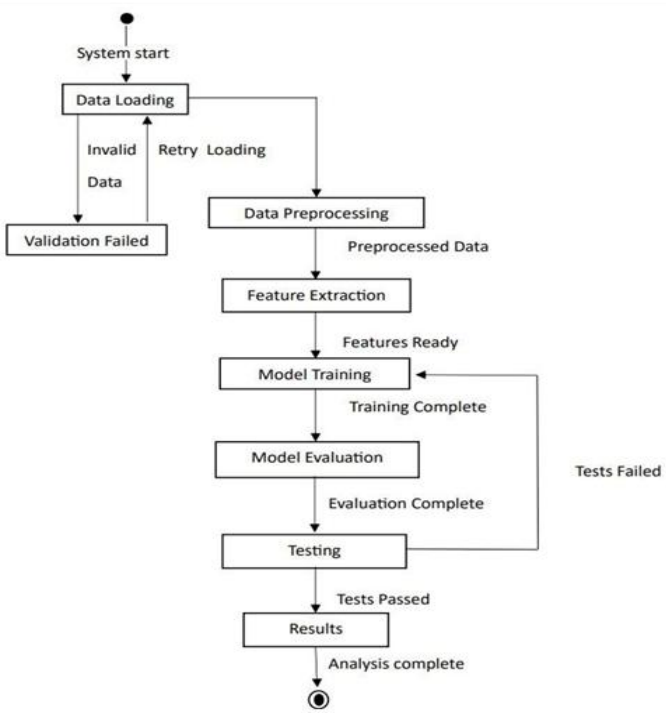

Figure 2: System Design for X Classification

To address this, the system performs a series of preprocessing steps to cleanse the data. The cleaning process includes removing irrelevant components like URLs, hashtags, and mentions, as well as special characters that do not contribute to sentiment analysis. Additionally, all text is converted to lowercase to standardize the data and prevent case-sensitive discrepancies. After cleaning, the system tokenizes the text, breaking it down into smaller units (tokens), and then applies vectorization techniques to convert the tokens into numerical representations. This transformation is achieved using tools like CountVectorizer for traditional models and Keras Tokenizer for deep learning, which map the tokens into a format that machine learning models can process. Vectorization enables the system to convert raw textual data into a structured form suitable for analysis by both traditional and deep learning algorithms.

Once the data is preprocessed, it passing into the model training phase, where two distinct approaches are employed. The first approach utilizes traditional machine learning models such as Naive Bayes and K-Nearest Neighbors (KNN). These models rely on frequency-based features derived from the preprocessed text. Naive Bayes uses probability theory to classify sentiment by evaluating the likelihood of a word occurring in different sentiment categories. KNN, on the other hand, is a distance-based model that classifies the sentiment based on the nearness of feature vectors in the training data. These traditional models are efficient and provide a quick baseline for sentiment analysis tasks

The second approach involves a more sophisticated deep learning model, specifically a Bidirectional Long Short-Term Memory (BiLSTM) network. Unlike traditional models, BiLSTM can process sequential data in both forward and backward directions, making it capable of learning long-term dependencies and context in the text. This event makes it particularly useful for capturing the complexities of natural language, where the meaning of a word can change based on its context. The BiLSTM model consists of layers designed to extract relevant features from the text, followed by a dense layer that classifies the sentiment based on these extracted features.

Finally, once the models are trained, the system enters the evaluation phase, where performance metrics such as accuracy, precision, recall, and F1 scores are calculated. These metrics allow for a comprehensive comparison of the models' effectiveness.The system's modular design allows for flexibility in model selection and fine-tuning, ensuring scalability and adaptability for different datasets or tasks. By leveraging both traditional machine learning methods and deep learning models, the system offers a robust and versatile solution for sentiment analysis

IV. METHODOLOGY

The methodology of the proposed sentiment analysis system is structured to efficiently process raw text data, transform it into meaningful representations, and classify sentiments using a hybrid approach that combines traditional machine learning models with deep learning techniques. This system is designed to ensure scalability and adaptability to various datasets, with a robust preprocessing pipeline and model integration.The process begins with data preprocessing, where raw textual data is cleaned to remove noise such as URLs, mentions, hashtags, and special characters. This step also involves normalizing the text to lowercase and handling missing or duplicate data. The cleaned text is then tokenized and transformed into numerical representations using techniques like CountVectorizer for traditional machine learning models or Keras Tokenizer for deep learning models. These vectorization methods allow the system to convert unstructured text into structured data that can be processed by machine learning algorithms.Once preprocessed, the data is split into training and testing sets to ensure unbiased evaluation of model performance. For traditional models, features are extracted using bag-of-words or TF-IDF methods, which represent word frequencies or importance. The machine learning models employed include Naive Bayes and K-Nearest Neighbors (KNN), chosen for their simplicity and effectiveness in baseline sentiment classification tasks.The deep learning component employs a Bidirectional Long Short-Term Memory (BiLSTM) network, which captures sequential dependencies in the text by processing data in both forward and backward directions. The BiLSTM model is built with an embedding layer to convert words into dense vector representations, followed by LSTM layers for feature extraction, dropout layers to prevent overfitting, and dense layers for classification.

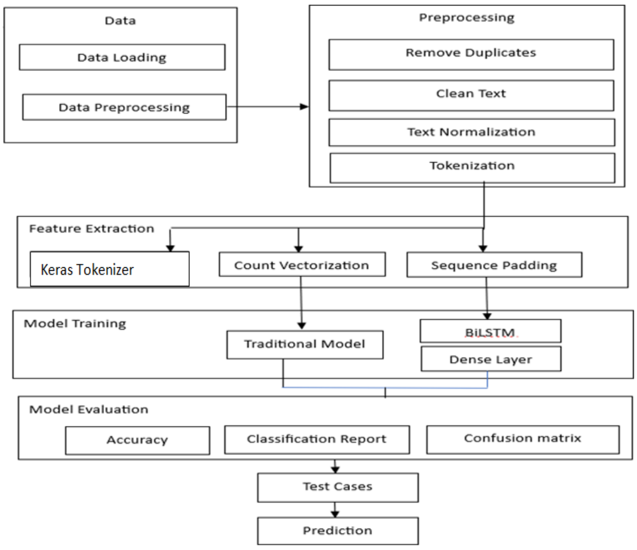

Figure 3: Work flow of methods in X Classification

The output layer uses a softmax activation function for multi-class sentiment classification. The integration of these models allows for a hybrid approach, leveraging the computational efficiency of traditional models and the contextual understanding provided by deep learning. During the evaluation phase, the system generates performance metrics such as accuracy, precision, recall, and F1 score for each model. This comparative analysis provides insights into the effectiveness of the hybrid approach. The following diagram illustrates the architecture of the proposed system, showcasing the flow of data from preprocessing to sentiment classification.

V. RESULT ANALYSIS

The evaluation of the Bidirectional Long Short-Term Memory (BiLSTM) model for sentiment analysis and classification on a X dataset demonstrated outstanding performance. The model achieved an impressive accuracy of 96.43%, surpassing baseline models like Naive Bayes and K-Nearest Neighbor (KNN). The performance metrics showed high precision, recall, and F1-scores across all sentiment categories. Specifically, for negative sentiments, the model achieved 89% precision, 90% recall, and a 90% F1-score. Neutral sentiments exhibited the highest performance with 97% precision, 98% recall, and a 98% F1-score, while positive sentiments recorded 97% precision, 96% recall, and a 97% F1-score.

```
Test Accuracy: 0.9643

Classification Report:
              precision    recall  f1-score   support

    negative       0.89      0.90      0.90      3160
     neutral       0.97      0.98      0.98     13671
    positive       0.97      0.96      0.97     12018

    accuracy                           0.96     28849
   macro avg       0.95      0.95      0.95     28849
weighted avg       0.96      0.96      0.96     28849
```

Figure 4: Classification Report

The comparison of model accuracies through a bar chart highlighted the BiLSTM's superior accuracy relative to Naive Bayes and KNN models.

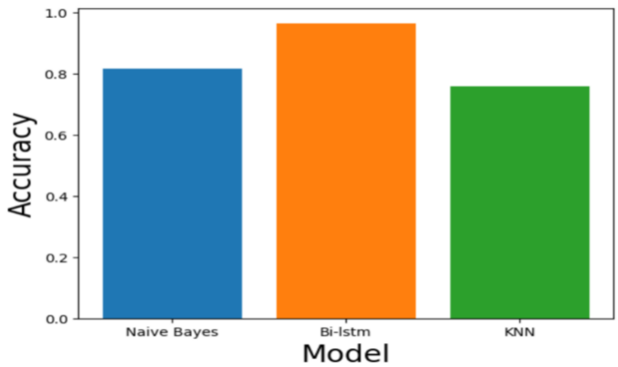

Figure 5: Comparison Model

This performance improvement is attributed to the model's ability to capture bidirectional dependencies, enabling a deeper contextual understanding of text sequences. The neutral sentiment category's exceptional metrics further validates the model's strength in analyzing tweets with subtle emotional nuances. Overall, the BiLSTM model exhibited balanced and consistent performance, making it a reliable tool for sentiment classification and recommendation tasks. Future Improvements may focus on extending the model's application to multilingual datasets to enhance its generalization capability. This result establishes the BiLSTM model as a robust framework for sentiment analysis and recommendation tasks on social media platforms.

VI. CONCLUSION

In this study, we developed a sentiment analysis system using a Bidirectional Long Short-Term Memory (BiLSTM) model for classifying and recommending tweets related to products. The BiLSTM model was integrated with preprocessing techniques and evaluated using performance metrics such as accuracy, precision, recall, and F1-score. The system effectively analyzed sentiments (positive, negative, neutral) and provided actionable recommendations based on the results. The results demonstrate that the BiLSTM model's ability to capture dependencies in both directions of text sequences significantly enhances classification accuracy. This framework provides a robust solution for abstracting insights from informal, contextually rich social media data like tweets. While the model performs well, improvements can be made to handle real-time tweet processing and multilingual data for broader applicability. Future work includes optimizing the system for real-time performance, expanding its capability to process multiple languages, and improving its ability to understand nuances such as sarcasm and irony in textual data. The system's design ensures scalability and adaptability, making it suitable for practical deployment in social media analytics and recommendation systems.

REFERENCES

[1] Xinyi Li, Yinchuan Li, Hongyang Yang, Liuqing Ypang, XiaoYang Liu. DP- LSTM:Differential Privacy- inspired LSTM for Stock Prediction Using Financial News. Last accessed 2020.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] Sidra Mehtab, Jaydip Sen. A Robust Predictive Model for Stock Price Prediction Using Deep Learning and
Natural Language Processing. Last accessed 2023.

[3] Dev Shah, Haruna Isah. Predicting the Effects of News Sentiments on the Stock Market. Last accessed 2023.

[4] Narayana Darapaneni, Anwesh Reddy Paduri, Himank Sharma, Milind Manjrekar, Pranali Bhagat, Yogesh Agarwal. Stock Price Prediction using Sentiment Analysis and Deep Learning for Indian Markets. Last accessed 2022.

[5] Mehar Vijha, Deeksha Chandola, Vinay Anand Tikkiwal, Arun Kumar. Stock CloPrice Prediction using Machine Learning Techniques. Last accessed 2023.

[6] Halyna Padalko , Vasyl Chomko and Dmytro Chumachenko, (2024) A novel approach to fake news classification using LSTM-based deep learning models

[7] Wiwik Anggraeni1, Moch Farrel Arrizal Kusuma1 Edwin Riksakomara1 Radityo P. Wibowo1 Pujiadi2 Surya Sumpeno3 .(2023).Combination of BERT and Hybrid CNN-LSTM Models for Indonesia Dengue Tweets Classification.

[8] Diksha Bansal ,Naveen Saini and Sriparana Saha.(2021). DCBRTS: A Classification-Summarization Approach for Evolving Tweet Streams in Multiobjective Optimization Framework. Digital Object Identifier 10.1109/ACCESS.2021.3120112 .

[9] Anuradha Khattar and S.M.K.Quadri .(2022). CAMM: Cross Attention Multimodal Classification of Disaster- Related Tweets. Digital Object Identifier 10.1109/ACCESS.2022.3202976 .

[10] Christopher Ifeanyi Eke , Mwenge Mulenga , Musa Phiri, Ahmad Sami Al-Shamayleh, Mohammad Imran and Dekera Kenneth Kwaghtyo.(December 11th, 2023). Effective Tweets Classification for Disaster Crisis based onEnsemble of Classifiers. https://doi.org/10.21203/rs.3.rs- 3721532/v1.

[11] Orlando Iparraguirre-Villanueva ,Melquiades Melgarejo-Graciano , Gloria Castro-Leon , Sandro Olaya- Cotera , John Ruiz-Alvarado , Andrés Epifanía-Huerta , Michael Cabanillas-Carbonell and Joselyn Zapata- Paulini.(2023).Classification of Tweets Related to Natural Disasters Using Machine Learning Algorithms https://doi.org/10.3991/ijim.v17i14.39907

[12] Jieli Chen , Kah Phooi Seng , Li Minn Ang , Jeremy Smith and Hanyue Xu. (2024). AI-Empowered Multimodal Hierarchical Graph-Based Learning for Situation Awareness on Enhancing Disaster Responses.

[13] Sanjiban Sekhar Roy, Akash Roy, Pijush Samui, Mostafa Gandomi, and Amir H. Gandomi, (2024), Hateful Sentiment Detection in Real-Time Tweets: An LSTM- Based Comparative Approach .

[14] Siva Krishna Dasari, Srinivas Gorla and Prasad Reddy P VG D.(2023). DPHFM: A Deep Parallel Hybrid Fusion Model for Disaster Tweet Classification on Tweeter Data. https://doi.org/10.21203/rs.3.rs3230920/v1.