Study And Implementation of Sentiment Analysis in Social Media

Shivani Sangwan[1], Akhilesh Kumar Bhardwaj[2], Dr. Tarun Kumar[3]
[1]*Research Scholar, Computer Science and Engineering, Shri Krishna Institute of Engineering and Technology, Kurukshetra*
[2]*Assistant Professor, Computer Science and Engineering, Shri Krishna Institute of Engineering and Technology, Kurukshetra*
[3]*Assistant Professor, Computer Science and Engineering, Global Research Institute of Management and Technology, Nachraun (Radaur)*

Abstract- Twitter has become a powerful place for expressing public opinion in the digital age, making it a useful tool for sentimental research. This study investigates how deep learning methods, especially Long Short-Term Memory (LSTM) networks, can be used to sort the feelings expressed in tweets into positive, negative, and neutral groups. Because tweets are short and casual, traditional machine learning algorithms often don't pick up on small differences in context. LSTM is a more advanced way to understand the emotional tone of short texts since it can keep track of long-term dependencies in sequential data. This study uses a model structure with an embedding layer, LSTM units, and a Softmax activation function in the output layer to classify multiple classes. The Softmax function makes sure that the model gives a probability distribution among sentiment classifications, which makes predictions easier to understand and more certain. It used the Adam optimizer to speed up the learning process since it has a learning rate that can be changed and updates based on momentum. The SoftMax model got a training accuracy of 93.02% and a validation accuracy of 96.53%, which were higher than the Sigmoid model's accuracies of 92.1% and 92.2%, respectively. This means that the SoftMax-based model not only learned from the training data faster, but it also did a better job of applying what it learned to new data.

Keywords: Sentiment Analysis, social media, Accuracy, Twitter etc.

1. INTRODUCTION

Sentiment analysis, often known as opinion mining, is a type of natural language processing (NLP) that looks for and sorts of emotions, opinions, or attitudes that are expressed in text. The main goal is to find out if a certain piece of writing, like a tweet, review, or comment, has a good, negative, or neutral mood. In the digital age, where a lot of user-generated content is made every day on sites like Twitter, Facebook, and online forums, this method has become more important [1].

The first step in sentiment analysis is to gather data. The next step is to use text preprocessing methods including tokenization, stop word removal, and normalization. After the text is cleaned up, it is turned into numbers using methods like Bag of Words, TF-IDF, or word embeddings. Next, labeled datasets are used to train machine learning or deep learning models to find patterns that are related to different types of sentiment. These models include both basic classifiers like Naïve Bayes and Support Vector Machines and more advanced designs like LSTM (Long Short-Term Memory) networks and transformer-based models like BERT [2].

Sentiment analysis can be used in many different fields. In business, it lets companies use customer feedback, brand reputation, and product reviews to make decisions based on facts. Politicians use it to find out how people feel about legislation, leaders, or events. It also influences financial markets since how investors feel about news or social media can change how they trade. Sentiment analysis is also being used more and more in mental health research, disaster response, and social scientific research to better understand how people feel and act [3].

Sentiment analysis is useful, but it faces a lot of problems. Language that is unclear, sarcastic, or full of slang, as well as cultural context, might make it hard for models to accurately read sentiment. The presence of mixed feelings in a single text or specialized language may make it harder to accurately classify things. However, as deep learning, contextual embeddings, and multilingual NLP continue to improve, sentiment analysis systems are becoming more accurate and flexible, making them more resilient and perceptive in a variety of real-world situations [4].

Twitter Sentiment Analysis

Twitter sentiment analysis uses natural language processing (NLP) and machine learning to figure out how people feel about tweets. Twitter's short, informal style, which includes a 280-character restriction, hashtags, emojis, and slang, makes it hard to classify emotions. The goal is to sort tweets into groups based on their feelings, such as good, negative, or neutral. This will help businesses, researchers, and legislators understand what people think in real time [5].

This type of analysis usually involves collecting tweets using keywords or hashtags, cleaning up the text to get rid of unnecessary parts (such URLs, mentions, and special characters), and then using a classification model, like an LSTM neural network, to predict sentiment. Twitter sentiment analysis is widely used in fields including brand monitoring, political forecasting, crisis management, and consumer behavior research. It gives important information about how people feel about events, products, or people [6].

2. REVIEW OF LITERATURE

Nip et al. (2024) [7] posited that social media sentiment analysis involved the algorithmic identification and extraction of human subjective assessments of entities available on social media platforms. Sentiment analysis had been performed on discrete written texts, generally categorizing sentiment into positive, negative, and neutral classifications. Sentiment analysis in social media encompassed multi-modal texts, temporal dynamics, interactions, network relationships, and sentiment propagation. Emotions and the intensity of sentiments were also identified. To evaluate the practical efficacy of machine learning and deep learning techniques for sentiment analysis, Briciu et al. (2024) [8] conducted a thorough investigation using Romanian reviews as a case study. According to the results, supervised deep learning techniques yield the best results at the document level. The dynamics of sentiment creation and spread within the public agenda were analyzed by Huh et al. (2024) [9] by looking at the impact of news sentiment on sentiment on social media. CNN's overall stance was influenced by Hillary Clinton's supporters' sentiments, while Fox News's harsh tone influenced Trump supporters' thoughts. The results demonstrated that the public agenda and the media have an impact on public mood, underscoring sentiment as a crucial component in understanding public opinion. Gupta et al. (2024) [10] examined the application of five traditional machine learning algorithms to categorize Twitter hate speech as neutral, racist, or sexist. The performance of the model had been evaluated using both raw twitter data and pre-processed tweets following data cleanup. Moreover, it emphasized two approaches for addressing imbalanced datasets that enhanced prediction accuracy. It achieved a 96% accuracy in classifying tweets into various classifications. To enhance understanding of the academic landscape of social media marketing research from the previous decade, Shaheen et al. (2025) [11] conducted a thorough bibliometric analysis. The findings emphasized the increasing importance of social media as a transformative marketing tool. A thorough analysis of the field's prominent countries, top institutions, prolific authors, and high-impact journals was among the key insights. Tourism marketing, AI-driven marketing, digital advertising, and virtual marketing emerged as research themes. The research demonstrated the necessity of localized research frameworks and the improved integration of Arabic databases into global academic discussions, particularly in the Arab region, due to regional disparities in research output. This analysis provided actionable insights for businesses, policymakers, and marketers who advocated for platform-specific marketing strategies, ethical AI regulations, and the necessity of localized marketing approaches that accounted for cultural nuances. Mohammed et al. (2025) [12] introduced a Generalized Sentiment Analytics Framework (GSAF) that enabled the comprehension of public sentiments regarding various critical societal issues in real time. The framework utilized natural language processing techniques to compute sentiments and presented them in a variety of emotions, utilizing publicly available social media data (i.e., X threads (formerly Twitter)). More than 3 million tweets were utilized to map, analyze, and visualize public sentiment on various societal issues by state in the United States as a case study of the developed framework. This study leveraged the extensive user base of X, a critical social media platform, to offer real-time insights into the affective responses to significant societal and political events. The platform offered users interactive visualizations of emotion-specific sentiments, including wrath, joy, and trust, which were displayed on a U.S. state-level choropleth map. This platform had been developed using R and the ShinyWeb framework. Advanced text-processing techniques were implemented to filter and clean tweet data for robust analysis, and keyword-based queries were permitted.

3. OBJECTIVES OF STUDY

The objectives of this work is to develop an algorithm for sentiment analysis of social media applications like twitter etc.

4. RESEARCH METHODOLOGY

Using an LSTM (Long Short-Term Memory) model to analyze Twitter sentiment means using deep learning techniques to sort tweets into three groups: positive, negative, or neutral. This is done by finding patterns in the order of words and small differences in meaning in short social media content. Unlike traditional machine learning methods that rely heavily on feature engineering, LSTM models can learn on their own and keep track of connections between words over time. This makes them especially good at figuring out the informal, emotional, and context-sensitive aspects of tweets.

The first step in the process is data preprocessing, which involves cleaning up the tweets by removing URLs, mentions, hashtags, emojis, and special characters. After that, the text is broken up into words and turned into sequences of numbers using embedding methods like Word2Vec, GloVe, or embedding layers that are part of the model itself. These embeddings change each word into a dense vector that contains its meaning. Because tweets are written in order, LSTM networks use memory cells and gating mechanisms to keep and understand important information from previous words when they guess the sentiment at the end of the series.

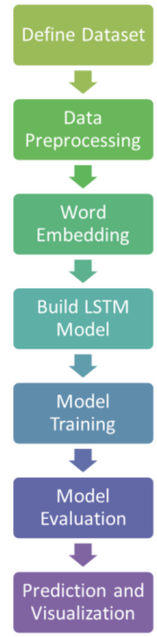

Fig 1: Twitter Sentiment Analysis using LSTM Model

1. Steps of Data Pre-Processing

- Take out URLs, usernames, hashtags, emojis, and special characters.
- Change the text to lowercase.
- Break the tweets down into separate phrases.
- Get rid of stop words like "is," "the," and "and."
- Used stemming or lemmatization.
- Ensure the LSTM model gets the same length of input every time, pad or cut off sequences.

The preprocessing of textual data for an LSTM network comprised multiple steps: tokenization, sequence transformation, and sequence padding. It examines each item individually:

- Tokenization: This procedure entails deconstructing the review content into discrete words or "tokens". It utilized the Tokenizer class from Keras, which also manages the conversion of words to lowercase and the elimination of punctuation.
- Token transformation: Subsequently, it transformed the tokens into sequences of integers. Every distinct word was allocated a specific integer. These sequences furnish a format compatible with our LSTM network.

- Sequence padding: LSTM networks necessitate input data of uniform length; hence, we employed padding to standardize the length of all sequences.
- Label encoding: Ultimately, it transformed our sentiment labels (positive, neutral, negative) into numerical representation, as models perform more effectively with numerical data.

After training, the model is tested on new twitter data using metrics including accuracy, precision, recall, and F1-score. Confusion matrices and ROC curves are examples of visualizations that can help check how well a model works. When it comes to capturing the emotional subtleties and grammatical variation in tweets, LSTM-based models work better than simpler models. Still, adding attention mechanisms or combining them with pretrained language models like BERT can make them work better. LSTM sentiment analysis is a good way to figure out how people feel about things on platforms like Twitter, especially where timing and nuance are important.

5. PERFORMANCE METRICS

1. Accuracy

A metric called accuracy quantifies the frequency with which a model's predictions come true. It is the proportion of accurate forecasts made relative to all predictions.

$$Accuracy = \frac{\text{True Positives} + \text{True Negatives}}{\text{True Positives} + \text{True Negatives} + \text{False Positives} + \text{False Negatives}} \quad (1)$$

2. Precision

Precision is the percentage of correctly categorized instances among all the cases that were classified.

$$Precision = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (2)$$

3. Recall

In deep learning, recall is a metric that evaluates a machine learning model's ability to accurately identify positive examples within a dataset. It is also referred to as the true positive rate (TPR), sensitivity, likelihood of detection, and hit rate.

$$Recall = \frac{\text{True Positive}}{\text{Predicted Results}} \quad (3)$$

4. F1-Score

The F1 score is a statistic utilized to evaluate the efficacy of machine learning models by reconciling precision and recall. It is frequently employed in binary classification tasks, including healthcare, information retrieval, and fraud detection.

$$F1\,Score = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

6. RESULTS OF SYSTEM

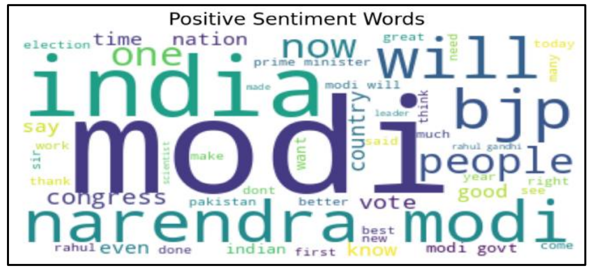

Fig 2: Positive Sentiment Words

The Bag of Words (BoW) model is a basic and widely used way in natural language processing (NLP) to mathematically represent text input. In sentiment analysis, especially when looking at Twitter data, it lets you count how many times certain phrases appear in tweets and how they relate to the feelings expressed in those tweets, whether they are good, negative, or neutral. The Bag of Words model describes text by:

- Not caring about how the syntax is set up
- Finding out how often words occur: Every tweet is turned into a vector, with each dimension representing a different word from the whole corpus. The values of the vector show how often each term appears in the tweet.
- Bag of Words (BoW) makes it easier to find common words in positive, negative, and neutral tweets based on how often they appear in sentiment analysis.

Positive Sentiment Result of the Bag of Words

- The following words were most often used in tweets that were rated as good. Possible situations could include:
- "Great," "love," "thank you," "remarkable," "help," "content," "achievement"
- A lot of the time, these words show excitement, gratitude, or appreciation.

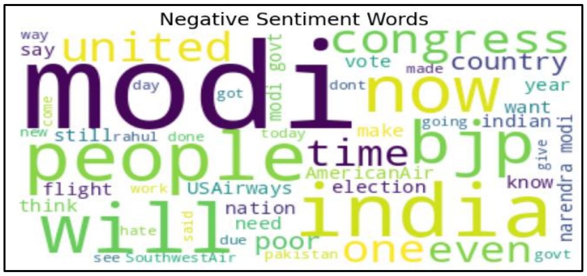

Fig 3: Negative Sentiment Words

Output of Negative Sentiment Bag of Words

- Some words that are often used in bad tweets are:
- "Detest," "falsehoods," "flounder," "issue," "depraved," "irate," and "dreadful"
- These show anger, irritation, or criticism.
- BoW shows how specialized, emotionally charged phrases are related to unhappiness.
- This helps brands or political analysts find language patterns that go along with anger or complaint.

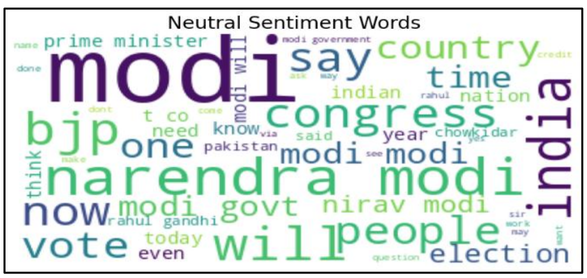

Fig 4: Neutral Sentiment Words

Output of the Neutral Sentiment Lexicon Most of the time, neutral tweets are about:

- Words that relate to information or facts include "meeting," "update," "report," "announced," "statement," and "details."
- These tweets show less emotion and more objectivity.
- The BoW for neutral sentiment is fair communication without strong positive or negative feelings.

Table 1: Performance Analysis using LSTM

Parameters	Values
Training Accuracy	0.9302
Validation Accuracy	0.9653
Precision	0.9331
Recall	0.9144
F1-Score	0.9237

- The performance evaluation of your LSTM model shows that it is a strong and well-calibrated classifier, notably good at handling Twitter sentiment data. The model is not overfitting the training data and generalizes well because the training accuracy is 93.02% and the validation accuracy is 96.53%. A high validation accuracy means that the model is still useful when it sees new tweets that it hasn't seen before. This is an important feature for any sentiment analysis application that is useful in the real world.
- A precision score of 0.9331 means that the model is right 93.31% of the time when it says a tweet has a certain sentiment (like positive). This is especially important when false positives, such wrongly labeling a neutral tweet as positive, cost a lot of money. A recall score of 0.9144 means that the model correctly identifies 91.44% of all real emotion cases. This shows that it is quite sensitive and has very few false negatives.
- The F1-score of 0.9237 is a harmonic mean of precision and recall, which means it balances both parts. The model not only makes good predictions, but it also fully captures sentiment classes, which is important in datasets with class imbalance, like Twitter, where neutral or positive comments may be more common. All of these tests show that the LSTM model is quite reliable and strong, and it can be used with confidence for real-time sentiment analysis or customer feedback monitoring.

7. CONCLUSION

Long Short-Term Memory (LSTM) networks have shown a lot of promise for accurately figuring out and classifying the emotional tone of tweets. Unlike traditional machine learning models, LSTM networks are designed to work with sequential data. This makes them quite good at understanding the meaning and flow of language in short, casual writing like tweets. The model can find subtle changes in emotion within a single tweet, like sarcasm or ambivalence, that less advanced models often miss. This is because it can keep track of and assess long-term dependencies. The LSTM model was quite good at telling the difference between positive, negative, and neutral attitudes, especially when it was trained on a large and varied dataset. The SoftMax model got a training accuracy of

93.02% and a validation accuracy of 96.53%, which were higher than the Sigmoid model's accuracies of 92.1% and 92.2%, respectively. This means that the SoftMax-based model not only learned from the training data faster, but it also did a better job of applying what it learned to new data. The improvement in validation accuracy is especially important since it shows that SoftMax helps the model make more confident and accurate predictions across a range of sentiment categories.

Future models may contain multilingual embeddings or translation layers, which would allow for analysis of Hindi, Tamil, Bengali, and other regional and global languages in addition to English. Adding attention mechanisms to LSTM topologies could help with understanding complex meanings like sarcasm, which simple models often have trouble with.

REFERENCE

[1] Alotaibi, A.; Nadeem, F. (2024) Leveraging social media and Deep Learning for Sentiment Analysis for Smart Governance: A Case Study of Public Reactions to Educational Reforms in Saudi Arabia. *Computers*, *13*, 280. https://doi.org/ 10.3390/computers13110280

[2] Vanaja S., Belwal M., (2018), "*Aspect-Level Sentiment Analysis on E-Commerce Data*", International Conference on Inventive Research in Computing Applications, pp. 245-251.

[3] Wang J., Zhao Y., (2010). "*Effective feature selection with particle swarm optimization based one dimensional searching*", IEEE,978(1):4244-6044.

[4] Woldemariam Y., (2016), " *Sentiment Analysis in A Cross-Media Analysis Framework*", IEEE Access Journal, pp. 355-359.

[5] Xue Z., Du P., (2013), " *A Novel Classification Technique for Hyper spectral Imagery Based on Harmonic Analysis, SVM and PSO*", IEEE Access Journal, pp. 1-4.

[6] Yuxi G., (2018), "*An Improved Hybrid Group Intelligent Algorithm Based on Artificial Bee Colony and Particle Swarm Optimization*", IEEE International Conference on Virtual Reality and Intelligent Systems, pp. 458-466.

[7] Nip, J.Y.M.; Berthelier, B. (2024), Social Media Sentiment Analysis. *Encyclopedia*, *4*, 1590–1598. https://doi.org/10.3390/ encyclopedia4040104

[8] Briciu, A.; C˘alin A.-D.; Miholca, D.-L.; Moroz-Dubenco, C.; Petras, cu, V.; Dasc˘alu, G. (2024), Machine Learning-Based Approaches for Multi-Level Sentiment Analysis of Romanian Reviews. *Mathematics*, *12*, 456. https://doi.org/ 10.3390/math12030456.

[9] Huh, C.U.; Park, H.W. (2024), Setting the Public Sentiment: Examining the Relationship between Social Media and News Sentiments. *Systems*, *12*, 105. https:// doi.org/10.3390/systems12030105

[10] Gupta, S., & Jawahar, P. (2024). *Twitter sentiment analysis using machine learning*. *Journal of Student Research, 13*(2). https://doi.org/10.47611/jsrhs.v13i2.6819.

[11] Shaheen, H. (2025). Social media marketing research: A bibliometric analysis from Scopus. *Future Business Journal, 11*(1), Article 46. https://doi.org/10.1186/s43093-025-00465-2.

[12] Mohammed, A.M.K.; Ali, G.G.M.N.; Khairunnesa, S.S. (2025), GSAF: An ML-Based Sentiment Analytics Framework for Understanding Contemporary Public Sentiment and Trends on Key Societal Issues. *Information*, *16*, 271. https:// doi.org/10.3390/info16040271