International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 06 | June - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

# Sentiment Analysis of Textual Data using Deep Learning

Prof.T. Ramya, J.Mary Ratna Manjari,Kanagandula Aravind,Dasari Aravind Sagar,
Daripally Akhilesh Goud,Yenugula Arun,
Artificial Intelligence & MachineLearning Department
Of Computer Science And Engineering Malla Reddy
University, Hyderabad,Telangana,India

**Abstract**: Textual data is generated in large volume everyday over internet which makes sentiment analysis important
. Sentiment analysis of any textual data denotes the feelings and attitudes of the individual on particular topics or products
. It extracts the Sentiment polarity (negative ,neutral or positive ) from textual data using deep learning algorithms .Sentiment analysis of textual data involves using neural network architecture to analyze and determine the sentiment expressed in text .It is a subfield of text classification which involves analyzing people's opinions , emotions , and attitudes towards entities and their characteristics as expressed in a written text . It utilizes three deep learning algorithms which are Neural Networks , Long Short Term Memory (LSTM) , and Gated Recurrent Unit (GRU). The results of RNN, LSTM, and GRU obtain an excellent rate of accuracy . It can be concluded from the outcomes that the used preprocessing stages made a positive impact on the accuracy rate

**Keywords**: These are the neural network architectures employed for sentiment analysis such as
Convolutional Neural Networks (CNNs),

customer feedback analysis, social media monitoring, market research, and more.

1. Fundamentals of Sentiment Analysis

Recurrent Neural Networks (RNNs),
Long Short-Term Memory (LSTM) networks,
Gated Recurrent Units (GRUs),
Transformer models

## I.INTRODUCTION

The primary objective is to develop deep learning model to accurately classify the sentiment expressed in a piece of text as positive,neutral or negative. This involves processing and analyzing large amounts of text data to understand the underlying sentiment and emotions conveyed by the language. Deep learning techniques are used to automatically learn and extract features from the text data to make accurate sentiment predictions. This typically requires training a deep learning model on a large dataset of labeled text samples to learn the patterns and nuances of sentiment in language. The goal is to create a system that can automatically analyze and understand the sentiment behind text data, enabling applications such as sentiment analysis in social media, customer reviews, and other text-based content
Sentiment analysis, a branch of natural language processing (NLP), has gained significant traction owing to its myriad applications across various domains, including

Sentiment analysis involves the automated extraction of sentiment or opinion from text data. Textual inputs, such as reviews, social media posts, and news articles, are analyzed to determine the sentiment expressed, typically categorized as positive, negative, or neutral.

2. Deep Learning Paradigm

Deep learning has emerged as a transformative paradigm in sentiment analysis, offering unparalleled capabilities in feature learning and representation

3. Key Deep Learning Models

Several deep learning architectures have been leveraged for sentiment analysis tasks, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Transformer models

## LITERATURE REVIEW

Sentiment Analysis (SA) is one of the active research field in text mining [2]. Sentiment analysis (SA) is also called Opinion analysis or Opinion mining, it is a subfield of natural language processing (NLP) that evaluates the degree of polarity in the sentence to analyze and extracts feelings from text data. [1].The advancement of technology produces an enormous number of social media users that generate huge data[1]. Furthermore, for every passing day the textual data is also increasing in amount which makes data mining especially sentiment analysis or opinion mining, a research hungry area.[5] This is mainly because of data is represented in the form of calculations about reviewers' comments, assessment, attitudes, behavior and emotions to individual issues, events, topics, services and attributes [5]. One of the main focuses of Natural Language processing (NLP) is to make use of Artificial Intelligence (AI) approaches to design and construct computational platforms. These platforms automate the process of extracting knowledge and previously unknown interesting patterns,from both structured and unstructured text sources [2] Sentiment Analysis (SA) is the method of extracting the contextual polarity (positive, negative, or neutral) of the text.[2]. Extracting such subtle emotions from texts present some unique problems as one single sentence may represent multiple related emotions. The Deep learning models based on NeuralNetwork, Long Short Term Memory,Recurrent Neural Network(for classifying textual data into one or more fine grained emotions) [3]. Emotions are crucial part of any social interaction.And hence it has been a subject of research to extract emotions from data. It can lead to imporvements in fields such as Mental Health Diagnosis and Human-computer interaction. It classifies text into one of five or six board emotions[3].It examines the impact of the utilized preprocessing stages on the model's accuracy and also on the heterogeneous data collected from different sources.[1] Data is collected from multiple social media sources[1] Current approaches to SA are mainly based on supervised learning which is based on manual labeled samples like commercial product reviews and other social networks product reviews in which the overall attitude of customers is explicitly specified in the reviews . Also deep learning and machine learning models can improve text based application[ 2] RNN, LSTM, and GRU obtain an excellent rate of

Literature reviews in the field of sentiment analysis of textual data using deep learning

1.Data Collection:

1. **Identify Data Sources**: Determine sources of textual data relevant to the sentiment analysis task, such as social

encompass a vast array of research endeavors, spanning from seminal works to the latest advancements. Here's a synthesized overview

## II. PROBLEM STATEMENT

The problem statement for sentiment analysis of textual data using deep learning typically encompasses several key components, including the context, objectives, data considerations, and potential challenges. Here's a structured breakdown. Sentiment analysis, a subfield of natural language processing (NLP), aims to automatically determine the sentiment or opinion expressed in textual data. With the exponential growth of online content, including social media posts, product reviews, and news articles, there's an increasing demand for efficient and accurate sentiment analysis techniques to extract valuable insights from unstructured text

**Developing Effective Models** The primary objective is to design and develop deep learning models capable of accurately analyzing sentiment in textual data. These models should be capable of handling the nuances of human language, including sarcasm, irony, and context-dependent sentiment expressions.

**Achieving High Accuracy**: The goal is to achieve high accuracy in sentiment prediction across different sentiment classes (positive, negative, neutral), enabling stakeholders to make informed decisions based on sentiment insights.

**Scalability and Efficiency**: Ensuring that the sentiment analysis models are scalable and efficient, capable of processing large volumes of text data in real-time, is crucial for practical deployment in various applications.

Commonly used datasets include IMDb movie reviews, Yelp reviews, Twitter sentiment datasets, and product review datasets.

**Data Preprocessing**: Preprocessing textual data involves tasks such as tokenization, removing stopwords, stemming or lemmatization, and handling special characters and emoticons to prepare the data for training deep learning models effectively.

**Handling Contextual Nuances**: Deep learning models need to be robust enough to capture and understand contextual nuances in language, including slang, colloquialisms, and cultural references, to avoid misinterpretation of sentiment.

**Domain Adaptation**: Adapting sentiment analysis models to specific domains or industries where language usage may vary significantly requires techniques for domain adaptation or fine-tuning on domain-specific data.

The problem statement for sentiment analysis of textual data using deep learning revolves around developing accurate, scalable, and efficient models capable of discerning sentiment from unstructured text data. Addressing data considerations, potential challenges, and aligning objectives with real-world applications are essential for advancing research and practical deployment of sentiment analysis solutions

## III. METHODOLOGY

The methodology for sentiment analysis of textual data using deep learning typically involves a systematic approach encompassing data collection, preprocessing, model selection and training, evaluation, and potentially fine-tuning or optimization. Here's a structured outline:

media platforms, review websites, or specific domain-related documents.

2. **Acquire Annotated Datasets**: Gather labeled datasets with sentiment annotations, ensuring diversity and representativeness across sentiment classes (positive, negative, neutral).
3. **Ethical Considerations**: Ensure compliance with data privacy regulations and ethical guidelines, especially when dealing with user-generated content.

2. Data Preprocessing:

1. **Text Cleaning**: Remove noise from the textual data by eliminating HTML tags, special characters, and irrelevant symbols.
2. **Tokenization**: Segment the text into individual tokens (words or subwords) to facilitate further processing.
3. **Normalization**: Convert text to lowercase, standardize spellings, and handle contractions to ensure uniformity.
4. **Stopword Removal**: Eliminate common stopwords (e.g., "and," "the," "is") that carry little semantic meaning.
5. **Stemming or Lemmatization**: Reduce words to their root form to consolidate variations of the same word.
6. **Encoding**: Convert tokens into numerical representations using techniques like one-hot encoding or word embeddings (e.g., Word2Vec, GloVe).

3. Model Selection and Training:

1. **Architecture Selection**: Choose a suitable deep learning architecture based on the characteristics of the data and the complexity of the sentiment analysis task (e.g., CNNs, RNNs, Transformer models).
2. **Hyperparameter Tuning**: Optimize hyperparameters such as learning rate, batch size, dropout rate, and model architecture to enhance performance.
3. **Model Training**: Train the selected deep learning model using the preprocessed data, utilizing optimization algorithms like SGD, Adam, or RMSprop.
4. **Transfer Learning (Optional)**: Utilize pre-trained language models (e.g., BERT, GPT) for transfer learning, either by fine-tuning on sentiment-specific tasks or feature extraction.

4. Evaluation:

1. **Selection of Metrics**: Choose appropriate evaluation metrics (e.g., accuracy, precision, recall, F1-score, ROC-AUC) based on the characteristics of the sentiment analysis task and the dataset.
2. **Validation Strategy**: Split the dataset into training, validation, and test sets to assess model performance effectively.
3. **Model Evaluation**: Evaluate the trained model on the test set using the selected metrics to measure its accuracy and generalization capability.

5. Fine-Tuning and Optimization (Optional):

1. **Fine-Tuning**: Refine the model parameters or architecture based on feedback from the evaluation phase to improve performance further.
2. **Regularization Techniques**: Implement regularization techniques such as dropout, batch normalization, or early stopping to prevent overfitting and enhance model robustness.
3. **Model Compression (Optional)**: Reduce the model size or complexity through techniques like pruning, quantization, or knowledge distillation to improve efficiency in deployment scenarios.

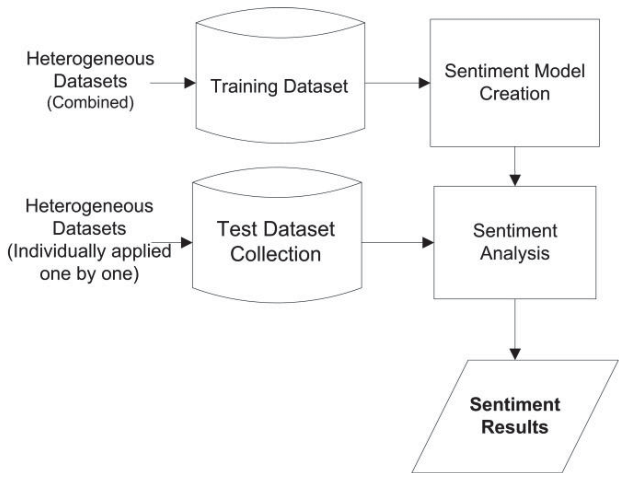
6. Deployment:

1. **Integration**: Integrate the trained model into the desired application or system

architecture, ensuring compatibility and scalability.

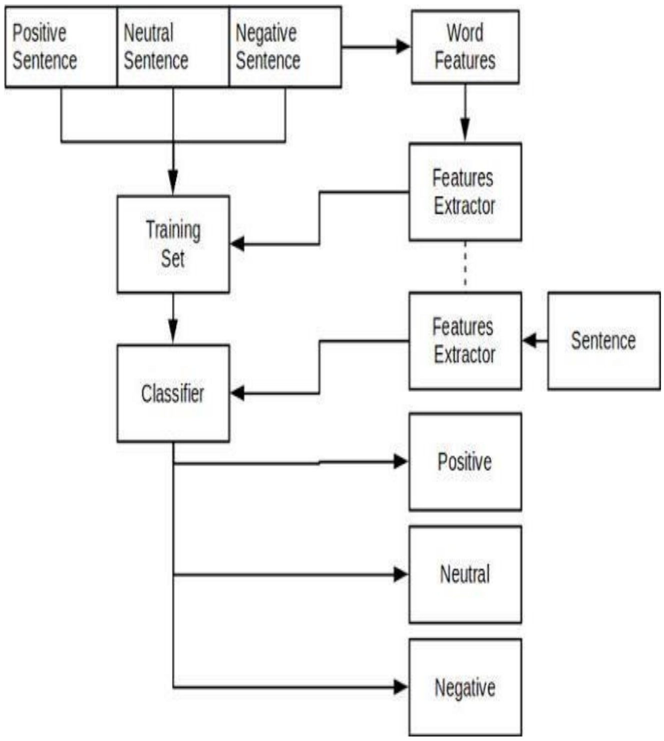
2. **Monitoring and Maintenance**: Continuously monitor model performance in production environments and update as necessary to adapt to evolving data distributions or requirements.

## ARCHITECTURE




A data set for sentiment analysis on textual data typically consists of a collection of text samples along with corresponding sentiment labels, such as positive, negative, or neutral. Each text sample could be a sentence, a review or any other form of text. The sentiment labels indicate the overall sentiment or emotion expressed in the text. This data set is used to train machine learning models to accurately predict the sentiment of new, unseen text data. The goal is to analyze and understand the sentiment oropinion conveyed in the text.

## ER DIAGRAM




Creating an Entity-Relationship (ER) diagram for sentiment analysis of textual data using deep learning involves identifying the key entities and their relationships within the system.This ER diagram provides a high-level overview of the entities involved in sentiment analysis of textual data using deep learning and their relationships. It serves as a foundational blueprint for designing and implementing the system architecture. Depending on specific requirements and system complexity, additional entities and relationships may beincluded in the diagram

## VI. EXPERIMENTAL RESULTS

The results of sentiment analysis of textual data using deep learning typically include insights derived from analyzing the sentiment expressed in the text data

**Positive**: Text expressing positive sentiment, such as satisfaction, happiness, or approval.

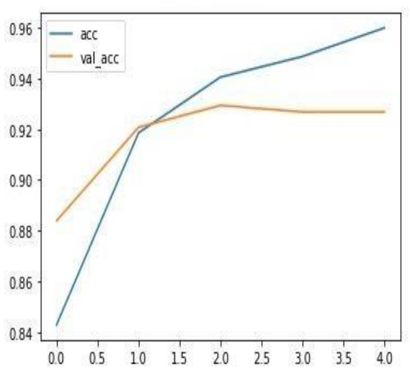
**Negative** : Text conveying negative sentiment, including dissatisfaction, sadness, or disapproval.

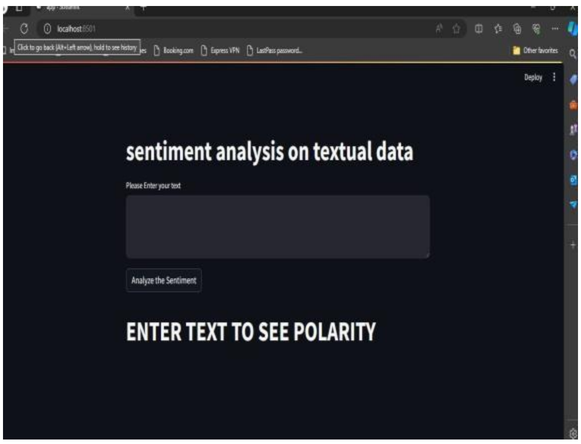
**Neutral**: Text that does not exhibit any strong sentiment polarity, often factual or objective in nature

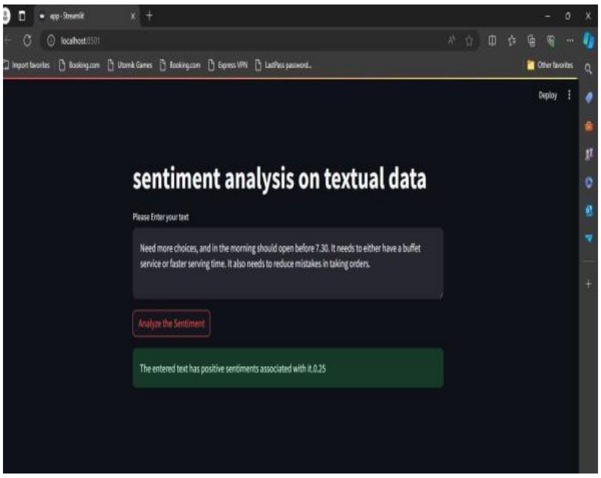
**Model Implementation and Training**

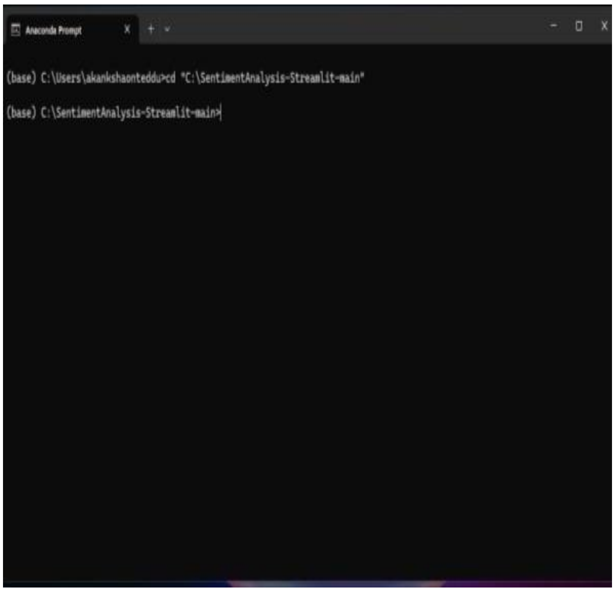
```
Model: "sequential"
_________________________________________________________________
 Layer (type)            Output Shape          Param #
========================================
 embedding (Embedding)   (None, 500, 35)
175000
 lstm (LSTM)             (None, 100)           54400
 dense (Dense)           (None, 1)             101
========================================
Total params: 229501 (896.49 KB)
Trainable params: 229501 (896.49 KB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

**Model Accuracy**



## VII. CONCLUSION

In conclusion, sentiment analysis on textual data using deep learning is successful in accurately categorizing text into positive, negative, or neutral sentiments. It helps in achieving high accuracy levels through the use of deep learning techniques such as LSTM and CNN, as well as through the use of pre-trained word embeddings .It demonstrats the effectiveness of deep learning models in sentiment analysis tasks, showing that they can handle complex text data and extract meaningful sentiment information.Overall, this project has provided valuable insights into the application of deep learning in sentiment analysis and has the potential for further research and development in the field.

## VIII. FUTURE WORK

Building a framework for sentiment analysis of textual data using deep learning involves structuring the workflow and processes involved in developing, training, and deploying deep learning models.

## IX. REFERENCES

[1] Sentiment Analysis of Financial Textual data Using Machine Learning and Deep Learning Models,Informatica 47(5):153-158

[2] A Survey on Concept-Level Sentiment Analysis Techniques of Textual Data

[3] Analysis of Underlying Emotions in Textual Data Using Sentiment Analysis Which Classifies Text In to Positive, Negative or Neutral Sentiments

[4] Financial sentiment analysis of tweets based on deeplearning approach

[5] Sentiment analysis using product review data ,2015

[6] Arif Ullah, Sundas Naqeeb Khan, Nazri Mohd Nawi, "Review on sentiment analysis for textclassification techniques from 2010 to 2021," Multimedia Tools and Applications, pp. 2-58, October 2022
.
[7] A Study of Sentiment Analysis: Concepts, Techniques and Challenges,
january 2019 DOI:10.1007/978-981-13-6459-4_16

[8] Sentiment Analysis using Deep Learning

[9] Decoding sentiment: A sentiment analysis model for movie reviews,2024