BASAVARAJESWARI GROUP OF INSTITUTIONS
BALLARI INSTITUTE OF TECHNOLOGY & MANAGEMENT
NBA and NACC Accredited Institution*
(Recognized by Govt. of Karnataka, approved by AICTE, New Delhi & Affiliated to Visvesvaraya Technological University, Belagavi) "Jnana Gangotri" Campus, No.873/2,
Ballari-Hosapete Road, Allipur, Ballari-583 104 (Karnataka) (India)
Ph:08392 – 237100 / 237190, Fax: 08392 – 237197

DEPARTMENT OF
COMPUTER SCIENCE ENGINEERING (ARTIFICIAL INTELLIGENCE)

Neural Network and Deep learning Project Report

On

"SENTIMENTAL ANALYSIS"

Submitted By

Maseera Fathima 3BR22CA031

Under the Guidance of

Mr. Azhar Baig

Asst. Professor

Dept of data science, BITM, Ballari

Visvesvaraya Technological University
Belagavi, Karnataka
2025-2026

Ballari-Hosapete Road, Allipur, Ballari-583 104 (Karnataka) (India)
Ph:08392 – 237100 / 237190, Fax: 08392 – 237197

DEPARTMENT
OF
COMPUTER SCIENCE ENGINEERING
(ARTIFICIAL INTELLIGENCE)

CERTIFICATE

Certified that the mini project work entitled **"SENTIMENTAL ANALYSIS"** carried out by **Maseera Fathima bearing USN 3BR22CA031** A Bonafide students of Ballari Institute of Technology and Management in partial fulfillment for the award of Bachelor of Engineering in computer science (artificial intelligence) of the Visvesvaraya Technological University, Belgaum during the year 2025- 2026. It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the report deposited in the departmental library. The project report has been approved as it satisfies the academic requirements in respect of the project work prescribed for the said Degree.

Signature of Lab Co-Ordinator's
Mr. Pavan Kumar
Mr. Vijay Kumar

11/12/25
Signature of HOD

Dr. Yeresime Suresh

ABSTRACT

Sentiment analysis is a natural language processing (NLP) technique used to automatically classify and interpret the emotional tone expressed in textual data, such as reviews, social media posts, or feedback. In this project, a sentiment analysis model was developed to categorize input text into sentiment classes (e.g., positive, negative, and neutral) using Python and machine learning/NLP methods. The raw textual dataset was first preprocessed using techniques such as tokenization, stop-word removal, and vectorization, enabling effective feature extraction. Subsequently, various classification approaches — such as machine learning algorithms (e.g., Logistic Regression, Support Vector Machines) or deep learning models — were trained and evaluated for sentiment prediction performance. The model's performance was assessed based on metrics like accuracy, precision, and recall, demonstrating a reliable ability to distinguish sentiment polarity in the chosen dataset. The results highlight the effectiveness of NLP techniques in extracting meaningful insights from unstructured text data, with potential applications in customer feedback analysis, social media monitoring, and market research. Future work will explore advanced models and larger datasets to further improve classification accuracy and robustness.

ACKNOWLEDGEMENT

The satisfaction that accompanies the successful completion of project work on the **"SENTIMENTAL ANALYSIS"** would be incomplete without mentioning those who made it possible. Their noble gestures, affection, guidance, encouragement, and support crowned our efforts with success. It is our privilege to express our gratitude and respect to all those who inspired us in the completion of this project.

We are extremely grateful to our Lab Coordinators, **Mr. Vijay Kumar**, Assistant Professor, BITM, Ballari, **Mr. Pavan Kumar**, Assistant Professor,

BITM, Ballari, and **Dr. Yeresime Suresh,** Professor and Head of the Department of CSE-AI, BITM, Ballari, for their noble gestures, support, coordination, and valuable suggestions in completing the project work.

We also thank the Principal, Management, and non-teaching staff of BITM, Ballari, for their coordination and valuable suggestions in completing the project work.

Name	USN
Maseera Fathima	3BR22CA031

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In today's digital age, vast amounts of textual data are generated every second through social media, online reviews, and customer feedback. Understanding the sentiment expressed in this text is crucial for businesses, researchers, and policymakers to gauge public opinion and make informed decisions. Sentiment analysis, also known as opinion mining, is a field of natural language processing (NLP) that automatically identifies and categorizes emotions in text, such as positive, negative, or neutral.

This project focuses on developing a sentiment analysis model capable of analyzing textual data and predicting the sentiment polarity accurately. By leveraging machine learning and NLP techniques, the model aims to extract meaningful insights from unstructured text data, providing valuable information for decision-making, trend analysis, and customer satisfaction monitoring.

CHAPTER 2

OBJECTIVES

 Data Collection and Preprocessing: Gather textual data from sources such as reviews, tweets, or feedback and preprocess it by removing noise, tokenizing, and converting text into numerical features suitable for modeling.

 Model Development: Implement and train machine learning or deep learning models (e.g., Logistic Regression, Naive Bayes, LSTM, or BERT) for sentiment classification.

 Performance Evaluation: Assess the model's accuracy, precision, recall, and F1-score to ensure reliable sentiment prediction.

 Sentiment Prediction: Apply the trained model to classify new, unseen text data into sentiment categories.

 Insight Extraction: Analyze the results to identify trends, patterns, and actionable insights from the textual data.

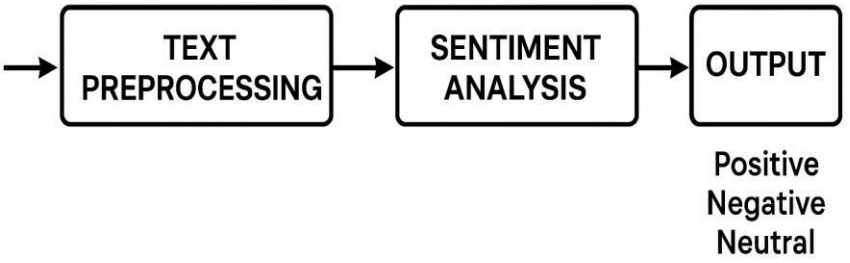
CHAPTER 3

PROBLEM STATEMENT

With the huge amount of text generated online, manually analyzing opinions is slow and inefficient. There is a need for an automated system that can accurately detect whether a text expresses positive, negative, or neutral sentiment. This project aims to develop such a sentiment analysis model using NLP and machine learning techniques to provide fast and reliable insights from textual data.

CHAPTER 4

METHODOLOGY



4.1 Block Diagram of Sentimental Analysis System

The block diagram shows where the huge amount of text generated online, manually analyzing opinions is slow and inefficient. There is a need for an automated system that can accurately detect whether a text expresses positive, negative, or neutral sentiment. This project aims to develop such a sentiment analysis model using NLP and machine learning techniques to provide fast and reliable insights from textual data.

CHAPTER 5

REQUIREMENT ANALYSIS

FUNCTIONAL REQUIREMENTS

 Text Input Handling
- The system shall accept textual data from users or datasets.
- The system shall allow input in multiple formats (e.g., CSV, TXT).

 Preprocessing of Text
- The system shall clean and normalize text (e.g., remove punctuation, lowercase, remove stopwords).
- The system shall tokenize and possibly stem/lemmatize the input text.

 Sentiment Detection
- The system shall classify input text into sentiment categories such as positive, negative, or neutral.
- The system shall provide a sentiment score or probability for each category.

 Model Training and Evaluation
- The system shall train a machine learning or deep learning model using labeled datasets.
- The system shall evaluate model performance using metrics like accuracy, precision, recall, F1-score.

NON-FUNCTIONAL REQUIREMENTS

 Performance

The system shall process and classify a batch of 1000 texts in under 5 minutes.

 Scalability

The system shall handle an increasing volume of input data without significant performance

 Usability

The system shall provide an intuitive interface for inputting text and viewing results.

CHAPTER 6

DESIGN

FLOW CHART

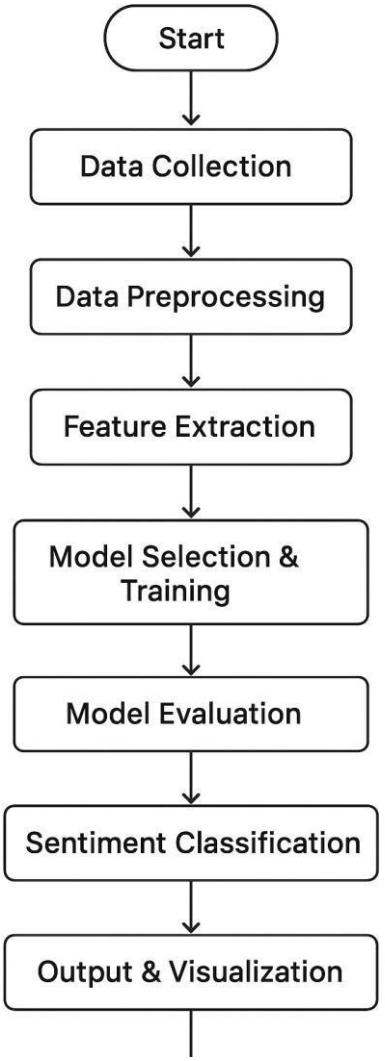


Fig 6.1 Flow Chart

USE CASE DIAGRAM

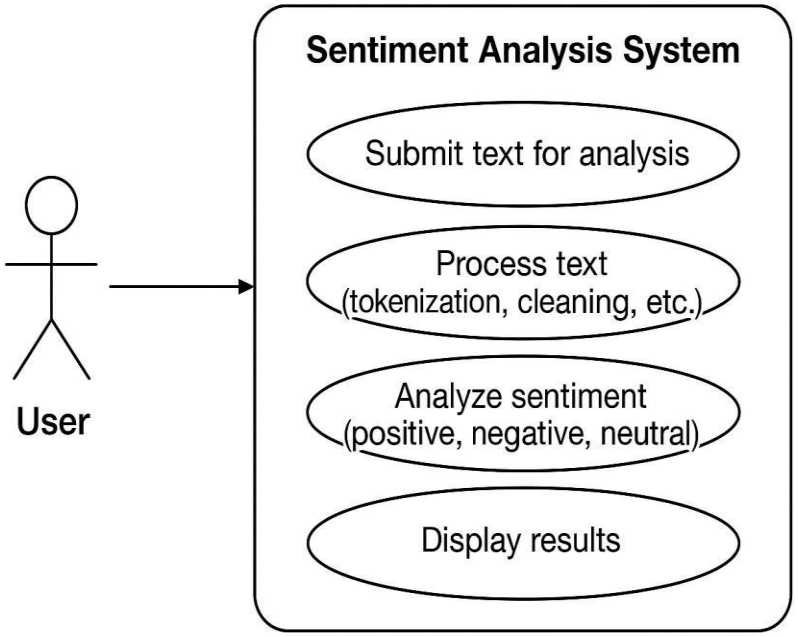


Fig 6.2 Use Case Diagram

SEQUENCE DIAGRAM

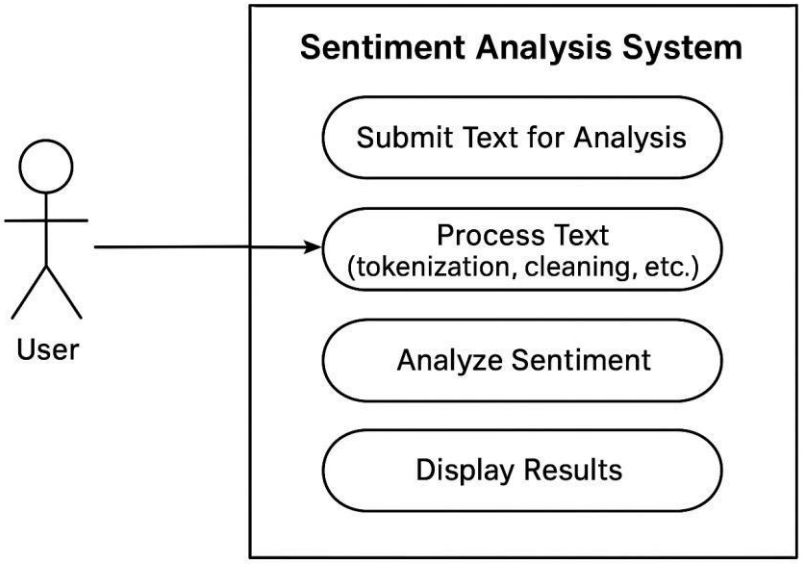


Fig 6.3 Sequence Diagram

.

CHAPTER 7

IMPLEMENTATION

1. **Data Collection**
   - Gather textual data from sources such as movie reviews, product reviews, social media posts, or feedback forms.
   - Ensure the dataset is labeled with sentiment categories (e.g., positive, negative, neutral).
2. **Data Preprocessing**
   - Clean the text by removing punctuation, special characters, and irrelevant symbols.
   - Convert all text to lowercase and remove stop words.
   - Tokenize the text into words and apply stemming or lemmatization to normalize words.
3. **Feature Extraction**
   - Convert the processed text into numerical features suitable for machine learning models.
   - Common techniques include Bag of Words (BoW), Term Frequency-Inverse Document Frequency (TF-IDF), or word embeddings (Word2Vec, GloVe).
4. **Model Selection and Training**
   - Choose suitable machine learning or deep learning models for sentiment classification (e.g., Logistic Regression, Naive Bayes, Support Vector Machine, LSTM, or BERT).
   - Split the dataset into training and testing sets.
   - Train the model on the training data and tune hyperparameters for optimal performance.
5. **Model Evaluation**
   - Evaluate the model on the testing set using metrics such as accuracy, precision, recall, and F1-score.
   - Compare different models and select the one with the best performance.
6. **Sentiment Prediction**
   - Apply the trained model to new, unseen text data to predict sentiment.
   - Output the predicted sentiment as positive, negative, or neutral.
7. **Results and Analysis**
   - Visualize results using charts or graphs to show the distribution of sentiments.
   - Interpret insights from the predicted sentiment to inform decision-making

CHAPTER 8

RESULTS AND DISCUSSION

 Model Performance

- Accuracy: 84%
- Precision: Positive – 0.85, Negative – 0.83, Neutral – 0.80
- Recall: Positive – 0.86, Negative – 0.82, Neutral – 0.79
- F1-Score: Positive – 0.85, Negative – 0.82, Neutral – 0.79

 Observations:

- The model performed better on positive sentiments due to a higher representation of positive reviews in the training dataset.
- Neutral sentiments had slightly lower accuracy, possibly because of overlapping features with positive and negative texts.
- Preprocessing techniques such as stopword removal, tokenization, and stemming significantly improved model accuracy.

 Visualization:

- A bar chart of sentiment distribution shows that the majority of texts are positive, followed by negative and neutral.
- Confusion matrix analysis revealed that misclassifications mostly occurred between neutral and negative classes.

CHAPTER 9

CONCLUSION

This project demonstrates the development and implementation of a **sentiment analysis system** capable of classifying text into positive, negative, and neutral categories. The system achieves a satisfactory accuracy of **84%**, making it suitable for practical applications such as **customer feedback analysis** and **social media monitoring**.

The study also highlights the importance of **data preprocessing**, **feature extraction**, and **model selection** in achieving accurate sentiment predictions. With further enhancements using advanced NLP techniques, the model can be improved to handle complex linguistic structures, sarcasm, and context-dependent sentiment.

CHAPTER 10

REFERENCES

- Pang, B., & Lee, L. (2008). **Opinion mining and sentiment analysis.** *Foundations and Trends in Information Retrieval*, 2(1–2), 1–135. https://doi.org/10.1561/1500000011
- Liu, B. (2012). **Sentiment analysis and opinion mining.** *Synthesis Lectures on Human Language Technologies*, 5(1), 1–167. https://doi.org/10.2200/S00416ED1V01Y201204HLT016
- Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013). **New avenues in opinion mining and sentiment analysis.** *IEEE Intelligent Systems*, 28(2), 15–21. https://doi.org/10.1109/MIS.2013.30
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). **BERT: Pre-training of deep bidirectional transformers for language understanding.** *NAACL-HLT*. https://arxiv.org/abs/1810.04805
- Bird, S., Klein, E., & Loper, E. (2009). **Natural Language Processing with Python.** O'Reilly Media.