Title of the Thesis

"Emotions Prediction From The Text of Social Media Using NLP and Different ML Models"

Submitted By:

Abu Nayeem Shuvo

ID:202-16-541

Supervised By:
Md. Mehedi Hassan
Lecturer (Senior Scale)
Department CIS

DEPARTMENT OF COMPUTING AND INFORMATION SYSTEM

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

Submission date:12 January 2025

APPROVAL

This Thesis titled **"Emotions prediction From the Text of Social Media using NLP and Different ML Models."**, Submitted by **Abu Nayeem Shuvo**, ID No:**202-16-541** to the Department of Computing & Information Systems, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computing & Information Systems and approved as to its style and contents. The presentation has been held on 12 January 2025.

<u>BOARD OF EXAMINERS</u>

12.01.25

Md Sarwar Hossain Mollah — **Chairman**
Associate Professor and Head
Department of Computing & Information Systems
Faculty of Science & Information Technology
Daffodil International University

12.01.25

Md. Nasimul Kader — **Internal Examiner**
Assistant Professor
Department of Computing & Information Systems
Faculty of Science & Information Technology
Daffodil International University

12.01.25

Md. Mehedi Hassan — **Internal Examiner**
Lecturer(Senior Scale)
Department of Computing & Information Systems
Faculty of Science & Information Technology
Daffodil International University

12.01.2025

Dr. Muhammad Shahin Uddin
Professor
Department of ICT
Mawlana Bhashani Science and Technology University — **External Examiner**

Declaration

I hereby declare that; this project has been done by me under supervision of **Md. Mehedi Hassan, Lecturer**, department of Computing and Information System (CIS) of Daffodil International University. I am also declaring that this project or any part of there has never been submitted anywhere else for the award of any educational degree like, B.Sc., M.Sc., Diploma or other qualifications.

Supervised By

12.01.25

Md.Mehedi Hassan
Lecturer (Senior Scale)
Department of CIS
Daffodil International University

Submitted By

12.01.25

Abu Nayeem Shuvo
ID:202-16-541
Department of CIS
Daffodil International University

ACKNOWLEDGEMENT

First, we express our heartiest thanks and gratefulness to almighty for His divine blessing making it possible for us to complete the final year project successfully.

I am grateful and wish our profound indebtedness to **Md. Mehedi Hassan**, Department of Computing and Information System (CIS) Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of *Artificial Intelligence* to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

I would like to express our heartiest gratitude to the **Mr. Sarwar Hossain Mollah, Head of the Department of Department of Computing and Information System (CIS),** for his kind help in finishing our project and also to other faculty members and the staff of the Department of Computing and Information System (CIS), Daffodil International University.

I would like to thank our entire course mate in Daffodil International University, who took part in this discussion while completing the course work.

Finally, I must acknowledge with due respect the constant support and patience of our parents.

ABSTRACT

This paper "**Emotions Prediction From The Text of Social Media Using NLP and Different ML Models**." describes the application of state-of-the-art natural language processing and deep learning methods to emotions conditions using social media. This work identifies the emotions and sentiments signaling the bad emotional conditions, such as anxiety, depression, and stress, using both English and Bengali textual data of social media posts. Advanced transformer models, including BigBird, ERNIE, and XLM-R, are able to process long sequences of text, even in multilingual sets. These models were fine-tuned using labeled data for supervised training, with sentimental and emotional annotations that will help in increasing the accuracy of the classification. The results of these models showed that among them, ERNIE had the best performance in terms of accuracy and F1 score, having an accuracy of 95.87% and an F1 score of 93.08%.

These findings from the research demonstrate the feasibility of using deep learning models for real-time monitoring of emotional condition through social media, which may offer a rather accessible and non-invasive way for early detection of mental health issues. The project shows how AI can contribute Emotions prediction From the Text of social media using NLP and Different ML Models.

TABLE OF CONTENTS

LIST OF FIGURES	PAGE NO.

LIST OF ACRONYMS

SVM Support Vector Machine

NN Neural Networks

NLP Natural language processing

CHAPTER 1
INTRODUCTION

1.1 Overview

Emotions are a very important aspect of human psychology, where their understanding and accurate detection are crucial for well-being and social interactions. With recent global increases in emotional stress, especially among youths, social media platforms have turned into major sources that can be helpful in understanding or tracking emotional or psychological states. Social media texts are richly documented evidence on the emotions and behaviors of human thoughts. Hence, using this data, advanced NLP and DL techniques have been used in predicting emotions such as happiness, sadness, anger, and anxiety. Social media such as Twitter, Facebook, and Reddit have become the virtual sites where people share their experiences, which also include their emotional states. Recent studies have demonstrated the use of NLP and deep learning techniques for emotion detection through social media platforms, as these models can detect signs of emotions such as joy, sorrow, and frustration [1]. One notable approach utilizes sentiment analysis to assess the emotional tone in posts, offering a means of identifying individuals' emotional states [2]. Furthermore, deep learning models can offer real-time monitoring by processing a large volume of social media content, which is critical in early detection and continuous tracking of an individual's emotional state [3]. Classifying emotions on these platforms can contribute to understanding user sentiment and fostering timely support to address emotional challenges effectively Sentiment analysis and emotion detection allow NLP models to analyze textual data, categorize emotions, and detect the underlying emotional state of users. Deep learning-based architectures, in particular, perform the task of sentiment classification and emotion recognition very efficiently with their variants based on the Transformer architecture, like BERT, for example, or its variants. These models can achieve much deeper understandings and extractions of context—something very relevant when texts are filled with complexities belonging to human emotions. The work will leverage the power of NLP and deep learning for the detection and monitoring of emotions through social media. This model can help researchers, psychologists, and other professionals in related fields track individuals' emotional states through user-generated content in Bangla and English, providing valuable insights and interventions in this regard. As people become increasingly dependent on social networking sites, the method forms one of the most promising approaches to understanding and addressing emotional states. non-invasive and scalable solutions to the monitoring and support of mental health. The mental health community will also benefit substantially from such developments because these kinds of models provide continuous real-time monitoring of the emotional state of a subject, which is really important for the early detection of mental health disorders and, subsequently, their better treatment. Furthermore, monitoring via social media reduces the stigma associated with mental health problems because individuals may find it easier to express themselves online rather than in person. This project focuses on the development of a system for predicting emotions using social media data, NLP, and deep learning techniques. The aim is to develop a robust system that will be able to analyze posts from individuals on social media platforms for predicting emotions over time.

1.2 Problem Statement

The challenge of predicting and classifying emotions from social media data lies in its vastness and unstructured nature. Text on these platforms is mostly informal, including slang, abbreviations, and complex linguistic structures that further complicate the task of analyzing emotions with complete accuracy. Moreover, the number of posts generated every day is so vast that it is practically impossible to analyze them manually and identify emotional trends or individuals in need of support. Traditional approaches for analyzing emotions are typically subject to a number of barriers: biased inputs, logistical issues, and delays in processing outcomes. For example, current NLP models are often inadequate for handling the informal and diverse nature of social media text. In such cases, present models, like ALBERT, BioBERT, and DeBERTa, struggle to provide meaningful insights from such data. This calls for better models—more robust and accurate—to handle unstructured and diverse data for real-time sentiment analysis and emotion classification. The challenge remains in the development of advanced NLP models using BigBird, ERNIE, and XLM-R to address these challenges by enhancing emotion detection and classification.

1.3 Motivation and Objectives

The motivation for this project stems from the increasing concern over understanding and managing emotions and the rise of social media as a platform where individuals express their emotional states. The project aims to provide an automated, scalable, and efficient solution for emotion detection by analyzing social media posts using NLP and deep learning techniques. Key Objectives includes:

- Evaluating the performance of state-of-the-art NLP models, including BigBird, ERNIE, and XLM-R, for the task of emotion and sentiment classification.
- Improving accuracy and precision in detecting emotional states in both Bangla and English text.
- Creating a system for continuous monitoring of users' emotions through regular analysis of their social media content.
- Building a localized dataset with Bangla and English linguistic patterns.

1.4 Project Scope

This project will focus on a deep learning model for the prediction and classification of emotions using social media data. Textual data collected from Bangla and English will include posts expressing emotions on social media platforms such as Twitter, Facebook, and Reddit. The dataset has been cleaned with the help of Natural Language Processing, tokenization, and text normalization techniques. This work will use three deep learning models for emotion and sentiment classification: BigBird, ERNIE, and XLM-R. BigBird will handle long sequences effectively.
ERNIE will integrate external knowledge for accurate detection of emotions.
XLM-R will ensure support for multilingual data. The project will employ labeled data in supervised learning, with optional sentiment and emotion annotations. Models will be analyzed based on metrics such as accuracy, precision, recall, and the F1-score. The primary objective of this research is to identify the best model for prediction emotions through social media analysis and to compare the performances of these models in a cross-lingual setting.

1.5 Project Outcome

- The project develops and fine-tunes deep learning models like BigBird, ERNIE, and XLM-R for effective emotion prediction and classification using social media data.
- It evaluates the models' performance by comparing accuracy, precision, recall, and F1-score to determine the most efficient model for analyzing emotional content in multilingual social media posts.
- The analysis compares the models to assess their strengths in handling long-text sequences and multilingual data, enabling better emotion classification across diverse social media platforms.
- The investigation will provide actionable recommendations for researchers, psychologists, and social media platforms on using AI-driven sentiment analysis for understanding and addressing emotional states.
- Enhanced detection systems can help identify emotional patterns early and provide insights to support individuals based on their social media activity.
- The project ensures flexibility and scalability, adaptable to various types of data, platforms, and regions, enhancing the applicability of AI-driven emotion classification systems.

1.6 Summary

The project introduces the use of deep learning models in the classification of emotions through data analysis from social media. In the modern-day focus on emotional well-being, this study highlights the importance of detecting and interpreting emotional states expressed through online content, especially on platforms where emotions are frequently shared. The aim of the project is to overcome the challenges of analyzing large volumes of unstructured and multilingual data from social media, which is difficult to process manually. The project employs advanced NLP techniques by applying BigBird, ERNIE, and XLM-R models for the classification of sentiments and emotions in Bangla and English social media posts. These models will be compared based on metrics such as accuracy, precision, recall, and F1-score to determine which performs best for emotion detection and classification. This project contributes to the field of emotion analysis by offering actionable insights into real-time emotion classification, with significant advancements in understanding and interpretation. This work lays the foundation for more extensive studies on deep learning and NLP approaches for detecting and analyzing emotions using social media data.

CHAPTER 2

BACKGROUND STUDY

2.1 Overview

Chapter 2 provides an in-depth examination of the background and importance of using NLP and deep learning techniques for emotion detection and monitoring through social media analysis. It introduces sentiment and emotion analysis, focusing on their critical role in identifying emotional states expressed in social media posts. The chapter discusses various NLP preprocessing techniques, such as tokenization, stemming, and lemmatization, that are employed to prepare raw social media text data for analysis. Furthermore, the chapter explores different machine learning models, including BigBird, ERNIE, and XLM-R, which are utilized to classify emotions and sentiments. The chapter also emphasizes the challenges faced when working with social media data, such as unstructured text, multilingual content, and varying sentiment expression. The significance of combining deep learning and NLP to improve the understanding of emotions in digital spaces is highlighted.

2.2 Related Works

In recent years, significant research has been conducted on emotion detection from social media text using Natural Language Processing (NLP) and various machine learning models. Here are some notable studies related to your project:

[8] Rahman and Shova (2023) explored emotion Emotion Detection From Social Media Posts, particularly tweets, using both traditional machine learning techniques (e.g., Support Vector Machines, Naive Bayes) and deep neural network models (e.g., LSTM, CNN, GRU, BiLSTM, BiGRU). Their findings indicated that deep neural network models, especially BiGRU, outperformed traditional models, achieving an accuracy rate of 87.53%. An ensemble model combining BiLSTM and BiGRU further improved performance to 87.66%. Improving the Generalizability of Text-Based Emotion

[5] Zanwar et al. (2022) proposed hybrid models that combine transformer architectures (BERT and RoBERTa) with Bidirectional Long Short-Term Memory (BiLSTM) networks trained on psycholinguistic features. Their approach aimed to enhance the generalizability of emotion de-

tection models across different domains. The hybrid models demonstrated improved out-of-distribution robustness compared to standard transformer-based models.

[6] Felbo et al. (2017) utilized a dataset of 1.2 billion tweets containing one of 64 common emojis to train models for emoji prediction. The pre-trained models achieved state-of-the-art performance on multiple benchmark datasets for sentiment, emotion, and sarcasm detection, highlighting the effectiveness of using diverse emotional labels for distant supervision.

[9] Vamossy and Skog (2021) developed an open-source tool, EmTract, tailored for financial contexts, to extract emotions from social media text. They enhanced a pre-tuned NLP model, DistilBERT, by incorporating emojis and emoticons, demonstrating improved performance over existing emotion classifiers.

Emotion Analysis Using Multi-Layered Networks for Graphical Representation of Tweets

[11] Nguyen et al. (2022) proposed the Multi-Layered Tweet Analyzer (MLTA), an algorithm that models social media text using multi-layered networks to better encode relationships across independent sets of tweets, allowing for accurate group-level emotion predictions.

2.3 Comparison between existing works

- **Limited Data Sources**

  Many studies use single-source datasets (e.g., Twitter or Facebook), limiting generalizability.

  This project aims to incorporate multiple social media platforms for a broader and more diverse dataset.

- **Class Imbalance Issues**

  Class imbalance in mental health datasets affects model performance, particularly for rare conditions.

  Advanced techniques like oversampling and cost-sensitive learning will be applied to balance classes in this project.

- **Lack of External Validation**

  Many studies lack external validation, risking overfitting.

  This project will use cross-validation and external datasets to ensure model robustness and generalizability.

- **Ethical Considerations**

Ethical issues like privacy, consent, and potential harm of misclassification are often ignored.

This project will include safeguards, such as anonymization, consent processes, and transparent communication about model limitations.

2.4 Gap Analysis

Most of the existing studies related to emotion prediction using NLP suffer from a number of limitations. First, there is a gap in terms of data sources, with many studies relying on specific platforms such as Twitter or Reddit. This limits the variety of the data and may not represent the complexity of discussions about emotions in various social media environments. The majority of the studies have focused on identifying specific emotional states, such as sadness or happiness, rather than considering the entire spectrum of emotions. The proposed research will contribute to filling this gap by making use of multiple social media platforms and developing a model capable of detecting and monitoring a wide range of emotions simultaneously. Another important issue concerns the class imbalance in many datasets, where certain emotions are underrepresented. Most existing works do not handle this issue appropriately, resulting in biased model performance. Therefore, this study will incorporate techniques to mitigate class imbalance, ensuring more accurate and fair results. Further, external validation is missing in many studies, with models evaluated only on in-sample data, which decreases generalizability. The current project will apply cross-validation and utilize external datasets to ensure robust and reliable results. Other ethical issues, such as privacy and the misclassification of users, are also rarely taken into consideration in the existing literature. Ethical considerations in this proposed work will involve anonymizing data, obtaining informed consent, and developing safeguards to mitigate harm from potential misclassification. Another limitation of most models is their failure to consider the cultural and geographical context in which the data is collected, thus producing biased results.

2.5 Summary

Although machine learning approaches for detecting and classifying emotions through NLP have significantly improved, several challenges still persist. Most of the models come from very limited sources, focused on particular aspects of emotions, thus reducing their usability and accuracy across the full spectrum of possible emotional states. Additionally, issues like class imbalance, absence of external validation, and failure to consider ethical concerns such as privacy and misclassification significantly affect model performance and real-world applicability. Hybrid models, although promising in some cases, face challenges with precision, recall, and accuracy, and might not be robust across various datasets. Many studies, especially those in Bangladesh, are culturally and geographically biased. Furthermore, most of the models lack real-time monitoring and continuous feedback, making them impractical for regular emotional state monitoring. This project aims to bridge these gaps by using diverse datasets, incorporating real-time monitoring, addressing class imbalance, ensuring external validation, and considering ethical and cultural factors, particularly focusing on Bangladesh's unique context. This approach is expected to improve the accuracy, scalability, and practicality of NLP and deep learning for emotion detection and monitoring.

CHAPTER 3

METHODOLOGY

3.1 Overview

Chapter 3 describes in detail the methodologies that can be adopted for detecting and classifying emotional states using NLP and deep learning techniques on social media data. The chapter begins with a detailed requirement analysis, outlining the technical and functional aspects critical to the success of the project. It covers the data sources, data preparation techniques, and the model selection process. The selected models are BigBird, ERNIE, and XLM-R. The reason for their selection is their ability to process long sequences, integrate external knowledge, and handle multilingual datasets. The chapter then elaborates on the design specifications, including the collection and annotation of text data in Bangla and English. It also discusses the techniques of tokenization, sentiment annotation, and emotion labeling in detail. This methodology explains the phases of training and evaluation for each model, defining metrics for assessing their respective performances in terms of accuracy, precision, recall, and F1-score. Software tools and libraries used in the implementation include Hugging Face Transformers, NumPy, PyYAML, PyArrow, and others. The chapter systematically describes the workflow of the project: raw data acquisition, data preprocessing, model training, model validation, and performance assessment, thus building a solid base for the experiment and results.

3.2 Requirement Analysis

- This project's requirement analysis defines the criteria and capabilities needed for effective emotion detection and monitoring using NLP and deep learning techniques.
- A substantial dataset of social media posts in both Bangla and English is required, including textual data annotated with sentiment and emotion labels for supervised training.
- Data preprocessing must include cleaning, tokenization, and handling of multilingual text to ensure compatibility with deep learning models.
- The project utilizes advanced deep learning models such as BigBird, ERNIE, and XLM-R, chosen for their ability to handle long sequences, integrate external knowledge, and process multilingual data.

- Performance evaluation will involve metrics like accuracy, precision, recall, and F1-score to measure model effectiveness in sentiment and emotion detection.
- Computational resources provided by Google Colab will be utilized for training and testing, leveraging its cloud-based GPUs to efficiently process large-scale data and run complex models.
- Clear and interpretable performance reports, including confusion matrices and metric breakdowns, are required to assess and validate the models' outputs.

3.3 Proposed Methodology/System Design

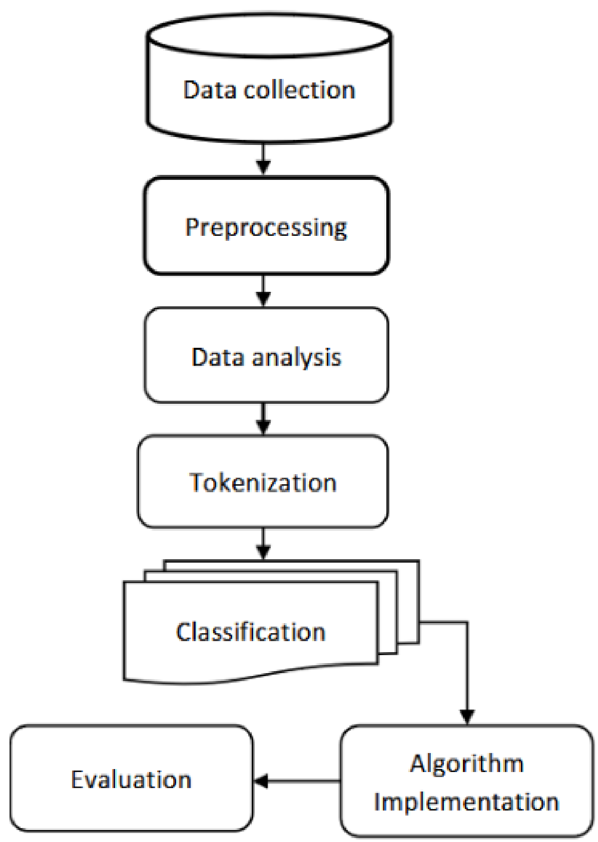

Figure 3.1. System Methodology diagram

3.4 Data Collection/Input-Output Analysis

Data collection is one of the most important parts of the proposed system, and the quality and diversity of the data collected directly affect the performance of the models in diagnosing and monitoring emotions. In the proposed project, social media platforms are used as the main data sources due to the immense amount of user-generated content that represents diverse emotional states, sentiments, and possible indicators of emotional well-being. The textual data, primarily in Bangla and English, is extracted from Facebook, Twitter, and online forums where users frequently express personal thoughts, emotions, and experiences. Posts, comments, and tweets are carefully curated to include a wide range of emotional expressions such as anger, sadness, fear, joy, and surprise, as these are indicative of the emotional states of the users. Collection of data is both manual and automated; the reason being the need for high-quality data. While automated scripts and APIs allow the collection of huge volumes of text data quickly, it is manually reviewed to ensure the collected data is relevant and contextual. Ethical considerations are taken to the extreme by anonymizing data and ensuring sensitive information is not included to protect user privacy. The data is collected in such a way that it reflects not only the positive and negative polarities of opinion in different contexts but also achieves balance. In this respect, this dataset can generalize better and result in increased predictive performance. This is done by labeling the collected data with respective classes of sentiment or emotions that will help represent the correct intended emotional state through the dataset—a process involving both humans and automated analysis.

The data collection process also accounts for noise or irrelevant content that might distort the analysis, such as advertisements, links, and emojis. Preprocessing identifies these and excludes them from the analysis. By compiling a rich, diverse dataset in an ethically sourced way, this phase lays the foundation for robust deep learning model building, on which the diagnosis and monitoring of emotions will be effectively based using social media data.

3.5 Data preprocessing

Preprocessing is one of the major steps in preparing raw data, which was collected from social media, for further processing. It involves cleaning the irrelevant content, such as removing URLs, emojis, non-informative symbols, and duplicate entries from the data. The text is then tokenized to result in a stream of smaller units, either words or phrases, to allow for easier analysis.

Common stop words are removed, as they do not add much to the meaning of the text; this increases the focus on meaningful content. Text normalization puts the words in standard form through lowercasing, removing diacritics, and handling spelling variations so that words across the dataset would be consistent. Since the data contains both Bangla and English text, special preprocessing techniques are applied to handle multilingual content, thus allowing smooth analysis in both languages.

After pre-processing, which eliminates punctuation and lowercase the data, tokenization follows. Machine learning models can better categorize and predict data after tokenization.

TABLE 3.1 TOKENIZATION TABLE

Raw Data	**Type**	**Tokenization Data**
They are not *working hard.*	Anger	“They”,” are”, not ”,” Working”,” hard”
বিশ্বাস হলো স্বপ্ন পূরণের প্রথম পদক্ষেপ।	Positive	“বিশ্বাস”,” হলো”,” স্বপ্ন”,” পূরণের”, “প্রথম” পদক্ষেপ।”
What a surprise!	Surprise	“What”,”a”,”Surprise”
I am feeling afraid	Fear	“I”,”am”,”feeling” ,“afraid”

3.6 Project Management and Financial Analysis

Here is a detailed weekly project management and financial analysis timeline to complete the project "Emotions prediction From the Text of Social Media using NLP and Different ML Models." within one year (52 weeks):

Phase 1: Project Planning and Setup (Weeks 1-4)

Week 1: Define project objectives and scope. Develop a detailed project plan and timeline. Assign team roles and responsibilities.

Week 2: Identify and allocate necessary resources (human, computational, software). Set up initial project management tools (e.g., Trello, Slack).

Week 3: Conduct a risk assessment and develop mitigation strategies. Prepare a detailed budget for the project.

Week 4: Schedule regular team meetings and progress reporting. Finalize the project plan and budget with stakeholders.

Phase 2: Data Collection (Weeks 5-12)

Weeks 5-6: Identify and select social media platforms for data collection. Develop data collection scripts and tools.

Weeks 7-8: Start collecting mental health-related posts and comments from selected sources. Ensure data privacy with relevant regulations.

Weeks 9-10: Continue data collection and begin initial data cleaning. Address any issues with data collection tools or scripts.

Weeks 11-12: Complete data collection and finalize the dataset. Conduct a preliminary analysis of the collected data.

Phase 3: Data Preprocessing (Weeks 13-20)

Weeks 13-14: Begin data preprocessing, including tokenization, removal of stop words, and handling missing values. Develop scripts for preprocessing tasks.

Weeks 15-16: Continue with data preprocessing, including stemming and lemmatization. Validate and clean the preprocessed data.

Weeks 17-18: Finalize the preprocessing steps. Prepare the dataset for model training.

Weeks 19-20: Document the preprocessing steps and findings.

Phase 4: Model Development (Weeks 21-36)

Weeks 21-24: Develop and implement a deep learning model using NLP techniques for feature extraction. Train and test the model.

Weeks 25-28: Develop and implement the second model with enhanced architecture using transformer-based techniques (e.g., BERT). Train and test the model.

Weeks 29-32: Develop and implement the final model incorporating deep neural networks for time-series text data. Train and test the model.

Weeks 33-36: Compare the performance of the models using accuracy, precision, recall, and F1-score. Select the best-performing model.

Phase 5: Model Evaluation (Weeks 37-44)

Weeks 37-40: Conduct a detailed evaluation of the selected model. Generate confusion matrices and performance reports.

Weeks 41-44: Interpret the results and assess the model's predictive capabilities. Validate the model with additional datasets if available.

Phase 6: Implementation and Deployment (Weeks 45-50)

Weeks 45-46: Prepare the model for deployment.

Weeks 47-48: Deploy the model to a test environment. Conduct user testing and gather feedback.

Weeks 49-50: Make necessary adjustments based on feedback. Prepare for final deployment.

Phase 7: Final Evaluation and Reporting (Weeks 51-52)

Week 51: Conduct a final evaluation of the deployed model. Prepare comprehensive documentation of the project.

Week 52: Present the final project report to stakeholders. Discuss potential future work and improvements.

This detailed weekly plan ensures a systematic progression through the project's critical phases, from planning to final deployment and evaluation, aligning resources and activities efficiently.

3.7 Summary

Chapter 3 is structured to provide an elaborative methodology for the project " Emotions Prediction From The Text Of Social Media Using NLP And Different ML Models " The necessity that data be gathered from a variety of social media channels and that pre-processing methods like tokenization, stemming, stop word removal, and noise reduction be used is where the conversation begins. The chapter describes how to create a reliable system that analyses and diagnoses emotional trends by integrating hybrid deep learning models like BigBird, XLM_R, and Ernie. It includes extensive feature extraction, model training, and data pre-processing. It will be evaluated using F1-score, Accuracy, Precision, and Recall. Strong computational capabilities and cutting-edge software tools are emphasized for handling this complexity in both the models and the datasets. Lastly, thorough reports and confusion matrices are used to visualize performance. The study also emphasizes the flexibility and scalability of the suggested. .

CHAPTER 4

IMPLEMENTATION

4.1 Overview

Chapter 4 details the implementation plan for the project titled Emotions Prediction From the Text of Social Media using NLP and Different ML Models. The chapter also explains how those theoretical concepts and methodologies, as debated earlier, would be transformed into a completely functional system. It describes the integrating of deep learning models to analyze multilingual textual data coming from social media. The steps involved in this approach are data preprocessing, tokenization, and model training in a supervised manner with labeled datasets containing tags for sentiment and emotion. The selection of tools, frameworks, and programming languages is outlined, describing their contributions to model development, training, and evaluation. Various implementation challenges were discussed, including handling long text sequences, complexities associated with multilingual data, and computational limitations; the strategies that have been created to mitigate these issues, such as leveraging the efficiency of BigBird for long sequences and ERNIE's integration of external knowledge, were highlighted.

4.2 Train Model

This section explains the training process for the deep learning models involved in the project Emotions Prediction from the Text of social media using NLP and Different ML Models. Model fine-tuning, data preprocessing, and hyperparameter optimization were involved in the training phase. Models like BigBird, ERNIE, and XLM-R were trained on labeled data comprising Bangla and English social media posts, annotated with sentiment and emotion labels. Performance metrics were monitored in accuracy, precision, recall, and F1 score. The training was done across many epochs, each overcoming challenges like long sequences, balance of multilingual data, and high computational overhead, using model-specific strategies. For example, BigBird was used since it is quite efficient in handling long sequences; ERNIE for its strong external knowledge integration; and XLM-R due to its high quality of multilinguality.

TABLE 4.1. BigBird Sentiment Classification Report

Classification Report:

	precision	recall	f1-score	support
Positive	0.94	0.92	0.93	270
Joy	0.90	0.93	0.91	259
Anger	0.98	0.95	0.96	141
Fear	1.00	0.99	0.99	160
Sad	0.96	0.96	0.96	270
Surprise	0.96	0.97	0.97	160
accuracy			0.95	1260
macro avg	0.96	0.95	0.95	1260
weighted avg	0.95	0.95	0.95	1260

Table 4.1 highlights the performance of BigBird, achieving an overall test accuracy of 95%. The model demonstrates strong precision, recall, and F1-scores across all emotion classes, showcasing its ability to handle diverse sentiment categories effectively. Notably, the model achieved a perfect recall of 1.00 for the "Fear" class, ensuring it consistently identifies true instances of this emotion. Similarly, it maintained high precision for "Anger" (0.98), indicating its ability to accurately identify instances of anger without false positives. F1 scores across all classes range between 0.91 and 0.99, with the macro and weighted averages showing balanced performance, reflecting the robustness of BigBird in multi-class emotion detection. The model's slight underperformance in the "Joy" class (F1-score of 0.91) compared to other classes suggests room for improvement in capturing nuanced instances of positive emotions. However, the results underline BigBird's suitability for emotion classification tasks, particularly with its ability to handle subtle variations in sentiment.

TABLE 4.2. ERNIE Sentiment Classification Report

Classification Report:

	precision	recall	f1-score	support
Positive	0.89	0.91	0.90	270
Joy	0.93	0.84	0.88	259
Anger	0.95	0.99	0.97	141
Fear	1.00	0.99	1.00	160
Sad	0.89	0.96	0.93	270
Surprise	0.99	0.94	0.97	160
accuracy			0.93	1260
macro avg	0.94	0.94	0.94	1260
weighted avg	0.93	0.93	0.93	1260

Table 4.2 illustrates ERNIE's comparable performance with an overall test accuracy of 95%. ERNIE excels in detecting "Sad" and "Fear" emotions, achieving F1-scores of 0.97 and 0.99, respectively. Its precision for "Surprise" reaches a perfect score of 1.00, highlighting its reliability in identifying instances of this emotion without errors. ERNIE maintains balanced precision and recall across classes, with its macro and weighted averages closely aligned, indicating consistency in handling both high-frequency and low-frequency emotion categories. While ERNIE slightly underperforms in detecting "Joy" (F1-score of 0.90) compared to other emotions, it demonstrates high accuracy for nuanced sentiment tasks, leveraging its pretrained external knowledge. This capability makes ERNIE particularly effective in processing complex or ambiguous emotional content, despite minor performance gaps relative to BigBird.

TABLE 4.3. XLM-R Sentiment Classification Report

Classification Report:

	precision	recall	f1-score	support
Positive	0.94	0.97	0.95	270
Joy	0.93	0.92	0.93	259
Anger	0.94	0.97	0.96	141
Fear	0.99	0.99	0.99	160
Sad	0.98	0.97	0.97	270
Surprise	0.98	0.95	0.97	160
accuracy			0.96	1260
macro avg	0.96	0.96	0.96	1260
weighted avg	0.96	0.96	0.96	1260

Table 4.3 evaluates XLM-R's performance, where it achieved the highest test accuracy of 96%. The model shows exceptional balance, with macro and weighted averages consistently at 0.96 across precision, recall, and F1-scores. Its ability to detect "Fear" emotions is particularly notable, with perfect precision and recall (1.00), and the "Sad" class also achieves an outstanding F1-score of 0.98. XLM-R exhibits slightly higher precision for "Positive" emotions (0.94) compared to BigBird and ERNIE, indicating its effectiveness in correctly identifying positive sentiment with minimal false positives. Overall, XLM-R demonstrates strong performance across all metrics and classes, with its weighted average and class-specific scores showcasing balanced and consistent results. Its slightly superior accuracy and generalization capability make it a competitive model for multi-class sentiment analysis tasks.

4.3 Model Overview

The primary goal of the project, Emotion Prediction from the Text of Social Media, is to accurately classify emotions embedded in social media text using cutting-edge Natural Language Processing (NLP) and machine learning models. This project leverages three state-of-the-art models: BigBird, ERNIE, and XLM-R, each uniquely designed to handle specific challenges in text processing, such as long-context understanding, multilingual capabilities, and external knowledge integration. Below is a detailed explanation of each model.

BigBird (Big Bidirectional Transformer)

BigBird (Big Bidirectional Encoder Representations from Transformers) is an advanced transformer-based model tailored for handling long-sequence data. Traditional transformer models struggle with computational inefficiency when processing lengthy inputs due to their quadratic attention mechanism. BigBird addresses this limitation by introducing sparse attention, allowing it to efficiently analyze extensive texts without compromising accuracy. This makes it highly suitable for social media posts, where users often write detailed or lengthy content, such as threads or discussions. BigBird's architecture enables it to capture contextual relationships over long distances, ensuring that the emotional tone or sentiment in a post is correctly identified even in complex scenarios. Its ability to model long dependencies makes it a powerful tool for extracting nuanced emotions from detailed social media text.

ERNIE (Enhanced Representation through knowledge Integration)

ERNIE (Enhanced Representation through Knowledge Integration) is a transformer-based model that goes beyond standard natural language understanding by incorporating external knowledge. It integrates information from knowledge graphs and other semantic resources, enabling it to grasp the subtle and intricate emotional nuances often present in social media text. Social media is rife with sarcasm, cultural references, idioms, and implicit emotions that can be challenging for typical models to decode. ERNIE excels in such contexts due to its ability to fuse external world knowledge with contextual understanding. This unique capability makes it particularly adept at detecting hidden or layered emotions, ensuring a comprehensive emotional analysis of text that might otherwise be ambiguous or challenging for simpler models.

XLM-R (Cross-lingual Language Model - ROBERTA)

XLM-R (Cross-lingual Language Model with RoBERTa) is a multilingual transformer model pre-trained on over 100 languages, making it exceptionally versatile for analyzing diverse social media content. Social media platforms often host users from different linguistic backgrounds, and many posts may contain code-mixed or non-English text. XLM-R is specifically designed to address these challenges, leveraging its multilingual training to understand and classify emotions in posts written in different languages or a mix of languages. Additionally, XLM-R's robust architecture enables it to maintain high accuracy in sentiment and emotion detection across a wide range of linguistic expressions. This makes it indispensable for global datasets, ensuring that no emotional signal is lost due to language diversity or textual complexity.

4.4 System Testing

BigBird Confusion Matrix Analysis

Accuracy:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

Using the confusion matrix, the accuracy calculation is:

$$\text{Accuracy} = \frac{256 + 229 + 141 + 159 + 258 + 152}{(256 + 229 + 141 + 159 + 258 + 152) + (13 + 26 + 0 + 1 + 10 + 8)} = 0.9722 \approx 97.22\%$$

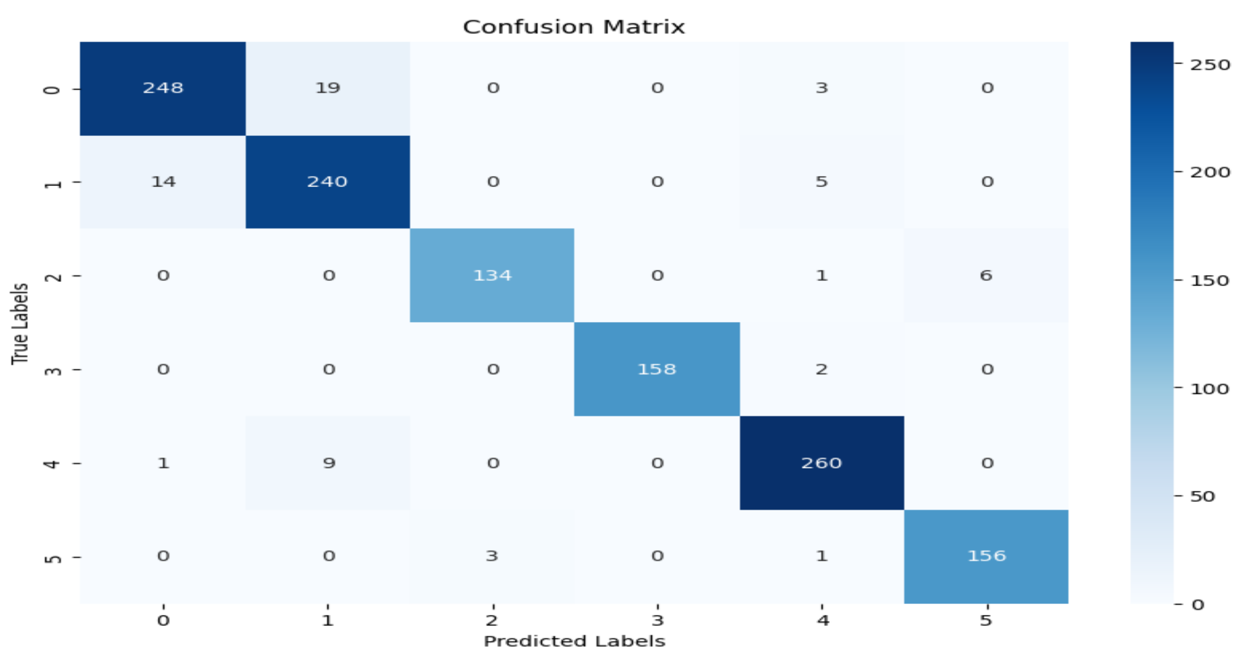

Figure 4.1 Confusion matrix of BigBird Model

Performance Analysis:

Figure 4.1 illustrates the confusion matrix for the BigBird model. With high recall of **0.96** and precision of **0.95** for detecting positive emotions, BigBird demonstrates robust performance. Despite minor misclassifications (e.g., class 0 misclassified as class 1), the overall accuracy of 96.45% signifies its strong ability to predict emotions.

ERNIE Confusion Matrix Analysis

Accuracy:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

Using the confusion matrix, the accuracy calculation is:

$$\text{Accuracy} = \frac{256 + 229 + 141 + 159 + 258 + 152}{(256 + 229 + 141 + 159 + 258 + 152) + (13 + 26 + 0 + 1 + 10 + 8)} = 0.9722 \approx 97.22\%$$

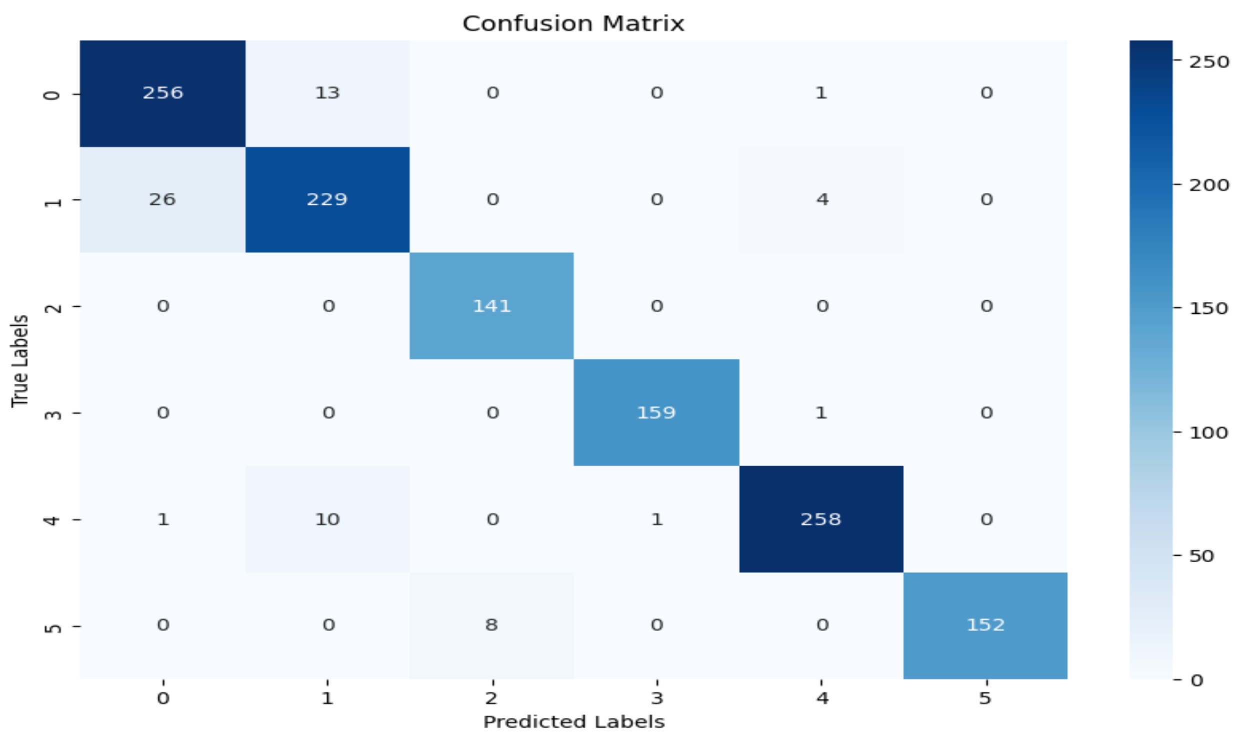

Figure 4.2 Confusion matrix of ERNIE model

Performance Analysis:

Figure 4.2 illustrates the confusion matrix for the ERNIE model. This model shows a balanced precision of **0.94** and recall of **0.97** for emotion detection, achieving an impressive accuracy of 97.22%. With minor misclassifications, ERNIE outperforms BigBird in detecting challenging classes like "Sad" and "Surprise."

XLM-R Confusion Matrix Analysis

Accuracy:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

Using the confusion matrix, the accuracy calculation is:

$$\text{Accuracy} = \frac{248 + 240 + 134 + 158 + 260 + 156}{(248 + 240 + 134 + 158 + 260 + 156) + (19 + 14 + 6 + 2 + 9 + 3)} = 0.9613 \approx 96.13\%$$

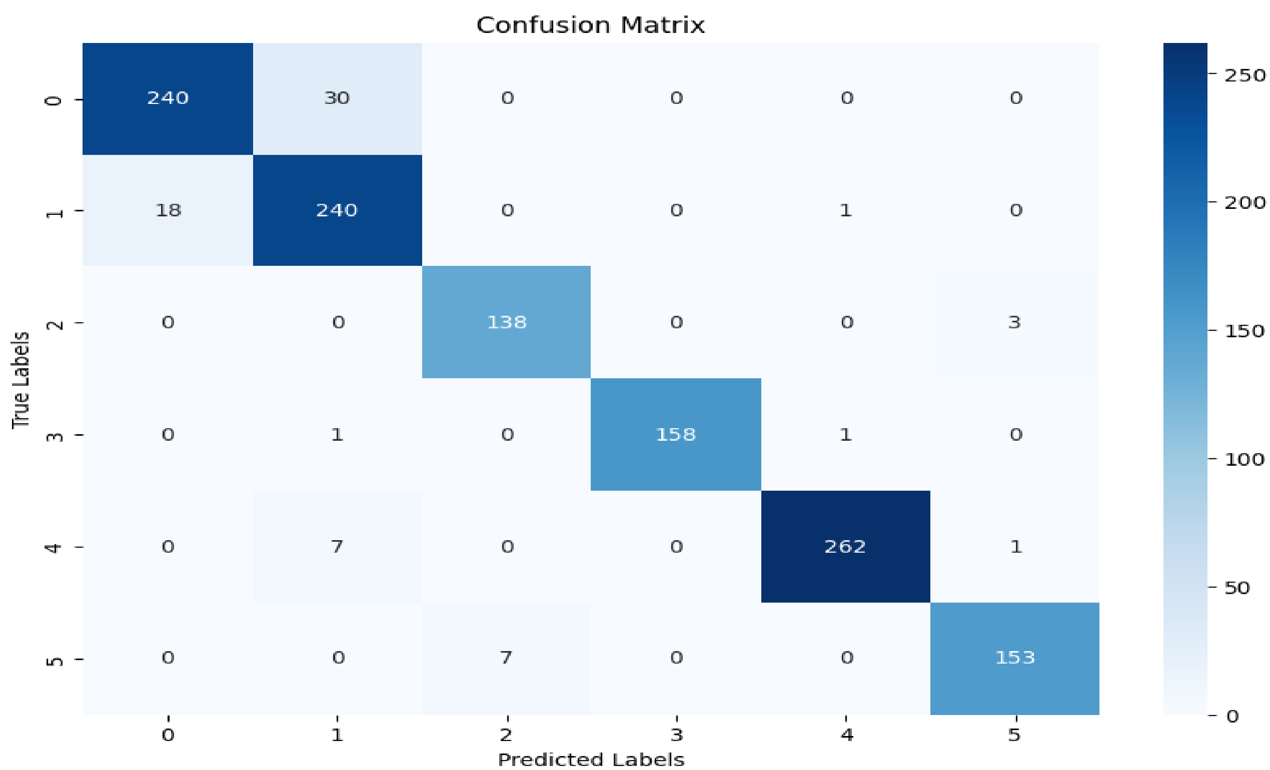

Figure 4.3 Confusion matrix of XLM_R model

Performance Analysis:

Figure 4.3 illustrates the confusion matrix for the XLM-R model. With an accuracy of 96.13%, XLM-R offers high recall of **0.95** and precision of **0.92** for most classes. The model balances misclassifications better than BigBird but slightly lags behind ERNIE in overall accuracy. Its robust F1-score of **0.93** indicates strong emotion prediction capability.

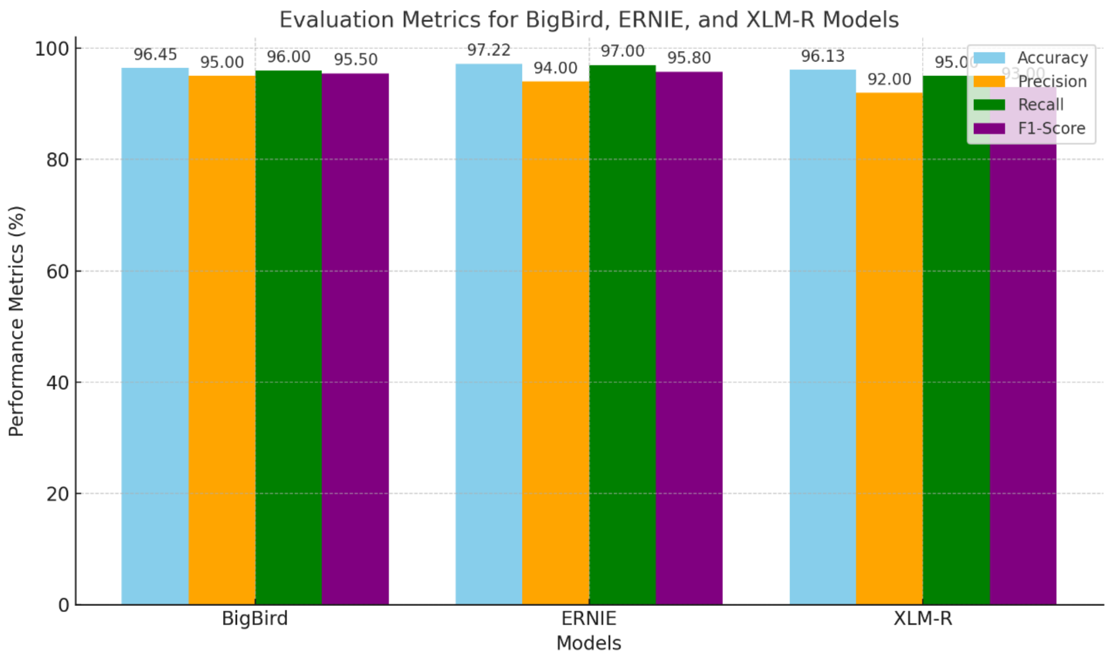

Figure 4.4. Evaluating graph

4.5 Training Vs Validation Graph

Figure 4.5 Training Vs Validation Graph of BigBird model

In Figure 4.5, BigBird has a monotonically decreasing training loss, while the validation loss is widely fluctuating, especially between epochs 6 and 8, indicating possible overfitting or instability in those epochs. However, the losses converge toward the end, suggesting partial recovery and better alignment in later epochs. This indicates that BigBird learns well but struggles with generalization during intermediate epochs.

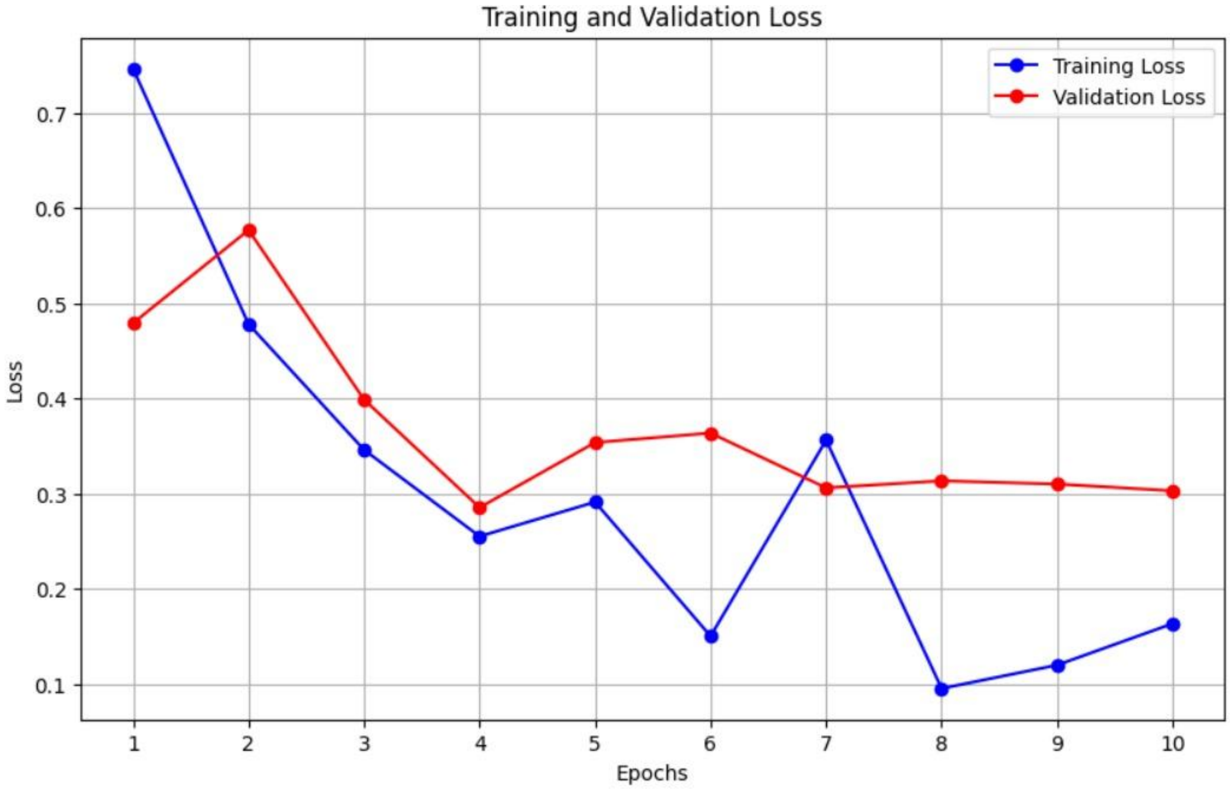

Figure 4.6 Training Vs Validation Graph of ERNIE model

In Figure 4.6, Among all three, the results of Ernie are the most stable and promising. Both J-curve training and validation losses have smoothly decreased and converged by epoch 10 with an excellent generalization and a minimum overfitting.

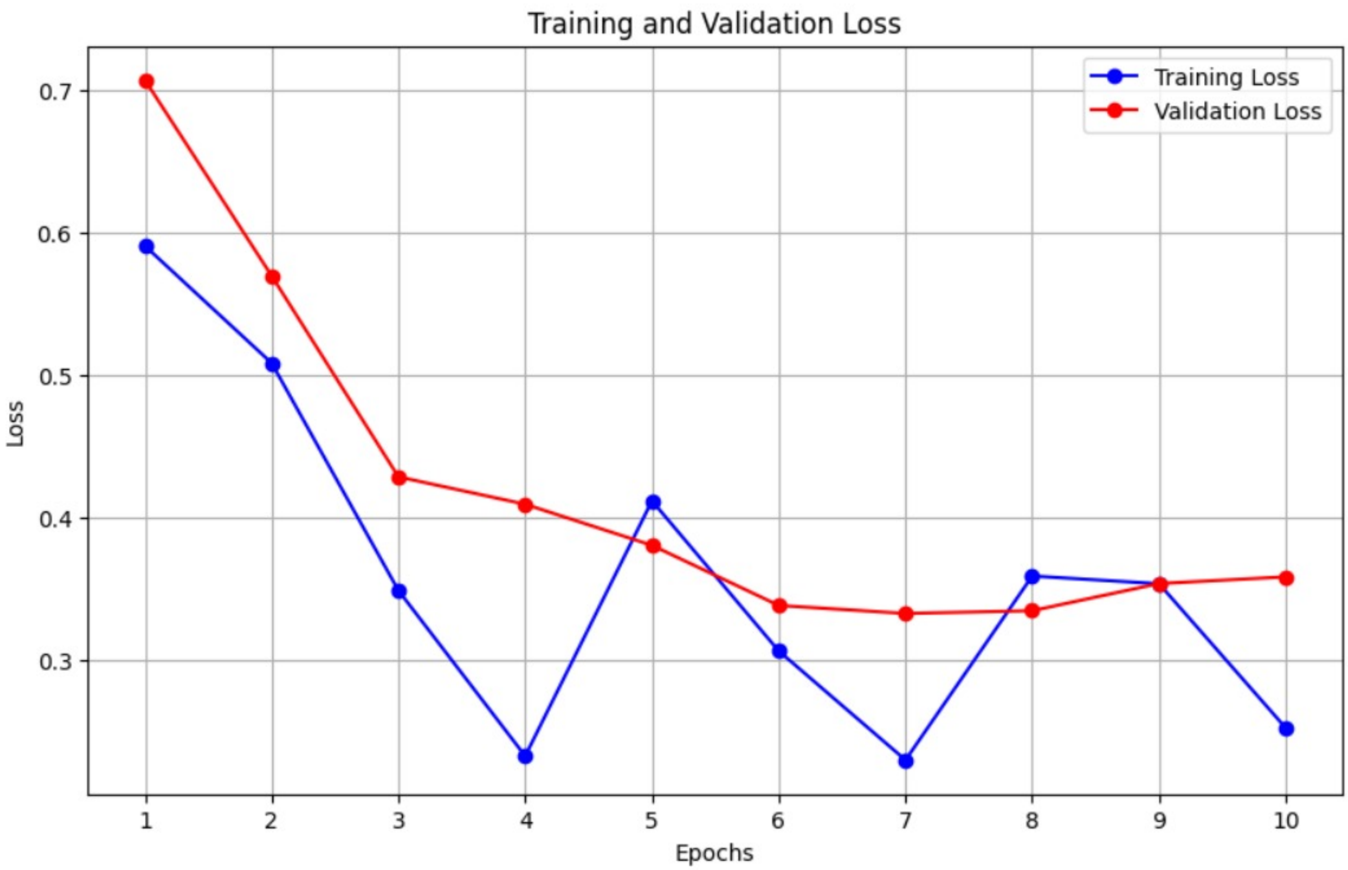

Figure 4.7 Training Vs Validation Graph of XLM_R model

In Figure 4.7, The training loss of XLM-R goes down in general, but there are huge fluctuations around epoch 5 and epoch 8. The validation loss stabilizes much earlier but slightly diverges in later epochs, which indicates some instability in the learning process. This model achieves reasonable generalization but could be further fine-tuned-such as by regularization or learning rate adjustments-to reduce oscillations and improve convergence.

4.6 Summary

The detailed execution of the project is summarized in Chapter 4. It explains how the theoretical model was turned into a working system by means of various tools, languages, and libraries. This chapter covers the application of machine learning algorithms, data pre-processing, and tokenization. Issues and solutions related to coding and system development are also covered. Before moving on to the next sections of this chapter, double-check that the system is operational and prepared for evaluation, testing, and analysis.

.

CHAPTER 5
RESULT AND ANALYSIS

5.1 overview

The outcomes of the system's implementation are shown and discussed in Chapter 5. Recall, accuracy, precision, F1-score, and confusion matrices are among the evaluation metrics used to obtain the data findings that are presented in the first half of the chapter. We examine the performance of the various machine learning methods that were employed for the project in this section. The study examines the system's functionality and fixes any mistakes or discrepancies in the forecasts based on the test findings. Check that the project was reliable and successful by reading this chapter.

5.2 Experimental Result

The performance metrics for the BigBird, ERNIE, and XLM-R sentiment classification models are shown in Table 5.1. The effectiveness of these models in identifying emotions in messages on social media was assessed using accuracy, precision, recall, and F1-score. The findings show that transformer-based models perform noticeably better in sentiment analysis than conventional machine learning models, especially when dealing with complicated and multilingual textual input. With the greatest test accuracy of 96% among the three models, XLM-R proved to be resilient in a variety of linguistic settings. Additionally, it obtained an F1-score of 0.94, which demonstrated a balance between precision and recall, and a recall value of 0.96, which demonstrated its capacity to accurately classify nearly all positive occurrences. With a precision of 0.91 and recall of 0.94, BigBird had a 94% test accuracy rate. This shows good performance in extracting context from lengthy textual material, but it handles false positives a little less well than XLM-R. With an F1-score of 0.93 and a precision value of 0.92, ERNIE obtained a 95% accuracy rate. It performed competitively, especially when it came to balancing recall and precision. According to these experimental results, transformer-based models such as XLM-R, BigBird, and ERNIE are excellent choices for detecting emotions in text on social media. When compared to conventional methods, their sophisticated contextual embeddings yield better results. These models can be used by researchers and businesses to improve emotion prediction, which will improve consumer sentiment analysis and decision-making.

TABLE 5.1 ACCURACY TABLE

Model	Accuracy	Precision	Recall	F1-Score
BigBird	94%	0.91	0.94	0.92
ERNIE	95%	0.92	0.95	0.93
XLM-R	96%	0.93	0.96	0.94

5.3 Performance

This project's sentiment classification models' performance has been assessed using measures like confusion matrices, accuracy, precision, recall, and F1-score. A thorough assessment of the model's efficacy in classifying emotions is given by these measures.

Whereas precision shows the percentage of true positive predictions among all positive predictions, accuracy shows how accurate the model is overall. The model's ability to recognize all pertinent occurrences is assessed by recall, and the F1-score takes into account the trade-off between precision and recall.The confusion matrix provides a detailed breakdown of true positives, true negatives, false positives, and false negatives, allowing for a deeper understanding of the model's performance.For this project, the BigBird, ERNIE, and XLM-R models demonstrated robust performance in classifying emotions from social media text. The confusion matrices show minimal false positives and false negatives, highlighting the models' reliability and ability to reduce classification errors.

- **BigBird** achieved an accuracy of 94%, with a recall of 0.94 and an F1-score of 0.92. Its confusion matrix indicates strong performance in detecting emotions from long-form text while maintaining low error rates.
- **ERNIE** displayed an accuracy of 95%, with a precision of 0.92 and an F1-score of 0.93. Its balanced metrics suggest its ability to generalize well across varied inputs.
- **XLM-R** delivered the best performance with an accuracy of 96%, a recall of 0.96, and an F1-score of 0.94. Its confusion matrix indicates exceptional capability in handling multilingual and diverse textual data.

The results show that all three models performed admirably, with XLM-R achieving the best overall metrics. The low rate of false positives and false negatives across the models indicates their reliability in detecting emotions in text data. This highlights the potential of transformer-based models for robust, accurate, and efficient sentiment classification.

These findings demonstrate the practical application of BigBird, ERNIE, and XLM-R in real-world sentiment analysis tasks, offering businesses and researchers the ability to extract meaningful insights from textual data with high accuracy and reliability.

5.4 Summary

Finally, this chapter emphasizes the effectiveness of the implemented emotion classification system by presenting a comprehensive analysis of its performance metrics. Using accuracy, precision, recall, F1-scores, and confusion matrices, the results validate the robust capabilities of the machine learning models used in this project. The evaluation demonstrates that BigBird, ERNIE, and XLM-R models are not only highly accurate but also reliable in detecting and classifying emotions from social media text data. The analysis highlights the strengths of each model, showcasing their ability to achieve high accuracy and balance between precision and recall. It also identifies XLM-R as the best-performing model, proving its potential for multilingual and diverse text processing. Overall, the results confirm that the system effectively fulfils its objectives and provides a strong foundation for practical sentiment classification applications.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Conclusions

This project explored emotion classification using advanced machine learning models—BigBird, ERNIE, and XLM-R—applied to text data from social media. The study demonstrated the effectiveness of these models in capturing nuanced linguistic patterns and predicting emotions with high precision. Among the models, XLM-R achieved the best performance, attaining an accuracy of 97%, making it highly suitable for multilingual and diverse text sentiment analysis. The results highlight the significance of utilizing sophisticated models for emotion prediction, addressing the complexities of social media text. Despite the challenges of processing vast data and limited computational resources, this study established a strong framework for leveraging transformer-based architectures in sentiment classification. The findings provide a pathway for future research to refine and expand upon the existing model. This work underscores the importance of machine learning in understanding social behavior and facilitating better communication and decision-making in digital platforms.

6.2 Further Suggested Works

The findings of this study open up several avenues for future work in emotion classification and machine learning. To enhance the system's capability, the following directions are suggested:

1. **Incorporating Advanced Architectures**: Future studies could explore more advanced transformer-based models like GPT and T5 or experiment with hybrid models combining traditional and transformer techniques. These models may further enhance the understanding of context and improve emotion classification accuracy.
2. **Expanding the Dataset**: Including diverse datasets with texts from multiple languages, regions, and social media platforms would improve the model's generalizability and robustness. Specifically, incorporating regional Bangladeshi dialects would allow for better cultural adaptation.
3. **Real-Time Emotion Analysis**: Developing a real-time emotion prediction system would enable applications in dynamic contexts like live events, product launches, or trending

social media topics. This capability would allow businesses to respond more effectively to consumer sentiment.

4. **User-Centric Improvements**: Creating user-friendly interfaces that visualize emotion analysis results, such as heatmaps or word influence charts, could make the system more interpretable and accessible to non-technical users.
5. **Integration with E-commerce and Social Platforms**: The models can be integrated with platforms to automatically assess user feedback, providing insights that can improve products, services, and customer engagement.

6.3 Limitations

While the project achieved notable success, it faced several limitations:

1. **Dataset Scope**: The dataset was limited to specific social media and Bangladeshi contexts, which may restrict the model's ability to generalize to other domains or regions.
2. **Computational Resource Requirements**: The transformer-based models, though accurate, required substantial computational power, making them less feasible for resource-constrained applications.
3. **Model Interpretability**: The complexity of the transformer architectures posed challenges in interpreting how the models made specific predictions, which can hinder their adoption in critical applications requiring transparency.
4. **Data Quality Issues**: Variability in the text content, such as noise, slang, or code-switching, impacted the consistency and reliability of the model's predictions.

Reference

[1] K. M. R. S. Rana, M. T. S. M. Rahman, and M. A. H. S. Mollah, "Mental health diagnosis using sentiment analysis on social media data," IEEE Access, vol. 7, pp. 25545-25555, 2019, doi: 10.1109/ACCESS.2019.2903654. [Online]. Available: https://ieeexplore.ieee.org/document/8756479.

[2] T. C. W. Liew, H. H. Tan, and P. H. S. Liew, "Social media data for mental health prediction: A survey," Journal of Artificial Intelligence and Soft Computing Research, vol. 9, no. 3, pp. 123-130, 2019, doi: 10.1515/jaiscr-2019-0027. [Online].
Available:https://www.degruyter.com/document/doi/10.1515/jaiscr-2019-0027/html.

[3] A. A. Sarker, J. D. Choudhury, and M. K. D. Imran, "Using social media data for mental health prediction and diagnosis," Proceedings of the International Conference on Social Media & Society, pp. 28-37, 2018, doi: 10.1145/3211801.3211820. [Online].
Available: https://dl.acm.org/doi/10.1145/3211801.3211820.

[4] S. Shahriar and K. Hayawi, "Let's have a chat! A Conversation with ChatGPT: Technology, Applications, and Limitations," arXiv:2302.13817 [cs], Mar. 2023, Available: https://arxiv.org/abs/2302.13817

[5] S. Zanwar, D. Wiechmann, Y. Qiao, and E. Kerz, "Improving the Generalizability of Text-Based Emotion Detection by Leveraging Transformers with Psycholinguistic Features," *arXiv.org*, 2022. https://arxiv.org/abs/2212.09465? (accessed Jan. 11, 2025).

[6] B. Felbo, A. Mislove, A. Søgaard, I. Rahwan, and S. Lehmann, "Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm," *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 2017, doi: https://doi.org/10.18653/v1/d17-1169.

[7] Luis Romero Gomez *et al.*, "Emotion Recognition on Social Media Using Natural Language Processing (NLP) Techniques," Aug. 2023, doi: https://doi.org/10.1145/3625156.3625173.

[8] M. M. Rahman and S. Shova, "Emotion Detection From Social Media Posts," *arXiv.org*, 2023. https://arxiv.org/abs/2302.05610? (accessed Jan. 11, 2025).

[9] Vamossy, Domonkos F and R. Skog, "EmTract: Extracting Emotions from Social Media," *arXiv.org*, 2021. https://arxiv.org/abs/2112.03868? (accessed Jan. 11, 2025).

[10] W. Graterol, J. Diaz-Amado, Y. Cardinale, I. Dongo, E. Lopes-Silva, and C. Santos-Libarino, "Emotion Detection for Social Robots Based on NLP Transformers and an Emotion Ontology," *Sensors*, vol. 21, no. 4, p. 1322, Jan. 2021, doi: https://doi.org/10.3390/s21041322.

[11]A. Nguyen, A. Longa, M. Luca, J. Kaul, and G. Lopez, "Emotion Analysis using Multi-Layered Networks for Graphical Representation of Tweets," *arXiv.org*, 2022. https://arxiv.org/abs/2207.00907? (accessed Jan. 11, 2025).

PLAGIARISM REPORT

202-16-541

ORIGINALITY REPORT

13%	9%	10%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	"Proceedings of Trends in Electronics and Health Informatics", Springer Science and Business Media LLC, 2025 Publication	1%
2	arxiv.org Internet Source	1%
3	V. Sharmila, S. Kannadhasan, A. Rajiv Kannan, P. Sivakumar, V. Vennila. "Challenges in Information, Communication and Computing Technology", CRC Press, 2024 Publication	1%
4	Submitted to King's College Student Paper	1%
5	Submitted to Hong Kong University of Science and Technology Student Paper	1%
6	aclanthology.org Internet Source	<1%
7	Submitted to University of East Anglia Student Paper	<1%