MULTI-LABEL EMOTIONS BANGLA TEXT CLASSIFICATION

BY

MD. SAJIB AHMED
ID: 202-15-14391
AND

AZMAIN HUDA ADIB
ID: 202-15-14378

FINAL YEAR PROJECT REPORT

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Mohammad Jahangir Alam
Lecturer (Sr. Scale)
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised By

Mr. Md Assaduzzaman
Lecturer
Department of Computer Science and Engineering
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JULY 2024

APPROVAL

This Project titled "**Multi-label Emotions Bangla Text Classification**", submitted by "**Md Sajib Ahmed**" and "**Azmain Huda Adib**" to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **14th July 2024.**

BOARD OF EXAMINERS

Dr. Md. Ismail Jabiullah (MIJ) — **Board Chairman**
Professor
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Mr. Saiful Islam (SI) — **Internal Examiner 1**
Assistant Professor
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Mr. Afjal Hossan Sarower (AHS) — **Internal Examiner 2**
Sr. Lecturer
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Dr. Ahmed Wasif Reza (DWR) — **External Examiner**
Professor
Department of Computer Science and Engineering
East West University

DECLARATION

We hereby declare that this project has been done by us under the supervision of **Mohammad Jahangir Alam, Lecturer (Sr. Scale), Department of Computer Science and Engineering, Daffodil International University**. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

Mohammad Jahangir Alam
Lecturer (Sr. Scale)
Department of CSE
Daffodil International University

Co-Supervised by:

Md Assaduzzaman
Lecturer (Sr. Scale)
Department of CSE
Daffodil International University

Submitted by:

Sajib

Md. Sajib Ahmed
ID: 202-15-14391
Department of CSE
Daffodil International University

Azmain

Azmain Huda Adib
ID: 202-15-14378
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

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

We are grateful and wish our profound indebtedness to **Mohammad Jahangir Alam**, **Lecturer (Sr. Scale)**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of Machine Learning and Deep Learning 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.

We would like to express our heartiest gratitude to the **Head of the Department of CSE**, for his kind help in finishing our project and also to other faculty members and the staff of the Department of CSE, Daffodil International University.

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

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

ABSTRACT

Emotion recognition in text is a challenging task, especially in languages with complex syntax and semantics such as the Bengali language. This paper presents a novel approach for multi-label emotions in Bangla text classification, aimed at accurately identifying and categorizing emotional content within Bengali texts especially encompassing emotion of happiness, sadness, anger, fear, and surprise, among others. The proposed method leverages advanced natural language processing (NLP) techniques, including word embeddings, deep learning models, and multi-label classification algorithms. The first step involves preprocessing the Bangla text data, including tokenization, stemming, and removal of stop words to enhance the quality of feature extraction. We then utilize pre-trained word embeddings to represent the semantic meaning of Bengali words effectively. Subsequently, a deep learning architecture, such as a Bi-LSTM(BiLSTM) network will employed to capture the contextual information and hierarchical relationships within the text. To address the multi-label nature of emotions in Bangla text, we will try to employ techniques such as binary relevance, classifier chains, and label powersets to enable the classification of multiple emotions simultaneously. We achieved the most optimal result at a training accuracy of 89.76%, a training loss of 29.88%, a validation loss of 35.22%, and a validation accuracy of 22.84%. We will also evaluate the performance of our proposed approach using standard evaluation metrics precision, recall, and F1-score. Experimental results will demonstrate the effectiveness of the proposed method in accurately classifying multi-label emotions in Bangla text. The model achieves competitive performance compared to existing approaches, indicating its potential for practical applications in sentiment analysis, opinion mining, and affective computing in Bengali-language domains. The findings of this research can contribute to the advancement of emotion recognition technologies in linguistically diverse contexts, paving the way for a more nuanced understanding and interpretation of textual emotions in the Bengali language.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
Introduction

1.1 Overview

Emotion recognition in text holds significant importance in understanding human behavior, emotion analysis, and natural language understanding tasks. With the exponential growth of digital content in diverse languages, there is a pressing need for effective techniques to recognize emotions expressed in multilingual texts. One of the most spoken languages in the world is Bengali (Bangla), which presents unique challenges and opportunities in this context. The recognition of emotions in Bengali text, often characterized by its rich vocabulary and complex grammatical structures, remains relatively underexplored compared to widely studied languages like English. Emotions in textual content can vary widely, encompassing emotions of happiness, sadness, anger, fear, and surprise, among others. Furthermore, texts often express multiple emotions simultaneously, requiring sophisticated techniques for accurate classification.

This paper addresses the challenge of multi-label emotions in Bangla text classification, aiming to develop robust models capable of identifying and categorizing diverse emotional expressions in Bengali language text. By leveraging advancements in NLP, ML, and DL, we seek to bridge the gap in emotion recognition technologies for Bengali text. The significance of this research extends beyond academic curiosity. Practical applications for emotion recognition in Bengali text include social media analysis, processing consumer feedback, mental health support systems, and instructional tools designed for Bengali-speaking communities. By enabling automated analysis of emotions in Bengali text, our work contributes to the development of more culturally sensitive and linguistically inclusive computational tools and applications. Here in this Report, we provide a comprehensive approach to multi-label emotions in Bangla text classification. We begin by discussing related work within the domains of emotion recognition, text classification, and NLP, highlighting existing methodologies and gaps in research concerning Bengali text. Subsequently, we describe the dataset used for training and evaluation, emphasizing its relevance and representativeness of emotions expressed in Bengali text corpora. Our methodology section outlines the preprocessing steps, feature extraction techniques, as well

as model architectures applied in our approach. We discuss the rationale behind our choices and elucidate how each component contributes to the total efficiency of the emotion classification system. Additionally, we delve into the challenges specific to Bengali text processing and explore strategies for mitigating linguistic complexities during model development. Furthermore, we present experimental results and performance evaluations, showcasing the efficacy of our proposed approach in accurately classifying multi-label emotions in Bengali text. Through comparative analyses with baseline models and state-of-the-art techniques, we highlight the strengths and limitations of our method while providing insights into areas for future research and improvement. This paper advances the state of the art in emotion recognition for Bengali text and underscores the importance of linguistic diversity in computational linguistics and natural language understanding. By offering a robust framework for multi-label emotions in Bangla text classification, we assist in the creation of more inclusive and culturally aware AI systems tailored for Bengali-speaking communities.

1.2 Background and Present State

One of the languages that is most commonly spoken is Bengali in South Asia and Africa, with a rich literary and cultural heritage. Recognizing Bengali text conveys feelings is essential for understanding the sentiments, opinions, and attitudes of Bengali-speaking populations across various domains. Bengali text exhibits diverse linguistic features, including complex grammar, idiomatic expressions, and cultural nuances. Emotion recognition in Bengali text presents unique challenges and opportunities for advancing natural language understanding in linguistically diverse contexts. These things drove us through this emotion-based Bangla text thesis. Accurate emotion recognition in Bengali text has practical implications in numerous domains, including social media analysis, customer feedback processing, mental health support systems, and educational technologies. By enabling automated analysis of emotions in Bengali text, we can enhance user experiences and insights across various applications and platforms. Emotions expressed in Bengali texts are often deeply rooted in cultural contexts and societal norms. Developing emotion recognition systems tailored for Bengali language and culture fosters greater cultural sensitivity and understanding in human-computer interaction and communication. The objectives are quite simple and the strategy is developed based on the objectives. Develop

robust classification models as well as design and implement ML and DL models capable of accurately recognizing and categorizing multiple emotions expressed in Bengali text. Develop techniques to handle the linguistic complexity of Bengali text, including syntactic variations, idiomatic expressions, and contextual nuances, to improve the accuracy of emotion classification. Devise strategies to effectively handle multi-label scenarios where text may express multiple emotions simultaneously, enabling the classification system to capture the nuanced emotional content comprehensively. Leverage available linguistic resources, annotated datasets, pre-trained word embeddings, and linguistic tools to enhance the quality and performance of the emotion recognition system for Bengali text. We will evaluate performance metrics precision, recall, F1-score, and Hamming loss for assessing the performance of the classification models and validate their effectiveness in capturing emotions expressed in Bengali text. We will ensure that the developed models generalize well across diverse domains and are scalable to process large volumes of Bengali text data efficiently in real-world applications. Also plan to share insights, methodologies, and findings with the research community through publications, workshops, and open-source contributions to foster collaboration and advance the state of the art in emotion recognition for Bengali text. By pursuing these objectives, we aim to contribute to the development of culturally sensitive and linguistically inclusive AI systems tailored for Bengali-speaking communities while advancing the broader field of natural language understanding and emotion recognition in multilingual contexts. Original CNNs are a class of deep CNNs particularly effective in examining pictures in the visual. They were motivated by how the animal visual cortex is structured, with a focus on hierarchical feature learning. The original CNN architecture can be traced back to the work of Yann LeCun, Yoshua Bengio, and others in the 1990s. The key components of original CNNs include Layers. First convolutional Layers, which are These layers use learnable filters or kernels to apply convolution operations on the input image. The input data's spatial feature hierarchies are captured by convolutional layers. Second Pooling Layers: These layers preserve the most significant information while shrinking the overall size of the convolutional map of features. For this, max pooling and typical pooling are popular methods. Next Activation Functions: By adding non-linearities to the network, activation functions enable the network to model intricate data interactions. Rectified Linear Units (ReLUs) are the most often utilized activation function in CNNs, while sigmoid and tanh have also been employed in the past. Next, Completely Connected Layers:

Usually, the network is expanded with fully connected layers after a number of convolutional and pooling layers. These layers use the features that were extracted by preceding layers to execute higher-level reasoning and decision-making. Finally, Layer of Softmax: A softmax layer is frequently used at the output of classification jobs to produce class probabilities. It translates the network's raw output scores into cumulative probabilities. The architecture of the original CNNs was relatively shallow compared to modern deep networks, primarily due to computational constraints at the time. LeNet-5, developed by Yann LeCun and his team, is one of the earliest and most influential CNN architectures. It was primarily designed for handwritten digit recognition tasks.

Bi-LSTM networks are based on the principles of regular LSTM networks, a kind of recurrent neural network (RNN) architecture created specially to solve the vanishing gradient problem and simulate sequential data. The shortcomings of conventional RNNs in capturing long-range relationships and avoiding the vanishing gradient issue are addressed by LSTM networks, a kind of RNN architecture. Memory cells and information-flow-controlling gating mechanisms are the means by which LSTMs accomplish this. A storage cell, a data input gate, a gate to forget, and a result gate are all present in each LSTM cell. Unlike traditional LSTM networks, which process sequences in one direction (typically from past to future), Bi-LSTMs process sequences in both directions simultaneously. This means that at each time step, the Bi-LSTM network processes the input sequence both forwards and backwards, allowing it to capture dependencies from both past and future contexts. In a Bi-LSTM, the outputs from both the forward and backward passes are typically concatenated or combined in some way before being passed to subsequent layers or used for downstream tasks. This combined representation captures information from both directions of the input sequence. SGD, Adam, or RMSprop are examples of gradient-based optimization techniques that can be used to optimize the network parameters (weights and biases) for training a Bi-LSTM. It is common practice to compute gradients and update parameters over a number of time steps using BPTT. The use of Bi-LSTMs has become prevalent in many NLP tasks, especially those involving sequential data where capturing bidirectional context is crucial for accurate predictions. However, it's worth noting that Bi-LSTMs have been largely overshadowed by more advanced architectures such as transformers, which have established the accepted norm in numerous applications and have demonstrated improved performance on a broad spectrum of NLP tasks.

1.3 Problem Statement

The Problem Statement for conducting a study on multi-label emotion classification of Bangla text is multifaceted and includes many considerations. Bengali, usually referred to as Bangla, is a language spoken by millions of people mostly in Bangladesh along with the Indian state amongst West Bengal. It is one of those most frequently spoken tongues in South Asia. Applications such as analysis of sentiment, customer feedback analysis, online communication monitoring, as well as mental health evaluation, among others, depend on an understanding of the emotions portrayed in Bangla text. Given the unique cultural and linguistic characteristics of Bangla, a dedicated study on emotion classification in this language can provide valuable insights and tools tailored to its specific nuances and expressions.

Despite the importance of emotion analysis in Bangla text, there may be a scarcity of resources, including labeled datasets, pre-trained models, and research studies, in this domain. Conducting a study on multi-label emotion classification in Bangla can help address this gap by creating annotated datasets, developing effective algorithms and models, and fostering further research and innovation in the field. Emotion classification in Bangla text has numerous practical applications across various domains, including social media analysis, customer service, marketing, education, and mental healthcare. By accurately identifying the emotions expressed in text data, organizations and individuals can better understand user sentiments, tailor their responses and interventions, and improve overall user experiences and outcomes.

Multi-label emotion classification poses unique challenges compared to single-label classification tasks, as text data may contain multiple emotions simultaneously or exhibit subtle nuances that require nuanced analysis. By investigating and resolving these issues in relation to Bangla literature, new approaches, strategies, and algorithms that improve the cutting-edge in emotion categorization and benefit the larger area of processing natural language can be developed. Studying emotion classification in Bangla text can also contribute to cross-cultural understanding and empathy by shedding light on the emotional expressions and experiences of Bangla-speaking populations. This can promote cultural sensitivity, diversity, and inclusivity in the development of AI and NLP technologies and foster greater global collaboration and cooperation in research and development efforts.

1.4 Objectives

The project's objectives for multi-label emotions Bangla text classification encompass both tangible deliverables and broader impacts on research, technology, and society. The following outlines the key outcomes:

Trained Models: The project delivers trained machine learning and deep learning models capable of accurately classifying multi-label emotions expressed in Bengali text across diverse domains. These models incorporate advanced techniques for feature extraction, model architecture design, and handling multi-label classification scenarios.

Bangla Datasets: Curated and annotated Bengali text datasets serve as valuable resources for training, testing, and benchmarking emotion recognition models. These datasets capture the nuances of emotional expressions in Bengali text across various domains, enabling researchers and practitioners to evaluate model performance and advance the state of the art in the field.

Documentation and Reports: Comprehensive documentation, research papers, and reports detailing the methodology, experiments, findings, and insights obtained throughout the project lifecycle are produced. These resources contribute to the academic and research community, facilitating knowledge dissemination, replication of experiments, and further exploration of related research topics.

Evaluation Metrics and Validation: The project establishes standard evaluation metrics and validation techniques for assessing the performance and robustness of multi-label emotions Bangla text classification models. Through rigorous evaluation, researchers and practitioners gain insights into model strengths, limitations, and areas for improvement.

Applications and Impact: The project outcomes have practical applications in various domains, including sentiment analysis, opinion mining, mental health support systems, educational technologies, and customer feedback analysis. Emotion recognition in Bengali text enables enhanced user experiences, a better understanding of user sentiments, and informed decision-making processes in diverse contexts.

1.5 Scope and Limitations

By addressing these points in this report, a comprehensive overview of the scope and limitations of using CNN, RNN, and BiLSTM models for multi-label emotion classification in Bangla text is provided.

Scope:

- Target Language, the primary language of interest is Bangla. This necessitates handling unique linguistic features such as script, syntax, and semantics specific to Bangla. Script and Orthography, focus on text written in Bengali script, which includes handling complex characters and diacritics.
- Multi-Label Classification, the classification system aims to identify multiple emotion labels such as hate speech, offensive language, abusive content, and other forms of emotional communication within a single instance of text.
- CNN, Utilizes CNNs for feature extraction from text sequences. RNN, Employs Recurrent Neural Networks to capture sequential dependencies in text. BiLSTM, Incorporates Bi-LSTM networks to capture context from both past and future states in the text sequence.
- Content Types, Include social media posts, comments, and other user-generated content in Bangla. Data Collection involves creating or sourcing datasets that are representative of the types of emotion to be classified.
- Normalization, standardizes text by handling spelling variations, colloquial language, and diacritics. Tokenization, Splits text into tokens or words while preserving the structure of Bangla. Embedding uses word embeddings or other representation techniques tailored to Bangla to convert text into numerical form.
- Performance Metrics, employs precision, recall, F1-score, and accuracy to evaluate the effectiveness of the models. Multi-Label Metrics, Uses specific metrics for multi-label classification such as Hamming loss, subset accuracy, and macro scores.

Limitations:

- Scarcity of Labeled Data, there is a limited availability of labeled datasets for emotion classification in Bangla, which may impact model training and evaluation.

Quality of Data, the quality and diversity of the data might be limited, affecting the model's ability to generalize across different contexts.

- In imbalanced Datasets, Emotional content may be less frequent than non-emotional content, leading to class imbalance issues that can skew model performance. Minority Classes, Difficulty in accurately detecting less frequent emotion categories due to the imbalance.
- Dialects and Informal Usage, Bangla has various dialects and informal usage that can affect the model's understanding and classification accuracy. Contextual Ambiguities, Emotion detection requires understanding the context, which can be challenging for models, especially in cases of sarcasm or indirect emotion.
- Computational Resources, training complex models like CNN, RNN, and BiLSTM requires substantial computational resources, which might not be accessible to all researchers. Time Consumption, the process of training and fine-tuning these models can be time-consuming.
- Model Complexity, Due to the complexity of the architectures and limited data, there is a risk of overfitting, a condition in which a model functions well on data used for training but badly on untested data. Generalization, ensuring that the model generalizes well to new, unseen data is a significant challenge.
- DL models are often seen as black boxes, making it hard to interpret the rationale behind classifications. explanations for why a particular text was classified as emotional can be difficult, impacting trust in the model.
- The continuous emergence of new emotional expressions and slang necessitates regular updates to the models and retraining with new data. Ensuring that the model adapts to new trends and variations in emotional language over time.

1.6 Report Organization

This Report Organization for Multi-label Emotions Bangla Text Classification. The report is organized into chapters. Our report organization is given below:

Chapter 1: Introduction: Overview of the project objectives, motivation, and significance. Problem statement highlighting the challenges in multi-label emotions Bangla text classification. Scope and objectives of the report.

Chapter 2: Background: Review of existing literature and research related to emotion recognition, text classification, and natural language processing. Discussion of methodologies, techniques, and models employed in emotion recognition for Bengali text and other languages. Identification of gaps, challenges, and opportunities for research in multi-label emotions Bangla text classification.

Chapter 3: Research Methodology: A detailed explanation of the methodology employed for multi-label emotions Bangla text classification. Discussion of feature extraction techniques, including word embeddings, contextual embeddings, and linguistic features. Description of model architectures explored. Strategies for handling multi-label classification scenarios and optimizing model performance. Description of the experimental setup, including hardware specifications, software environment, and libraries used. Overview of the training, validation, and testing procedures.

Chapter 4: Experimental Result: Present the results of the study, including quantitative and qualitative findings. Describe the performance metrics of the trained CNN, Bi-LSTM, and LSTM models, accuracy, loss,val_loss, and val_accuracy. Provide detailed analysis and interpretation of the results, highlighting the strengths and limitations of the CNN- based approach. Discuss the implications of the findings and their alignment with the research objectives. Compare the performance of the CNN-based approach with existing methods. Analyze the factors contributing to the effectiveness or limitations of the approach. Address any challenges encountered during the research.

Chapter 5: Impact on Society: Provide guidelines and recommendations for implementing the CNN-based sensitive speech detection and classification system in real- world scenarios and content moderation frameworks. Discuss practical considerations, scalability, and potential challenges in integrating the system into online platforms.

Chapter 6: Synopsis, Findings, Suggestion, and Research Implications for the Future: Give a brief summary of the study's main conclusions and their consequences. Talk about the research's limits and contributions. Provide recommendations for future research and development pertaining to CNNs' use in sensitive conversation detection and categorization.

References: A comprehensive list of references cited throughout the report, including academic papers, articles, and relevant literature.

Appendices: Additional resources that provide information on the methodology and experimentation include Complex Engineering Problems (EP) and Complex Engineering

Activities (EA) Analysis. Anything more that is thought to be required in order to comprehend and duplicate the research findings.

1.7 Summary

Emotion recognition in text is crucial for understanding human behavior, emotion analysis, and natural language understanding tasks. Bengali (Bangla), one of the world's most spoken languages, presents unique challenges and opportunities in this context. The recognition of emotions in Bengali text, often characterized by its rich vocabulary and complex grammatical structures, remains relatively underexplored compared to widely studied languages like English. Emotions in textual content can vary widely, encompassing emotions of happiness, sadness, anger, fear, and surprise, among others. Furthermore, texts often express multiple emotions simultaneously, requiring sophisticated techniques for accurate classification.

This paper addresses the challenge of multi-label emotions in Bangla text classification, aiming to develop robust models capable of identifying and categorizing diverse emotional expressions in Bengali language text. By leveraging advancements in NLP, ML, and deep learning, the researchers seek to bridge the gap in emotion recognition technologies for Bengali text. The significance of this research extends beyond academic curiosity, as it has practical implications in various domains, including social media analysis, customer feedback processing, mental health support systems, and educational technologies tailored for Bengali-speaking populations. The methodology section outlines the preprocessing steps, feature extraction techniques, and model architectures employed in the approach. The rationale behind our choices and elucidate how each component contributes to the overall effectiveness of the emotion classification system. Additionally, the paper delves into the challenges specific to Bengali text processing and explores strategies for mitigating linguistic complexities during model development. Experimental results and performance evaluations showcase the efficacy of our proposed approach in accurately classifying multi-label emotions in Bengali text. The objectives of this paper are simple: develop robust classification models, and design and implement machine learning and deep learning models capable of accurately recognizing and categorizing multiple emotions expressed in Bengali text. Develop techniques to handle the linguistic complexity of Bengali text, including syntactic variations, idiomatic expressions, and contextual nuances, to improve the accuracy of emotion classification. Devise strategies to effectively handle multi-label scenarios where

text may express multiple emotions simultaneously, enabling the classification system to capture the nuanced emotional content comprehensively.

The developed models will ensure that they generalize well across diverse domains and are scalable to process large volumes of Bengali text data efficiently in real-world applications. By pursuing these objectives, the researchers aim to contribute to the development of culturally sensitive and linguistically inclusive AI systems tailored for Bengali-speaking communities while advancing the broader field of natural language understanding and emotion recognition in multilingual contexts.

CHAPTER 2

Literature Review

2.1 Overview

The review of the literature on multi-label emotions an extensive overview of current research, methodology, and strategies pertinent to the subject of Bangla text classification is provided. It provides as a basis for comprehending the state of the art at the time, pointing out areas in need of research, and guiding the project's technique and strategy. The technical description of emotion recognition is covered in this literature review, along with its importance in the processing of natural languages and human-computer interaction. Overview of approaches to emotion recognition, including lexicon-based, machine learning, and deep learning methods. Summary of existing research on emotion recognition in Bengali text, including sentiment analysis and emotion classification. Discussion of datasets, annotation schemes, and challenges specific to emotion recognition in Bengali text. Overview of multi-label classification techniques and algorithms, including binary relevance, classifier chains, and label powerset methods. Discussion of strategies for handling multi-label scenarios and evaluating model performance. Review of natural language processing techniques and resources available for Bengali text processing, including tokenization, stemming, and part- of-speech tagging. Summary of existing tools, libraries, and datasets for Bengali language processing and analysis. Overview of deep learning architectures used for text classification tasks, including RNNs, and CNNs -based models. Discussion of pre-trained language models and transfer learning approaches for emotion recognition in text. Explanation of evaluation metrics used for assessing the performance of emotion recognition models, including accuracy, precision, recall, F1-score, and loss. Overview of datasets used for evaluating emotion recognition models in Bengali text. Identification of challenges and limitations in existing approaches to multi-label emotions Bangla text classification. Comparison of emotion recognition techniques and models across different languages, highlighting similarities, differences, and lessons learned. Summary of key findings from the literature review and their implications for the project. Identification of gaps in existing research and justification for the proposed methodology and approach. By conducting a thorough literature review, the project gains insights into the current state of research, leverages existing knowledge and methodologies, and identifies opportunities for innovation and advancement in multi-label emotions Bangla text classification.

2.2 Related Works

Nafis Irtiza Tripto et al. [1] Natural language processing relies heavily on sentiment analysis, which has applications in opinion mining, emotion extraction, and social media trend forecasting. An extensive collection of methods for recognizing sentiment and obtaining emotions from Bangla texts is presented in this work. Sentences are classified using three-class and five-class sentiment labels using deep learning-based models, which also extract emotions as one of six fundamental emotions. The model's performance is better for domain and language-specific texts. Md Sakib Ullah Sourav et al. [2] Using transformer-based models, suggest a Bangla emotional classifier for six categories (Anger, Disgust, Fear, Joy, Sadness, as well as Surprise) found in Bangla literature. The model has demonstrated impressive results, particularly for languages with abundant resources. The models' performance is evaluated using the Unified Bangla Multi-class Emotional Corpus (UBMEC), which combines two previously published datasets with newly manually annotated Bangla comments. Tanvirul Alam et al. [3] Optimize multilingual transformer models to perform Bangla text classification tasks in several domains, such as authorship attribution, sentiment analysis, emotion recognition, and news classification. The state-of- the-art results show a 52.95% accuracy, demonstrating the potential of these models in tackling new areas like noisy social media content in Natural Language Processing (NLP). Sara Azmin et al. [4] Emotions are crucial in human interaction, especially on social media and blogs. This paper investigates emotion detection in the Bangla language using a Multinomial Naïve Bayes classifier and features like stemmer, parts-of-speech tagger, n- grams, and term frequency-inverse document frequency. With an overall accuracy of 78.6%, the model successfully categorizes text into three emotion classes: happy, sad, and angry. A hybrid approach for multilabel emotion classification in the Bangla language is presented by Ahasan Kabir et al. [5], combining transformers and lexicon features. The model outperforms pre-trained transformers, with the highest performance achieved using BanglaBERT (large) as the pre-trained transformer. Avishek Das et al. [6] introduce TEmoX, a transformer-based ensemble method for textual emotion classification in Bengali. The method divides Bengali text into six emotional categories: surprise, terror, dread, anger, and disgust. The model performs better than baseline models and current techniques. Asif Iqbal, MD, and others [7] For many applications, it is essential to classify emotions in Bengali texts, but creating an automatic system for a language with limited resources like Bengali is difficult. An emotional corpus

(BEmoC) for the classification of six emotions is presented in this work, showing a high degree of agreement amongst annotators and specialists. Noman Ashraf et al. [8] multi-label emotion dataset, consisting of 6,043 tweets and six basic emotions in the Urdu Nastalíq script, was created to detect emotions from Urdu. The paper presents baseline classifiers using machine learning algorithms, deep-learning algorithms, and transformer-based baselines. The paper focuses on the annotation criteria, features of the dataset, and techniques for Urdu emotion categorization. Anger, fear, disgust, sadness, joy, and surprise are the six fundamental emotions that can be identified in Bengali literature using the transformer-based method proposed by Avishek Das et al. [9]. The method achieves a weighted f1-score of 69.73% on the test data, outperforming other approaches using a Bengali emotion corpus of 6243 texts. Using an 8047-emotion corpus, Tanzia Parvin et al. [10] demonstrate an ML-based ensemble technique for classifying six main textual emotions from Bengali texts. The ensemble approach yielded the greatest weighted f1-score of 62.39% when paired with tf-idf feature extraction approaches. A detailed emotional analysis of Bangla text is presented by Md. Ataur Rahman et al. [11], who examine user comments on political and economic topics. Classical machine learning techniques such as Naïve Bayes, Decision Tree, K-Means Clustering, Support Vector Machine (SVM), and k-nearest Neighbor were employed in the study. Using a non-linear radio-basis function kernel, the best model, SVM, had an average accuracy score of 52.98%. This study Sumit Kumar Banshal et al. [12] presents a method for extracting multiple sentiments in the Bangla language using scrapped Facebook data. It uses context-based annotation and Bidirectional Encoder Representations from Transformers, resulting in a web application for multi-label ER. Laurence Vidrascu et al. [13] The study explores affective systems in computer science, focusing on real-world audio data from a French Human-Human call center. Using acoustic features and speech transcription knowledge, the researchers achieved a detection rate of 45%. By adding orthographic transcription knowledge and selecting the best 25 features, they achieved 56% good detection for the task of discriminating five emotions.

2.3 Comparison between existing works

Existing works on multi-label emotions in Bangla text classification have primarily focused on leveraging machine learning and deep learning techniques to discern complex emotional expressions in Bangla textual data.

TABLE 2.3: COMPARATIVE ANALYSIS USING EARLIER RESEARCH

SL No	Author Name	Used Algorithm	Best Accuracy with Algorithm
1.	Nafis Irtiza Tripto etal. [1]	LSTM, CNN, NB, SVM	65.97% and 54.24% accuracy
2.	Md Sakib UllahSourav et al. [2]	LR. NB, XGBoost	highest f1-score 76%
3.	Tanvirul Alam et al.[3]	BERT, XLM-RoBERTa	71.7%, 76%
4.	Ahasan Kabir et al.[4]	BEmoLexBERT	an impressive F1 scoreof 82.7%
5.	Avishek Das et al.[5]	LR, RF, MNB, BiLSTM, CNN+BiLSTM,	highest weighted F1-score of 80.24 percent
6.	MD. Asif Iqbal et al.[6]	BEmoC, cross-validation	similarity scores 97.47%
7.	Noman Ashraf et al.[7]	LSTM, and LSTM with CNN	F1 score of 56.10%
8.	Avishek Das et al.[8]	CNN, BiLSTM, CNN+BiLSTM	weighted f1-score of 59.73%
9.	Tanzia Parvin et al.[9]	KNN, AdaBoost, tf-idf and Bag of Words (BoW) feature extraction methods in LR, RF, and SVM.	f1-score of 62.39%

10.	Sumit Kumar Banshal et al. [12]	LR, RF, MNB, SVM, KNN	LR=80.03%,MNB=82.64%, SVM=80.72%, RF=79.11%, KNN=70.87%

2.4 Open Issues

The open Issues for multi-label emotions Bangla text classification encompasses various aspects related to understanding and analyzing the emotional content expressed in Bangla language text data. Here's an overview of the scope:

Language and Cultural Context: Spoken by millions of people, mainly in Bangladesh as well as the Indian state of West Bengal, Bangla is a rich and diversified language. The scope involves understanding the linguistic and cultural nuances of Bangla text, including variations in expressions, idiomatic phrases, and regional dialects, which influence the interpretation of emotions.

Emotion Representation: Emotions are complex and multifaceted constructs that manifest differently across languages and cultures. The scope includes identifying and representing a wide range of emotions expressed in Bangla text, including basic emotions anger, happiness, surprise, fear, sadness, confused.

Multi-Label Classification: Unlike single-label classification tasks where each instance is assigned only one label, multi-label classification involves predicting multiple labels or categories for each input instance. The scope encompasses developing algorithms and models capable of accurately predicting multiple emotion labels simultaneously for Bangla text samples, reflecting the nuanced and multifaceted nature of emotions.

Dataset Collection and Annotation: The scope involves collecting or creating annotated datasets of Bangla text samples labeled with multiple emotion categories. This includes selecting appropriate sources of text data, annotating the data with relevant emotion labels, ensuring annotation consistency and quality, and addressing biases or imbalances in the dataset.

Feature Engineering: Feature engineering involves extracting informative features from Bangla text data to represent emotional content effectively. The scope includes exploring various feature representation techniques, such as lexical, syntactic, semantic, and contextual features, as well as advanced embedding methods tailored to Bangla text.

Model Development and Evaluation: The range of work includes creating and assessing deep learning and machine learning models for multi-label emotions. Bengali text categorization. This includes selecting appropriate model architectures, training and fine-tuning the models on annotated data, evaluating model performance using suitable metrics, and iteratively refining the models based on evaluation results.

Cross-Domain Generalization: Emotions are expressed across various domains, including social media, news articles, literature, and personal communications. The scope involves investigating the generalization ability of models trained on one domain to predict.

emotions in other domains of Bangla text, considering domain-specific linguistic styles and contextual differences.

Deployment and Application: The ultimate goal of multi-label emotions Bangla text classification is to deploy the developed models for practical applications. The scope encompasses deploying models in real-world settings, integrating them into software applications or platforms, and evaluating their performance and usability in relevant use cases, such as analysis of consumer feedback and sentiment, mental health assessment, and social media monitoring.

TABLE 2.4: ISSUES FOR ANALYZING BANGLA TEXT

S.No.	Issues	Strategies or Plans
1	Data Scarcity andQuality	Dataset Creation and Enhancement: Encourage the creation of open-access datasets through community-driven efforts or academic collaborations. Focus on creating diverse datasets that include various text types and emotions.
2	Complexity ofLanguage	Invest in developing NLP tools specifically tailored to Bangla, such as custom tokenizers, stemmers, and morphological analyzers.

3	Model Generalization andOverfitting	To prevent overfitting, use strategies like early halting, L2 regularization, and dropout during model training.
4	Real-World Application and Integration	Deploy models using a microservices approach, allowing easy integration with different applications and scalable maintenance. For applications requiring low latency, use edge computing to process data on local devices instead ofrelying on cloud services.

2.5 Summary

Prior research on multi-label emotion in Bangla text classification has mostly concentrated on comprehending complicated emotional expressions in Bangla textual data through the application of ML and DL techniques. The scope of the problem for multi-label emotions Bangla text classification encompasses various aspects related to understanding and analyzing the emotional content expressed in Bangla language text data. The scope includes understanding the linguistic and cultural nuances of Bangla text, including variations in expressions, idiomatic phrases, and regional dialects that influence the interpretation of emotions. Emotions are complex and multifaceted constructs that manifest differently across languages and cultures. The scope also includes identifying and representing a wide range of emotions expressed in Bangla text, including basic emotions such as anger, happiness, surprise, fear, sadness, and confused. Multi-label classification involves predicting multiple labels or categories for each input instance, which requires developing algorithms and models capable of accurately predicting multiple emotion labels simultaneously for Bangla text samples. Dataset collection and annotation involve selecting appropriate sources of text data, annotating the data with relevant emotion labels, ensuring annotation consistency and quality, and addressing biases or imbalances in the dataset. Feature engineering involves extracting informative features from Bangla text data to represent emotional content effectively. This involves exploring various feature representation techniques, such as lexical, syntactic, semantic, and contextual features, as well as advanced embedding methods tailored to Bangla text.The ultimate goal of multi-label emotions Bangla text classification is to deploy the

developed models for practical applications, such as analysis of consumer feedback and sentiment, mental health assessment, and social media monitoring. Strategies or plans for analyzing Bangla text include dataset creation and enhancement, investing in NLP tools specifically tailored to Bangla, model generalization and overfitting, and real-world application and integration using a microservices approach.

CHAPTER 3

Methodology

3.1 Overview

Previous studies have focused on using ML and DL techniques to classify multi-label emotions in Bangla text. These studies have achieved high accuracy rates, with some researchers achieving an F1 score of 82.7%. Other researchers have achieved similarity scores of 97.47%. However, there are still open issues, such as a lack of a comprehensive understanding of the complex emotional expressions in Bangla textual data. Subsequent investigations ought to concentrate on enhancing these models' accuracy and resolving the fundamental problems. The multi-label emotions Bangla text classification problem involves understanding and analyzing the emotional content expressed in Bangla language text data. This includes understanding the linguistic and cultural context of Bangla, which influences the interpretation of emotions. Emotions are complex and multifaceted constructs that manifest differently across languages and cultures, including basic emotions like anger, happiness, surprise, fear, sadness, and confused. Multi-label classification involves predicting multiple labels or categories for each input instance, reflecting the nuanced and multifaceted nature of emotions. Data collection and annotation involve collecting or creating annotated datasets of Bangla text samples labeled with multiple emotion categories. Feature engineering involves extracting informative features from Bangla text data to represent emotional content effectively. Model development and evaluation involve developing and evaluating machine learning and deep learning models for multi-label emotions Bangla text classification. Cross-domain generalization involves investigating the generalization ability of models trained on one domain to predict emotions in other domains of Bangla text, considering domain-specific linguistic styles and contextual differences. The ultimate goal of multi-label emotions Bangla text classification is to deploy the developed models for practical applications, such as analysis of consumer feedback and sentiment, mental health assessment, and social media monitoring. Strategies or plans for analyzing Bangla text include data scarcity and quality, the complexity of language, model generalization and overfitting, and real-world application and integration. The research subject for multi-label emotions Bangla text classification focuses on creating models and algorithms that can reliably identify various emotion labels for text samples written in the Bangla language. Data

collection and annotation involve collecting Bangla language text data from diverse sources, annotating it with multiple emotion labels, and ensuring annotation quality and consistency. Feature engineering involves extracting informative features from Bangla text data, using natural language processing techniques, word embeddings, and contextual embeddings. Model development involves using machine learning frameworks, deep learning libraries, and evaluation metrics. Ethical considerations include adhering to ethical guidelines, privacy regulations, and informed consent forms.

3.2 Proposed Methodology

Using a Bi-LSTM model for multi-label emotions Bangla text classification is a promising approach due to its ability to capture sequential dependencies and contextual information from textual data. Here's how we will integrate a Bi-LSTM model into our methodology:

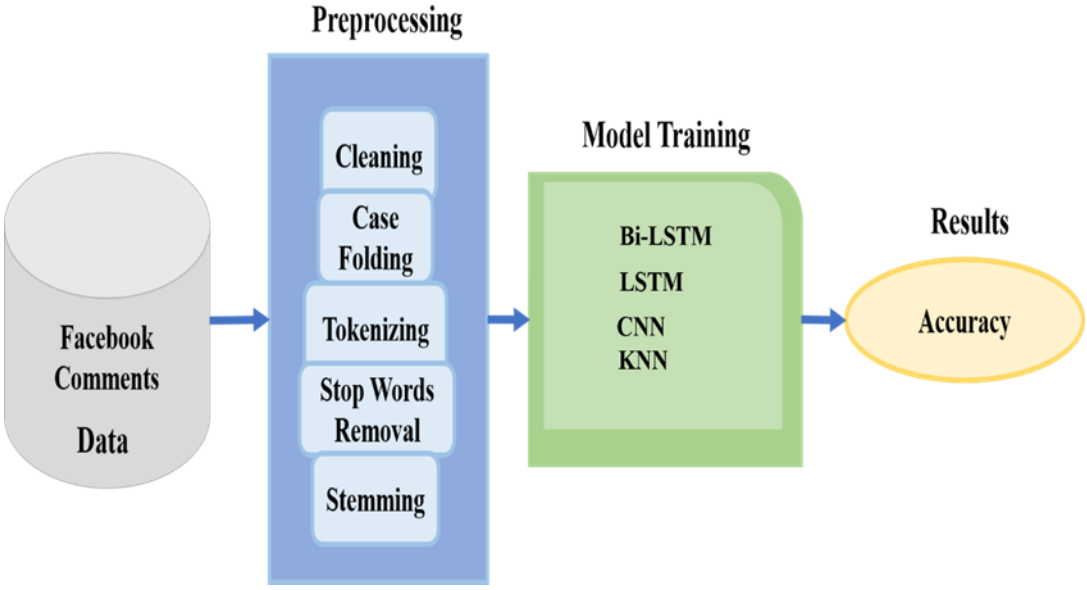

Figure 3.2.1: Workflow diagram for Bangla Emotion analysis

We will prepare our Bangla text dataset, ensuring it is annotated with multiple emotion labels for each text sample. To make model training, tuning, and evaluation easier, divide the dataset into training, validation, and testing sets. Then tokenize Bangla text into individual words or characters. Convert tokens into numerical representations, possibly using word embeddings like Word2Vec, FastText, or pre-trained language models. Design a BiLSTM architecture capable of processing sequential data bidirectionally. Stack multiple BiLSTM layers to capture increasingly abstract representations of the input text. Optionally, incorporate additional layers such as convolutional layers or attention mechanisms to

enhance model performance and interpretability. Configure the output layer to produce multi-label predictions, typically using a sigmoid activation function to enable independent classification of each emotion category. Evaluate the trained Bi-LSTM model on the held-out test set using evaluation metrics suitable for multi-label classification tasks. Common metrics include F1 Score, validation loss, Precision, Recall, and m-AP. Interpret model performance across different emotion categories and analyze error patterns to identify areas for improvement. By incorporating a Bi-LSTM model into our methodology, we will leverage its ability to capture long-range dependencies and contextual information in Bangla text, thereby enhancing the accuracy and effectiveness of multi-label emotion classification tasks.

Data Collection

Storing Bangla emotion-related comments in an Excel file and leveraging our supervisor's guidance for annotation is a practical approach to managing our dataset. Here's a step-by-step guide:

Create an Excel Spreadsheet: Open Microsoft Excel or any spreadsheet software in our storage. Create a new Excel sheet with columns to accommodate relevant information such as:

- Comment ID: A unique identifier for each comment.
- Comment Text: The actual Bangla text content of the comment.
- Emotion Labels: Columns to be filled with emotion labels assigned to each comment by our supervisor.

Copy Comments into the Excel Spreadsheet: Copy and paste the Bangla emotion-related comments collected from Facebook into the appropriate column ("Comment Text") in the Excel spreadsheet.

Share the Spreadsheet with our Supervisor: We will share the Excel file with our supervisor through a file-sharing platform email. Provide clear instructions and guidelines on the annotation task, including the emotion categories to be assigned to each comment.

Annotation Process with Supervisor: We create a Google form and survey among common Bengali people through a Facebook group [20]. Then we collaborate with our supervisor to annotate those comments with emotion labels based on the defined categories. [21] Our supervisor reviewed each comment and assigned appropriate emotion labels happiness, sadness, fear, anger, surprise, and confused, according to the content. Maintain open communication with our supervisor to

address any questions or uncertainties during the annotation process. The survey form is given below:

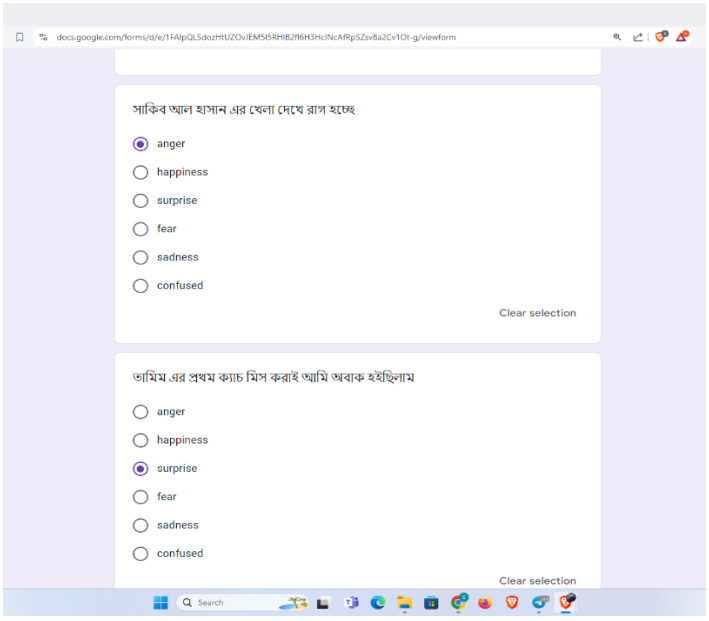

Figure 3.2.2: Survey form for Bangla emotion

Document Annotation Decisions: Record the emotion labels assigned by our supervisor in the corresponding columns of the Excel spreadsheet. Document any annotation guidelines, criteria, or disagreements encountered during the labeling process for future reference.

Review and Iterate: Review the annotated dataset with our supervisor to ensure consistency and accuracy in emotion labeling. Address discrepancies in the annotations through discussion and clarification with our supervisor. Iterate on the annotation process as needed to refine the dataset and enhance its quality.

Finalize and Store the Dataset: Once the annotation process is complete and consensus is reached on the emotion labels, finalize the dataset. Save the Excel file containing the annotated dataset in a secure location for future analysis and modeling tasks.

By storing Bangla emotion-related comments in an Excel file and leveraging our supervisor's guidance for annotation, we will effectively manage the dataset creation process and ensure the quality and reliability of the annotated data for multi-label emotions in Bangla text classification.

Preprocessing Steps:

Here we discuss our data preprocessing step by step:

text_to_word_list(text) function: In this function we take a string text as input and splitsit into a list of words. It uses the split () method to split the text based on whitespace and returns

the resulting list of words.

```
def text_to_word_list(text):
    text = text.split()
    return text

def replace_strings(text):
    emoji_pattern = re.compile("["
                               u"\U0001F600-\U0001F64F"  # emoticons
                               u"\U0001F300-\U0001F5FF"  # symbols & pictographs
                               u"\U0001F680-\U0001F6FF"  # transport & map symbols
                               u"\U0001F1E0-\U0001F1FF"  # flags (iOS)
                               u"\U00002702-\U000027B0"
                               u"\U000024C2-\U0001F251"
                               u"\u00C0-\u017F"          #latin
                               u"\u2000-\u206F"          #generalPunctuations

                               "]+", flags=re.UNICODE)
    english_pattern=re.compile('[a-zA-Z0-9]+', flags=re.I)
    #latin_pattern=re.compile('[A-Za-z\u00C0-\u00D6\u00D8-\u00f6\u00f8-\u00ff\s]*',)

    text=emoji_pattern.sub(r'', text)
    text=english_pattern.sub(r'', text)

    return text

def remove_punctuations(my_str):
    # define punctuation
    punctuations = '''`[illegible]'''

    no_punct = ""
    for char in my_str:
        if char not in punctuations:
            no_punct = no_punct + char

    # display the unpunctuated string
    return no_punct


def joining(text):
    out=' '.join(text)
    return out

def preprocessing(text):
    out=remove_punctuations(replace_strings(text))
    return out
```

Figure 3.2.3: Preprocessing steps

replace_strings(text) function: In this function, we replace certain patterns in the input text with an empty string. Specifically, it removes emojis and non-English characters. It uses regular expressions (re module) to define patterns for emojis and English characters and then applies sub () method to replace these patterns with an empty string. **remove_punctuations(my_str) function:** This function removes punctuations from the input string my_str. It defines a string punctuation containing all punctuation characters, then iterates over each character in my_str and appends it to no_punct if it's not in the punctuations string.

joining(text) function: This function takes a list of words `text` and joins them back intoa single string using whitespace as the separator.

preprocessing(text) function: This function applies preprocessing steps to the input text text. It first removes punctuations and then replaces certain strings (emojis and non-English characters) using the replace_strings () function.

3.3 Hardware and Software Requirement

To ensure the successful implementation of multi-label emotion classification of Bangla text using CNN, RNN, and BiLSTM models, the following hardware and software specifications are recommended.

Hardware Requirements

- CPU: A multi-core processor as Intel i7 to facilitate efficient data preprocessing and model training.

- Graphics Processing Unit (GPU): A dedicated GPU, preferably NVIDIA GeForce GTX 1080, RTX 2070, or higher (e.g., NVIDIA RTX 3080) to expedite the training of deep learning models, especially for larger datasets.
- Random Access Memory (RAM): A minimum of 16GB of RAM is essential for handling large datasets and model parameters, though 32GB or more is advisable for enhanced performance.
- Storage: At least 500GB of SSD storage is recommended to accelerate data loading and saving of model checkpoints.
- Peripherals: Quality peripherals including a monitor, keyboard, and mouse for an efficient and comfortable working environment.

Software Requirements

- Operating System: Linux distributions are used for better compatibility with deep learning libraries, though Windows and macOS are also supported.
- Programming Language: Python 3.7 or higher.
- Essential Python Libraries
- TensorFlow and PyTorch: For constructing and training DL models.
- tensorflow==2.x and pytorch==1.x
- Keras: A top-level TensorFlow-based neural network application program.
- keras==2.x
- Scikit-learn: For data preprocessing and evaluation metrics.
- scikit-learn==0.24.x
- Pandas: For data manipulation and analysis.
- pandas==1.x
- NumPy: For numerical operations.
- numpy==1.19.x
- NLTK or spaCy: For natural language processing tasks.
- nltk==3.5 or spacy==3.x
- Matplotlib or Seaborn: For data visualization.
- matplotlib==3.x or seaborn==0.11.x
- Additional Tools
- Jupyter Notebook and CoLab: For an interactive coding environment.

3.4 Project Management and Financial Analysis

Project Management: After defining the scope, objectives, and deliverables of the projectfor multi-label emotions Bangla text classification. We will establish clear goals and success criteria to measure the project's progress and outcomes. Identify key team members with expertise in natural language processing, machine learning, Bengali language, and emotion recognition. Define roles and responsibilities, including project manager, data annotators, developers, and researchers. Develop a detailed project timeline with milestones and deadlines for each phase of the project, including data collection, model development, evaluation, and documentation. Monitor progress regularly and adjust timelines as needed to ensure timely completion of tasks. Allocating resources including personnel, computing infrastructure, and budget for data acquisition, model training, and experimentation. Optimize resource utilization and efficiency to meet project objectives within budget constraints. We can identify potential risks and uncertainties associated withdata availability, model performance, technical challenges, and project timelines. Developmitigation strategies and contingency plans to address risks and minimize their impact onproject outcomes. And establish regular communication channels and collaboration platforms enabling team members to talk about progress, exchange updates, and resolve issues. Facilitate stakeholder engagement and feedback throughout the project lifecycle to ensurealignment with project goals and objectives.

Project Management Plan: The multi-label classification of emotions in the Bangla tongue is the focus of this study plan. We will adhere to these measures in our project management strategy.

Project Objectives:

- Develop a multi-label emotion classification system for Bangla Language.
- Make it possible to accurately identify and categorize offensive material conveyedin Bangla Language.
- Offer a useful tool for sentiment analysis, online emotion, and content control onBangla-language platforms and communities.

Project Timeline:

Phase 1: Research and Project Planning (2-4 weeks)

Phase 2: Data Collection and Labeling (4-6 weeks)

Phase 3: Model Development and Training (6-8

weeks)Phase 4: Evaluation and Fine-tuning (4-6

weeks)

Phase 5: Integration along with Deployment (2-4

weeks)Phase 6: Testing and Quality Assurance (4-6

weeks Phase 7: Post Support and Review (Ongoing)

Resource Planning:

Equipment and Tools

- High-performance computing resources for model training.
- Software tools for data preprocessing, machine learning, and evaluation.
- Collaboration platforms for team communication and coordination.

Software and Technologies

- NLP libraries and frameworks suitable for Bangla text.
- Machine learning, DL algorithms for text classification.
- Google Drive platforms for scalable computing and storage.

Data and Content

- Bangla text datasets with annotated emotional labels.
- Annotation tools for manual labeling and validation.

Training and Skill Development

- Ensure team members possess expertise in Bangla language, NLP, machine learning,Deep learning, and software development.
- Provide training on specific tools, techniques, and methodologies required for theproject.

Risk Management:

Perform a comprehensive risk assessment, considering any issues with data accessibility, annotation quality, model bias, and regulatory compliance. Put risk mitigation techniquesinto practice, such as varying the sources of data, carrying out sensitivity studies, and including measures for fairness and transparency in the model. To protect project goals andstakeholder interests, regularly monitor project progress and take immediate action to mitigate any developing risks. Here is a SWOT analysis for the e-commerce platform for emotion classification for the Bangla language.

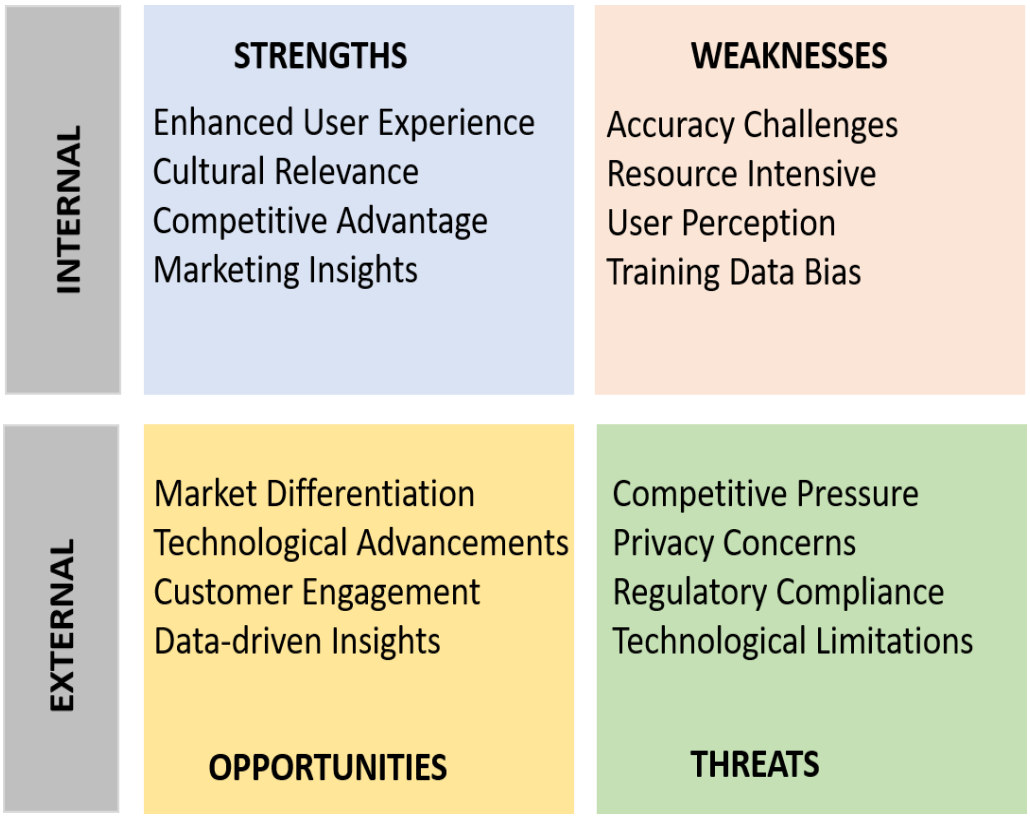

Figure 3.2.3: SWOT analysis for emotion classification for the Bangla language.

Communication Plan:

- Regular Team Meetings: Weekly or biweekly gatherings to review developments, resolve issues, and coordinate priorities.
- Stakeholder Announcements Update project stakeholders on a frequent basis with milestones reached, difficulties faced, and future plans.
- Recordkeeping and Summarizing Keep thorough records of all project activities, including data sources, procedures, outcomes of experiments, and choices taken. Whennecessary, create summaries and progress reports for stakeholders.
- Channels for Feedback Create procedures for gathering input from stakeholders, teammembers, and subject matter experts at various stages of the project to help with decision-making and enhance project results.

Financial Analysis: We will estimate the costs associated with data acquisition, annotation, computing resources, software licenses, and personnel. Here is our table for the cost component.

TABLE 3.4 FINANCIAL ANALYSIS FOR BANGLA EMOTION TEXT

SN	**Components**	**EstimatedCost (BDT)**
01.	Infrastructure	25000
02.	Visiting Stakeholders	8000
03	Software and Tools	10000
04.	Data Collection and Processing	7000
05.	Documentation and Report Writing	7000
06.	Software licenses, tools licenses	3000
07.	Data scientists	16000
08.	Annotators	9000
09.	Model Training	7000
10.	Contingency (10% of total)	9000
	Total Estimated Cost	1,01,000

This table provides an overview of the estimated costs for each expense category involved in the research project. Actual costs may vary depending on factors such as project duration, specific requirements, and market rates. Adjustments can be made based on available resources and funding opportunities.

3.5 Summary

The Bi-LSTM model is promising approach for multi-label emotions Bangla text classification. It captures sequential dependencies and contextual information from textual data, enhancing the accuracy and effectiveness of multi-label emotion classification tasks. The methodology involves preparing a Bangla text dataset, splitting it into training, validation, and testing sets, tokenizing Bangla text into individual words or characters, and converting tokens into numerical representations. The BiLSTM architecture can process sequential data bidirectionally, stacking multiple layers to capture abstract representations of the input text. The output layer is configured to produce multi-label predictions using a sigmoid activation function. The BiLSTM model is evaluated using evaluation metrics such as F1 Score, validation loss, Precision, Recall. The model is then analyzed for improvement across different emotion categories and error patterns. Hardware and software requirements for successful implementation include a multi-core processor Intel i7, a dedicated GPU, 16GB of RAM, 500GB of SSD storage, quality peripherals, and Python 3.7 or higher. Python libraries such as TensorFlow, Keras, Scikit-learn, Pandas, Numpy, NLTK, Matplotlib, and CoLab are essential for data manipulation and analysis. Project management and financial analysis are also essential for the success of the project. A multi-label emotion classification system for Bangla, recognize objectionable content with accuracy, and provide a tool for sentiment analysis and content control for Bangla-language communities and platforms. A detailed project timeline is developed, with milestones and deadlines for each phase. Resources are allocated, and mitigation strategies are developed to minimize risks. Regular communication channels and collaboration platforms are established for team members to share updates and address challenges. Stakeholder engagement is facilitated throughout the project lifecycle to ensure alignment with project goals and objectives. The project will involve a comprehensive risk assessment, mitigation techniques, and regular monitoring to protect project goals and stakeholder interests.

CHAPTER 4

Implementation

4.1 Overview

Emotion detection from Bangla text requires specific tools, technologies, and resources. Key implementation requirements include choosing a suitable programming language, such as Python, for Bi-LSTM and LSTM models, and selecting a deep learning framework like TensorFlow, PyTorch, or Keras. GPU acceleration is also essential for efficient model training, as CNNs and Bi-LSTM are computationally intensive. Access to GPUs can be achieved through cloud-based services. Data storage and management are crucial for efficient data management, such as setting up a database or file system for training, validation, and testing. The Bi-LSTM and LSTM model architecture are implemented, defining layers, filters, activation functions, pooling strategies, and other components. Optimization techniques like learning rate scheduling, regularization techniques, and batch normalization are explored to improve performance and generalization.

It is essential to plan the implementation process carefully, ensuring compatibility between the chosen framework, programming language, and available computational resources. Regular updates and maintenance are necessary as new research and advancements in Bi-LSTM and LSTM models and technologies emerge.

4.2 Train Model

We will prepare our Bangla text dataset, ensuring it is annotated with multiple emotion labels for each text sample. To make model training easier, divide the dataset among testing, validation, and training sets, tuning, and evaluation. Then tokenize Bangla text into individualwords or characters. Convert tokens into numerical representations, possibly using word embeddings like Word2Vec, FastText, or pre-trained language models. Design a BiLSTMarchitecture capable of processing sequential data bidirectionally. Stack multiple BiLSTMlayers to capture increasingly abstract representations of the input text. Optionally, incorporate additional layers such as convolutional layers or attention mechanisms to enhance model performance and interpretability. Configure the output layer to produce multi-label predictions, typically using a sigmoid activation function to enable independent classification of each emotion category. Evaluate the trained BiLSTM model on the held-out test set using evaluation metrics suitable for multi-label classification tasks. Common metrics include F1 Score, validation loss, Precision, Recall, and Mean Average Precision (mAP). Interpret model

performance across different emotion categories and analyze error patterns to identify areas for improvement. By incorporating a BiLSTM model into our methodology, we will leverage its ability to capture long-range dependencies and contextual information in Bangla text, thereby enhancing the accuracy and effectiveness of multi-label emotion classification tasks.

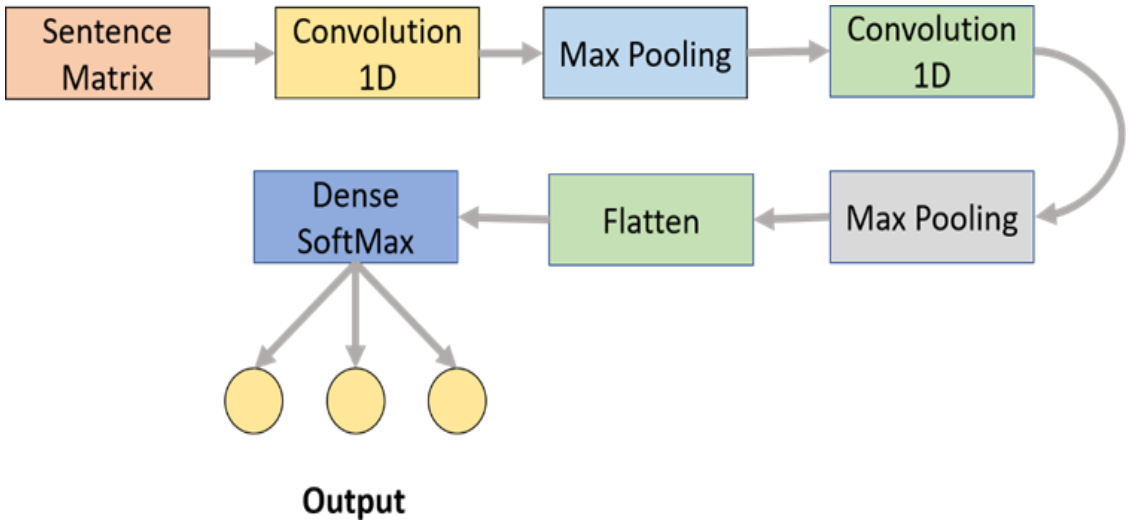

Figure 4.2.1: CNN speech classifier

In Figure 4.2.1 the CNN speech classifier architecture. A multilayer perceptron modelcalled a CNN (CNN) makes use of convolution processes. The CNN technique was created first for text image recognition and it has been demonstrated to produce very accurate results. An overview of the architecture of the hate speech classification system is shown in Figure 3.4.3 The one-dimensional convolutional layer of the Bi-LSTM and LSTM architecture uses the sentences as input. For the first convolutional layer, I specified 3 filters, each witha batch size of 128. The dimensionality of the convolution output is decreased by a max pooling layer of size 3 without sacrificing feature quality. the softmax dense layer, then asecond 3-layer max pooling layer. The second max pooling layer is followed by the flattened layer, which is then followed by a fully linked layer made up of softmax activationfunction layers.

The results from the flattened pooling layer are input into the fully connected layer, which outputs the probability for each class depending on the input to determine which attributesare most correlated with a specific class. The dense procedure adds flattened data with theReLU activation function into the fully linked layer. To transfer the flattened values into outputs that fall between 0 and 1, the sigmoid activation function is used as a nonlinear function.

$$S(x) = \frac{1}{1+ex} \quad \text{.....(1)}$$

5

A significant drawback of the softmax activation function as stated in Eq. (1) is that when

multiple weight values approach the extreme values of 0,1, and 3, the function gradient tends to attain the value 0. Due to the neuron's inability to produce large changes as a result,the backpropagation process will be less efficient. The softmax function, which only exists between 0 and 1, can nonetheless forecast the likelihood between two classes despite this drawback.

The conventional neural network model is inadequate for handling sequence learning dueto its inability to capture the association between the beginning and end of a series. Recurrent Neural Networks (RNN) are a type of model used for sequence learning. They consist of nodes that are connected between hidden layers and have the ability to dynamically learn sequence features. Figure 3.4.4 illustrates the application of Recurrent Neural Network (RNN) in Chinese text emotion analysis. The input text in the figure statesthat the hotel's environment is favorable. Following the process of word segmentation, it transforms into. Every word is transformed into its respective word vector (w1, w2, w3, w4), which is subsequently fed into the RNN in a sequential manner as the matching wordvector (wt, wt, wt, wt).

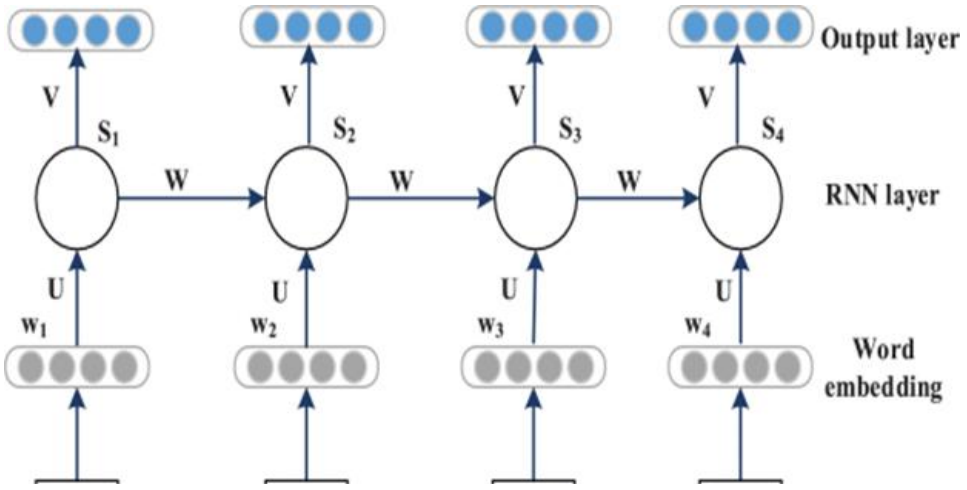

Figure 4.2.2: The RNN architecture for the emotion analysis model.

The conventional recurrent neural network model is unable to capture semantic connections that span large distances, despite its ability to transmit semantic information between words. During the parameter training procedure, the gradient diminishes progressively until it vanishes. Consequently, the extent of sequential data is restricted. LSTM addresses the issue of gradient vanishing by incorporating an Input gate (i), Outputgate (o), Forget gate (f), and Memory cell. The structure of the LSTM network, as described in reference, is depicted in Figure 4.2.3.

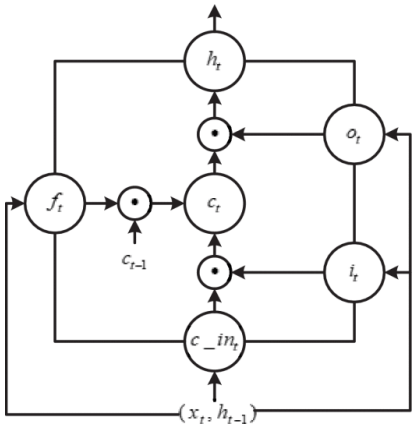

Figure 4.2.3: Diagram of LSTM structure for emotion.

This research proposes an emotion analysis approach for comments based on BiLSTM to address the limitations of current comment emotion analysis methods. In both the traditional recurrent neural network model and LSTM model, information is only able to propagate in a forward direction. As a consequence, the state at time t is solelyinfluenced by the information preceding time t. To ensure that all moments have contextualinformation, a Bi-LSTM model is employed. This model combines bidirectional recurrent neural network (BiRNN) models with LSTM units to effectively capture the context information. The Bi-LSTM model assigns equal importance to all inputs. The emotional polarity of thetext in emotion analysis is mostly determined by the presence of words that convey emotional information. This paper implements emotion reinforcement of emotion word vectors. emotion analysis tasks are considered text classification tasks, and distributed word vectors do not include the individual contributions of distinct words to the categorization process. Section A of Research Methods involves the construction of weighted word vectors that incorporate both emotional information and categorization contribution. Initially, the weighted word vectors serve as the inputs for the Bi-LSTM model, and the resulting outputs of the Bi-LSTM model are utilized as the representationsof the comment texts. Next, the comment text vectors are sent into the feedforward neuralnetwork classifier. Ultimately, the emotional inclination of the comments is acquired. Theactivation function utilized in feedforward neural networks is the Rectified Linear Unit (ReLU) function. To mitigate the issue of over-fitting during the training phase, the dropoutmechanism was included, with a dropout discard rate of 0.5.

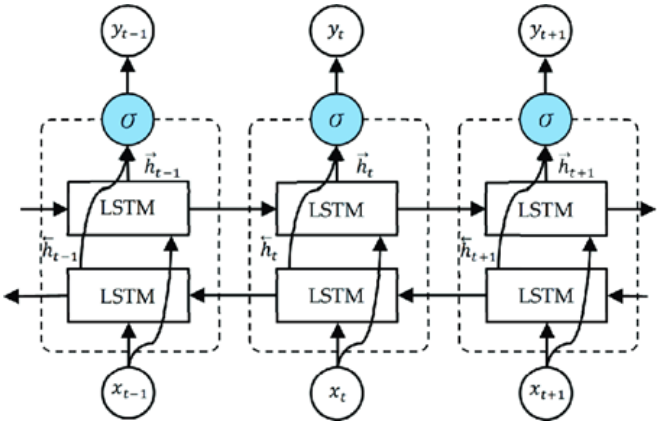

Figure 4.2.4: Bi-LSTM architecture for emotion

The schematic diagram illustrating the emotion technique described in this paper can be found in Figure 4.2.4. The procedure for extracting features from comments text is shown in the left subgraph. The procedure of figuring out the psychological polarity of the remark text is discussed in the right subgraph. The number network nodes in the Bi-LSTM's hidden layer is indicated by NodeNum.

4.3 System Testing

Bangla emotion detection and classification using CNNs and Bi-LSTM is an important research area with various implications for social media platform Facebook, online communities, and content moderation. In this section, I discuss the key findings, implications, limitations, and future directions of the study. Discuss the performance of the CNN, Bi-LSTM, and LSTM models in detecting and classifying emotion. Present the achieved accuracy, loss, val_Acc, val_Loss, and metrics. Compare the performance with baseline models or existing approaches. Highlight any significant improvements achieved by the Bi-LSTM model, showcasing its effectiveness in addressingthe challenges of emotion detection. Address the issue of model interpretability in the context of emotion detection. CNN models are often considered black-box models, makingit challenging to understand the decision-making process. Discuss the generalizationperformance of the CNN Base model. Assess its ability to handle different contexts, languages, and social media platforms. Address the potential challenges of adapting the model to new or unseen data and propose strategies to improve the generalization and robustness of the model. Acknowledge the limitations and challenges encountered during the research. Discuss factors of data quality, class imbalance, noise in social media data, and limitations of the CNN Base architecture itself. Highlight the

impact of these limitations on the model's performance and suggest areas for future improvement.

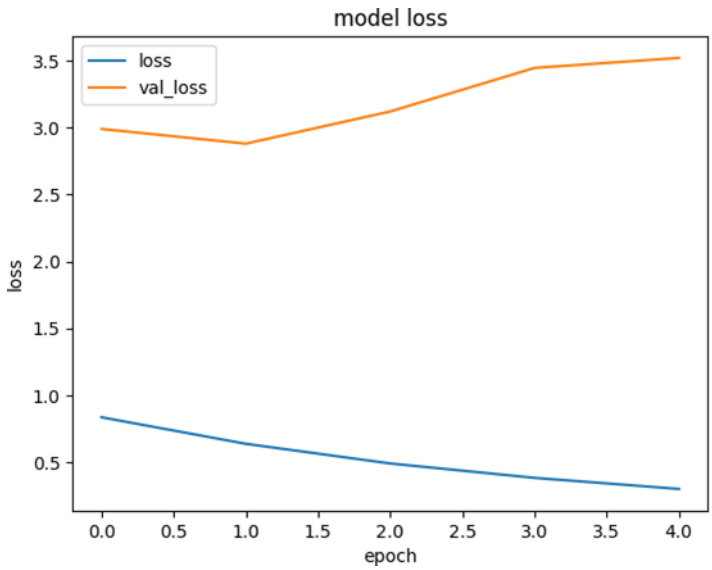

Figure 4.3.1: Loss and val_loss comparison graph

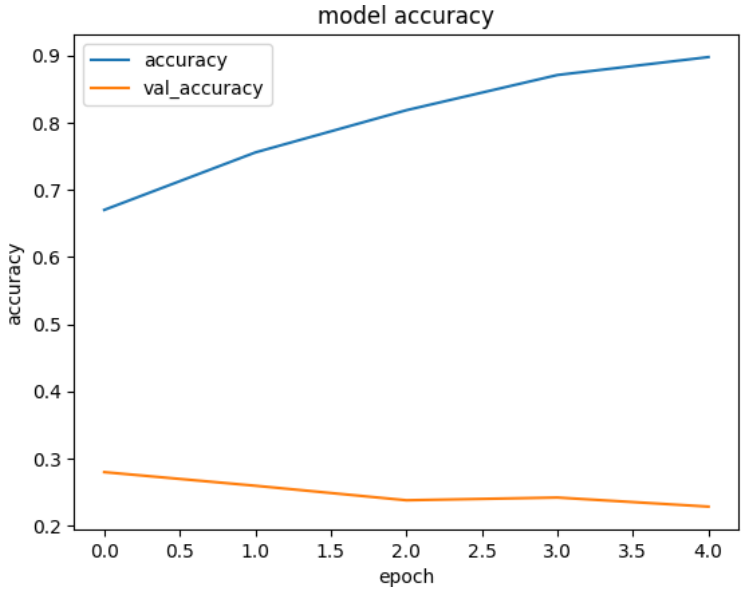

Figure 4.3.2: Accuracy and val_accuracy comparison graph

Figure 4.3.2 shows a comparison line between the accuracy and validation accuracy. Consider the scalability and real-time deployment of the CNN Base model for sensitive speech detection. Discuss the computational requirements, inference speed, and potential challenges when implementing the model in a production environment. Address scalability concerns and propose approaches to handle large-scale data and real-time processing. Discuss the role of user feedback and human-in-the-loop approaches in enhancing the performance of sensitive speech detection systems. Consider the importance of continuousfeedback loops,

user reporting mechanisms, and human moderation to improve the accuracy and adaptability of the system. Highlight the need for ongoing collaboration between AI models and human moderators. Discuss the practical applications of sensitivespeech detection and classification using the CNN Base model. Highlight the potential impact on content moderation, hate speech mitigation, online safety, and fostering healthy online communities. Address the implications for social media platforms, policymakers, and society as a whole.

4.4 Summary

Emotion detection from Bangla text requires specific tools, technologies, and resources. Key implementation requirements include choosing a suitable programming language, deep learning framework, GPU acceleration, and data storage and management. The Bi-LSTM and LSTM model architecture are implemented, defining layers, filters, activation functions, pooling strategies, and other components. Optimization techniques like learning rate scheduling, regularization techniques, and batch normalization are explored to improve performance and generalization. The Bangla text dataset is prepared, tokenized, and converted into numerical representations. A BiLSTM architecture is designed to process sequential data bidirectionally, stacking multiple layers to capture abstract representations of the input text. The output layer is configured to produce multi-label predictions using a sigmoid activation function. The CNN speech classifier architecture uses a multilayer perceptron model called a CNN (CNN) to capture long-range dependencies and contextual information in Bangla text. The one-dimensional convolutional layer of the Bi-LSTM and LSTM architecture uses sentences as input, with a max pooling layer and flattened layer. The fully connected layer outputs the probability for each class depending on the input, determining which attributes are most correlated with a specific class. The conventional neural network model is insufficient for sequence learning due to its inability to capture semantic connections between words. Recurrent Neural Networks (RNN) are used for sequence learning, but they struggle to capture semantic connections that span large distances. The LSTM model addresses this issue by incorporating input, output, forget, and memory gates. This research proposes an emotion analysis approach for comments based on Bi-LSTM to address the limitations of current methods. The Bi-LSTM model combines Bi-RNN models with LSTM units to capture context information. The emotional polarity of the text is determined by the presence of words conveying emotional information. The paper implements emotion reinforcement of emotion word vectors, which are used as inputs for the Bi-LSTM model and used as outputs in the feedforward neural network classifier. The activation function is the ReLU function.

CHAPTER 5

Result and Analysis

5.1 Overview

The experimental setup is crucial in emotion detection and classification using Bi-LSTM. It involves defining the research objective, preprocessing the dataset, selecting an appropriate architecture for sensitive speech detection and classification, training the model, and evaluating the performance. The experiment was conducted on a dataset of 3698 Facebook Bangla speech comments, with six emotion class labels (anger, happiness, surprise, fear, sadness, confused). The dataset was split into training and validation sets, and the experiment was conducted for 1, 2, 3, 4, and 5 epochs. The results showed that the 5th epoch had the highest accuracy and validation accuracy results, but the validation loss increased after 2 epochs. By the 3rd epoch, both training and validation loss reached the lowest percentage. At 89.76% training accuracy, 29.88% training loss, 35.22% validation loss, and 22.84% validation accuracy, the model produced the best results. The dataset's f1-score, precision, and recall were displayed in the classification report based on class labels.

5.2 Experimental Analysis

In this paper, we use a dataset of a total of 3698 Facebook Bangla speech comments and collected each data from real-time Facebook comments. The CNN technique for classifying speech in Indonesian was developed using the Python Keras package. The dataset has 6 emotion class labels (anger, happiness, surprise, fear, sadness, confused).

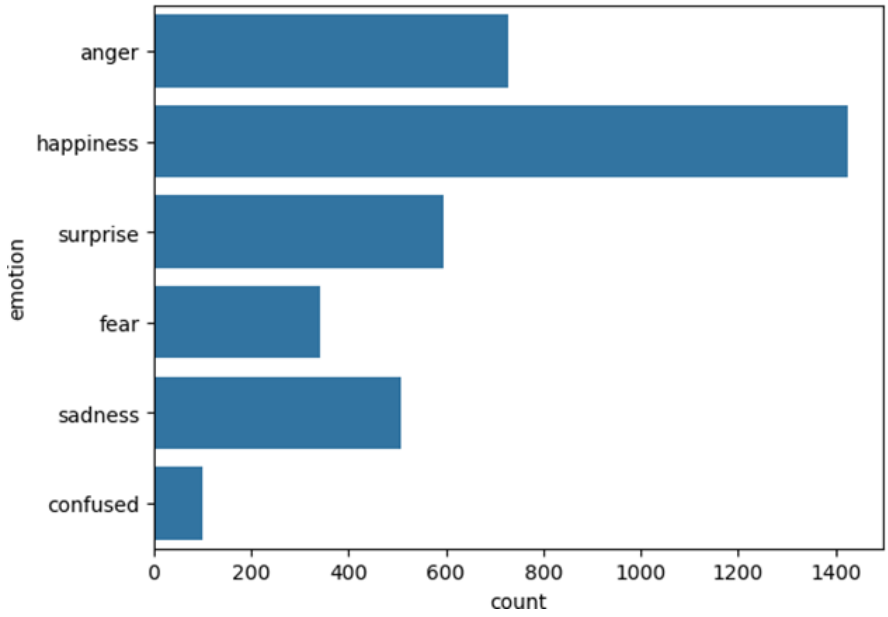

Figure 5.2.1: Emotion class labels of our dataset

The dataset used for training was split into 80% (2958 comments) for the training set and 20% (740 comments) for the validation set. The experiment was conducted for 1, 2, 3, 4, and 5 epochs. From the experimental results shown in figure 5.2.2.

```
x_val=testing_padded
y_val=test_labels

history=model.fit(x=padded,y=train_labels,epochs=5,batch_size=64,validation_data=(x_val,y_val))

Epoch 1/5
47/47 [==============================] - 142s 3s/step - loss: 0.8347 - accuracy: 0.6700 - val_loss: 2.9910 - val_accuracy: 0.2797
Epoch 2/5
47/47 [==============================] - 141s 3s/step - loss: 0.6362 - accuracy: 0.7556 - val_loss: 2.8811 - val_accuracy: 0.2595
Epoch 3/5
47/47 [==============================] - 137s 3s/step - loss: 0.4892 - accuracy: 0.8185 - val_loss: 3.1219 - val_accuracy: 0.2378
Epoch 4/5
47/47 [==============================] - 141s 3s/step - loss: 0.3815 - accuracy: 0.8709 - val_loss: 3.4476 - val_accuracy: 0.2419
Epoch 5/5
47/47 [==============================] - 141s 3s/step - loss: 0.2988 - accuracy: 0.8976 - val_loss: 3.5223 - val_accuracy: 0.2284
```

Figure 5.2.2: Prediction results of our dataset

The fifth epoch had the best accuracy and the reliability of validation result, as seen in Figure 5.2.2. But after two epochs, the validation loss began to steadily rise while the training loss continued to decline, leading to overfitting of the model with more training. By the third epoch, the percentage of both training loss and validation loss was at its lowest. The best result was obtained with an accuracy in training of 89.76%, a training loss of 29.88%, a test loss of 35.22%, and an accuracy for validation of 22.84% since the model checkpoint stopped saving after 5 epochs.

```
from sklearn.metrics import classification_report

predictions = model.predict(testing_padded)

predicted_labels = (predictions > 0.5).astype(int)

print(classification_report(test_labels, predicted_labels, target_names=class_names))
```

```
24/24 [==============================] - 14s 580ms/step
              precision    recall  f1-score   support

       Anger       0.20      0.63      0.30       147
   Happiness       0.34      0.66      0.45       260
    Surprise       0.18      0.63      0.28       124
        Fear       0.11      0.24      0.15        80
     Sadness       0.14      0.28      0.19       105
    Confused       0.06      0.21      0.09        24

   micro avg       0.21      0.54      0.30       740
   macro avg       0.17      0.44      0.24       740
weighted avg       0.22      0.54      0.31       740
 samples avg       0.22      0.54      0.31       740
```

Figure 5.2.3: Classification report of our dataset

5.3 Performance Analysis

In this section, I discuss the performance, implications, and future directions of the study. Discuss the performance of the CNN, Bi-LSTM, and LSTM models in detecting and classifying emotion. Present the achieved accuracy, loss, val_Acc, val_Loss, and metrics. Compare the performance with baseline models or existing approaches. Highlight any significant improvements achieved by the Bi-LSTM model, showcasing its effectiveness in addressing the challenges of emotion detection. Address the issue of model interpretability in the context of sensitive emotion. CNN models are often considered black-box models, making it challenging to understand the decision-making process. Discuss the generalization performance of the CNN Base model. Assess its ability to handle different contexts, languages, and social media platforms. Address the potential challenges of adapting the model to new or unseen data and propose strategies to improve the generalization and robustness of the model. Acknowledge the limitations and challenges encountered during the research. Discuss factors such as data quality, class imbalance, noise in social media data, and limitations of the CNN Base architecture itself. Highlight the impact of these limitations on the model's performance and suggest areas for future improvement. The experimental results are shown in Figure 5.2.2.

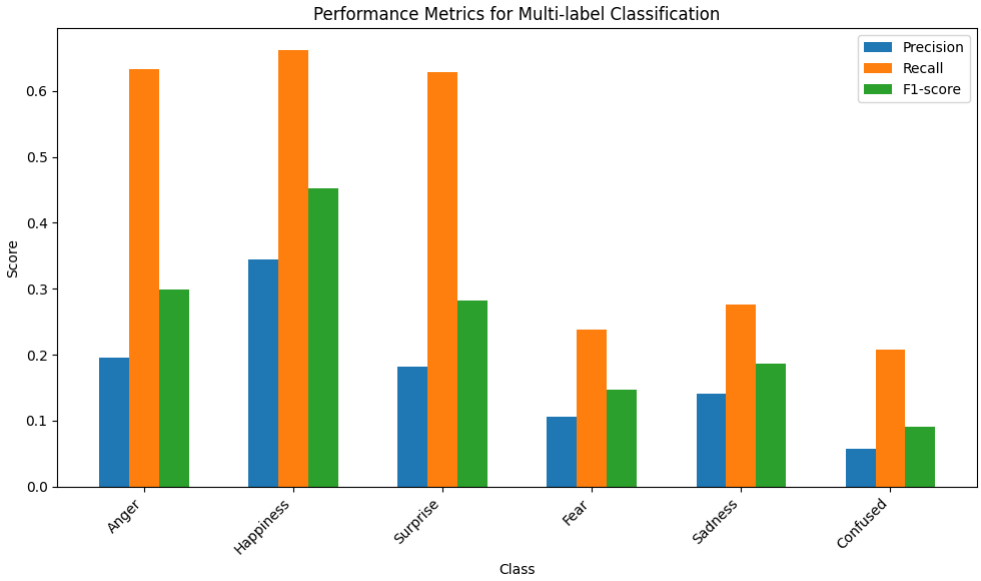

Figure 5.3.1: Performance Metrics for Multi-label Classification

Figure 5.3.1 shows the classification report of our dataset. Here we show precision, recall and f1-score of our dataset according to our class labels.

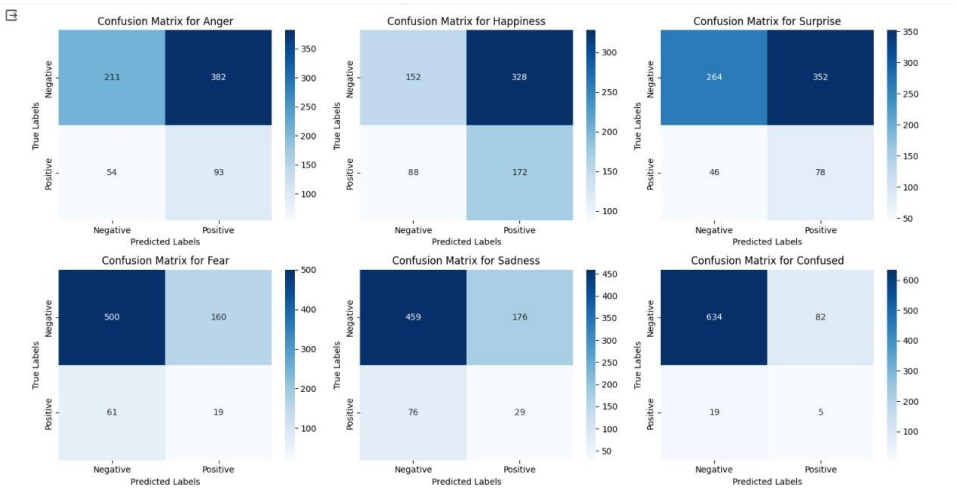

Figure 5.3.2: Confusion Matrix

5.4 Summary

Developing The experimental setup is crucial in emotion detection and classification using Bi-LSTM. It involves defining the research objective, preprocessing the dataset, selecting an appropriate architecture for sensitive speech detection and classification, training the model, and evaluating the performance. In this paper, a dataset of 3698 Facebook Bangla speech

comments was used to train the method for speech classification in the Indonesian language using the dataset had six emotion class labels (anger, happiness, surprise, fear, sadness, confused). The experiment was conducted for 1, 2, 3, 4, and 5 epochs. The results showed that the 5th epoch had the highest accuracy and validation accuracy results, but the validation loss increased after 2 epochs, while the training loss kept decreasing, causing the model to overfit when trained longer. Both training and validation loss reached the lowest percentage by the 3rd epoch. By the 3rd epoch, both training and validation loss reached the lowest percentage. The model achieved the most optimal results at a training accuracy of 89.76%, a training loss of 29.88%, a validation loss of 35.22%, and a validation accuracy of 22.84%. The performance of the CNN, Bi-LSTM, and LSTM models in detecting and classifying emotion was discussed. The Bi-LSTM model showed significant improvements in addressing the challenges of emotion detection, while the CNN Base model was assessed for its ability to handle different contexts, languages, and social media platforms. The limitations and challenges encountered during the research were acknowledged, such as data quality, class imbalance, noise in social media data, and limitations of the CNN Base architecture itself. In conclusion, the experimental setup plays a critical role in emotion detection and classification using Bi-LSTM. The study highlights the importance of understanding the decision-making process and adapting the model to new or unseen data to improve its generalization and robustness.

CHAPTER 6

Impact on Society, Environment and Sustainability

6.1 Impact on Life

The impact of multi-label emotions Bangla text classification on society can be profound and multifaceted, influencing various aspects of human interaction, communication, and well-being. Here are some ways in which our technology can impact society:

Enhanced Communication Understanding: Multi-label emotions Bangla text classification can help individuals and organizations better understand the emotional content of Bangla language text, including social media posts, customer feedback, news articles, and literary works. This understanding can facilitate more empathetic and effective communication strategies, leading to improved relationships and interactions.

Mental Health Assessment and Support: Emotions play a crucial role in mental health, and accurate identification of emotional states in Bangla text can aid in early detection and intervention for individuals experiencing mental health issues. Multi-label emotions classification can be used in mental health applications to analyze online conversations, social media posts, and personal messages, providing insights into individuals' emotional well-being and directing them to appropriate support resources.

Cultural Preservation and Promotion: The Bangla language and culture are rich in emotional expression, poetry, literature, and art. Multi-label emotions Bangla text classification can contribute to the preservation and promotion of Bangla cultural heritageby analyzing and understanding emotional nuances in Bangla language text, including classical literature, folk songs, and contemporary media.

Customer Feedback Analysis: Businesses and organizations can use multi-label emotionsin Bangla text classification to analyze customer feedback, reviews, and comments in Bangla language platforms. By understanding customers' emotional responses, organizations can identify areas for improvement, tailor their products or services to bettermeet customer needs, and enhance customer satisfaction and loyalty.

Social Media Monitoring and Crisis Response: Social media platforms are often used as channels for expressing emotions and opinions on various social and political issues. Multi-label emotions Bangla text classification can enable real-time monitoring of social media conversations in Bangla, helping governments, NGOs, and emergency responders assess

public sentiment, detect emerging crises, and respond effectively to events such as natural disasters, political unrest, or public health emergencies.

Educational Applications: In educational settings, multi-label emotions Bangla text classification can be used to analyze students' written assignments, essays, or forum discussions, providing teachers and educators with insights into students' emotional engagement, motivation, and learning experiences. This information can informpersonalized teaching strategies and interventions to support students' socio-emotional development and academic success.

Ethical Considerations: It's essential to consider ethical implications associated with multi-label emotions in Bangla text classification, including privacy, bias, fairness, and transparency. Ensuring that classification models are deployed responsibly and ethically, with proper consent, data protection measures, and safeguards against algorithmic biases, is crucial to mitigating potential negative impacts on society.

6.2 Impact on Society & Environment

The impact on the environment for multi-label emotions Bangla text classification primarily stems from the computational resources required to develop, train, and deploy classification models. Here's how this technology can affect the environment:

Energy Consumption: Training deep learning models for multi-label emotions Bangla text classification often requires significant computational resources, including high-performance GPUs or TPUs. These hardware components consume substantial amounts of electrical power, contributing to increased energy consumption and carbon emissions, especially in data centers where large-scale model training occurs.

Carbon Footprint: The energy consumption associated with training and running deep learning models for multi-label emotions Bangla text classification contributes to the carbon footprint of AI systems. The electricity used to power data centers and computational infrastructure is often generated from fossil fuels, leading to greenhouse gas emissions that contribute to climate change and environmental degradation.

E-waste Generation: The rapid pace of technological advancement in AI hardware leads to frequent upgrades and replacements of computational devices GPUs, CPUs, and servers. This results in the generation of electronic waste as obsolete or outdated hardware is discarded, contributing to environmental pollution and resource depletion if not properly

managed through recycling and responsible disposal practices.

Data Center Infrastructure: The infrastructure required to support large-scale model training and deployment, including data centers, cooling systems, and networking equipment, consumes resources such as land, water, and materials. The construction and operation of data centers can have environmental impacts such as habitat disruption, water consumption, and air pollution, particularly in areas with high concentrations of data center facilities.

Efficient Computing Practices: Adopting efficient computing practices, such as optimizing algorithms, reducing model complexity, and implementing hardware acceleration techniques, can help mitigate the environmental impact of multi-label emotions Bangla text classification. By improving computational efficiency and reducing energy consumption, these practices can lower the carbon footprint associated with AI model development and deployment.

Cloud Computing Solutions: Leveraging cloud computing platforms for model training and inference can offer environmental benefits compared to on-premises infrastructure, as cloud providers often use energy-efficient data centers and renewable energy sources to power their operations. By utilizing cloud-based AI services, organizations can reduce their reliance on local hardware and minimize their environmental footprint.

Sustainable AI Development: Promoting sustainable AI development involves considering environmental factors alongside technical and economic considerations when designing and deploying AI systems. This includes incorporating environmental impact assessments into AI project planning, adopting energy-efficient hardware and software solutions, and advocating for policies and regulations that incentivize sustainable AI practices.

6.3 Ethical Aspects

Ethical considerations are crucial in the development, deployment, and use of multi-label emotions in Bangla text classification systems. Here are some key ethical aspects to consider:

Privacy and Data Protection: Ensure that user data, including Bangla text samples annotated with emotion labels, is collected, stored, and processed in compliance with privacy regulations and data protection laws. Obtain informed consent from users before collecting and using their data, and implement measures to safeguard sensitive information. **Bias and Fairness:** Address potential biases in the dataset used for training multi-label emotions classification models, such as underrepresentation of certain demographic groups or

overrepresentation of specific emotions. Employ techniques like data augmentation, bias detection, and fairness-aware algorithms to mitigate bias and ensure fair treatment for all individuals.

Transparency and Accountability: Promote transparency in the development and deployment of multi-label emotions classification systems by documenting model architectures, training data sources, and evaluation methodologies. Provide users with clear explanations of how their data is used and how classification decisions are made, enabling them to understand and trust the system's behavior.

Mitigation of Harm: Take proactive measures to mitigate potential harms arising from the use of multi-label emotion classification systems, such as misclassification errors, privacy breaches, or unintended consequences. Implement safeguards such as error monitoring, user feedback mechanisms, and regular audits to identify and address issues promptly.

Cultural Sensitivity: Recognize and respect cultural differences in emotional expression and interpretation, particularly in the context of Bangla language and culture. Avoid imposing Western-centric or universalistic perspectives on emotions and instead incorporate culturally relevant frameworks and linguistic nuances into classification models and algorithms.

Ethical Review and Oversight: Subject multi-label emotions classification projects to ethical review and oversight by independent bodies or institutional review boards (IRBs). Conduct ethical impact assessments to identify potential ethical risks and ensure that appropriate safeguards are in place to protect the interests and rights of all stakeholders involved.

6.4 Sustainability Plan

Creating a Multi-Label Emotions Sustainability Plan In order to reduce negative effects and improve long-term viability, Bangla text classification entails incorporating environmental, social, and economic issues throughout the project's lifecycle. Here's a comprehensive plan:

Energy-efficient Computing: Optimize algorithms, model architectures, and hardware configurations to reduce energy consumption during model training and inference. Utilize energy-efficient computing resources, such as GPUs with low power consumption or cloud computing services powered by renewable energy sources.

Data Efficiency: Implement data-efficient approaches to reduce the amount of data required for model training and validation. Use techniques like transfer learning, data augmentation,

and active learning to leverage existing datasets effectively and minimize the need for additional data collection.

Resource Recycling and Reuse: Minimize waste generation by recycling and reusing computational resources, GPUs, CPUs, and servers, whenever possible. Adopt circular economy principles by refurbishing outdated hardware, repurposing idle resources, and extending the lifespan of computing equipment through upgrades and maintenance.

Ethical Data Practices: Adhere to ethical data practices to ensure the responsible collection, usage, and sharing of Bangla text data for multi-label emotions classification. Obtain informed consent from data subjects, anonymize sensitive information, and implement data protection measures to safeguard user privacy and rights.

Community Engagement: Engage with the Bangla-speaking community, including language experts, researchers, and end-users, to solicit feedback, gather insights, and foster collaboration in multi-label emotion classification projects. Organize workshops, seminars, and community forums to promote dialogue, knowledge exchange, and capacity building.

Open Access and Collaboration: Foster an open and collaborative research culture by sharing datasets, code repositories, and research findings with the wider community. Embrace open access principles to promote transparency, reproducibility, and knowledge dissemination in multi-label emotions Bangla text classification research.

Skills Development: Invest in skills development and capacity-building initiatives to empower researchers, practitioners, and students with the knowledge and expertise needed to contribute to multi-label emotions Bangla text classification projects. Offer training programs, mentorship opportunities, and educational resources to support professional development and career advancement.

Continuous Improvement: Regularly monitor and evaluate the sustainability performance of multi-label emotions Bangla text classification projects, identify areas for improvement, and implement corrective actions to enhance sustainability outcomes. Embrace a culture of continuous improvement and innovation to adapt to evolving environmental, social, and economic challenges.

6.5 Summary

The Multi-label emotions Bangla text classification can significantly impact society by improving communication understanding, aiding in mental health assessment, preserving

cultural heritage, and analyzing customer feedback. It can help individuals and organizations understand the emotional content of Bangla language text, leading to better relationships and interactions. Emotions play a crucial role in mental health, and accurate identification of emotional states in Bangla text can aid in early detection and intervention. It can also help businesses analyze customer feedback, identifying areas for improvement and tailoring products to meet customer needs. It can also enable real-time monitoring of social media conversations, aiding in crisis response. In education, it can analyze students' writings, providing insights into emotional engagement and motivation. However, ethical considerations, such as privacy, bias, fairness, and transparency, must be considered. Multi-label emotions in Bangla text classification has a significant environmental impact due to the computational resources needed for its development, training, and deployment. This includes energy consumption, carbon footprint, e-waste generation, and data center infrastructure. To mitigate these impacts, efficient computing practices, such as optimizing algorithms and reducing model complexity, can be adopted. Cloud computing solutions can also offer environmental benefits by reducing reliance on local hardware. Sustainable AI development involves considering environmental factors alongside technical and economic considerations when designing and deploying AI systems. This includes incorporating environmental impact assessments into project planning, adopting energy-efficient hardware and software solutions, and advocating for policies and regulations that incentivize sustainable AI practices. Ethical considerations are crucial in the development, deployment, and use of multi-label emotions in Bangla text classification systems. Key ethical aspects include privacy and data protection, bias and fairness, transparency and accountability, mitigation of harm, cultural sensitivity, and ethical review and oversight.

Developing a sustainability plan for multi-label emotions Bangla text classification involves integrating environmental, social, and economic considerations into the project's lifecycle to minimize negative impacts and promote long-term viability. This comprehensive plan includes optimizing algorithms, model architectures, and hardware configurations, ensuring data security, and promoting sustainable practices.

CHAPTER 7

Conclusion and Future Work

7.1 Conclusions

The multi-label emotions Bangla text classification methodology involves a systematic approach to develop models capable of accurately identifying and categorizing multiple emotions expressed in Bengali text. The methodology involves collecting diverse Bengali text datasets containing emotional expressions across various domains, preprocessing the raw text data, annotating the datasets with multiple emotion labels, designing and implementing machine learning and deep learning architectures, and experimenting with different model architectures. The developed models are trained using appropriate training algorithms, loss functions, optimizers, and regularization techniques. The performance of the trained models is evaluated using standard evaluation metrics such as accuracy, precision, recall, F1-score, and Hamming loss. Ethical considerations such as data privacy, consent, and bias mitigation are also considered throughout the system's lifecycle.

Non-functional requirements include scalability, performance, robustness, and usability. The system should be capable of processing and classifying Bengali text data in real-time or near-real-time settings, and designed to be robust to noise, variations in language usage, and domain-specific contexts. The system will also provide user-friendly interfaces and documentation for researchers and practitioners to interact with effectively.

The proposed methodology for multi-label emotions Bangla text classification involves integrating a Bi-LSTM model. The process involves preparing a Bangla text dataset with multiple emotion labels, splitting it into training, validation, and testing sets, tokenizing the text into individual words or characters, and designing a BiLSTM architecture capable of processing sequential data bi- directionally. The output layer is configured to produce multi-label predictions using a sigmoid activation function. The trained BiLSTM model is evaluated using metrics such as F1 Score, validation loss, Precision, Recall, and Mean Average Precision. The data collection involves storing Bangla emotion-related comments in an Excel file and collaborating with the supervisor to assign appropriate emotion labels based on the content. The annotation process is documented for future reference. The process of creating an annotated dataset for multi-label emotions in Bangla text classification involves several steps. First, the dataset is reviewed with the supervisor to ensure consistency and accuracy in emotion labeling. Once the annotation process is complete, the dataset is

finalized and stored in an Excel file for future analysis and modeling tasks. Project management involves defining the scope, objectives, and deliverables, identifying key team members, and developing a detailed project timeline. Resources are allocated for data acquisition, model training, and experimentation, and potential risks and uncertainties are identified. Mitigation strategies and contingency plans are developed to minimize their impact on project outcomes. Regular communication channels and collaboration platforms are established to ensure alignment with project goals and objectives. Financial analysis is conducted to estimate costs, assess the potential return on investment, and consider the sustainability and long-term viability of the project

7.2 Further Suggested Works

The Future considerations for advancing multi-label emotions Bangla text classification involve addressing several key areas for improvement and expansion. As the field of NLP and AI continues to evolve, focusing on these areas can significantly enhance the effectiveness and applicability of emotion classification systems in the Bangla language. Here are several future directions to consider:

Enhanced Dataset Quality and Size: Building larger and more diverse corpora that include a wider range of text sources such as books, social media, news articles, and spoken language transcriptions to train more robust models. Ensuring high-quality, nuanced annotations of emotions by employing expert linguists and psychologists, particularly for complex emotions that are harder to identify.

Advanced Modeling Techniques: Utilizing deep learning innovations, such as Transformer models, which have demonstrated notable performance in a range of NLP tasks. Developing or fine-tuning existing pre-trained models specifically for Bangla can capture the nuances of the language better. Exploring cross-lingual and multilingual models that can understand and process multiple languages, including Bangla. This approach can help leverage knowledge learned from data-rich languages to improve performance in data- sparse languages like Bangla.

Improved Language Understanding: Improving the model's ability to understand context and semantics in text, which is crucial for accurately identifying emotions, especially where they are implied rather than explicitly expressed. Developing methods to better detect sarcasm and irony, which can significantly change the emotional tone of a text.

Real-World Applications: Expanding the integration of emotion detection systems into real-world applications such as customer service bots, and mental health assessments Utilizing emotion classification to enhance user experience by personalizing content and responses according to the emotional tone of user inputs.

Technological Advancements: Utilizing advancements in processing power and storage to handle increasingly complex models and larger datasets without compromising performance. Deploying emotion detection models on edge devices, enabling faster and privacy-preserving real-time processing of data.

7.3 Limitations

By acknowledging these limitations, our research provides a clear understanding of the constraints and areas for future work, such as expanding the dataset, exploring ensemble methods, and testing on multiple platforms to enhance model robustness and generalizability.

Dataset Quantity: The availability of large, labeled datasets for Bangla text is limited. Training effective deep learning models requires substantial amounts of data, and the lack of sufficient labeled examples can hinder model performance. Besides quantity, the quality of available data is crucial. Incomplete or noisy data can negatively impact the training process and lead to poor generalization.

Classification of a Single Model: Relying on a single type of model, such as CNN, RNN, or BiLSTM, may not capture all the nuances of emotion in Bangla text. Different models excel at different aspects CNNs are good for feature extraction, RNNs for sequence modeling, and a single model might not be sufficient. Not using ensemble methods or hybrid models can limit performance. Combining multiple models often leads to better accuracy and robustness, as it leverages the strengths of different architectures.

Just One Platform: Developing and testing models on data from a single platform only Facebook can introduce platform-specific biases. The language, tone, and type of content can vary significantly across different platforms. Models trained on data from one platform might not generalize well to others. This limitation reduces the applicability and effectiveness of the models in broader contexts or on new platforms.

REFERENCES

[1] Tripto, N. I., & Ali, M. E. (2018, September). Detecting multilabel sentiment and emotions from bangla youtube comments. In *2018 international conference on Bangla speech and language processing (ICBSLP)* (pp. 1-6). IEEE.

[2] .Sourav, M. S. U., Wang, H., Mahmud, M. S., & Zheng, H. (2022). Transformer-based text classification on unified Bangla multi-class emotion corpus. *arXiv preprint arXiv:2210.06405*..

[3] Kabir, A., Roy, A., & Taheri, Z. (2023, December). BEmoLexBERT: A Hybrid Model for Multilabel Textual Emotion Classification in Bangla by Combining Transformers with Lexicon Features. In Proceedings of the First Workshop on Bangla Language Processing (BLP-2023) (pp. 56-61).

[4] Kabir, A., Roy, A., & Taheri, Z. (2023, December). BEmoLexBERT: A Hybrid Model for Multilabel Textual Emotion Classification in Bangla by Combining Transformers with Lexicon Features. In Proceedings of the First Workshop on Bangla Language Processing (BLP-2023) (pp. 56-61).

[5] Das, A., Hoque, M. M., Sharif, O., Dewan, M. A. A., & Siddique, N. (2023). Temox: Classification of textual emotion using ensemble of transformers. *IEEE Access*.

[6] Iqbal, M. A., Das, A., Sharif, O., Hoque, M. M., & Sarker, I. H. (2022). Bemoc: A corpus for identifying emotion in bengali texts. SN Computer Science, 3(2), 135.

[7] Iqbal, M. A., Sharif, O., Hoque, M. M., & Sarker, I. H. (2021). Word embedding based textual semantic similarity measure in Bengali. *Procedia Computer Science*, *193*, 92-101.

[8] Das, A., Sharif, O., Hoque, M. M., & Sarker, I. H. (2021). Emotion classification in a resource constrained language using transformer-based approach. arXiv preprint arXiv:2104.08613.

[9] Parvin, T., & Hoque, M. M. (2021). An ensemble technique to classify multi-class textual emotion. *Procedia Computer Science*, *193*, 72-81.

[10] Banshal, S. K., Das, S., Shammi, S. A., & Chakraborty, N. R. (2023). MONOVAB: An Annotated Corpus for Bangla Multi-label Emotion Detection. arXiv preprint arXiv:2309.15670.

[11] Islam, K. I., Yuvraz, T., Islam, M. S., & Hassan, E. (2022, November). Emonoba: A dataset for analyzing fine-grained emotions on noisy bangla texts. In Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (Volume 2: Short Papers) (pp. 128-134).

[12] Sinha, S., Saxena, K., & Joshi, N. (2021). Detecting Multi-label emotions from code-mixed Facebook Status Updates. Indian Journal of Science and Technology, 14(31), 2542-2549.

[13] Hossain, P. S., Chakrabarty, A., Kim, K., & Piran, M. J. (2022). Multi-label extreme learning machine (MLELMs) for Bangla regional speech recognition. Applied Sciences, 12(11), 5463.

[14] Surolia, A., Mehta, S., & Kumaraguru, P. (2023). Understanding Emotions: A PoliEMO Datasetand Multi-label Classification in Indian Elections.

[15] Sultana, S., Iqbal, M. Z., Selim, M. R., Rashid, M. M., & Rahman, M. S. (2021). Bangla speech emotion recognition and cross-lingual study using deep CNN and BLSTM networks. IEEE Access, 10, 564-578.

[16] Liu, X., Shi, T., Zhou, G., Liu, M., Yin, Z., Yin, L., & Zheng, W. (2023). Emotion classification for short texts: an improved multi-label method. Humanities and Social Sciences Communications, 10(1), 1-9.

[17] Jabreel, M., & Moreno, A. (2019). A deep learning-based approach for multi-label emotion classification in tweets. Applied Sciences, 9(6), 1123.

[18] Huang, C., Trabelsi, A., Qin, X., Farruque, N., & Zaïane, O. R. (2019). Seq2emo for multi-label emotion classification based on latent variable chains transformation. arXiv preprint arXiv:1911.02147.

[19] Ameer, I., Bölücü, N., Siddiqui, M. H. F., Can, B., Sidorov, G., & Gelbukh, A. (2023). Multi-label emotion classification in texts using transfer learning. Expert Systems with Applications, 213, 118534.

[20] Islam, K. I., Kar, S., Islam, M. S., & Amin, M. R. (2021, November). SentNoB: A dataset for analysing sentiment on noisy Bangla texts. In *Findings of the Association for Computational Linguistics: EMNLP 2021* (pp. 3265-3271).

[21] Islam, K. I., Islam, M. S., & Amin, M. R. (2020, December). Sentiment analysis in Bengali via transfer learning using multi-lingual BERT. In *2020 23rd International Conference on Computer and Information Technology (ICCIT)* (pp. 1-5). IEEE.

[22] Hossain, M. M., Akhi, I. A. H., Raisa, S. R. S., Onni, T. S., Rifat, A. I., & Islam, A. (2023, December). Critical Analysis of BERT and LSTM Model for Bengali Sentiment Analysis Across Varied Datasets. In *2023 IEEE 15th International Conference on Computational Intelligence and Communication Networks (CICN)* (pp. 663-668). IEEE.

[23] Elahi, K. T., Rahman, T. B., Shahriar, S., Sarker, S., Shawon, M. T. R., & Shahariar, G. M. (2024). A Comparative Analysis of Noise Reduction Methods in Sentiment Analysis on Noisy Bengali Texts. *arXiv preprint arXiv:2401.14360*.

[24] Bhowmik, N. R., Arifuzzaman, M., & Mondal, M. R. H. (2022). Sentiment analysis on Bangla text using extended lexicon dictionary and deep learning algorithms. *Array*, *13*, 100123.

[25] Hossain, T., Kabir, A. A. N., Ratul, M. A. H., & Sattar, A. (2022, March). Sentence level sentiment classification using machine learning approach in the bengali language. In *2022 International Conference on Decision Aid Sciences and Applications (DASA)* (pp. 1286-1289). IEEE.

Appendix A

Course Outcomes, Complex Engineering Problems (EP) and Complex Engineering Activities (EA) Addressing

Title: MULTI-LABEL EMOTIONS BANGLA TEXT CLASSIFICATION

Student ID: 202-15-14391 And 202-15-14378

CO Description for FYDP

CO	CO Descriptions	PO
Phase -I		
CO1	Integrate recently gained and previously acquired knowledge to identify **emotions detection** problem for the Final Year Design Project (FYDP)	**PO1**
CO2	Analyze different aspects of the goals in designing a solution for this FYDP	**PO2**
CO3	Explore diverse problem domains through a literature review, delineate the issues, and establish this goals for the FYDP	**PO4**
CO4	Perform economic evaluation and cost estimation and employ suitable project management procedures throughout the development life cycle of the FYDP	**PO11**
Phase -II		
CO5	Design and develop technical solutions and system components or processes that meet specified requirements, ensuring compliance with public health and safety standards, as well as considering cultural, socioeconomic, and environmental factors in this FYDP	**PO3**
CO6	Choose and apply appropriate methodologies, resources, and contemporary engineering and IT technologies to address complex engineering processes, encompassing prediction and modeling, while adhering to relevant constraints in this FYDP	**PO5**
CO7	Analyze societal, health, safety, legal, and cultural considerations, along with associated responsibilities, in the context of professional engineering practice and the resolution of this problem, employing logical reasoning guided by contextual understanding.	**PO6**
CO8	Comprehend and evaluate the enduring sustainability and impact of professional engineering endeavors in addressing intricate engineering challenges within social and environmental frameworks.	**PO7**
CO9	Implement ethical principles and adhere to professional standards and norms in this FYDP	**PO8**
CO10	Capable of operating proficiently both individually and as a team member or leader across diverse teams and interdisciplinary settings in this FYDP.	**PO9**
CO11	Proficiently communicate with the engineering community and broader society regarding complex engineering endeavors, including the ability to comprehend and generate comprehensive reports and design documentation, as well as provide and receive clear instructions throughout this FYDP.	**PO10**

CO12	Acknowledge the importance of self-directed and life-long learning within the evolving landscape of technology, and possess the readiness and capability to engage in lifelong learning endeavors.	**PO12**

Addressing CO (1 to 8), Knowledge Profile (K), Attainment of Complex Engineering Problems (EP), and Attainment of Complex Engineering Activities (EA)

Addressing CO (1 to 8), Knowledge Profile (K), Attainment of Complex Engineering Problems (EP):

SN	EP Definition	Attainment	CO	Justification (with Knowledge Profile)	References
1.	**EP1: Depth of Knowledge required**	Yes	CO1, CO2, CO3, CO5, CO6, CO7 and CO8	Through the use of NLP, data augmentation methods, and several Bi-LSTM structures for image analysis and classification tasks, our project **illustrates fundamental engineering (K3)** principles. Through transfer learning using cutting-edge RNN designs like Bi-LSTM and NLP to improve pneumonia diagnosis precision in the Bangla test, the project displays **specialized knowledge (K4).**	**Page no:** [21-24] **Section:** [3.2] **Page no:** [1] **Section:** [1.1]
				Our research uses the figure of the experimentation process to apply **engineering practice & design (K5**). The project uses CNN and Bi-LSTM models to address **engineering practice & technology (K6).**	**Page no:** [22] **Section:** [3.2(B)] **Page no:** [31-35] **Section:** [4.2]

				With the use of machine learning and the synthesis of ideas from recent studies, our study **ensures K8 (Research Literature)** and demonstrates a thorough comprehension of current approaches for Bangla NLP text identification.	**Page no:** [12-16] **Section:** [2.2, 2.3]
2.	**EP2: Range of Conflicting Requirements**	Yes	CO2, and CO7	In order to handle **EP-2**, our project takes into account Bangla raw text data, as well as the drawbacks of conventional techniques and the complexity of CNNs and Bi-LSTMs. It addresses difficulties in comprehending geographical distributions through comparative analysis, providing information for improving detection techniques.	**Page no:** [16-19] **Section:** [2.4, 2.5] **Page no:** [26-30] **Section:** [3.4]
3.	**EP3: Depth of analysis required**	Yes	CO2, and CO6	In order to improve Bangla emotion recognition, our project carefully compares experimental results to address **EP-3**, identifying natural language processing (NLP) as the most important answer.	**Page no:** [38-42] **Section:** [5.2,5.3]
4.	**EP4: Familiarity of Issues**	Yes	CO8	The multidisciplinary approach of this project goes beyond the fields of engineering and computer science, influencing everything from advances in online health and protection **(EP-4)** to social feeling such as “Anger”,”Happiness”,”Sadness”,”Surprise”,”Fear”, ”Confused”.	**Page no:** [13-14,16-18] **Section:** [2.2, 2.4]

5.	**EP5: Extends of application codes**	No	CO5	N/A	N/A
6.	**EP6: Extends of stakeholders involved and**	No	CO8	N/A	N/A
7.	**EP7: Interdependence**	Yes	CO5	Through the integration of several components from data gathering, statistical analysis, and recommended methodology, our project's complete approach tackles high-level problems and ensures a comprehensive solution to complicated challenges in Bengali emotion recognition, guaranteeing **EP-7.**	**Page no:** [21-25] **Section:** [3.2, 3.3]

Addressing CO11 with Complex Engineering Activities (EA) [Some or all of the following]:

SN	EA Definition	Attainment	CO	Justification	References
1.	**EA1: Range of resources**	Yes	CO11	To assure methodical study and advance Bangla emotion text data, our project makes use of a variety of resources, including high-performance computer facilities GPUs, artificial intelligence systems, annotated data sets, and ethical considerations.	**Page no:** [24-25, 22-23] **Section:** [3.2, 3.3]

2.	**EA2: Level of interaction**	No		N/A	N/A
3.	**EA3: Innovation**	No		N/A	N/A
4.	**EA4: Consequences for society and the environment**	Yes		Through the use of cutting-edge Bangla emotion detection techniques, this project improves social media and promotes environmental sustainability by using efficient computing resources and abiding by ethical standards for text data privacy.	**Page no:** [43-48] **Section:** [6.1, 6.2, 6.4]
5.	**EA-5: Familiarity**	Yes		By investigating a novel method in Bengali emotion detection using NLP and CNNs, as shown through a thorough comparison analysis, this project builds on previous research and provides fresh perspectives on the subject of Bangla text.	**Page no:** [12-16] **Section:** [2.1, 2.3]

Addressing CO (4, 9, 10, and 12):

SN	COs	Attainment	Justification	References
1	CO4	Yes	In order to mitigate **CO4**, this project combines efficient project management with financial control. Careful planning, resource distribution, and budget estimation are guaranteed to provide the best possible utilization of resources throughout the duration of the study project.	**Page no:** [26-29] **Section:** [3.4]
2	CO9	Yes	By protecting the privacy of Bangle people, obtaining informed consent, and openly recording the research process, the project	**Page no:** [45-46] **Section:**

			illustrates adherence to ethical principles. It also ensures responsible dissemination of information and the well-being of society by applying cutting-edge, ethically compliant **CO9** social communication technologies.	[6.3]
3	CO10	No	N/A	N/A
4	CO12	Yes	The project's extensive Bangla data collection, rigorous statistical analysis, careful methodology growth, and thorough experimental results and analysis all demonstrate a dedication to ongoing education **(CO12)** and adaptation within the dynamic technological landscape. This shows a commitment to staying current and improving techniques to address modern challenges.	**Page no:** [35-36, 38-41] **Section:** [4.3, 5.1, 5.2, 5.3]

APPENDIX B

Multi-label Emotions Bangla Text Classification

ORIGINALITY REPORT

16%	12%	9%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	3%
2	aclanthology.org Internet Source	1%
3	peerj.com Internet Source	<1%
4	"Proceedings of the 2nd International Conference on Big Data, IoT and Machine Learning", Springer Science and Business Media LLC, 2024 Publication	<1%
5	arxiv.org Internet Source	<1%
6	preview.aclanthology.org Internet Source	<1%
7	Submitted to Daffodil International University Student Paper	<1%
8	pure.ulster.ac.uk Internet Source	<1%