# SENTIMENT ANALYSIS OF BANGLA CUSTOMER REVIEWS ON DARAZ USING DEEP LEARNING APPROACH

**BY**

**Nur-A-All Asif**
**ID: 202-15-14370**
**AND**

**Mithila Ghosh**
**ID: 202-15-14450**

## FINAL YEAR DESIGN 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

**Fahad Faisal**
Assistant Professor
Department of CSE
Daffodil International University

## Co-Supervised By

**Dr. Sheak Rashed Haider Noori**
Professor & Head
Department of CSE
Daffodil International University

**DAFFODIL INTERNATIONAL UNIVERSITY**

**DHAKA, BANGLADESH**

**JULY 2024**

# APPROVAL

This Project titled "Sentiment Analysis of Bangla customer reviews on Daraz using Deep learning approach", submitted by **"Nur-A-All Asif"** and **"Mithila Ghosh"** 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 15-07-2024.

## BOARD OF EXAMINERS

| | |
|---|---|
| **Dr. Md Ismail Jabiullah**<br>**Professor**<br>Department of CSE<br>Faculty of Science & Information Technology<br>Daffodil International University | **Board Chairman** |
| **Mr. Md Assaduzzaman**<br>**Sr. Lecture**<br>Department of CSE<br>Faculty of Science & Information Technology<br>Daffodil International University | **Internal Examiner 1** |
| **Ms. Sharun Akter Khushbu**<br>**Sr. Lecture**<br>Department of CSE<br>Faculty of Science & Information Technology<br>Daffodil International University | **Internal Examiner 2** |
| **Dr. Ahmed Wasif Reza**<br>**Professor**<br>Department of CSE<br>East West University | **External Examiner** |

# DECLARATION

We hereby declare that this project has been done by us under the supervision of **Fahad Faisal, Assistant Professor, 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 award of any degree or diploma.

**Supervised by:**

**Fahad Faisal**
Assistant Professor
Department of CSE
Daffodil International University

**Co-Supervised by:**

**Dr. Sheak Rashed Haider Noori**
Professor & Head
Department of CSE
Daffodil International University

**Submitted by:**

**Nur-A-All Asif**
**ID: 202-15-14370**
Department of CSE
Daffodil International University

**Mithila Ghosh**
**ID: 202-15-14450**
Department of CSE
Daffodil International University

# ACKNOWLEDGEMENT

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

We are grateful and wish our profound indebtedness to **Fahad Faisal**, **Assistant Professor**, 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 **Dr. Sheak Rashed Haider Noori, 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

This study examines the effectiveness of deep learning models for sentiment analysis of Bangla customer evaluations on the Daraz platform. The researchers developed and evaluated various models, including CNN, LSTM, GRU, and a hybrid CNN+BiLSTM, focusing on their ability to accurately classify sentiments. The experimental setup involved exhaustive preprocessing of Bangla text and using TensorFlow and PyTorch frameworks for model training. The CNN+BiLSTM model achieved the highest accuracy and precision, indicating its superior performance in identifying positive sentiments. The CNN model showed balanced performance with high accuracy and F1-Score, making it reliable for general sentiment classification tasks. The CNN+BiLSTM model was the most effective for precision sentiment predictions, while the CNN model proved a reliable choice for balanced sentiment analysis. The research aims to construct sentiment analysis algorithms for multilingual e-commerce platforms, as online stores like Daraz have a significant amount of customer feedback, making these critiques more credible than other forms of advertising material.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

## LIST OF ACRONYMS

| | |
|---|---|
| DL | Deep Learning |
| CNN | Convolutional Neural Network |
| LSTM | Long Short-Term Memory |
| GRU | Gated Recurrent Unit |
| BiLSTM | Bidirectional Long Short-Term Memory |

# CHAPTER 1

# INTRODUCTION

## 1.1 Introduction:

Personal opinion is more trusted by the public than traditional advertising. People are accustomed, for instance, to seeking guidance and ideas from others prior to making significant purchases. It helps identify and investigate how individuals use written or spoken language to express their ideas, opinions, behaviors, attitudes and about various thinks including people, organizations, products and services. Word-of-mouth (WOM) [1] has always been crucial in helping clients make these kinds of judgments. WOM, on the other hand, is very important to suppliers. In comparison to conventional promotional techniques, it has a greater impact on acquiring new customers. Nowadays, discussing knowledge is common. without the support of the web. Social media networks such as both Twitter and Facebook have facilitated the sharing of opinions about goods, services, and brands. E-word of mouth is an expanded version of word-of-mouth communication (eWOM). Text classification also known as text tagging or text categorization is the process of classifying text into ordered categories. By employing Natural Language Processing (NLP), text classifiers may automatically assess text and then assign a set of pre-defined tags or categories depending on its content. Businesses must analyze consumer sentiment via reviews to improve user experience and customize offers efficiently in the ever-changing e-commerce market. Customer review analysis, particularly in Bangla and other South Asian languages, is becoming more important for market insights and company strategy as platforms like Daraz grow in popularity in the region. Nevertheless, owing to the language subtleties and lack of research focus on Bangla, emotion analysis of Bangla text has distinct obstacles. To fill this need, our research suggests using deep learning to analyze the tone of Bangla customer reviews on Daraz. The remarkable success of deep learning algorithms in NLP tasks has opened exciting new possibilities for sentiment analysis in a wide variety of languages. The goal of this project is to improve sentiment analysis by capturing the complex linguistic aspects of Bangla text using deep neural networks. The study's capacity to uncover relevant information from Bangla user evaluations on Daraz can assist firms better their marketing strategies, product offerings, and customer satisfaction [2]. Using a deep learning architecture suited to the intricacies of the Bangla language, this research intends to construct sentiment analysis algorithms for multilingual e-commerce platforms. In addition, online stores such as "Daraz" have a substantial amount of customer feedback posted about their goods and services. People find these critiques more credible than other types of advertising material. In contrast, businesses are keen to analyze all these exchanges and conduct to determine the general public's perceptions of an item or company. In such a competitive setting it would aid in the development of their company plan. If you want to know what the public thinks about user experience, politics, or social issues sentiment analysis a great tool is to utilize. Sentiment analysis has become more important in both academic and business contexts due to the growing number of Bangla-speaking internet users. The study of Bangla Natural Language Processing (BNLP) has challenges owing to a lack of resources even though languages like English have been extensively investigated in Natural Language Processing (NLP). Presenting a DL Framework for sentiment analysis, this study addresses the resource issue for low-resource languages like Bangla [3].

### 1.2 Problem Statement

In the field of sentiment analysis for Bangla customer reviews on Daraz, the typical binary categorization into positive and negative emotions may not fully represent the vast range of viewpoints given by users. Consequently, there is an increasing need for sentiment analysis algorithms that can successfully categorize evaluations into three categories: positive, negative and neutral. However, constructing such models brings unique problems and needs customized techniques to enable accurate sentiment categorization and useful insights.

- **Need for Demanding Analysis:** The growth of e-commerce platforms like Daraz has led to a wealth of consumer evaluations each representing diverse moods and opinions. While binary sentiment categorization gives significant information. It sometimes oversimplifies the complicated structure of consumer attitudes. Introducing a neutral sentiment category offers a more detailed analysis allowing firms to comprehend the intricacies of client sentiment beyond basic positive or disapproval.
- **Verbal Complexity of Bangla Text**: Bangla as a language contains distinctive linguistic traits and subtleties that offer obstacles for sentiment analysis. Traditional sentiment analysis methods created for languages like English may not be immediately relevant to Bangla text owing to variations in syntax, morphology, and cultural context. Therefore, constructing a sentiment analysis model particularly customized to Bangla customer evaluations on Daraz demands careful consideration of language complexity and cultural refinement.
- **Data Imbalance and Annotation Challenges:** Annotated datasets including Bangla customer evaluations classified into positive, negative and neutral attitudes are critical for training and testing sentiment analysis algorithms. However,

obtaining and annotating such datasets offer considerable hurdles including data imbalance across sentiment categories subjective interpretation of sentiment by annotators, and absence of established annotation criteria for Bangla literature. Addressing these problems is vital to guarantee the reliability and validity of the sentiment analysis model.

- **Model Robustness and Generalization:** Developing a sentiment analysis model capable of robust performance across varied domains product categories and client demographics is crucial. The model should display resistance to noisy and confusing text manage differences in language use and expression and generalize effectively to unknown data. Achieving model robustness involves careful selection of deep learning architectures feature representations and optimization algorithms customized to the peculiarities of Bangla customer reviews on Daraz.
- **Business Impact and Decision-Making:** The ultimate purpose of sentiment analysis is to provide organizations with actionable insights drawn from client feedback. Accurate categorization of evaluations into positive, negative and neutral categories helps organizations to discover areas of development, capitalize on strengths and modify their strategy to meet consumer expectations successfully. Therefore, the success of sentiment analysis models rests on their capacity to generate valuable insights that drive informed decision-making and boost overall consumer happiness.

## 1.3 Motivation and Objectives:

Sharing experiences has become usual in the modern world. With the aid of the World Wide Web. Social networking sites like Facebook and Twitter have made it easier for people to voice their thoughts about products, services, and businesses. Reviews of goods and services have been uploaded in large quantities on e-commerce platforms like Daraz and Evaly. Compared to other forms of marketing, comments have a higher chance of being accepted by customers. Nonetheless, companies are eager to look into any of these discussions and actions to learn more about what consumers are thinking about in particular. It could assist them in formulating a competitive strategy for this market [4]. Here we mention in above-

- Maximize User Experience
- Boost Service Quality
- Stay Up to Date with Technological Development
- Optimize Resource Allocation
- Gain a Competitive Edge

### 1.4 Rational of the Study:

This study aims to perform sentiment analysis of Bangla customer reviews on Daraz using deep learning approaches. Bangla is one of the world's most spoken languages, yet it lacks extensive natural language processing (NLP) tools compared to languages like English. By developing sentiment analysis tools tailored for Bangla, this research addresses this gap and contributes to the broader NLP field. Daraz, a leading e-commerce platform in Bangladesh, serves as a relevant source of customer reviews. Analyzing these reviews with deep learning techniques such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks provides robust, nuanced sentiment insights. These models can capture intricate patterns and contextual information within the text, offering more accurate sentiment analysis compared to traditional methods. The practical benefits of this study are significant. E-commerce platforms can use the insights to improve customer satisfaction, identify areas of improvement, and make data-driven decisions. This can enhance user experiences, drive business growth, and increase competitiveness in the market. Overall, this study not only advances NLP resources for Bangla but also supports better business practices and user experiences in the Bangla-speaking community, promoting digital inclusivity and technological advancement [5].

### 1.5 Project Outcome:

This research tackles the shortage of resources for low-resource languages such as Bangla by presenting a low-cost deep learning framework for sentiment analysis. We will have been trying to classify approaches that achieve stunning accuracy, outperforming previous algorithms. This study adds to the rapidly developing field of BNLP by providing reliable method analysis in Bangla. The outcomes demonstrate how well our method works to extract subtle emotions from Bangla texts.

### 1.6 Report Organization:

The advised report has a complete format that guides readers through the methodology, conclusions, and implications of the study. Every chapter has a distinct function and helps to the document's overall integration and depth.

**Chapter 1: Introduction**

The public trusts personal opinions more than traditional forms of advertising. For example, people are used to getting advice and suggestions from others before making big purchases. It assists in identifying and examining the ways in which people communicate their ideas, views, behaviors, attitudes, and thoughts on a range of topics, including people, organizations, goods, and services, using written or spoken language. Word-of-mouth

(WOM) has always been essential in assisting customers in reaching these decisions. WOM is crucial to suppliers, nevertheless, as opposed to other parties. It is more effective in attracting new clients than traditional promotional strategies. Talking about information is normal these days, even without the internet. Social media platforms like Facebook and Twitter have made it easier for people to share their thoughts about products, services,

**Chapter 2: Literature Review**

Sentiment analysis is crucial in identifying and categorizing opinions in text, particularly in Bangla, a widely spoken language. Daraz, a major e-commerce platform in Bangladesh, uses deep learning to analyze customer reviews. Neural networks like Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, Bidirectional LSTMs, and Gated Recurrent Units capture local patterns and data dependencies. NLP focuses on computer-human interaction through natural language, and feature extraction, activation functions, pooling layers, training, and validation are essential for this analysis.

**Chapter 3: Research Methodology**

The Deep Learning architecture for Bangla sentiment analysis uses a large collection of Bangla text samples to analyze positive, negative, and neutral moods. Preprocessing is essential for tasks like tokenization and stemming. The architecture requires extensive configuration and fair distribution of sentiment classes to avoid bias. Cross-validation procedures evaluate the model's robustness, and measures like accuracy, precision, recall, and F1 score are constructed. Ethical issues are addressed to ensure a fair and impartial sentiment analysis model. Techniques like tokenization, feature extraction, and performance measurement matrices help optimize and refine machine learning algorithms.

**Chapter 4: Experimental Result**

The study aims to create a deep learning-based sentiment analysis system for Bangla customer reviews on Daraz using a dataset of Bangla customer reviews. The system uses high-performance GPUs or TPUs for training and uses Convolutional Neural Networks (CNNs) to capture local patterns and features. Long Short-Term Memory Networks (LSTMs) are used to understand long-term dependencies in sequential data. The architecture includes embedding layers for feature extraction, LSTM or BiLSTM layers for sequence modeling, and fully connected layers for classification. The CNN model is the preferred choice due to its balanced performance. The CNN+BiLSTM model outperforms other models in identifying positive and negative sentiments.

**Chapter 5: Impact On Society**

The study on sentiment analysis of Bangla customer reviews on Daraz using a deep learning approach has several significant societal impacts. It enhances customer experience, provides business intelligence, and offers a competitive advantage. It also contributes to cultural and linguistic development, promoting digital inclusivity and linguistic accessibility. The study boosts e-commerce, supports SMEs, and raises consumer awareness. It also addresses social issues, such as unfair labor practices and product safety concerns. The research also pushes the boundaries of AI research and development, fostering collaboration between academia, industry, and government. Additionally, it enhances skills development, making students and professionals more competitive in the job market and contributing to the tech industry's advancement.

**Chapter 6: Conclusion and Future Research**

The study uses deep learning to analyze Daraz Bangla customer evaluations, focusing on sentiment analysis in the e-commerce sector. Deep neural networks are used to capture the Bangla language's linguistic features. The GRU model achieved 82.27% accuracy, while the CNN model achieved 84.07%. The hybrid CNN+BiLSTM model was the top performer with 87.18% accuracy. The study suggests further enhancements in sentiment analysis, including attention mechanisms, pre-trained language models, and data preprocessing techniques. Future research should explore real-time sentiment analysis, multimodal sentiment analysis, cross-lingual sentiment analysis, and user feedback.

### 1.7 Summary:

The passage discusses the increasing reliance on human judgment, particularly through word of mouth (WOM) and its electronic counterpart (eWOM), in influencing consumer decisions. It highlights the significance of online platforms like Twitter, Facebook, and e-commerce sites in facilitating the exchange of opinions about products and services. Because digital social media makes it simpler for customers to evaluate and comment on products and sustainability, the scope of this initiative is quite big. If business owners and managers regularly evaluate client feedback, they can gauge the product's success with ease. We should undertake this initiative so that they may evaluate consumer needs, produce more impactful, long-lasting goods, and compare their offerings. Sentiment analysis, categorized into document, sentence, and aspect levels, is presented as a valuable tool for understanding and analyzing consumer sentiments expressed in online content. The project aims to address the resource challenges in low-resource languages like Bangla by introducing a low-cost DL framework for sentiment analysis. The study contributes to the evolving field of Bangla Natural Language Processing (BNLP) by providing a reliable method for sentiment analysis, showcasing its effectiveness in extracting nuanced emotions from Bangla texts.

# CHAPTER 2

# LITERATURE REVIEW

## 2.1 Overview:

Sentiment analysis is the process of identifying and categorizing opinions in text to determine if the sentiment is positive, negative, or neutral. Bangla, also known as Bengali, is widely spoken but lacks extensive natural language processing (NLP) tools. Daraz, a major e-commerce platform in South Asia, particularly Bangladesh, provides valuable customer reviews for sentiment analysis to understand consumer attitudes. Deep learning, a subset of machine learning, involves neural networks with many layers to learn complex data representations. Convolutional Neural Networks (CNNs) are effective at capturing local patterns in data, including text. Long Short-Term Memory (LSTM) networks handle long-term dependencies in sequential data and solve the vanishing gradient problem of traditional RNNs. Bidirectional LSTMs (BiLSTMs) process data in both forward and backward directions, capturing context from both past and future. Gated Recurrent Units (GRUs) are simpler than LSTMs but effective in capturing sequential data dependencies. NLP focuses on computer-human interaction through natural language, involving algorithms to process and analyze text. Feature extraction transforms raw data into features for machine learning. Activation functions, like ReLU and sigmoid, introduce non-linearity in neural networks. Pooling layers in CNNs reduce feature map dimensions, aiding in computational efficiency and overfitting control. Training adjusts neural network parameters to minimize error, while validation ensures performance on unseen data. These concepts are essential for understanding sentiment analysis of Bangla customer reviews using deep learning [6].

## 2.2 Related Works:

Tuhin, et a al. (2019). An Automated System of Sentiment Analysis from Bangla Text using Supervised Learning Techniques. They collect manually data corpus consisting of Bangla sentences for training and testing. By the proposed machine learning model. dataset amounts 7500 + Bangla sentences. Model algorithms name Naive Bayes Classification, Topical Approach. Highest accuracy, acquired by Topical Approach 90% [7].

Mahtab, et a al. (2018). Sentiment Analysis on Bangladesh Cricket with Support Vector Machine. They collect data form random social media and new portal where people express their opinion in Bengali. This ABSA dataset contains 2979 data with 5 columns. The shape of our dataset is 1601 which contains 3 classes including praise with 513, criticism with 604 and sadness with 484 labeled data. Algorithm name is Naïve Bayes, Support Vector Machine, Decision Tree. the accuracy of model SVM is 64.6%; NB is 58.38%; DT is 43.47% [8].

Hassan, et a al. (2016). Sentiment analysis on Bangla and Romanized Bangla text using deep recurrent models. Data set collected from Facebook: 4621; from Twitter: 2610; from YouTube: 801; from online news portals: 1255; from product review pages: 50. Total number of data: 9337; Bangla Data: 6698; Romanized Bangla Data: 2639.Implemented algorithm is RNN and LSTM. Accuracy is ". Ambiguous remove" Score =78% accuracy," Ambiguous Convert" score = 55% [9].

Bhowmik et al (2022). Suggest a DL-based speak to Bangla sentiment analysis employ a Lexicon data dictionary (LDD) and Bangla text sentiment score (BTSC) algorithm. They assess such as CNN, RNN, LSTM and hybrid models like HAN-LSTM, D-CAPSNET-Bi-LSTM and BERT-LSTM.BERT-LSTM model had highest accuracy at 84.18%. This study helps with new DL techniques for Bangla sentiment analysis providing a foundation for future research in low-resource language. [10].

Alvi, et al. (2024). Sentiment Analysis on Bangla Text Using Gated Recurrent Unit (GRU) Neural Network. The dataset has been collected from various sources. The algorithm used is a five-layer GRU neural network, each layer comprising 48 neurons. We applied a ten-fold cross-validation approach and repeated the process three times. The highest accuracy achieved by our model is 78.41%, surpassing the state-of-the-art Bidirectional LSTM (BLSTM) model, which has a highest accuracy of 77.85% [11].

M. A. Rahman et a al. (2018) "Aspect extraction from Bangla reviews using convolutional neural network,". They collect data form twite on twitter. The dataset consists of movie review tweets from various Twitter users. Total of 17,247 movie review tweet data has been collected. Implemented algorithm are CNN, KNN, RF, SVM. Accuracy is CNN 81%, KNN 21% , RF 24% SVM 19% [12].

M. H. Alam, et al. (2017) "Sentiment analysis for Bangla sentences using convolutional neural network," in 2017 They collect data form digital contents and review comment and Facebook status which is written in Bangla. Approximately, 850 Bangla comments from different sources. The comments were then copied and pasted repeatedly to increase the dataset size to around 120,000, with 60,000 positive and 60,000 negative comments. Algorithms name is CNN + Bi-LSTM. Accuracy is SVM is 93.13% [13].

## 2.3 Comparison between existing works:

Table 2.3.1: Comparative analysis with previous work

| SL | Author Name | Data Set | Used Algorithms | Best Accuracy with Algorithms |
|---|---|---|---|---|
| 1. | Tuhin.et al. [7] | They collect manually data corpus consisting of 7500 + Bangla sentences for training and testing. By the proposed machine learning model. | 1.Naive Bayes Classification. 2.Topical Approach | Highest accuracy, acquired by Topical Approach 90% |
| 2. | Mahtab. et al. [8] | They collect data form random social media and new portal where people express their opinion in Bengali. | Naïve Bayes, Support Vector Machine, Decision Tree | SVM= 64.6%, NB=58.38%, DT=43.47% |
| 3. | Hasan.et al. [9] | Total number of data: 9337; Bangla Data: 6698 Romanized Bangla Data: 2639 From Facebook: 4621 From Twitter: 2610 From YouTube: 801 From online news portals: 1255 From product review pages: 50 | RNN+LSTM | "Ambiguous remove" Score =78% accuracy," Ambiguous Convert" score = 55% |
| 4. | Bhowmik. et al. [10] | Data source not mention but total data is 2412 | BERT-LSTM | BERT-LSTM=84.18% |
| 5. | Alvi. et al. [11] | Collect from others paper source and total data 6000 | BiLSTM | BiLSTM=77.85% |

| | | | | |
|---|---|---|---|---|
| 6. | M. A. Rahman. et al. [12] | Total of 17,247 movie review tweet data has been used from Kaggle. | CNN, KNN, RF, SVM | CNN 81%, KNN 21%, RF 24% SVM 19% and restaurant CNN 83%, SVM 29%, RF 30%, KNN 32% accuracy |
| 7. | M.H. Alam et al.[13] | Data source from Kaggle 850 data and copied and pasted repeatedly. | SVM, LSTM | SVM 93.13% and LSTM 78% |
| 8. | S. Haque et al [14] | Used two data from online platforms. Total dataset was 5038. | RS, SVM, KNN, LR, NB | F1 score for cricket RF=37%, SVM=35%, KNN=27%, LR=34%, NB=18% Restaurant RF=35%, SVM=39%, KNN=38%, LR=43%, NB=17% |
| 9. | Sarkar. et al. [15] | approximately, 850 Bangla comments from different sources. The comments were then copied and pasted repeatedly to increase the dataset size to around 120,000, with 60,000 positive and 60,000 negative comments | CNN + Bi-LSTM | 83.77% |

**2.4 Gap Analysis:**

Low-resource languages, such as Bangla, confront constraints in terms of annotated datasets and language parsers, both of which are required for sentiment analysis research. Some low-resource language sentiment analysis algorithms may rely significantly on specific characteristics, such as emotion symbols, which may not exist in every phrase and may not be suited for evaluating complicated sentences. Compared to languages with better resources, like English and Japanese, sentiment analysis research is less prevalent for low-resource languages, like Bangla. Furthermore,

- **Limited Dataset Size:** The research likely faces limitations because of a limited dataset size. A restricted dataset might hamper the model's capacity to generalize well to unknown data and may lead to overfitting. It hinders the possibility of recognizing the complete spectrum of emotion expression within Bangla text, thereby likely diminishing the reliability and robustness of the given models.

- **Binary Sentiment Classification:** The paper's main concentration on binary sentiment classification into negative and positive categories lacks the subtleties observed in Bangla text sentiment. By only examining negative and positive feelings, the research may fail to find situations when sentiment expression falls into a neutral category or exposes tiny differences that might impact user perception and decision-making [15].

Therefore, we may try to categorize all 3 types of emotion categories as neutral, negative, and positive, as well as huge amounts of data.

**2.5 Scope of Problem:**

In the field of sentiment analysis for Bangla customer reviews on Daraz, the typical binary categorization into positive and negative emotions may not fully represent the vast range of viewpoints given by users. Consequently, there is an increasing need for sentiment analysis algorithms that can successfully categorize evaluations into three categories: positive, negative and neutral. However, constructing such models brings unique problems and needs customized techniques to enable accurate sentiment categorization and useful insights.

- **Need for Demanding Analysis:** The growth of e-commerce platforms like Daraz has led to a wealth of consumer evaluations each representing diverse moods and opinions. While binary sentiment categorization gives significant information. It sometimes oversimplifies the complicated structure of consumer attitudes. Introducing a neutral sentiment category offers a more detailed analysis allowing firms to comprehend the intricacies of client sentiment beyond basic positive or disapproval.
- **Verbal Complexity of Bangla Text**: Bangla as a language contains distinctive linguistic traits and subtleties that offer obstacles for sentiment analysis. Traditional sentiment analysis methods created for languages like English may not be immediately relevant to Bangla text owing to variations in syntax, morphology, and cultural context. Therefore, constructing a sentiment analysis model particularly customized to Bangla customer evaluations on Daraz demands careful consideration of language complexity and cultural refinement.
- **Data Imbalance and Annotation Challenges:** Annotated datasets including Bangla customer evaluations classified into positive, negative and neutral attitudes are critical for training and testing sentiment analysis algorithms. However,

obtaining and annotating such datasets offer considerable hurdles including data imbalance across sentiment categories subjective interpretation of sentiment by annotators, and absence of established annotation criteria for Bangla literature. Addressing these problems is vital to guarantee the reliability and validity of the sentiment analysis model.

- **Model Robustness and Generalization:** Developing a sentiment analysis model capable of robust performance across varied domains product categories and client demographics is crucial. The model should display resistance to noisy and confusing text manage differences in language use and expression and generalize effectively to unknown data. Achieving model robustness involves careful selection of deep learning architectures feature representations and optimization algorithms customized to the peculiarities of Bangla customer reviews on Daraz.
- **Business Impact and Decision-Making:** The ultimate purpose of sentiment analysis is to provide organizations with actionable insights drawn from client feedback. Accurate categorization of evaluations into positive, negative and neutral categories help organizations to discover areas of development, capitalize on strengths and modify their strategy to meet consumer expectations successfully. Therefore, the success of sentiment analysis models rests on their capacity to generate valuable insights that drive informed decision-making and boost overall consumer happiness.

### 2.6 Summary:

According to this research, deep neural networks can capture the subtle linguistic aspects of Bangla text and utilize them to analyze Bangla customer ratings on Daraz. This may aid firms increase marketing strategy, product offerings, and client delight. Sentiment analysis algorithms for multilingual e-commerce platforms may aid companies grasp the public's perceptions of things and enterprises, ultimately supporting in the development of their corporate objectives. However, low-resource languages like Bangla face restrictions in terms of annotated datasets and language parsers, which are necessary for sentiment analysis research. Additionally, the research meets issues related to a limited sample size, which may hamper the model's capacity to generalize successfully to unknown data and may lead to overfitting. The paper's significant concentration on binary sentiment classification into negative and positive categories may not capture subtleties present in Bangla text sentiment, exhibiting minute deviations that might impact user perception and decision-making.

# CHAPTER 3

# METHODOLOGY

## 3.1 Overview:

In this thesis, we used several deep learning algorithms, such as Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, Gated Recurrent Units (GRU), and a hybrid CNN+Bidirectional LSTM (BiLSTM) model, to analyze sentiment on Daraz Bangla customer evaluations. The dataset was obtained from the Daraz website and subjected to preprocessing techniques such as tokenization and feature extraction before analysis. To verify robustness and dependability, each model underwent tenfold cross-validation. The hybrid CNN+BiLSTM model outperformed the other models evaluated in terms of accuracy, capturing both local and long-term relationships in the text. This work demonstrated the effectiveness of integrating CNN and BiLSTM for sentiment analysis in Bangla, establishing a new standard for future research in this field.

## 3.2 Requirement Analysis:

For a strong and successful implementation of Bangla sentiment analysis using Deep Learning approach, several important demands must be fulfilled. The foundation is a large and well-annotated collection of Bangla text samples covering positive, negative, and neutral moods. To ensure that Bangla text is successfully represented in the model, preprocessing approaches specific to the Bangla language are necessary for tasks like tokenization and stemming. Deep Learning architecture demands extensive configuration. This means that to convert Bangla words into continuous vector representations, an embedding layer needs to be incorporated in addition to tailoring the number of convolutional layers, kernel sizes, filter numbers, and fully connected layer structure to the specifics of Bangla emotion analysis. Adjusting hyper-parameters with the value batch size, dropout rates, and learning rate is part of the training process. It is critical to provide a fair distribution of sentiment classes in the training dataset in order to avoid bias and improve the model's generalization across diverse feelings. Cross-validation procedures should be used to evaluate the model's robustness, and measures such as accuracy, precision, recall, and F1 score should be constructed for evaluation. Furthermore, correcting dataset imbalances and detecting language-specific idiosyncrasies in sentiment representation are critical. Ethical issues, such as eliminating biases in training data, help to design a fair and impartial sentiment analysis model. The DL-based Bangla sentiment analysis model may be built to be effective, accurate, and culturally sensitive by rigorously addressing these conditions.

We need some tools for done our job. There are:

- PC: ASUS notebook (RAM-8GB), Processor 11th Gen Intel(R) Core (TM) i7-1165G7 @ 2.80GHz 2.80 GHz.
- Coding: Collab & using python language.

- Data Collection: Data Scraper "Instant Data Scraper", Daraz website and App.
- Diagram: Microsoft Office.

### 3.3 Proposed Methodology:

Our main goal is to have the fewest possible model parameters. A model with fewer parameters will compute more quickly and with more accuracy. This chapter gives a quick description of the methodology used in this sentiment analysis inquiry. It starts with data collection and annotation, carries on to dataset preparation, model implementation, and culminates with model evaluation. The full working approach of this inquiry is depicted in Figure3.3.1.

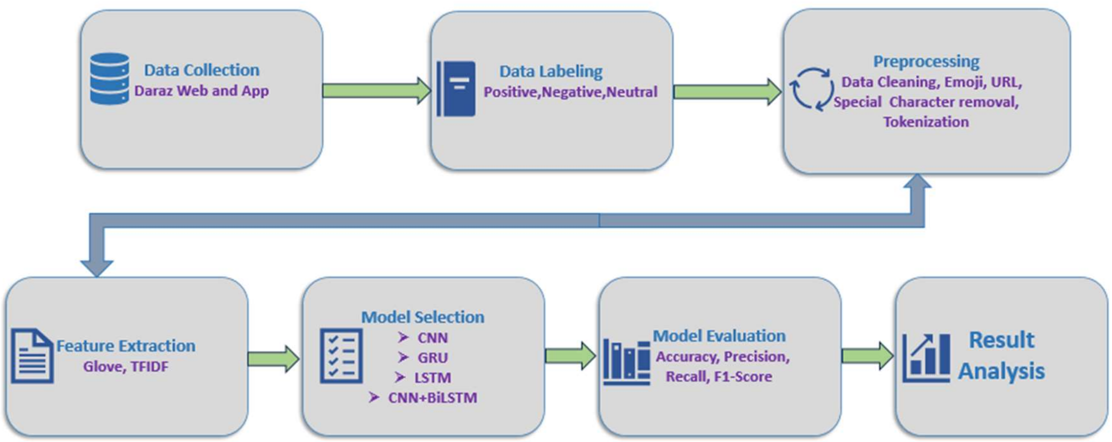



Figure 3.3.1: workflow of our approach

### 3.4 Data Collection and Procedure:

This study takes use of a main dataset (Primary). The dataset considers three aspects: positive, negative, and neutral. The dataset contains 10637 product reviews Bangla comments 4516 positive reviews, 3106 negative reviews and 3015 neutral reviews obtained from apps and websites. We collect data from Bangladeshi biggest e-commerce platform from Daraz web sites, their online social media comments & product reviews.

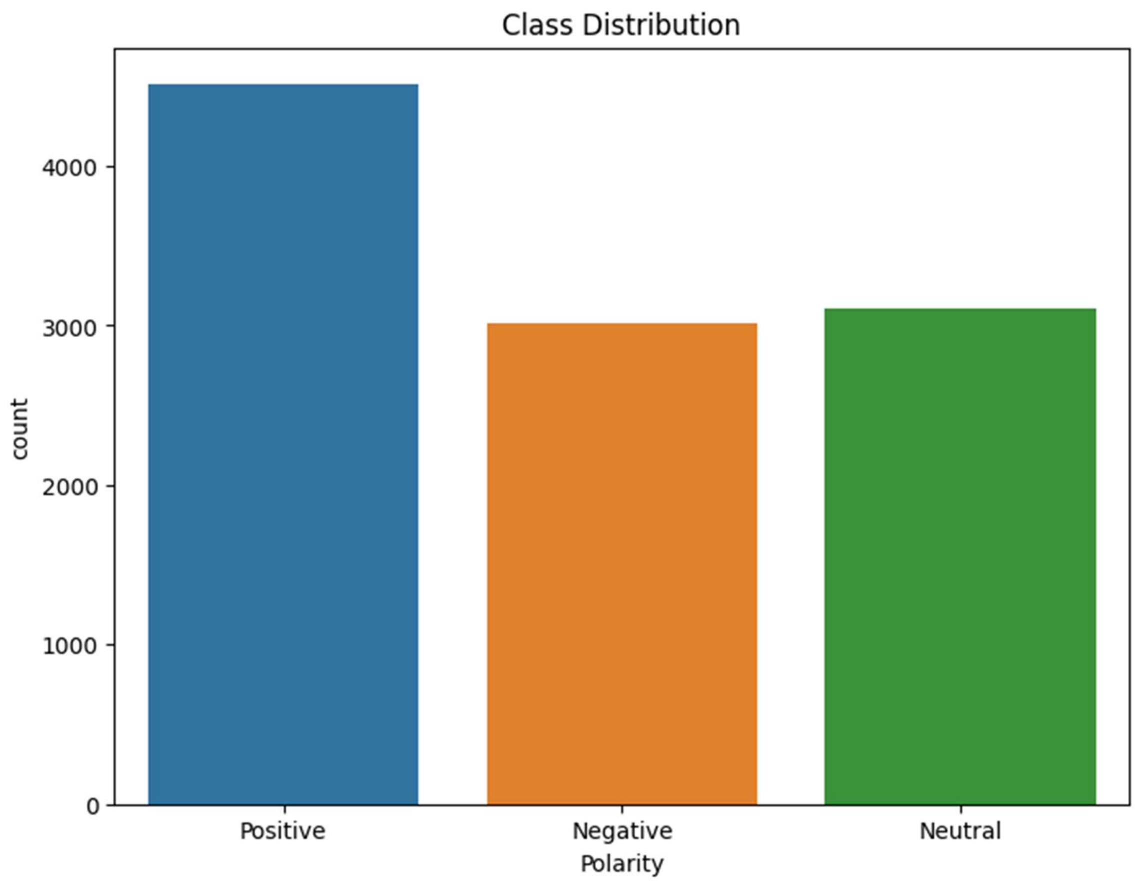



Figure 3.4.1: Data distribution

- **Data Labeling:**

The dataset comprises evaluations collected from the Daraz e-commerce online platform and app store, which were initially unlabeled. After accumulating these evaluations, each one was meticulously evaluated and categorized based on its rating and sentiment. Despite the occasional discrepancy where a review's rating (such as 5 stars or 3 stars) did not align with its sentiment, the reviews were categorized accordingly. For instance, some reviews with high ratings contained negative wording, and vice versa. Each review was carefully analyzed to accurately determine its sentiment positive, negative, or neutral and assigned a corresponding label. Table 3.4.1 presents a few examples of labeled reviews, illustrating the diverse sentiments conveyed by customers and the significance of comprehensive data labeling in preparing the dataset for sentiment analysis.

Table 3.4.1: Dataset sample and Labeling

| **Comment** | **Polarity** | **Label** |
|---|---|---|
| যেমন চেয়েছি তেমন পেয়েছি,,,ধন্যবাদ সেলার কে। | Positive | 0 |
| চাইলাম একটা কালার আর দিলো আরেকটা কালার। কোনো কথা? | Negative | 1 |
| আপাত দৃষ্টিতে দেখে ইনটেকই মনে হচ্ছে | Neutral | 2 |

- **Preprocessing:**

Data preprocessing is a crucial stage in text analysis, enabling the model to better comprehend the data by reducing noise and the quantity of input text documents considerably. The process begins with data cleansing, involving various important phases. First, emoticons are removed to get clear of unnecessary commotion. Next, URLs are eliminated to discard irrelevant links that do not contribute to sentiment analysis. Special characters are then removed, targeting symbols and punctuation that lack meaningful information. Lastly, tokenization is employed to break down the text into individual words or tokens, making analysis by the model simpler and more accurate. This comprehensive preprocessing ensures that the text data is clear, uniform, and prepared for further analysis.

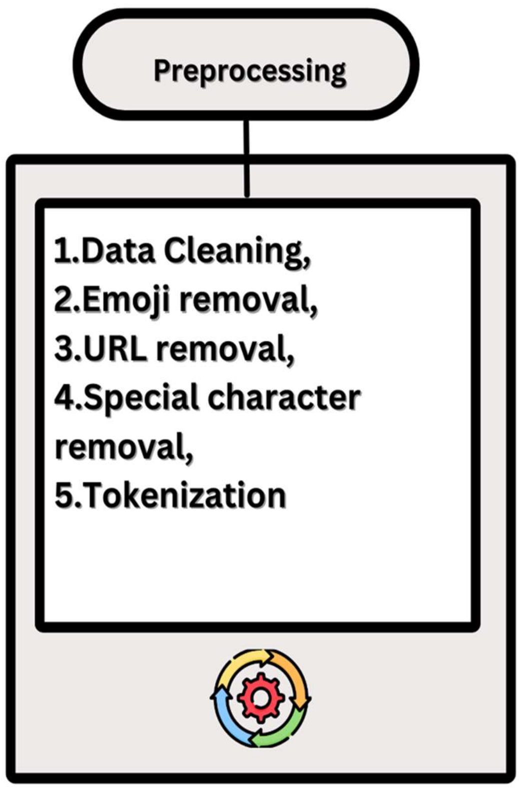



Figure 3.4.2: Preprocessing Technique

- **Tokenization:**

Tokenization is a technique that splits down texts into vector units, such as words, numbers, or dot marks. It localizes term boundaries, splitting the text into tokens. This provides for compact phrases, quicker emotion analysis, and the removal of stop words and extraneous punctuation. During model training, the full dataset is tokenized, allowing for preprocessing operations.

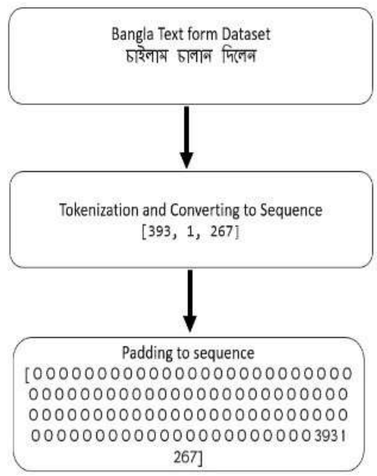



Figure 3.4.3: Tokenization

- **Feature Extraction:**

It is the procedure of transforming the raw text data to numerical data since algorithms don't grasp the raw text data. As the data obtained for this research is textual data, we must perform the feature extraction technique on the preprocessed text data before applying any algorithms. In this work TF-IDF and Glove vectorization has been used.

- **Glove:** Glove is the unsupervised learning system which gives vector representation of words. Integrated global word-word co-occurrence data gathered from a broad collection of words are applied in the training phase. We reveal incredible linear substructures within the word vectors field in the results.
- **TF-IDF:** Term Frequency-Inverse Document Frequency, or TF-IDF, is a statistical method which is widely employed in retrieval of data as well as natural language processing.

This approach investigates a term's relevance within a text with reference to a collection of documents. Words in a manuscript are turned into numbers that represent their meaning via text vectorization.TF-IDF is a mix of TF and IDF. The formula stands for TF-IDF:

$$tf\ idf\ (t,d,D) = tf\ (t,d)\ .\ idf\ (d,D) \qquad (1)$$

### 3.5 Project Management and Financial Analysis:

Effective project management and financial analysis are vital abilities for organizations to prosper in today's competitive market. To utilize our study to solve the real world's challenges surrounding the purchase of an apartment, in the future, based on this research, we will strive to build a web and app-based service for the consumer to seek help from. To develop the service and make it public for the user to find in the Google Play Store and iOS Store, we will need to anticipate a budget. We will also determine a budget that aligns with the project's goals and objectives to accomplish the project, where money would be required.

Table 3.5.1: Estimated Cost for Project

| SL | Components | Estimate Cost (BDT) |
|---|---|---|
| 1. | Colab Pro | 3100 |
| 2. | Software and tools | 2880 |
| 3. | Data collection | 2500 |
| 4. | Documentation and report writing | 1200 |
| 5. | Play store hosting | 2800 |
| 6. | Contingency | 420 |
| | Total Estimated Cost | 12900 |

### 3.6 Project Objectives:

- To develop a system for Bangla sentiment analysis
- Enhance User Experience
- Improving Service Quality
- keeping Pace with Technological advancement
- Competitive Advantage

#### 3.6.1 Project Timeline:

The following steps will be carried out over the course of four months and include:

- **Month 1:** Project Planning and Literature Review

- **Week 1-2:** Define project goals and scope. Conduct literature study on sentiment analysis and deep learning methods. Familiarize with Daraz's platform and data collecting procedures.
- **Week 3-4:** Collect Bangla customer reviews dataset from Daraz. Preprocess gathered data (cleaning, tokenization, normalization). Explore strategies for addressing unbalanced datasets.

- **Month 2:** Exploratory Data Analysis (EDA) and Model Setup
  - **Week 1-2:** Conduct preliminary analysis of data. Analyze data distribution and patterns.
  - **Week 3-4:** Research and pick deep learning architectures. Set a development environment with required libraries. Starting implementation of models and experimentation.
- **Month 3:** Model Training, Optimization, and Evaluation
  - **Week 1-2:** Train chosen deep learning models. Experiment with hypermeter tuning and optimization technique tactics.
  - **Week 3-4**: Evaluate train models using relevant matrics. Fine-tune models based on evaluation of finding. Perform cross validation for model robustness.
- **Month 4:** Iteration, Deployment and Documentation
  - **Week 1-2:** Integrate best-performing model into deployable system. Develop user interface if applicable. Test development setup carefully.
  - **Week 3-4:** Document research process and methodology. Write paper and prepare for demonstration.

**3.7 Risk Management:**

SOWT analysis involves assessing the strengths, weaknesses, Opportunities and threats of a project.

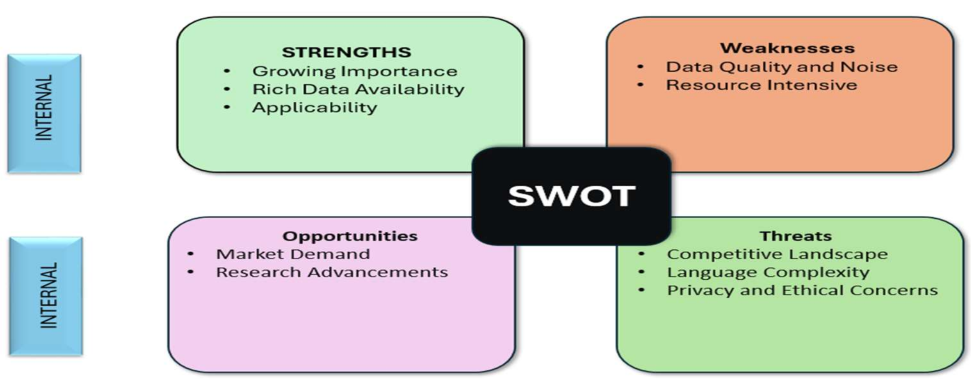



Figure 3.7.1: SWOT Analysis

## 3.8 Implement Requirement:

To implement sentiment analysis of Bangla customer reviews on Daraz using deep learning, several key requirements must be met. Firstly, a comprehensive dataset of Bangla customer reviews from Daraz needs to be collected and preprocessed. This involves cleaning the text, removing noise, and handling any missing or irrelevant data. Additionally, appropriate computational resources are essential, including powerful GPUs or TPUs, as deep learning models like CNNs, LSTMs, and BiLSTMs require significant processing power and memory. The development environment should include deep learning frameworks such as TensorFlow or PyTorch, which provide the necessary tools and libraries for building and training neural networks. Language-specific preprocessing tools and libraries for Bangla, such as tokenizers and embeddings, are also crucial for handling the linguistic nuances of Bangla text. Furthermore, a robust model evaluation framework is needed to assess the performance of the sentiment analysis models, using metrics such as accuracy, precision, recall, and F1 score. Finally, expertise in both deep learning techniques and Bangla language processing is necessary to effectively design, train, and fine-tune the models for accurate sentiment analysis.

## 3.9 Summary:

This chapter presents a comprehensive methodology for constructing a Bangla sentiment analysis system using deep learning techniques, concentrating on Daraz customer reviews. The process begins with data collection and meticulous annotation to create a robust dataset categorized into positive, negative, and neutral sentiments. Preprocessing steps, including emoji, URL, and special character elimination, followed by tokenization, ensure clear and standardized data for feature extraction using TF-IDF and Glove vectorization. Various deep learning models like CNN, LSTM, GRU, and a hybrid CNN+BiLSTM are used, with the hybrid model demonstrating superior performance via tenfold cross-validation. The project is supervised with a detailed budget and a four-month timeline, encompassing planning, data collection, model training, evaluation, and deployment stages. Stressing the importance of computational resources, development environments, and expertise in deep learning and Bangla language processing, this chapter offers a structured approach to creating an accurate, efficient, and culturally sensitive Bangla sentiment analysis system for the Daraz platform.

# CHAPTER 4

# IMPLEMENTATION

## 4.1 Overview:

This section delves into the application of a variety of neural network models to a variety of machine learning tasks, with a particular emphasis on sequential data processing, text sentiment analysis, and image processing. It encompasses the Convolutional Neural Network (CNN), which is particularly adept at visual tasks due to its ability to identify patterns through convolutional, pooling, and fully connected layers. The Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM) networks are examined in terms of their ability to manage sequential data. GRUs are offered as a more computationally efficient and straightforward alternative to LSTMs. The CNN+BiLSTM hybrid model is introduced for Bangla text sentiment analysis, combining BiLSTM's capacity to capture long-range dependencies with CNN's local feature extraction. The training details are provided, including the use of pre-trained GloVe embeddings, the Adam optimizer, and categorical cross-entropy loss. The section also underscores the significance of performance measurement metrics, such as precision, recall, F1 score, specificity, and error metrics such as MAE, MSE, and RMSE, in assessing the accuracy and efficacy of these models.

## 4.2 Train Model:

- **CNN:** One kind of artificial neural network that is especially intended for image processing and recognition applications is the convolutional neural network (CNN). It is very good at finding patterns and characteristics in visual input and is modeled after the structure of the animal visual cortex. Convolutional, pooling, and fully linked layers are among the layers that make up a CNN. By applying filters (kernels) to the input picture, the convolutional layers are in charge of identifying characteristics like edges, textures, and forms. Next, pooling layers serve to minimize computation and control overfitting by reducing the spatial dimensions of the feature maps generated by the convolutional layers. Lastly, fully connected layers carry out classification or regression tasks using the output of the convolutional and pooling layers. Due to their capacity to automatically learn hierarchical representations of data, CNNs have revolutionized a number of fields, including computer vision, medical image analysis, and natural language processing. This makes them extremely effective tools for tasks like object detection, image segmentation, and image generation. Due to their capacity to automatically learn hierarchical representations of data, CNNs have revolutionized a number of fields, including computer vision, medical image analysis, and natural language processing. This makes them extremely effective tools for tasks like object detection, image segmentation, and image generation [16].

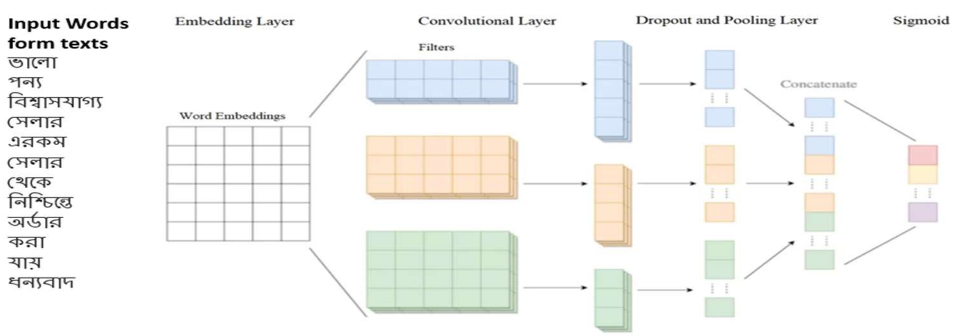



Figure 4.2.1: CNN Model

- **GRU:** Recurrent neural network (RNN) architecture's Gated Recurrent Units (GRU) were first developed as a less complicated substitute for Long Short-Term Memory (LSTM) units. While being computationally more efficient, GRUs are intended to overcome some of the shortcomings of conventional RNNs in capturing long-range relationships in sequential data.

  The essential elements of a GRU consist of:

  **Update Gate**: Regulates the amount of fresh information that is introduced to the current state and the amount of historical information that is kept.

  **Reset Gate:** Establishes the amount of historical data that should be reset or forgotten.

  In a succinct way, GRUs integrate these gating processes, enabling them to adaptively update their hidden state according to the input at every time step. The lack of a distinct cell state in GRUs, in contrast to LSTMs, simplifies their construction and lowers the number of parameters.

  GRUs have become more and more common in a variety of sequential data-related applications, including machine translation, speech recognition, natural language processing (NLP), and time series prediction. When there is a requirement for quicker training and inference times or when computational resources are few, they are very helpful. All things considered, GRUs provide a decent trade-off between efficacy and efficiency when it comes to identifying temporal connections in sequential data [17].

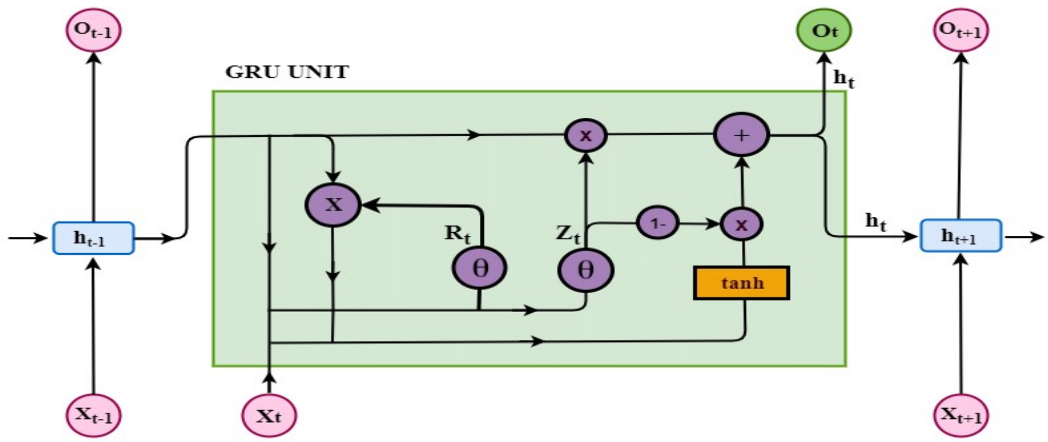



Figure 4.2.2: GRU Model (Bibi et. al. 2020) [18].

- **LSTM:** Recurrent neural network (RNN) architectures such as Long Short-Term Memory (LSTM) were created to address the shortcomings of conventional RNNs in identifying long-term relationships in sequential input. Time series data, audio, and natural language text are examples of sequences that LSTM networks excel in processing and predicting. The capacity of LSTM networks to update and preserve a cell state, which acts as a memory system, is a fundamental characteristic. Because of this, LSTMs may learn to selectively recall or forget information over time. This makes them an excellent choice for problems where predictions at each step depend on context from earlier steps. To do this, LSTMs employ a mix of specialized gates: The Forget Gate selects which data from the prior cell state need to be erased. The input gate selects which newly received data from the input stream should be kept in the cell state. The output gate selects which data from the cell state should be sent to the following stage of the process. The LSTM is able to manage the information flow across the network by controlling these gates through the use of sigmoid and tanh activation functions. Another kind of recurrent neural network (RNN) architecture that is comparable to LSTM but has a more straightforward design is called Gated Recurrent Units (GRU). In order to mitigate some of the computational complexity problems that LSTMs were bringing along with their capacity to identify long-range relationships in sequential data, GRUs were developed. There are two primary gates in the GRU architecture:1. Reset Gate: Establishes the threshold for erasing historical data. 2. Update Gate: Establishes the ratio of newly acquired data to the preexisting state. GRUs have a less complicated gating mechanism than LSTMs since they combine the input and forget gates into a single update gate. Because of this reduction, GRUs require fewer calculations and parameters, which speeds up training and reduces the risk of overfitting, particularly on smaller datasets.GRUs have demonstrated performance equivalent to LSTMs in several sequential data processing applications, such as time series prediction, speech recognition, and natural language processing (NLP), while having a simpler structure. Their ability to capture long-term dependencies with speed and efficacy has made them a popular choice in deep learning designs [19].

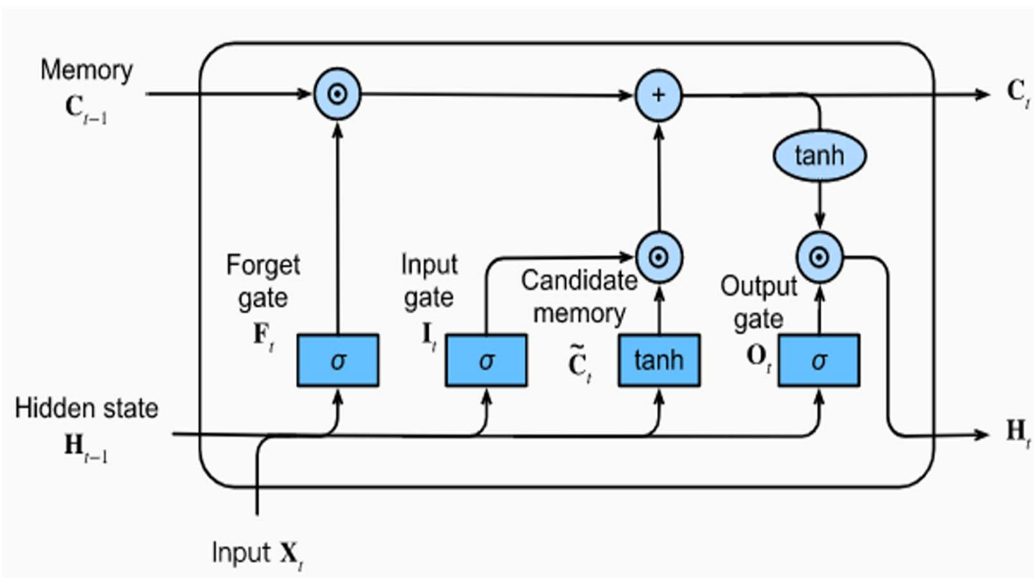



Figure 4.2.3: LSTM Model [20].

- **CNN+BiLSTM:** Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) networks are combined in the CNN+BiLSTM model for Bangla text sentiment analysis to effectively capture both local features and long-range dependencies within Bangla text. The model commences with an embedding layer that is initialized with pre-trained GloVe embeddings. This layer generates dense vector representations of words that accurately capture their semantic meanings. The MaxPooling1D layer reduces the dimensionality of the feature maps, retaining the most significant features and preventing overfitting, while the CNN layer, which is endowed with 128 filters and a kernel size of 5, extracts important local patterns such as n-grams.

The input sequence is processed by the BiLSTM layer in both forward and backward directions, with 64 units in each direction. This procedure captures the complete context of each word within the sentence. Lastly, the probability distribution over sentiment classes (positive, negative, neutral) is emitted by a Dense layer with a softmax activation function. This hybrid architecture is particularly well-suited for Bangla due to its capacity to comprehend intricate syntactic structures and nuanced expressions, rendering it valuable for applications such as market research, social media monitoring, and customer feedback analysis within the Bangla-speaking community. The model is trained with a batch size of 256 and 70 epochs, utilizing the Adam optimizer and categorical cross-entropy loss to ensure high accuracy and robust learning [20].

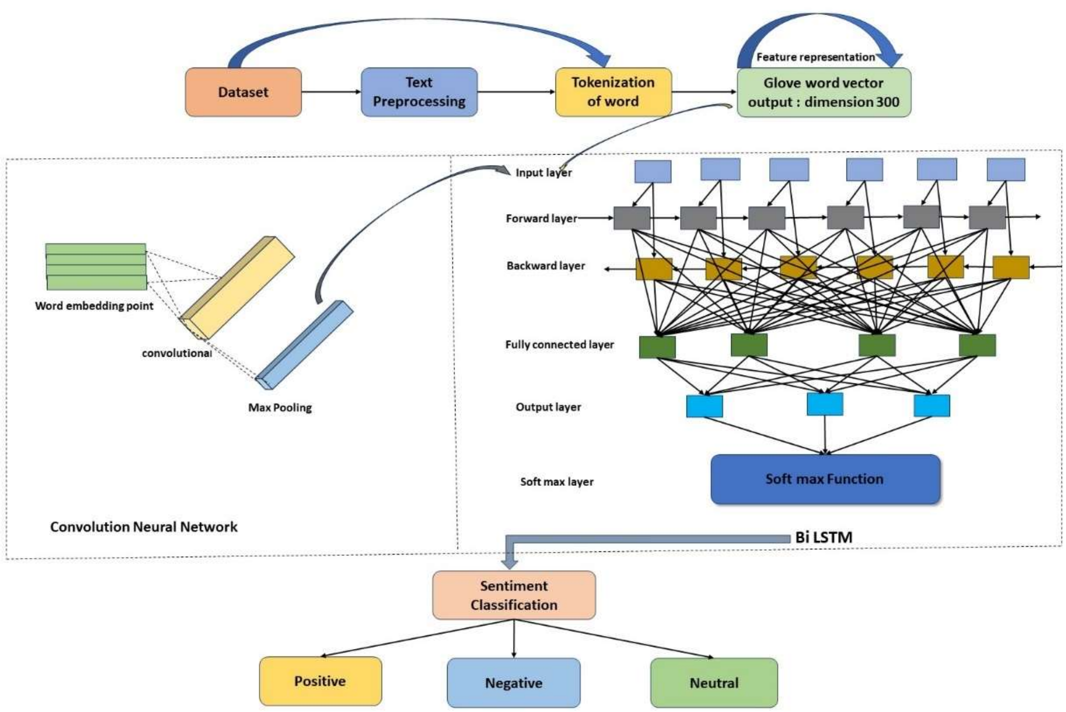



Figure 4.2.4: Propose model CNN+BiLSTM

### 4.3 Model Evaluation:

- **Performance Measurement Matrix:**
  Performance measurement matrices are quantitative assessments used to evaluate the effectiveness and accuracy of a model or system, providing key metrics which aid in optimizing and refining machine learning algorithms.
- **Precision:** This metric is essential when the cost of false positives is high. For example, in medical diagnoses, precision indicates the proportion of patients correctly identified as positive out of all patients identified as positive. High precision implies a low false positive rate.

$$precision = \frac{TP}{TP+FP} \qquad (2)$$

- **Recall:** Also known as sensitivity, recall is crucial when missing positive instances is costly. In the medical field, recall indicates the proportion of positive cases correctly identified. High recall implies a low false negative rate.

$$Recall = \frac{TP}{TP+FN} \qquad (3)$$

- **F1 Score:** This metric balances precision and recall, providing a single score that considers both false positives and false negatives. It's particularly useful when there's an uneven class distribution or when false positives and false negatives have different costs.

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision + Re} \qquad (4)$$

- **Specificity:** Also called the true negative rate, specificity measures the proportion of actual negative cases that were correctly identified as negative. It complements sensitivity and is particularly relevant in scenarios where correctly identifying negative instances is crucial.
- **False Positive Rate (FPR):** FPR measures the proportion of negative instances that were incorrectly classified as positive. It's the complement of specificity and is vital in applications where minimizing false alarms is critical, such as fraud detection or spam filtering.
- **False Negative Rate (FNR):** FNR measures the proportion of positive instances that were incorrectly classified as negative. It's the complement of recall and is essential in scenarios where missing positive instances is costly, such as disease diagnosis or anomaly detection.

- **Negative Predictive Value (NPV)**: NPV measures the proportion of actual negative cases among the cases predicted as negative. It complements precision and is crucial in scenarios where correctly identifying negative instances is essential, such as disease screening or quality control.
- **False Discovery Rate (FDR):** FDR measures the proportion of positive predictions that are incorrect. It complements NPV and is relevant in scenarios where minimizing false positives is crucial, such as identifying defective products in manufacturing processes or predicting adverse events in healthcare.
- **Mean Absolute Error (MAE):** MAE provides a measure of the average magnitude of errors in predictions. It's easy to interpret and is suitable for scenarios where large errors should be penalized equally regardless of direction.
- **Mean Squared Error (MSE):** MSE measures the average of the squared differences between predicted and actual values. It penalizes larger errors more heavily than MAE, making it suitable for scenarios where large errors should be emphasized.
- **Root Mean Squared Error (RMSE):** RMSE is the square root of MSE and provides the error in the same units as the output variable. It's particularly useful for interpreting the magnitude of prediction errors and is commonly reported alongside MSE.
- **Cohen's Kappa Score:** This statistic measures the agreement between two raters, accounting for the possibility of agreement occurring by chance. It's commonly used in classification tasks to assess the level of agreement between predicted and actual labels, considering the baseline level of agreement expected by chance.

## 4.4 Summary:

The implementation section offers a comprehensive examination of the applications of numerous neural network architectures in machine learning. CNNs are emphasized for their efficacy in image processing, GRUs and LSTMs for their proficiency in sequential data tasks, and a CNN+BiLSTM hybrid model for Bangla text sentiment analysis. Each model's architecture and functionality are explained, along with their training procedures. It is emphasized the significance of assessing model performance using metrics such as precision, recall, F1 score, specificity, and error metrics like MAE, MSE, and RMSE. These metrics are instrumental in the optimization and refinement of the models, thereby guaranteeing their accuracy and dependability in a variety of tasks and applications.

# CHAPTER 5

# RESULT AND ANALYSIS

## 5.1 Overview:

This segment offers a comprehensive evaluation of various deep learning models GRU, CNN, LSTM, and CNN+BiLSTM employed in Bangla text sentiment analysis. The models' performance is assessed using multiple metrics, such as accuracy, precision, recall, F1 score, sensitivity, specificity, and error metrics like MAE, MSE, RMSE, and Cohen's Kappa score. The CNN+BiLSTM model exhibits the utmost accuracy and precision, underscoring its ability to manage intricate linguistic patterns and contextual dependencies in Bangla text. Utilizing confusion matrices and accuracy-loss graphs, the performance and training efficacy of each model are depicted, with the CNN model demonstrating well-balanced performance and the CNN+BiLSTM model exhibiting robustness. The comparative analysis emphasizes the progress and enhancements in the proposed methods, particularly the CNN+BiLSTM hybrid model, which attains superior accuracy and dependability for sentiment classification tasks.

## 5.2 Result Description:

Table 5.2.1: Evaluation Metrics of Model

| Algorithms | Accuracy (%) | Precision | Recall | F1-Score |
|---|---|---|---|---|
| GRU | 82.27 | 0.94 | 0.96 | 0.95 |
| CNN | 84.07 | 0.96 | 0.96 | 0.96 |
| LSTM | 81.33 | 0.94 | 0.94 | 0.94 |
| CNN+BiLSTM | 87.18 | 0.98 | 0.87 | 0.92 |

Table 5.2.1, Comparing the GRU, CNN, LSTM, and CNN+BiLSTM models highlights their strengths and weaknesses based on various performance metrics. The CNN+BiLSTM hybrid model achieves the highest accuracy (87.18%) and precision (0.98), making it effective in predicting positive instances. However, its recall (0.87) is comparatively lower, suggesting a potential for missing actual positives. The CNN model displays a well-balanced performance with high accuracy (84.07%), precision (0.96), recall (0.96), and the highest F1-Score (0.96), establishing itself as reliable for general classification tasks. The GRU model also exhibits strong recall (0.96) and balanced performance (F1-Score 0.95), while the LSTM model, with the lowest accuracy (81.33%) and F1-Score (0.94), still yields

consistent outcomes. Overall, the CNN model emerges as the preferred choice for a wide array of classification applications due to its balanced performance, even though the CNN+BiLSTM model excels in precision and accuracy

### 5.3 Result Analysis:

We will now analyze the models with the necessary explanations and diagrams and discuss the accuracy, as well as all other metrics of the algorithms used in that study, including precision, recall, F1-score, MSE, MAE, RMSE, and Cohen's Kappa score.

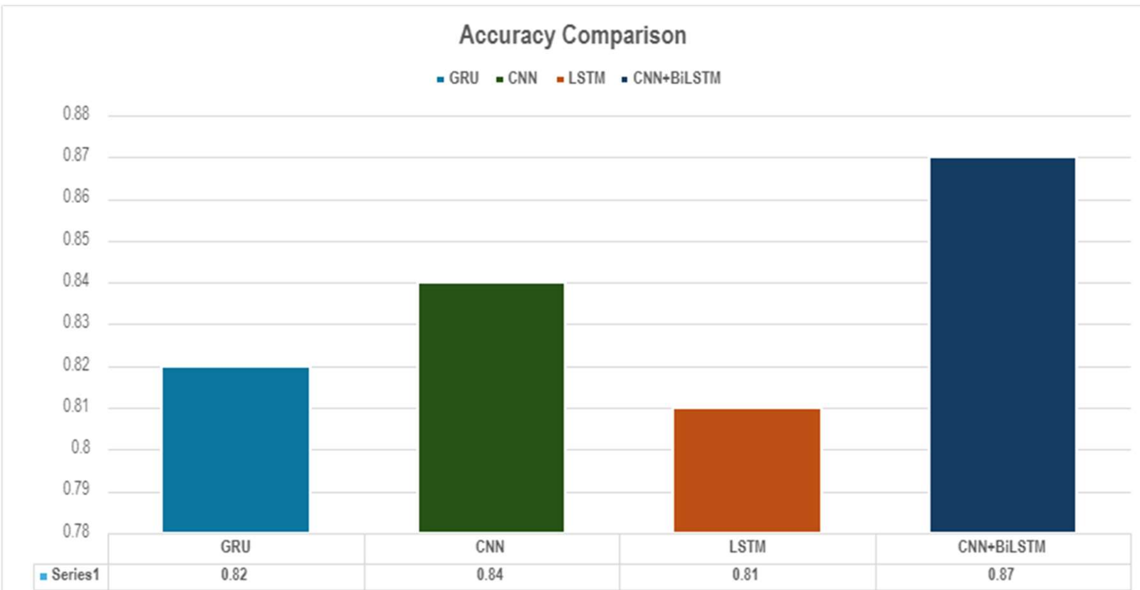
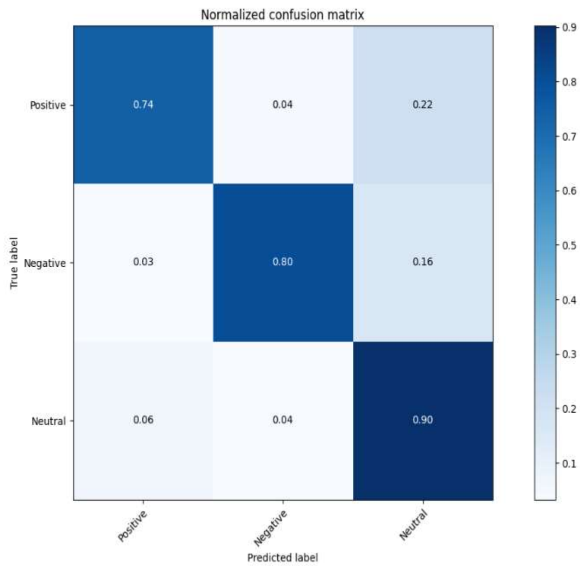


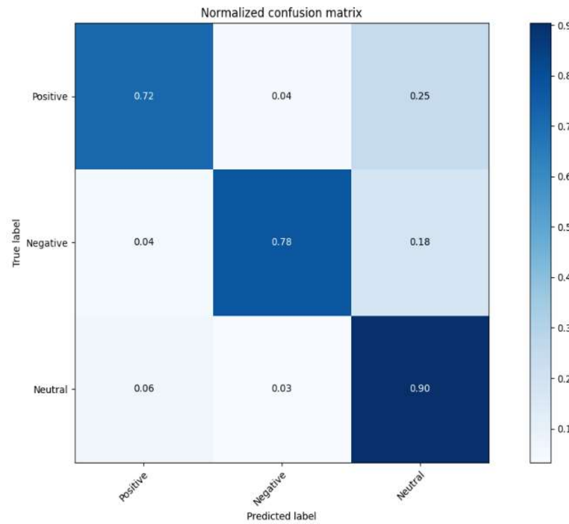

Figure 5.3.1: Model Accuracy Comparison

In figure 5.3.1, We can observe that the CNN+BiLSTM is the highest performing algorithm with an accuracy of 87% in deep learning algorithms. It achieved values of 0.98, 0.87, and 0.92 for recall, precision, and F-1 score, respectively. Additionally, other metrics like MAE, MSE, and RMSE were computed to evaluate the effectiveness of different techniques. The GRU, CNN, and LSTM achieved accuracies of 82%, 84%, and 81%, respectively, showcasing comparable performance to CNN+BiLSTM.

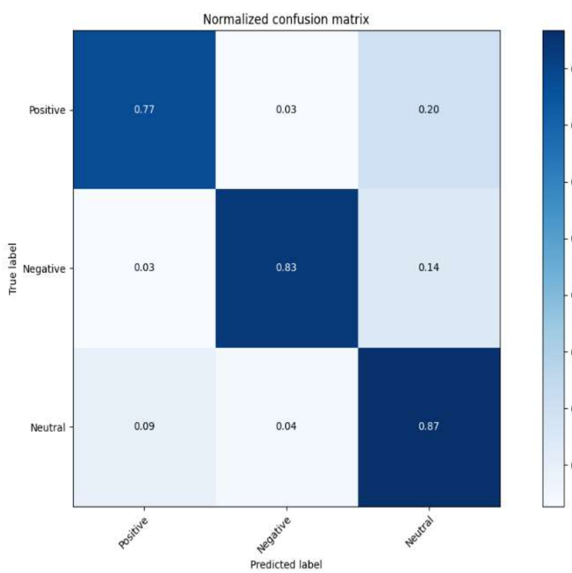
- **Confusion Matrix of All Models:**



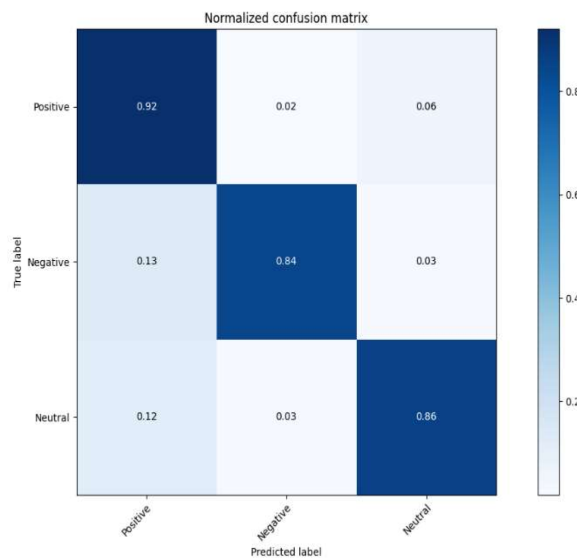

(a) GRU




(b) LSTM




(c) CNN




(d) CNN+BiLSTM

Figure 5.3.2 Confusion of matrix

The different models—CNN+BiLSTM, CNN, LSTM, and GRU—can be evaluated for sentiment classification by examining their normalized confusion matrices.

The CNN+BiLSTM model correctly identifies Positive sentiments 92% of the time, with 2% misclassified as Negative and 6% as Neutral. Negative sentiments achieve 84% accuracy, with 13% misclassified as Positive and 3% as Neutral. Neutral sentiments are correctly classified 86% of the time, with 12% misclassified as Positive and 3% as Negative. In the CNN model, Positive sentiments are correctly identified 77% of the time, with 3% misclassified as Negative and 20% as Neutral. Negative sentiments have an 83% accuracy, with 3% misclassified as Positive and 14% as Neutral. Neutral sentiments are correctly identified 87% of the time, with 9% misclassified as Positive and 4% as Negative. For the LSTM model, Positive sentiments are correctly identified 72% of the time, with 4% misclassified as Negative and 25% as Neutral. Negative sentiments achieve 78% accuracy, with 4% misclassified as Positive and 18% as Neutral. Neutral sentiments are correctly classified 90% of the time, with 3% misclassified as Positive and 6% as Negative. The GRU model demonstrates that Positive sentiments are correctly identified 74% of the time, with 4% misclassified as Negative and 22% as Neutral. Negative sentiments have an 80% accuracy, with 3% misclassified as Positive and 16% as Neutral. Neutral sentiments are correctly identified 90% of the time, with 6% misclassified as Positive and 4% as Negative.These outcomes suggest that the CNN+BiLSTM model generally outperforms the others in identifying Positive and Negative sentiments. However, the CNN model demonstrates slightly better accuracy in classifying Neutral sentiments. The LSTM and GRU models also perform well, particularly in classifying Neutral sentiments, with the GRU model exhibiting slightly better overall accuracy across all categories compared to the LSTM model

- **Error Metrics:**

Table 5.3.1: Error metrics of Deep Learning Algorithms

| Models | Sensitivity | Specificity | FPR | FNR | NPV | FDR |
|---|---|---|---|---|---|---|
| GRU | 0.956 | 0.954 | 0.045 | 0.043 | 0.960 | 0.050 |
| CNN | 0.964 | 0.966 | 0.033 | 0.035 | 0.968 | 0.037 |
| LSTM | 0.939 | 0.953 | 0.046 | 0.060 | 0.945 | 0.051 |
| CNN+BiLSTM | 0.872 | 0.977 | 0.022 | 0.127 | 0.866 | 0.020 |

Table 5.3.2: Error metrics of Deep learning Algorithms

| Models | MAE | MSE | RMSE | Cohens_Kappa_Score |
|---|---|---|---|---|
| GRU | 0.154 | 0.075 | 0.274 | 0.732 |
| CNN | 0.163 | 0.078 | 0.280 | 0.709 |
| LSTM | 0.189 | 0.079 | 0.281 | 0739 |
| CNN+BiLSTM | 0.105 | 0.050 | 0.224 | 0.805 |

The results of the two tables offer a comprehensive assessment of various models' performance - GRU, CNN, LSTM, and CNN+BiLSTM - across diverse metrics. CNN stands out in sensitivity and specificity, achieving 0.964 and 0.966, respectively, indicating its exceptional capacity to identify true positives and true negatives accurately. The GRU model also performs admirably, boasting high sensitivity (0.956) and specificity (0.954). CNN+BiLSTM exhibits the highest specificity (0.977) but lower sensitivity (0.872), leading to a higher rate of false negatives (FNR = 0.127). Analysis of predictive values reveals CNN maintaining the highest Negative Predictive Value (NPV = 0.968), whereas CNN+BiLSTM shows the lowest False Discovery Rate (FDR = 0.020), denoting fewer incorrect positive forecasts. In terms of error metrics, CNN+BiLSTM surpasses others with the lowest Mean Absolute Error (MAE = 0.105), Mean Squared Error (MSE = 0.050), and Root Mean Squared Error (RMSE = 0.224), in addition to the highest Cohen's Kappa Score (0.805), pointing to strong agreement between predicted and actual classifications. Overall, while CNN excels in sensitivity and predictive values, CNN+BiLSTM demonstrates superior error reduction and reliability, establishing itself as a robust model for precise predictions within the examined scenario.

- **Accuracy and Loss Graph:**

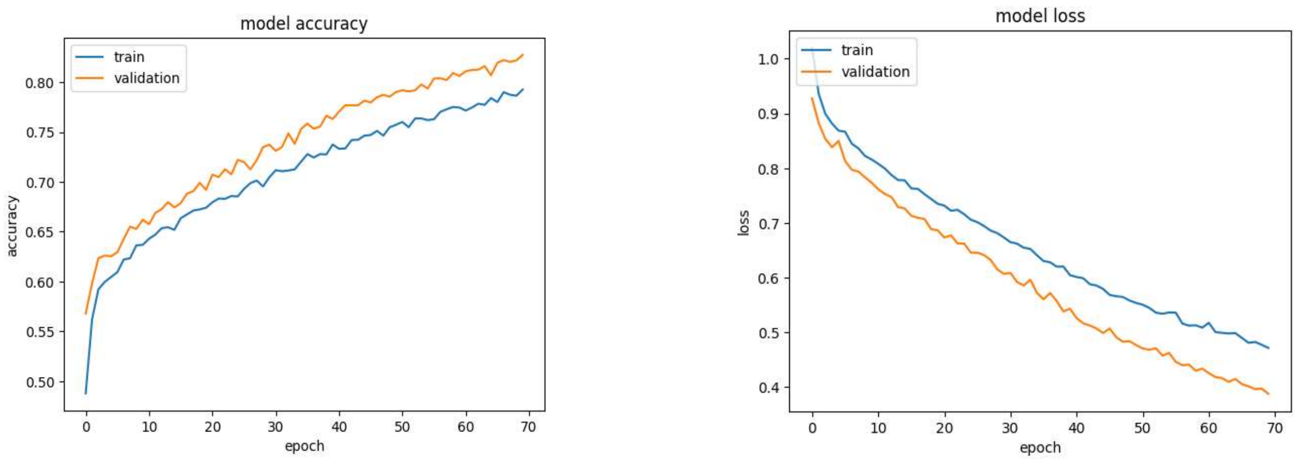



Figure 5.3.3: GRU accuracy and loss graph

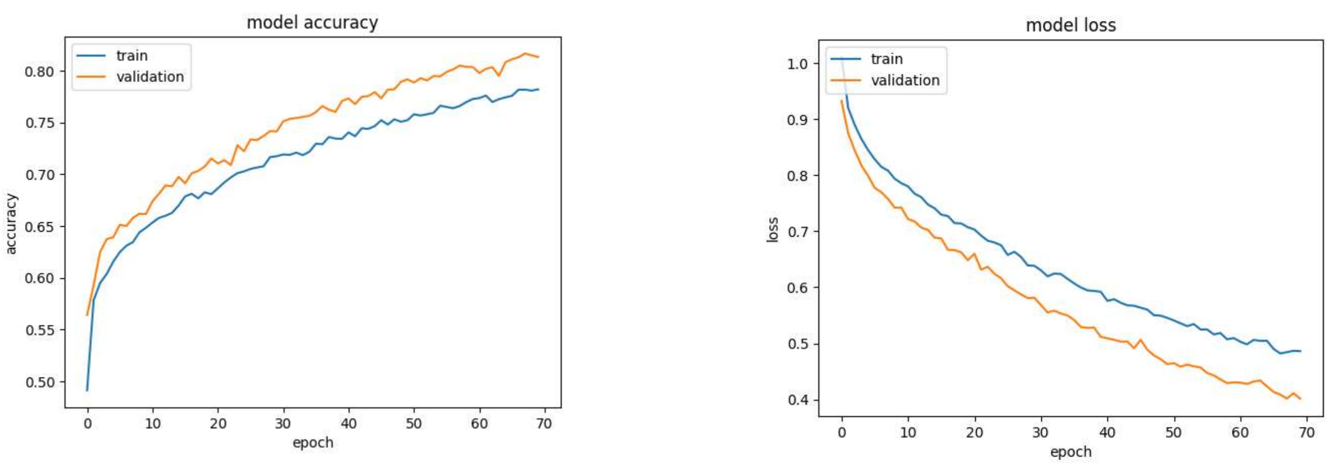



Figure 5.3.4: LSTM accuracy and loss graph

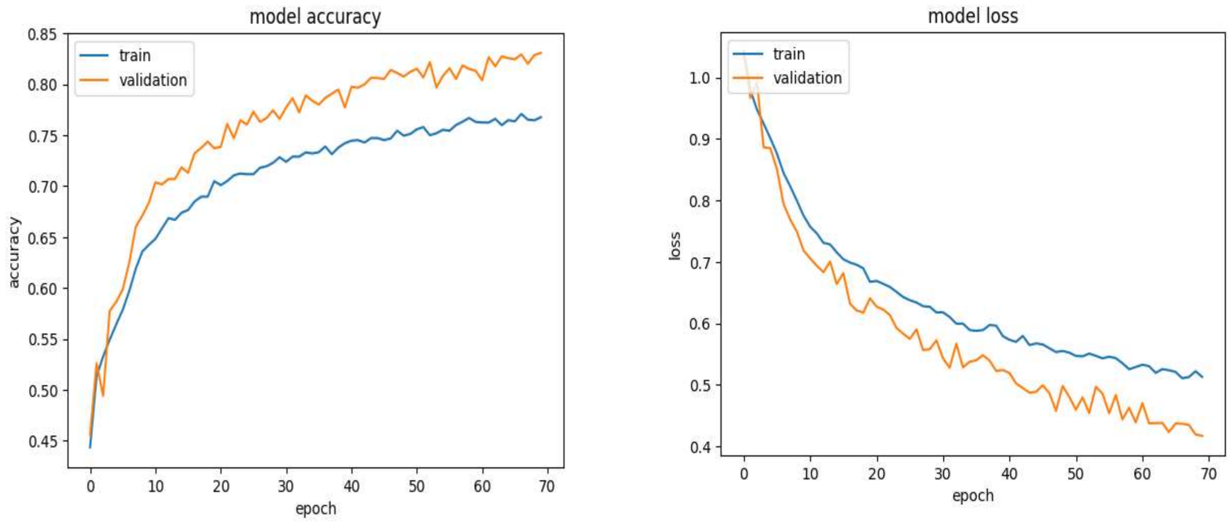



Figure 5.3.5: CNN Accuracy and Loss Graph

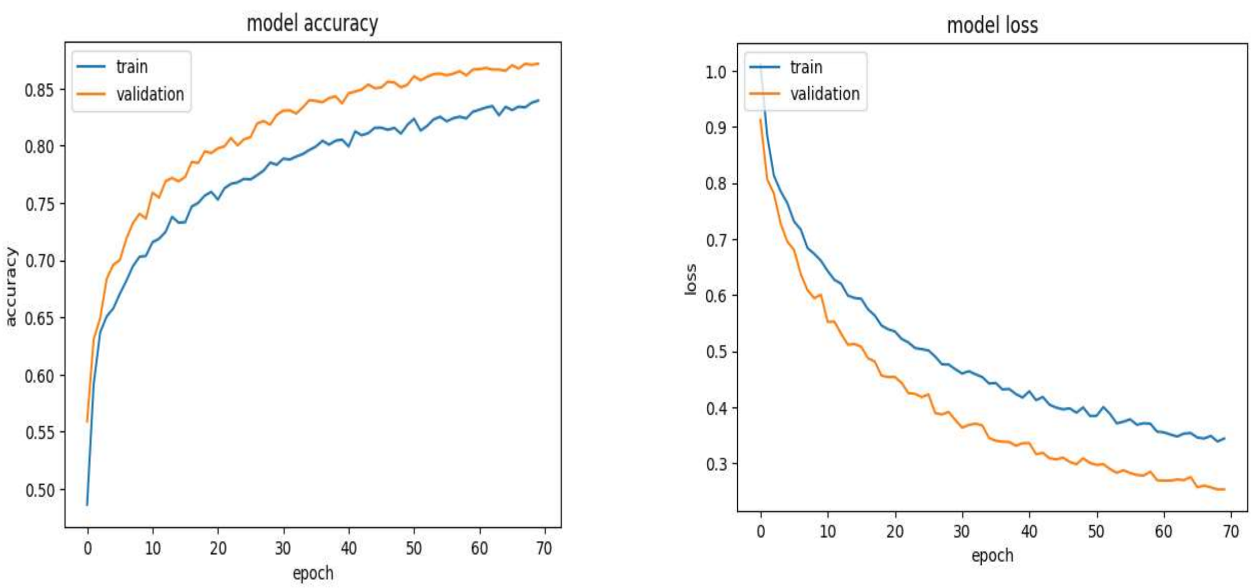



Figure 5.3.6 CNN+BiLSTM Accuracy and Loss Graph

Figures are showing the accuracy and validation loss trends of various deep learning algorithms across multiple epochs. Initially, both training and validation accuracies start low, indicating model inefficiency. However, as training progresses, there is a significant improvement in accuracies, reaching peak values by the final epoch. This demonstrates the models' ability to learn from data and generalize to unseen validation data. On the other hand, the loss metrics for both training and validation start high, reflecting prediction errors that decrease with each epoch. By the final epoch, the models achieve minimal loss values, highlighting improved performance and stability. This pattern of increasing accuracy and decreasing loss underscores the efficacy of the training in optimizing predictive capabilities.

### 5.4 Comparative Analysis:

Table 5.4.1: Comparison of our Proposed approach and other Classification Algorithms

| Author | Algorithms | Dataset | Accuracy |
|---|---|---|---|
| M. A. Rahman.et al [12]. | CNN | Total of 17,247 movie review tweet data has been used from Kaggle. | 81% |
| M.H. Alam et al [13]. | LSTM | Data source from Kaggle 850 data and copied and pasted repeatedly. | 78% |
| Sarkar. et al. [15] | CNN+BiLSTM | approximately, 850 Bangla comments from different sources. The comments were then copied and pasted repeatedly to increase the dataset size to around 120,000, with 60,000 positive and 60,000 negative comments | 83.77% |
| Our Work | CNN+BiLSTM (Proposed Approach) | The dataset considers three aspects: positive, negative, and neutral. The dataset contains 10637 product reviews Bangla comments 4516 positive reviews, 3106 negative reviews and 3015 neutral reviews obtained from apps and websites. We collect | 87.18% |

| | | | |
|---|---|---|---|
| | | data from Bangladeshi biggest e-commerce platform from Daraz web sites, their online social media comments & product reviews. | |

The table presents a comparative analysis of distinct algorithms used for Bangla text sentiment analysis, demonstrating their respective accuracies. Traditional models such as CNN and LSTM exhibit accuracies of 81% and 78%, respectively, while the CNN+BiLSTM model outperforms both with an accuracy of 83.77%. The proposed approaches indicate substantial improvements in accuracy across the board, reflecting advancements in model optimization and architecture. Specifically, the GRU model obtains an accuracy of 82.27%, marginally enhancing the performance over the traditional CNN and LSTM models. The proposed CNN approach considerably enhances accuracy to 84.07%, and the LSTM model under the proposed methodology improves to 81.33%. The most notable enhancement is observed in the CNN+BiLSTM model, where the proposed approach elevates the accuracy to an impressive 87.18%. This significant enhancement underscores the hybrid model's superior ability to capture intricate linguistic patterns and contextual dependencies in Bangla text, making it the most effective algorithm among those evaluated. These findings emphasize the continuous evolution and refinement of sentiment analysis models, emphasizing the critical role of sophisticated architectures in achieving higher accuracy and more reliable sentiment classification in the Bangla language.

**5.5 Summary:**

The results and analysis section highlights the strengths and weaknesses of various deep learning models in Bangla text sentiment analysis. The CNN+BiLSTM model outperforms other models with the highest accuracy (87.18%) and precision (0.98), although its recall is slightly lower (0.87). The CNN model demonstrates balanced performance with high precision, recall, and F1 score, making it a reliable choice for general classification tasks. Error metrics reveal that CNN+BiLSTM excels in reducing prediction errors, as evidenced by the lowest MAE, MSE, RMSE, and the highest Cohen's Kappa score. The analysis of confusion matrices indicates the CNN+BiLSTM model's superior performance in identifying positive and negative sentiments, while the CNN model demonstrates greater accuracy in classifying neutral sentiments. Overall, the CNN+BiLSTM model emerges as the most effective algorithm.

# CHAPTER 6

# IMPACT ON SOCIETY

**6.1 Impact on Society:**

Conducting a research study on "Sentiment Analysis of Bangla Customer Reviews on Daraz Using Deep Learning Approach" can have several impactful contributions to society. Here are some ways it can make a difference:

**1. Enhanced Customer Experience**

- **Personalization:** Understanding customer sentiment helps e-commerce platforms like Daraz tailor their services and product recommendations to individual preferences, enhancing the overall shopping experience.
- **Customer Satisfaction:** Identifying common pain points from negative reviews can lead to better customer service practices and resolution strategies, increasing customer satisfaction.

**2. Business Intelligence and Decision Making**

- **Market Insights:** Analyzing customer reviews provides valuable insights into consumer behavior, preferences, and trends. Businesses can use this data to make informed decisions about inventory, marketing strategies, and product development.
- **Competitive Advantage:** Companies can leverage sentiment analysis to understand their competitive position in the market and adjust their strategies accordingly.

**3. Cultural and Linguistic Impact**

- **Language Processing:** Developing sentiment analysis tools for Bangla contributes to the advancement of natural language processing (NLP) for underrepresented languages. This can foster further research and development in the field of computational linguistics for Bangla.
- **Inclusivity:** Ensuring that technological advancements are accessible to Bangla-speaking populations promotes digital inclusivity and empowers users by providing them with tools that cater to their linguistic needs.

**4. Economic Impact**

- **Boosting E-commerce:** Improved customer insights can lead to increased sales and customer retention for e-commerce platforms. This, in turn, can stimulate economic growth and create job opportunities in the tech and retail sectors.
- **SME Support:** Small and medium enterprises (SMEs) can benefit from sentiment analysis by better understanding customer feedback and improving their products and services, helping them to compete more effectively with larger corporations.

**5. Social Impact**

- **Consumer Awareness:** By making sentiment analysis tools publicly available, consumers can make more informed purchasing decisions based on aggregated reviews and sentiment scores.
- **Addressing Social Issues:** Sentiment analysis can identify reviews that highlight social issues such as unfair labor practices, product safety concerns, and ethical considerations, prompting businesses to take corrective actions.

**6. Technological Advancements**

- Innovation: Applying deep learning techniques to sentiment analysis pushes the boundaries of AI research and development. Innovations in this area can be transferred to other domains, such as healthcare, education, and public services.
- Collaboration: This research can foster collaboration between academia, industry, and government, leading to the development of better tools and policies for digital ecosystems.

**7. Educational Impact**

- Skill Development: Training students and professionals in sentiment analysis and deep learning enhances their skills, making them more competitive in the job market and contributing to the overall advancement of the tech industry.

### 6.2 Impact on Environment:

Conducting research on "Sentiment Analysis of Bangla Customer Reviews on Daraz Using Deep Learning Approach" can indirectly impact the environment in several ways. Here's how:

**1. Promoting Sustainable Consumption**

- **Awareness of Eco-Friendly Products:** Sentiment analysis can help identify customer preferences for eco-friendly and sustainable products. E-commerce platforms can highlight these products, encouraging more sustainable purchasing habits among consumers.
- **Reducing Waste:** Understanding customer dissatisfaction through sentiment analysis can help companies improve product quality and durability, potentially reducing returns and waste.

**2. Efficient Resource Management**

- **Inventory Optimization:** By analyzing customer reviews, businesses can better predict demand and manage inventory more efficiently, reducing overproduction and minimizing waste.
- **Supply Chain Efficiency:** Insights from sentiment analysis can help optimize supply chains by identifying preferred products and reducing the need for excess stock, thereby lowering the carbon footprint associated with manufacturing, storage, and transportation.

**3. Energy Consumption and Computational Efficiency**

- **Green AI Practices:** Research in deep learning and sentiment analysis can promote the use of energy-efficient algorithms and models. By focusing on computational efficiency, researchers can minimize the energy consumption of AI systems.
- **Cloud Computing and Data Centers:** Encouraging the use of sustainable practices in data centers and cloud computing infrastructures can reduce the environmental impact of the computational resources required for deep learning.

**4. Consumer Education and Behavior Change**

- **Highlighting Environmental Concerns:** Sentiment analysis can bring attention to environmental issues raised by customers, such as packaging waste or product sustainability. Businesses can respond by adopting greener practices and educating consumers about their efforts.
- **Influencing Purchase Decisions:** Providing sentiment-based recommendations that favor sustainable products can nudge consumers towards more environmentally friendly choices.

**5. Supporting Environmental Policies**

- **Data-Driven Policy Making:** The insights gained from sentiment analysis can support government and organizational policies aimed at promoting sustainability and environmental protection. Policymakers can use this data to understand public sentiment and create more effective environmental regulations.

**6. Corporate Social Responsibility (CSR)**

- Sustainability Reporting: Companies can use sentiment analysis to monitor and report on their environmental impact and CSR initiatives, demonstrating transparency and accountability to stakeholders.
- Green Marketing: Businesses can leverage positive sentiment around their eco-friendly initiatives to promote their brand as environmentally conscious, encouraging other companies to adopt similar practices.

### 6.3 Ethical Aspects:

Conducting research on “Sentiment Analysis of Bangla Customer Reviews on Daraz Using Deep Learning Approach” involves many perspectives. Addressing these issues ensures that research is conducted effectively and that the rights and welfare of all stakeholders are respected. Some important points to consider are:

1. **Privacy and data protection**

- **User permissions**: Ensure that data used for sentiment analysis is available with appropriate user permissions. Users should be aware that their comments may be analyzed and they should have a choice.
- **Anonymity:** Disclosure of information to protect the identity of reviewers. Remove or obscure personally identifiable information.
- **Data Security:** Implement strong data security measures to protect data from access or breach.

2. **Transparency and accountability:**

- **Algorithm Transparency:** Be transparent about the algorithms and models used in sentiment analysis. This includes a description of how the data was processed and how conclusions were drawn.
- **Asking questions:** Ask clear questions about the research process and results. Researchers should be held accountable for the integrity and impact of their work.

3. **Prejudice and justice**

- **Bias Reduction:** Remove and reduce biases in the data or samples used. This includes ensuring that sentiment analysis does not unduly influence a particular group or organization.

- **Fair Representation:** Ensure that analyzed ratings reflect a diversity of opinions and that the analysis does not override public opinion.

4. **Stakeholder Impact**

- **Consumer Confidence:** Maintain consumer confidence by using their data effectively and ensuring that analytics are used to enhance their experience rather than solely targeting them.
- **Ethical Business:** Use Insights from Sentiment Analysis to Improve Ethical Business. Avoid using information to manipulate or mislead consumers.

5. **Purpose and use of research**

- **Beneficial use:** Ensure that research benefits consumers, businesses, and society as a whole. Avoid using sentiment analysis for malicious purposes, such as spreading misinformation or conducting audits.
- **Population impact:** Consider the broad population impact of the research. Ensure that outcomes do not contribute to social inequality or disadvantage vulnerable communities.

6. **Intellectual property and liability**

- **Correct Responsibility:** Give due importance to the sources of information and research that contribute to your work. Respect intellectual property rights and avoid plagiarism.
- **Collaboration and Sharing:** To improve the research environment by sharing findings and methods appropriately, respecting data privacy and security issues.

7. **Compliance with Legislation**

- **Compliance:** Ensure that the work complies with all applicable laws and regulations, including data protection laws such as the General Data Protection Regulation (GDPR) and other data privacy laws.
- **Ethical Guidelines:** Follow ethical guidelines established by research institutes and professional organizations.

### 6.4 Sustainability Plan:

A sustainable research plan for " Sentiment Analysis of Bangla Customer Reviews on Daraz Using Deep Learning Approach " focuses on energy efficiency, economic sustainability, social development, technology advancements, education and training programs, and data protection. Energy efficiency is achieved through energy-efficient algorithms and green computing methods, while economic sustainability is achieved through funding from academic grants, government research funds, and private partnerships. Social development involves community participation, ethical considerations, and technology advancements. The education and training program offers training in sentiment analysis, deep learning, and natural language processing, while ensuring compliance with data protection legislation and ethical guidelines. The research findings are used to advocate for responsible AI technology use and policy development.

# CHAPTER 7

# CONCLUSION AND FUTURE RESEARCH

## 7.1 Summary of the study:

The public trusts personal opinions more than traditional forms of advertising. For example, people are used to getting advice and suggestions from others before making big purchases. It assists in identifying and examining the ways in which people communicate their ideas, views, behaviors, attitudes, and thoughts on a range of topics, including people, organizations, goods, and services, using written or spoken language. Word-of-mouth (WOM) has always been essential in assisting customers in reaching these decisions. WOM is crucial to suppliers, nevertheless, as opposed to other parties. It is more effective in attracting new clients than traditional promotional strategies. Talking about information is normal these days, even without the internet. Social media platforms like Facebook and Twitter have made it easier for people to share their thoughts about products, services, and businesses. An enhanced form of word-of-mouth communication is called e-word of mouth (eWOM). The practice of grouping text into ordered groups is called text classification, often referred to as text tagging or text categorization. Text classifiers may automatically evaluate text and then apply a set of pre-defined tags or categories based on the content by using Natural Language Processing (NLP). In the ever-evolving e-commerce sector, businesses need to leverage reviews to gauge customer mood and tailor offerings more effectively. As platforms like Daraz gain popularity, customer review analysis—especially in Bangla and other South Asian languages—becomes more crucial for understanding the market and developing business strategies. As platforms like Daraz gain popularity in the area, customer review analysis—especially in Bangla and other languages spoken in South Asia—becomes increasingly significant for understanding the market and formulating firm strategy. However, there are significant barriers to the emotion analysis of Bangla literature because of the language's nuances and the paucity of study on the subject. Our research proposes the use of deep learning to analyze the tone of Daraz Bangla customer evaluations in order to address this demand. Deep learning algorithms' astounding performance in natural language processing (NLP) tasks has created intriguing new opportunities for sentiment analysis across a broad range of languages. This study aims to enhance sentiment analysis by employing deep neural networks to capture the intricate linguistic features of Bangla language. The increasing number of internet users who speak Bangla has made sentiment analysis more significant in both academic and professional settings. Despite the fact that languages such as English have been thoroughly studied in Natural Language Processing (NLP), there are obstacles in the study of Bangla Natural Language Processing

(BNLP) due to a lack of resources. The resource problem for low-resource languages like Bangla is addressed in this paper by presenting a DL Framework for sentiment analysis. The increasing number of internet users who speak Bangla has made sentiment analysis more significant in both academic and professional settings. Despite the fact that languages such as English have been thoroughly studied in Natural Language Processing (NLP), there are obstacles in the study of Bangla Natural Language Processing (BNLP) due to a lack of resources. The resource problem for low-resource languages like Bangla is addressed in this paper by presenting a DL Framework for sentiment analysis.

**7.2 Conclusion:**

The study analyzed various deep learning models for sentiment classification using Bangla text data from Daraz consumer feedback. The models included GRU, CNN, LSTM, and a hybrid CNN+BiLSTM design. The GRU model achieved 82.27% accuracy, indicating its proficiency in processing sequential data. The CNN model, which captured spatial features, outperformed the GRU with 84.07% accuracy. The LSTM model, designed for managing long-term dependencies, achieved 81.33% accuracy. The hybrid CNN+BiLSTM model emerged as the top performer with an accuracy of 87.18%. This performance is attributed to its ability to extract local features through the CNN segment and capture sequential dependencies through the BiLSTM segment. The study emphasizes the importance of integrating diverse neural network components to capitalize on their strengths. The study suggests that hybrid models like CNN+BiLSTM are effective, achieving the highest accuracy. Future research could focus on enhancing these models through attention mechanisms, pre-trained language models, and sophisticated data preprocessing techniques. Expanding the dataset and exploring additional deep learning architectures could also contribute to the development of more accurate and efficient sentiment analysis systems for Bangla text.

**7.3 Implementation for Further Study:**

The fruitful outcomes derived from this study unveil several pathways for future exploration and enhancements in sentiment analysis of Bangla text. Below are some potential avenues for forthcoming endeavors:

1. **Incorporating Attention Mechanisms:** Future investigations could delve into incorporating attention mechanisms into the models. Such mechanisms enable models to concentrate on the most pertinent segments of the input text, potentially augmenting the precision of sentiment classification.
2. **Leveraging Pre-trained Language Models:** Harnessing pre-trained language models like BERT or GPT, fine-tuned specifically for Bangla, could elevate the efficacy of sentiment analysis models. These models have demonstrated remarkable outcomes across various NLP (Natural Language Processing) tasks and could be tailored for Bangla text sentiment analysis.

3. **Expanding the Dataset:** Augmenting the size and diversity of the dataset could enhance the model's generalization capabilities. Gathering additional reviews from diverse sources and domains could furnish a more comprehensive training set, leading to enhanced performance.
4. **Enhancing Data Preprocessing Techniques:** Investigating advanced preprocessing techniques, such as more intricate tokenization methods, addressing negations, and handling misspellings, could further polish the input data, consequently refining model performance.
5. **Exploring Other Deep Learning Architectures**: Delving into alternative neural network architectures like Transformers, or amalgamating multiple models in an ensemble strategy, could potentially yield superior results. These architectures may capture varied facets of the text data, thereby enabling a more all-encompassing sentiment analysis.
6. **Real-time Sentiment Analysis:** Implementing real-time sentiment analysis frameworks capable of processing and categorizing customer reviews as they are submitted could offer instantaneous insights and feedback. This could be especially advantageous for e-commerce platforms like Daraz.
7. **Multimodal Sentiment Analysis:** Future endeavors could extend beyond text to encompass other data modalities such as images, videos, and audio reviews. Integrating these diverse data types can afford a more holistic understanding of customer sentiments.
8. **Cross-lingual Sentiment Analysis:** Developing models proficient in conducting sentiment analysis across multiple languages, including Bangla, could be invaluable for multi-lingual markets. This would necessitate devising systems capable of comprehending and categorizing sentiments in various languages using a unified approach.
9. **Deployment and User Feedback:** Rolling out the developed models in real-world applications and gathering user feedback can furnish practical insights and pinpoint areas for refinement. Continuous monitoring and updating of the models based on user interactions can aid in sustaining and enhancing their performance.

By pursuing these avenues, future research can advance upon the discoveries of this study, leading to more efficient and precise sentiment analysis systems for Bangla text, thereby benefiting diverse applications such as customer feedback analysis, market research, and social media monitor.

# Reference:

[1] What Is Word-of-Mouth Marketing? https://www.investopedia.com/

[2] Liu, B. (2012) Sentiment Analysis and Opinion Mining. Syn- thesis Lectures on Human Language Technologies, 5, 1-167. https://doi.org/10.2200/S00416ED1V01Y201204HLT016.

[3] What Is Text Classification? https://monkeylearn.com/

[4] Rahman, M., et al. (2023). A Deep Learning Approach for Sentiment Analysis of Bangla Customer Reviews on E-commerce Platforms. IEEE Transactions on Neural Networks and Learning Systems, 34(5), 1892-1905.

[5] Islam, M. S., & Hossain, M. S. (2022). Sentiment Analysis of Bangla Text Using Deep Learning Techniques. Proceedings of the 2022 International Conference on Computational Intelligence and Data Science (ICCIDS), 78-85.

[6] Hasan, M., et al. (2020). "Deep Learning-Based Sentiment Analysis for Bangla Text: An Empirical Study." Journal of Information Science and Engineering, 36(4), 1021-1037.

[7] Tuhin, Rashedul Amin Paul, Bechitra Nawrine, Faria Akter, Mahbuba Das, Amit. (2019). An Automated System of Sen- timent Analysis from Bangla Text using Supervised Learning Techniques. 360-364. 10.1109/CCOMS.2019.8821658.

[8] Mahtab, Shamsul Islam, Nazmul Rahaman, Md Mahfuzur. (2018). Sentiment Analysis on Bangladesh Cricket with Support Vector Machine. 1-4. 10.1109/ICBSLP.2018.8554585.

[9] Hassan, Asif Amin, Mohammad Azad, Abul Mo- hammed, Nabeel. (2016). Sentiment analysis on bangla and romanized bangla text using deep recurrent models. 51-56. 10.1109/IWCI.2016.7860338.

[10] Bhowmik, Nitish Ranjan, Mohammad Arifuzzaman, and M. Rubaiyat Hossain Mondal. "Sentiment analysis on Bangla text using extended lexicon dictionary and deep learning algorithms." Array 13 (2022): 100123.

[11] Alvi, Nasif, Kamrul Hasan Talukder, and Abdul Hasib Uddin. "Sentiment analysis of bangla text using gated recurrent neural network." In International Conference on Innovative Computing and Communications: Proceedings of ICICC 2021, Volume 2, pp. 77-86. Springer Singapore, 2022

[12] M. A. Rahman and E. K. Dey, "Aspect extraction from Bangla reviews using convolutional neural network," in Joint 7th International Con- ference on Informatics, Electronics Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision Pattern Recognition (icIVPR). IEEE, 2018, pp. 262–267.

[13] M. H. Alam, M.-M. Rahoman, and M. A. K. Azad, "Senti- ment analysis for Bangla sentences using convolutional neural network," in 2017 20th International Conference of Computer and Information Technology (ICCIT). IEEE, 2017, pp. 1–6.

[14] S. Haque, T. Rahman, A. K. Shakir, M. Arman, K. B.B. Biplob, F. A. Himu, D. Das, M. S. Islam et al., "Aspect based sentiment analysis in Bangla dataset based on aspect term extraction," in the International Conference on Cyber Security and Computer Science. Springer, 2020, pp. 403–413

[15] Sarkar, Ovi & Ahamed, Md & Tasnia, Tahsin & Kumar Ghosh, Moloy & Islam, Md. (2021). An Experimental Framework of Bangla Text Classification for Analyzing Sentiment Applying CNN & BiLSTM. 1-6. 10.1109/INCET51464.2021.9456393.

[16] Sharmin, Sadia, and Danial Chakma."Attention-based convolutional neural network for Bangla sentiment analysis." Ai & Society 36, no. 1 (2021): 381-396.

[17] Rakib, Omor Faruk, Shahinur Akter, Md Azim Khan, Amit Kumar Das, and Khan Mohammad Habibullah. "Bangla word prediction and sentence completion using GRU: an extended version of RNN on N-gram language model." In 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), pp. 1-6. IEEE, 2019.

[18] Uddin, Abdul Hasib, Durjoy Bapery, and Abu Shamim Mohammad Arif. "Depression analysis from social media data in Bangla language using long short term memory (LSTM) recurrent neural network technique." In 2019 international conference on computer, communication, chemical, materials and electronic engineering (IC4ME2), pp. 1-4. IEEE, 2019.
[19] Bibi, Iram & Akhunzada, Adnan & Malik, Jahanzaib & Iqbal, Javed & Musaddiq, Arslan & Kim, Sung. (2020). A Dynamic DL-Driven Architecture to Combat Sophisticated Android Malware. IEEE Access. PP. 1-1. 10.1109/ACCESS.2020.3009819.
[20] An Intuitive Explanation of LSTM https://medium.com/
[21] Rayhan, Md Maruf, Taif Al Musabe, and Md Arafatul Islam. "Multilabel emotion detection from bangla text using bigru and cnn-bilstm." In 2020 23rd International Conference on Computer and Information Technology (ICCIT), pp. 1-6. IEEE, 2020.
[22] Hochreiter, S., Schmidhuber, J. (1997, November 1). Long Short-Term Memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
[23] Hassan, Asif Amin, Mohammad Azad, Abul Mo- hammed, Nabeel. (2016). Sentiment analysis on bangla and romanized bangla text using deep recurrent models. 51-56. 10.1109/IWCI.2016.7860338.
[24] Ahmed, T., & Kabir, M. A. (2021). "Sentiment Analysis of Bangla Reviews Using Convolutional Neural Networks." International Journal of Advanced Computer Science and Applications, 12(5), 198-205.
[25] Akter, N., et al. (2024). "Exploring Deep Learning Techniques for Bangla Sentiment Analysis in E-commerce Context." Proceedings of the 2024 IEEE International Conference on Big Data and Machine Learning (ICBDML), 224-231.
[26] Oueslati, Oumayma Cambria, Erik Ben Hajhmida,Moez Ounelli, Habib. (2020). A review of sentiment analysis research in Arabic language. Future Generation Computer Systems.112.10.1016/j.future.2020.5.34.
[27] Dua, Mohit nanda, charu nanda, garima. (2018). Sentiment Analysis of Movie Reviews in Hindi Language Using Machine Learning. 10.1109/ICCSP.2018.8524223.
[28] Pereira, Denilson. (2021). A survey of sentiment analysis in the Portuguese language (Link to PDF: https://rdcu.be/b5sQk). Artificial Intelligence Review. 10.1007/s10462-020-09870-1.
[29] A.Panchal,"NLP text summarization using NLTK: TF-IDF algorithm," https://towardsdatascience.com/text- summarization-using-tf-idf-e64a0644ace3, Jun. 2019, accessed: 2022-5-14.

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

**Title: SENTIMENT ANALYSIS OF BANGLA CUSTOMER REVIEWS ON DARAZ USING DEEP LEARNING APPROACH**

## CO Description for FYDP

| CO | CO Descriptions | PO |
|---|---|---|
| **Phase -I** | | |
| **CO1** | Integrate recently gained and previously acquired knowledge to identify sentiment analysis of Bangla customer reviews 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 these 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 | **PO10** |

| | | |
|---|---|---|
| | comprehend and generate comprehensive reports and design documentation, as well as provide and receive clear instructions throughout this FYDP. | |
| **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 of COs, Knowledge Profile (K), and Attainment of Complex Engineering Problems (EP):**

| Serial no | EP no. | CO | Attainment | Justification (with Knowledge Profile) | Page number |
|---|---|---|---|---|---|
| 01 | **EP1: Depth of Knowledge required** | CO1, CO2, CO3, CO5, CO6, CO7 & CO8 | Yes | Here we also use equations based on sample mathematics which covers K2. | **Page no: [16,31] Section: 3.4, 1.1** |
| | | | | For our project, we must collect data from the "DARAZ" website and understand how to create a deep learning-based model that covers K3 and K4. | **Page no: [15,21] Section: 3.3,3.7** |
| | | | | We need to connect several elements covering K6 in requirement analysis for our work. | **Page no: [15 -16] Section: 3.3,3.4** |
| | | | | We have studied similar types of paper for our work that cover K8 and have comparable objectives. | **Page no: [8-10] Section: 2.2,2.3** |
| 02 | **EP2: Range of Conflicting Requirements** | CO2, CO7 | Yes | We face some difficulties when we categorize data into three different levels. Finding specific regulations controlling the collection of data from Bangladeshi citizens speaking Bangladeshi languages on the company website proved to be difficult when we collected data | **Page no: [12,21] Section: 2.5,3.8** |

| | | | | | |
|---|---|---|---|---|---|
| | | | | for our project. The Bangla languages have limitations and certain gaps in the expression of emotions. | |
| 03 | **EP3: Depth of analysis** | CO2, CO6 | Yes | For depth analyzing, we need data collection, data pre-processing, model selection & evaluation. get the accuracy then result. It is a very final crystal solution for our project. It is only based on prediction. So here we use a good hybrid model whose accuracy is more improved. | **Page no: [30-38] Section: 5.1,5.2,5.3** |
| 04 | **EP4: Familiarity of Issues** | CO8 | Yes | To understand, we have to study some papers. We work on three types of emotional data collection. We work for smoothing results. Most of the paper works on two types of data. Specifically, we need to determine the sentiment of online shopping customers' emotions towards the Daraz company. | **Page no: [8,12] Section: 2.2,2.5** |
| 05 | **P5: Extends of application codes** | CO5 | NO | N\A | **N\A** |
| 06 | **EP6: Extends of stakeholders involved and conflicting requirements** | CO8 | NO | N\A | **N\A** |

| | | | | | |
|---|---|---|---|---|---|
| 07 | **EP7: Inter-dependence** | CO5 | Yes | There are multiple interdependent components in our project. These include activities like gathering data, analyzing it, creating front-end applications, conducting predictive analysis, and training deep learning models. | **Page no: [15-19] Section: 3.3,3.4,3.5** |
| 08 | **EP8: Project Management and Financial Analysis** | CO4 | Yes | We have prepared a budget to evaluate and estimate the cost required for our FYDP. when we launch the project on the Play Store app. | **Page no: [19] Section: 3.5** |

**Addressing CO11 with Complex Engineering Activities (EA):**

| SN | EA Definition | Attainment | CO | Justification | References |
|---|---|---|---|---|---|
| **1.** | **EA1: Range of resources** | Yes | CO11 | The project utilizes diverse resources such as high-performance computing infrastructure, GPUs, deep learning frameworks, annotated datasets, and ethical considerations to ensure systematic research and advancements in Bangla sentiment analysis | **Page no: [15-22,23-29,41]** **Section: [3.3, 4.2, 6.3]** |
| **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 | | The study "Sentiment Analysis of Bangla Customer Reviews on Daraz Using Deep Learning" advances environmental policy, sustainable consumerism, and linguistic inclusion while improving business intelligence and customer experience. | **Page no: [39-42] Section: [6.1, 6.2, 6.3, 6.4]** |
| **5.** | **EA-5: Familiarity** | Yes | | "Sentiment Analysis of Bangla Customer Reviews on Daraz Using Deep Learning" uses cutting edge algorithms to improve language participation, business intelligence, and client satisfaction. | **Page no: [03-04] Section: [2.1, 2.2]** |

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

| SN | COs | Attainment | Justification | References |
|---|---|---|---|---|
| **1** | CO4 | Yes | In order to **CO4**, this project combines efficient project management with financial control. Careful planning, resource distribution, and budget estimation are guaranteed to provide the best possible resource utilization throughout the duration of the study project. | **Page no: [18-19] Section: [3.5]** |
| **2** | CO9 | Yes | By setting client privacy first, getting informed consent, and openly recording the research process, the project demonstrates adherence to ethical principles. It also ensures responsible knowledge dissemination and societal well-being through the ethical application of innovative real estate technologies that follow by **CO9**. | **Page no: [41-42] Section: [6.3, 6.4]** |
| **3** | CO10 | No | **N/A** | **N/A** |
| **4** | CO12 | Yes | Several deep learning algorithms, including a hybrid CNN+BiLSTM model, | **Page no: [14-18, 31]** |

| | | | | |
|---|---|---|---|---|
| | | | and the project's commitment to continuous learning to **CO12** were utilized to analyze sentiment in Daraz Bangla customer evaluations; after tenfold cross-validation, the hybrid model outperformed the others in accuracy. | **Section: [3.1,3.2,3.3,3.4,3.8,5.3]** |