PRODUCT REVIEW ANALYSIS
USING DEEP LEARNING APPROACHES

AYESHA JULEKHA (C163237)

PROMA MUTSUDDI (C163215)

INTERNATIONAL ISLAMIC UNIVERSITY
CHITTAGONG (IIUC)

PRODUCT REVIEW ANALYSIS USING DEEP LEARNING APPROACHES

AYESHA JULEKHA (C163237)
PROMA MUTSUDDI (C163215)

THESIS SUBMITTED IN FULFILMENT FOR THE DEGREE OF
B. SC. IN COMPUTER SCIENCE AND ENGINEERING

DEPARTMRNT OF COMPUTER SCIENCE AND ENGINEERING (CSE)
INTERNATIONAL ISLAMIC UNIVERSITY CHITTAGONG (IIUC)
CHITTAGONH, BANGLADESH

2020

DECLARATION

We hereby declare that the work in this thesis is our own except for quotations and summaries which have been duly acknowledged.

29 May 2021

AYESHA JULEKHA(C163237)
PROMA MUTSUDDI(C163215)

SUPERVISOR'S DECLARATION

I hereby declare that i have read this thesis and in my opinion this thesis is sufficient in term of scope and quality for the award of the degree of B. Sc. in Computer Science & Engineering

29 May 2021

DR. ABDUL KADAR MUHAMMAD MASUM
Associate Professor, Dept. CSE, IIUC

DECLARATION OF THESIS / PROJECT REPORT AND COPYRIGHT

THESIS / PROJECT REPORT TITLE:

Product Review Analysis Using Deep Learning Approaches

AUTHORS:

SL NO.	AUTHOR'S NAME	STUDENT ID	SIGNATURE
1	Ayesha Julekha	C163237	Ayesha
2	Proma Mutsuddi	C163215	Proma
Name OF SUPERVISOR: DR. ABDUL KADAR MUHAMMAD MASUM			
SIGNATURE OF SUPERVISOR: 			

we declare

1. Our thesis to be published as online open access (full text) at IIUC database or archive.

2. Dept. of CSE, IIUC reserves the right as follows:

i. The thesis is the property of Dept. of CSE, IIUC

ii.The Library of IIUC has the right to make copies for the purpose of research only.

iii. The Library has the right to make copies of the thesis for academic exchange.

ACKNOWLEDGEMENT

Much obliged, Almighty, the Merciful and the Passionate, for giving us the chance to venture in the brilliant universe of science. We might want to say thanks to Almighty for giving us the quality, learning, capacity and chance to attempt this exploration to scrutinize and to endure and finish it acceptably.

The joy that follows the successful completion of any mission is incomplete without bringing up such names whom helped to come up with such works. As achievement is the epitome of hard work, perseverance, undeterred anger, enthusiasm, dedication, and the most inspiring advice and guidance that acted as the guiding light and successfully crowned our efforts. This task would not be structured without the general support of the Dept. of CSE, International Islamic University Chittagong. That rendered us the opportunity to fulfill our dreams and achieve Bachelor of Science in Computer Science and Engineering program. We are also grateful to our chairman Tanveer Ahsan, Associate professor, Dept. of CSE.

We are wholeheartedly thankful to our supervisor, **Dr. Abdul Kadar Muhammad Masum** for his persistent guidance, valuable advice, fruitful suggestions, morale-boosting, relentless motivation and supervision throughout our thesis work. His endless support helped us in channeling our efforts and ideas in proper direction

We should also like to thank to all of the faculties of Science and Engineering Department who helped us a lot of during the 4 years of Bachelor’s program. Without their inspiration, teachings, guidelines, suggestion and advice we were unable to learn a little bit of what we have learned in these years.

ABSTRACT

In today's commercial market, product reviews play an important role in online shopping. Most people verify product reviews before buying them online. Reviews can be positive or negative for any product. Positive product reviews attract customers more than negative product reviews. Reviews bring benefits and losses to any business. Customers share their product reviews on social media after purchasing the product. Product reviews are usually presented through e-commerce websites like Amazon, Flipkart, and other applications. Sentiment analysis is a rapidly thriving space in research in the space of natural language processing (NLP). It has gained a lot of attention in recent years. The data used in this study is a review of online products collected from the Amazon website we created. With the rapid increase of internet users, there is a huge number of Bangla Content generation. Though there are novel efforts in researches of resource-rich languages like English, Chinese, etc. Less work have been done yet for the Bangla language in this regard. In our proposed system, several deep learning approaches are used. With the deep learning approach, a good accuracy of 91.49% has been achieved with the CNN approach in the case of sentiment analysis and an accuracy of 91.01% has been achieved with the convolutional neural network with LSTM. Using the Bidirectional LSTM approach a fair accuracy of 92.13% has been achieved.

Keywords: **Sentiment Analysis, Deep Learning, Review Document, NLP task, Bangla NLP.**

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

LIST OF ABBREVIATIONS

NLP	Natural Language Processing
ML	Machine Learning
SVM	Support Vector Machine
RF	Random Forest
NB	Naïve Bayes
DT	Decision Tree
LR	Logistic Regression
KNN	K-Nearest Neighbors
BI-LSTM	Bidirectional Long–Short Term Memory
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network

CHAPTER I

INTRODUCTION

1.1 OVERVIEW

E-commerce is growing at an unprecedented rate all over the globe. With its growth, the impact of online reviews is increasing day by day. Reviews can influence people's purchasing decisions. Customers post their opinion about a product they purchase which may be positive or negative. Therefore, An NLP task of our proposed system identifies the expression of the review based on the opinions type to be considered. This is the process of deducing opinions is known as sentiment analysis. This chapter discusses in succession the field of sentiment analysis on the product review, its importance, and existing approaches in this field. In this context, the further problems that undertake our efforts are also explained in detail. Then motivation of our work and aims that needed to be procured in our study are highlighted. The research gap is described in the next along with our contributions to eradicate those issues. Finally, this whole report summarizes the overall structure of how our work is organized in the name structure of the thesis. Chapter summary at the end sums up the several contents mentioned in this chapter.

1.2 BACKGROUND OF THE STUDY:

There are five basic needs for our livelihood, but nowadays shopping plays an necessary role in our daily life. For daily needs every day we purchase various things. In the era of innovation, we very much rely on technology. Before purchasing anything we prefer to check reviews on the specific item because review helps consumers buy from a trustworthy platform. This pandemic COVID-19 people can’t go out for manual purchase. So, in terms of the previous situation nowadays people much more depend on online shopping, any kind of grocery items, fast food, cloth, shoes, etc. are purchased from an online platform. Bangla is a kind of language which has so many variations. So, a large number of people give their opinion on social media and any other site by using this language. But there is less research work done in the Bangla language. The Sentiment is a way by which opinions, emotions are expressed. Sentiment analysis is a broad part of data mining that is implemented with text mining. Sentimental analysis is an elaborate field of opinion mining that deals with people's opinions, sentiments, attitudes such as products, services issues and even their features. Product reviews express the people perceptions of the product either it is positive or negative. Product review helps the customer to buy a product on the basis of recommendation can be prepared by sentiment analysis. But problems arise when people give negative reviews about products which makes a hassle for customers. Sentiment analysis can be accomplished with the help of Natural language processing and Deep learning. Fake reviews also detect a major part of sentiment analysis. The area of study dealing with such sorts of topics is known as “NLP (Natural Language Processing)”. In our proposed work, we will utilize a number of deep learning techniques on the data of product review for review analysis. In our research, we will work with our native language (Bangla) with natural language processing techniques.

1.2.1 Sentiment Analysis

Sentiment analysis is a process that finds opinions, emotions from texts, tweets ,and other sources of natural languages [2]. All the opinion/emotion is captured using natural language processing [1]. Sentiment analysis in the field of information retrieval computationally identifying and categorizing opinions from a piece of text in the form of positive or negative[18]. In a real sense, sentiment analysis indicates that the summing up of underlying effective information exploring through the millions and millions of documents like reviews, opinions, news, interviews, and so on. Manually excavating this huge document and identifying the published opinions by arranging the data system can be both tedious and labor-intensive. With a view to solving this problem, diving through one of the most important branches of NLP i.e sentiment analysis is conspicuous. This territory is a large-scale problem domain. The fundamental challenging task in this field is to understand the complex semantic structure of languages and inspecting through vague statements in which positive words may point out negative meanings or vice versa. Such examples in Bengali are shown in the table. In sentence 1 the word “দক্ষ” infers to positive meaning whereas in sentence 2 it proclaims negative meaning without the use of any negation words.

Table 1 Illustration of Bangla Language Complexity

Root Word	SENTENCE 1	SENTENCE 2
দক্ষ	সে ছবি আকতে দক্ষ।	তার কোনো কাজে দক্ষতা নেই।

Previously works on sentiment analysis in Bangla included supervised and unsupervised techniques of machine learning. But with the advancement of natural language processing approaches tremendous scrupulous results have been achieved by deep learning methods. Therefore, a deep learning-based technique will be deployed for our proposed methodology focusing on sentiment analysis in a product review in the Bangla language.

1.3 PROBLEM STATEMENT

Language is a way by which people can express their feelings and opinions. Bangla is a language that ranks 5th in the world, around 265 million people use it as a First language and 37 million people use it as a secondary language. Bangla is a kind of language which has so many variations. Bangla texts still have a new field and opportunities for improvement. Nowadays, we see that more than 90% of the consumers read online reviews before making a purchase. Because of textual complexity, it is very difficult to identify the intent of Bangla text from the Bangla language. Extensive research has been done to analyze product reviews of English text which showed promising results. But very few initiatives for researches have been taken for the Bangla language. With the increase of technology, the e-commerce site is growing day by day, the use of an online e-commerce site and online reviews for products is also increasing. People not only shop on the e-commerce website but also share their reviews about the product. But, there is a challenge to work with the Bengali language due to a lack of Bengali resources. There is less work done in product review with Bangla language and less number of Bangla review datasets being available in the online platform. So, This triggered us to build up a system by NLP task on a Bangla dataset of product reviews.

1.4 MOTIVATION

There are several reasons why we choose this research area. First of all, we are very fond of shopping. Nowadays, every generation of people loves to shop. The important fact is in this era of technology everyone checks reviews before shopping, just like whenever we are going to buy a phone. At first, we check reviews then we will go to purchase it. Also, an online shopping review is an essential thing for purchase. Because most of the time good reviews make attraction towards the product. Before starting research on this topic we make a survey in shopping on the basis of the review. We took this survey by google form. Around 200 hundred people make responses, where 78.5% people claim that they check reviews before shopping, 11.7% claim that they check reviews sometime before purchase, and 9.5% claim that they don’t check reviews before purchase. In our survey, we also collect data about negative reviews. In the survey, 63.4% of people claim that they experienced negative reviews, and 35.6% of people claim that they don’t have experienced negative reviews. From the survey, we can say that there is a big portion standing by negative reviews. And there is less research work done before on the Bangla language. So we decided to make research on this topic.

1.5 RESEARCH OBJECTIVE

Nowadays people more rely on online shopping because it provides facilities for time-saving, no traffic and so on. So, the review is an important fact in terms of purchase because the review makes an impression on the consumer for buying the product or not. The goal of our proposed system is to make an approach that can analyze product reviews and extract the sentiments illustrated in those reviews in a large number of online product reviews. This system uses deep learning techniques for this purpose. The main points in this regard are highlighted below:

- To build a dataset of Bangla product reviews which is collected from the Amazon platform.

- To classify product reviews using deep learning approach.

1.6 RESEARCH GAP

There are some previous works on product review analysis in other languages but less work exists in the context of e-commerce websites on the Bangla language reviews. Some research gap given bellow:

- Lack of large volume of publicly available Bangla dataset.

- Less work has been done on Bangla review dataset with Natural language approach.

- Text classification application involving deep learning approach with advanced learning approaches are small in number.

1.6 KEY CONTRIBUTION

The key contributions to our proposed system are outlined below:

- Data has been collected manually from the Amazon Platform. Building a huge amount of the Bangla dataset was surely a challenge because there the reviews were present in the English alphabet.

- Product review analysis using a deep learning approaches.

- Labelling of the huge amount of Bangla dataset in two different fields which are positive and negative.

1.7 SCOPE OF STUDY

In the proposed work of our thesis, the main objective is to contribute to the build of a large amount of dataset for the Bengali language and to make it available to the public for future use in other researches in this field. We work with 15,000 product review datasets. More data will be collected for future work. The goal of our proposed system is to classify product reviews using a deep learning approach.

1.8 STRUCTURE OF THE THESIS

The thesis comprises five chapters. The thesis organization is generally described as follows: This thesis work has been sequentially outlined out into five different chapters. Chapter 1 discusses the background of our research. Importance of product review analysis. Further problems that encouraged or work has been stated and next the motivation or significance of our work and aims that need to be procured in our study are highlighted. Chapter 2 discusses Literature Review related to our thesis embarks on giving an overview of the basic terminologies of sentiment analysis and text classification. Then this chapter discusses different existing implemented approaches and additionally, the type of methods chosen by different authors with their brief illustration of work has been presented. The next phase of the chapter adds up the theoretical concept of the classifiers that are used to implement our system with necessary figures and equations. The chapter ends up with a concise concept of the chapter in the summary section. Chapter 3 discusses the proposed model including the algorithms and techniques. Chapter 4 presents the results and analysis and lastly, Chapter 5 gives the conclusion and future work.

1.9 CHAPTER SUMMARY

This chapter gives an intro to the brief discussion of our system and also analyzes the topics. Further problems that encouraged or work has been stated and next the motivation or significance of our work and aims that need to be procured in our study are highlighted. The research gap is pointed out in net sections with contributions. Finally, an overall structure of how our work has been organized in this whole report is summarized with the section named structure of the thesis. Chapter summary at the end sums up the several contents mentioned in this chapter.

CHAPTER II

LITERATURE REVIEW

2.1 INTRODUCTION

This chapter gives a clear idea about the researches and works that have been done so far which paves the way for future work and encouraged the urge for innovative progress in this field of natural language processing. As well as the challenges and limitations are briefly figured out and discussed in this section.

2.2 RELATED WORKS

Experiencing an investigation of different writing studies, it tends to be presumed that much of the research work has been done in English not only the Bengali language. There is less research work done in the Bengali language. On the other hand, other groups of researchers focus on the behavior of reviewers rather than review content. There has been a lot of effort in NLP work like sentiment analysis. This section provides an overview of literature reviews.

2.2.1 Related works in the field of Sentiment Analysis:

From the literature survey being carried out, it is seen that researchers have given incredible consideration in this field of sentiment analysis on online product reviews. Researchers have done their research in this field for the English language but very few works have been performed for the Bangla language. For example, Peter D. Turney [1] utilized an Unsupervised learning algorithm for classifying reviews. He classifies the reviews as thumbs up or thumbs down by the average semantic orientation. The semantic orientation of a phrase is calculated as the mutual information between the given phrase and the word “excellent” minus the mutual information between the given phrase and the word “poor”.

Fang and Bi Chen [2] utilized two typical approaches to sentiment analysis -lexicon look-up and machine learning approaches such as SVM. Lexicon look-up begins with a lexicon of positive and negative words. Current sentiment lexicons do not capture such domain and context sensitivities of sentiment expressions.

Hu and Liu [4] summarized a list of positive words and a list of negative words based on customer reviews. The positive list contains 2006 words and the negative list has 4783 words. Both lists also include some misspelled words that are frequently present in social media content

Istiaq et al. [17], proposed a hybrid approach to detect review spam (HDRS). At first, they detected duplicate reviews and then created hybrid dataset with the help of active learning. Lastly, they used a supervised approach to detect fake reviews.

Research from the University of Pittsburgh demonstrated that humans can only agree on whether a sentence has positive or negative sentiment, up to 80% of the time. Because of this any Natural Language Processing technique which scores up to 80% is working greatly with high accuracy [5] This mainly focuses on the feature extraction for sentiment analysis and also modifies the way in which the product sentiment analysis is done. A comparative study is given below on the sentiment analysis works done so far:

Table 2 Analysis of related work in the field of Sentiment analysis

Authors Name	Methodology	Used Classifier	Data Set	Result/ Findings
Tripathy, Agrawal, & Rath[29]	METHOD: Several supervised learning algorithms with TF-IDF features such as • SVM • Naïve Bayes • Maximum Entropy • Stochastic Gradient Descent • N-gram based techniques TOOLS: • Python Scikit-learn library	Support Vector Machine, Maximum Entropy, Stochastic Gradient Descent (SGD)	Acl IMDb movie review dataset	ACCURACY: NB: 86.23% ME: 88.48% SVM: 88.94% SGD: 85.11%
Zhang, Xu, Su, & Xu[28]	METHOD: Both word2vec and SVM perf method used for training. TOOLS: Python 3.6	word2vec and SVM perf	Chinese comment on clothing products	ACCURACY: Word2vec:87.10% SVM: 90.30%
Abinash Tripathy, Santanu Kumar Rath[27]	METHOD: Four supervised learning algorithm • NB • SVM • RF • LDA TOOLS: • Python 3.6	Naïve Bayes, Support Vector Machine, Random Forest, Linear Discriminant Analysis	Acl IMDb Dataset, Polarity Dataset	ACCURACY: NB:83% SVM:88% RF:88% LDA:86.5%

Authors Name	Methodology	Used Classifier	Data Set	Result/ Findings
Xiu Li, Lulu Xie, Fan Zhang, Huimin Wang[26]	▪ METHOD: ▪ Different Machine Learning approaches. ▪ • SVM ▪ • LR ▪ TOOLS: ▪ • Python 3.6 ▪ Scikit-learn library	bag-of-word, word embedding, and user features are used.	English deceptive review dataset comes from Ott. This dataset is manually constructed and composed of 800 positive reviews and 800 negative reviews.	ACCURACY: Logistic Regression and Support Vector Machine both achieves up to 93.6% accuracy in deceptive review detection.
Xing Fang* and Justin Zhan[25]	METHOD: Three supervised learning algorithms are used. • NB • RF • SVM TOOLS: • Python 3.6	Classification models selected for categorization are: Naïve Bayesian, Random Forest, and Support Vector Machine.	From Amazon.com	ACCURACY: Achieves accuracy Up to 85%
Raheesa Safrin1, K.R.Sharmila2, T.S.Shri Subangi3, E.A.Vimal4[24]	METHOD: Used unsupervised learning approach. TOOLS: • Python 3.6	K-means clustering is used to classify the retrieved dataset	Dataset collect from the e-commerce site in the form of feedback.	ACCURACY: K-means:90.47%
Reshan Ahmed Rabbi, Md. Walid Amin Khan,Md. Shaown ,Md. Shajidul Alam[23]	Method: Naïve Bayes TOOLS: • Python v3.6.2	Naïve Bayes classifier algorithm in TextBlob	Amazon phone review	not defined
Ashis Kumar Mandal1and Rikta Sen[22]	METHOD: Four supervised learning algorithm • KNN • NB • DT • SVM	Supervised Learning Methods implementation for Bangla Web Document Categorization	1000 documents from different Bangla web site	Achieved accuracy of about 70 %

Authors Name	Methodology	Used Classifier	Data Set	Result/ Findings
Munirul Mansur[21]	METHOD: N-gram based technique implementation Tools: Python 3.6	Analysis OF N-GRAM based Text Categorization f or Bangla in a newspaper corpus	Online news documents of about 2000	Accuracy to a satisfactory level of 70% has been achieved
Ankita Dhar, Niladri Sekhar Dash, Himadri Mukherjee, Kaushik Roy[20]	METHOD: Different classifiers applied on the basis of „term association‟ & „term aggregation‟such as: • Multi-layer Perceptron(MLP) • Random Forest(RF) • Support Vector machine(SVM) • Naïve Bayes Multinomial (NBM) • KStar(K*) TOOLS: • WEKA • 5-fold cross validation	Analysis of performance of different classifiers for categorization of Bangla text	8000 text documents from different leading online newspapers	A reasonable amount of average precision of 0.987 has been achieved using MLP classifier
Sadek Al Mostakim, Faiza Ehsan, Syeda Mahdiea Hasan, Sadia Islam & Swakkhar Shatabda [19]	METHOD: Several supervised learning algorithms with TF-IDF features such as • SVM • Naïve Bayes • Logistic regression • Random Forest • KNN TOOLS: • Python 3.6 • Scikit-learn library	Bangla online document classification	5870 no of documents has been collected from prothom alo & bd news 24	An online tool for Bangla content categorization showing good accuracy of more than 90% using logistic regression algorithm

2.3 CLASSIFICATION ALGORITHMS

This section focuses on theoretical and mathematical interpretations of the involved classification algorithms and methods. Going through a study of various literature surveys, it is observed that for product review analysis various types of types of Machine learning algorithms and Deep learning models have been used.

2.3.1 Supervised Learning Algorithms

Different types of supervised learning algorithm used for this NLP processes. The most famous ones of these are Naïve Bayes, Support Vector Machine, Logistic Regression, K- Nearest Neighbor, Decision Tree and Random Forest. This section presents a brief description about them in the following:

i. Naive Bayes

One of the supervised learning algorithms used for classification is Naive Bayes. It is based on the Bayes theorem which gives the conditional probability of an event A given B. So, the following equation:

$$P(A|B) = P(B|A)\,\frac{P(A)}{P(B)}$$

where:

P(A|B) = conditional probability of A given B

P(B|A) = conditional probability of B given A

P(A) = probability of event A

P(B) = probability of event B

This particular equation is called as bayes equation. In the bayes theorem, on the left the probability of A occurring given B has occurred equals the probability of B occurring given A has occurred times of probability of A over the probability of B.

ii. Support Vector Machine

Support Vector Machine (SVM), also known as Support Vector Network. It is a supervised learning technique used for classification and regression. In direct, given a training example set, each marked belonging to one of two categories. Thus the SVM classifier is represented by a separating hyperplane. This hyperplane generated from the training set then classifies the data from the test set. [7] [8].

Suppose two classes shown in Figure 2.1, denoted by square and circle and two axis x and y denoting features. SVM finds a hyperplane that classify all the training set into two classes.

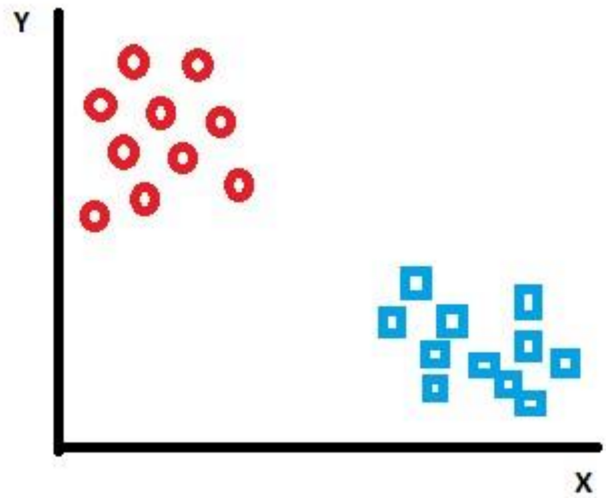

Figure 2.1 SVM Classifier with Two Classes [16]

Figure 2.2 denotes some separable hyperplane according to SVM classifier.

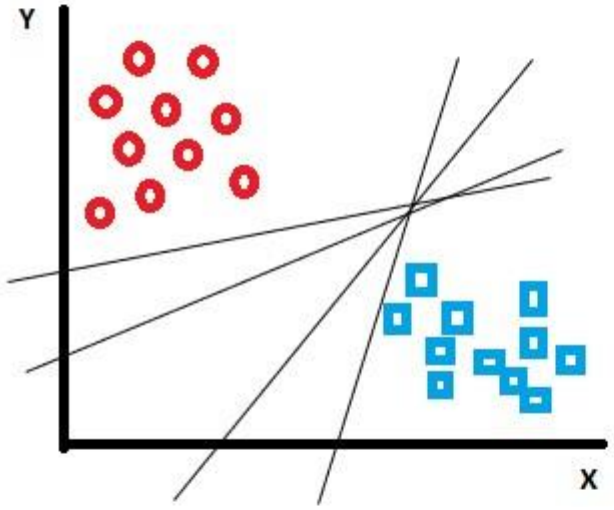

Figure 2.2 SVM Classifier with Hyperplane [16]

iii. Decision Tree

Decision tree is a supervised learning algorithm that is most used for classification problem. It represents a function that as input a vector of attribute values and returns a “decision”- a single output value. Decision tree algorithm falls under the category of supervised learning. They can be solve both regression and classification problems. A decision tree reaches its decision by performing a sequence of steps.

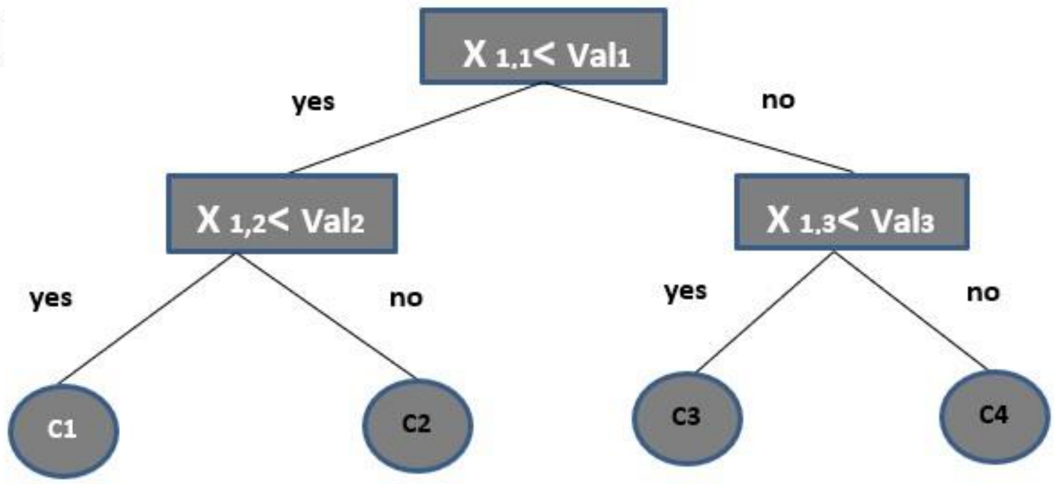

Figure 2.3 Decision tree architecture [16]

iv. Logistic Regression:

Logistic regression, also known as logit regression, is a very popular technology used for classification and regression. It provides simple and good performance. It is a differential probability model that operates on top of vector inputs that are truly valuable and predict the probability of an outcome that can only have two values (i.e. dichotomy). The dimension of input vectors is featured having without any restrictions against them being correlated. Logistic Regression produces a logistic curve, which values lies between 0 and 1 as shown in Figure 2.4.

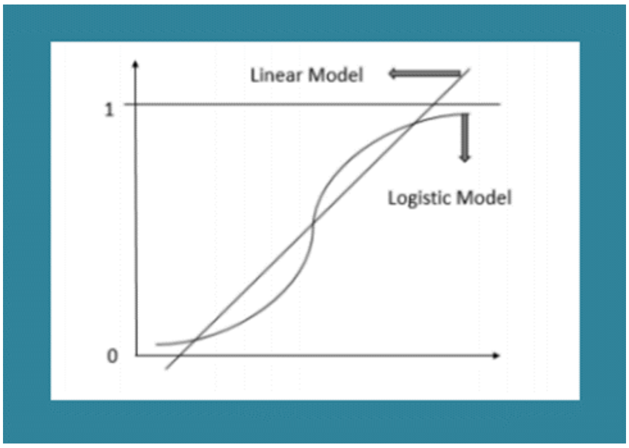

Figure 2.4 Logistic Model [16]

v. Random Forest

Random forest is a collection of decision trees. Random Forest classifier works where Decision Tree fails. In other words, if the trees are grown too deeply or taken up in irregular shapes i.e overfit training set then for averaging multiple deep decision trees, random forests work on a different parts of the same training set by generating a multitude of decision trees during training time. Figure 2.5 shows that three having node Y provide correct prediction because of their majority and tree having node N provide the wrong prediction.

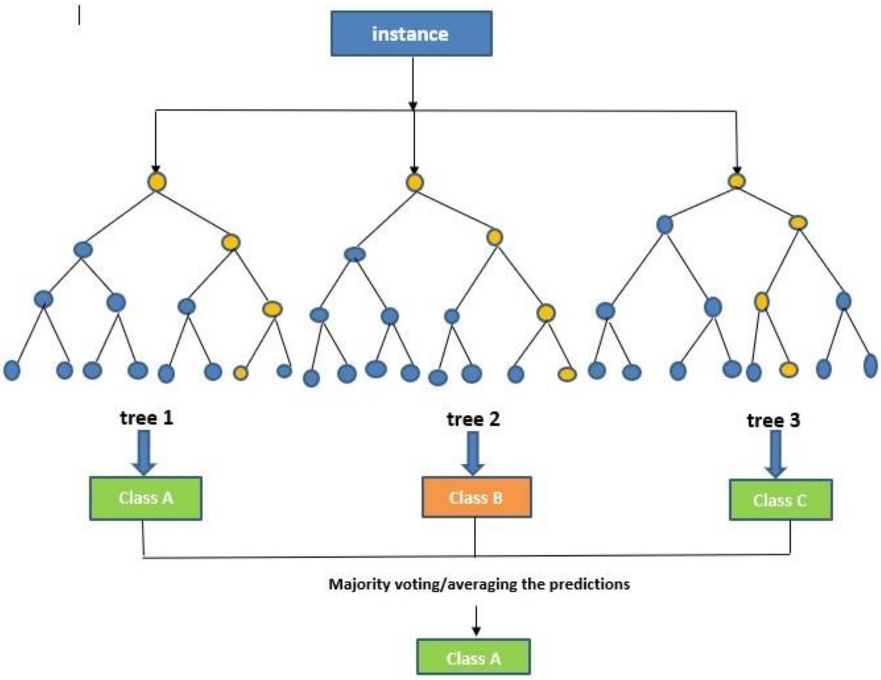

Figure 2.5 Random Forest Classifier [16]

vi. K-Nearest Neighbor

k-Nearest Neighbor (k-NN) classifier is the simplest among all the classifiers and is used for both classification and regression. In this, input consists of k closest training sets in the feature space. Its output is class membership.

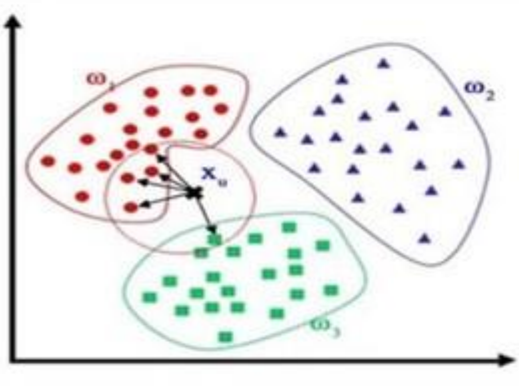

Figure 2.6 KNN Classifier [16]

2.4 DEEP LEARNING METHODS

Deep learning is a key technology behind driverless vehicles that enables them to detect a stop sign or separate a pedestrian from a lamppost. In deep learning, a computer model learns to perform classification tasks directly from images, text, or sound. Some models are illustrated below:

2.4.1 Convolutional Neural Networks (CNN)

CNN is an artificial neural network that uses convolution operations, and CNN models has shown excellent performance in image recognition because the feature patterns of the images can be extracted by training through convolutional filters.[30]. CNN (convolutional neural network) includes pooling layers and sophistication as it gives a standard architecture to map the sentences of variable length into sentences of fixed size scattered vectors. This study has proposed a novel convolutional neural network (CNN) framework for visual sentiment analysis to predict sentiments of visual content. CNN has been implemented using Caffe and Python on a Linux machine. Transfer learning approach and hyper-parameter have been used in biases and weights are utilized from pre-trained GoogLeNet. As CNN enhances its performance by increasing its size and depth, so a very deep CNN model, inspired by GoogLeNet is proposed with 22 layers for sentiment analysis. It is optimized by using SGD (Stochastic gradient descent) algorithm. The strategy with 60 epochs has been performed for training the network as GoogLeNet has performed 250 epochs. For experimental work, a dataset of Twitter containing 1269 images is selected and backpropagation is applied. Amazon Mechanical Turk (MTurk) and popular crowd intelligence are used to label the images. Five workers were involved to generate a sentiment label in favor of every image. The proposed model was evaluated on this dataset and acquired better performance than existing systems. Results show that the proposed system achieves high performance without fine-tuning on the Flickr dataset.

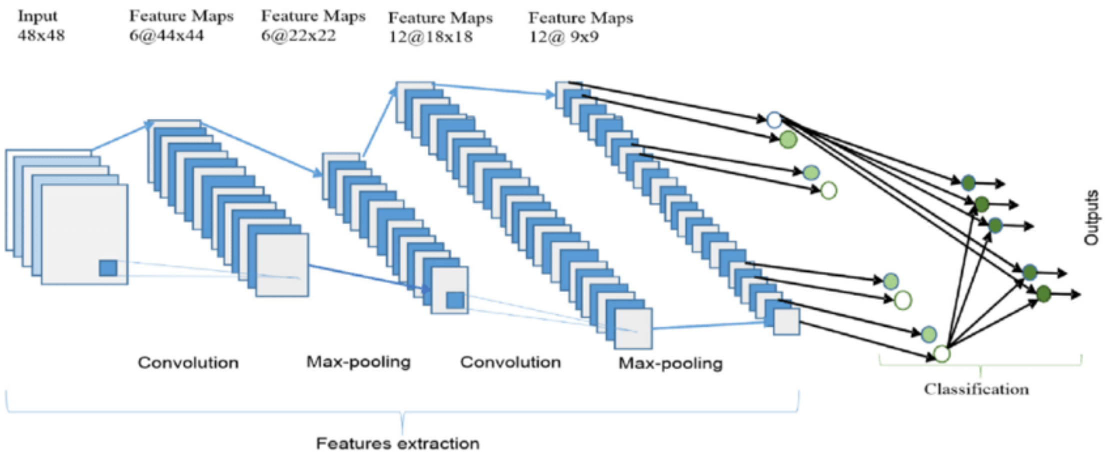

Figure 2.7 Convolutional Neural Network architecture

2.4.2 Recurrent Neural Networks (RNN)

Recurrent Neural Networks (RNN) is another neural network architecture that is addressed by researchers for text mining and classification. It has been widely used in speech recognition, handwriting recognition, natural language processing, and others. Moreover, RNN is the precursor to LSTM. While traditional neural networks failed to create a persistent model that would somewhat mimic the way our memory cells work for learning and remembering information, RNN – a class of ANN, has an interesting model design with a loop used as a feed-back connection which makes the information persistent [11, 12]. Influential model in language modeling because it doesn't represent the context of fixed-length that contaminates all history words. It has been widely used in speech recognition, handwriting recognition, natural language processing, and others. Moreover, RNN is the precursor to LSTM. While traditional neural networks failed to create a persistent model that would somewhat mimic the way our memory cells work for learning and remembering information, RNN – a class of ANN, has an interesting model design with a loop used as feedback. RNN assigns more weights to the previous data points of sequence. Therefore, this technique is a powerful method for text, string, and sequential data classification.

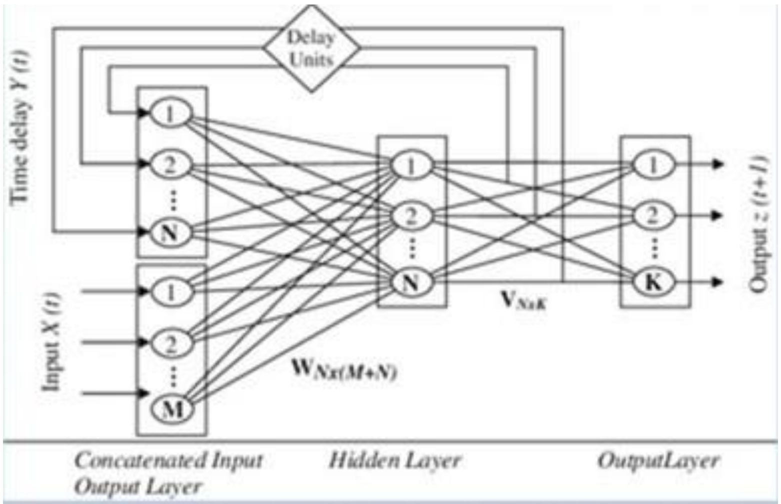

Figure 2.8 Recurrent Neural Networks

2.4.3 Long Short-Term Memory (LSTM)

Long Short-Term Memory~(LSTM) was introduced by S. Hochreiter and J. Schmidhuber and developed by many research scientists. In this network, the memory cell had linear dependence of its present activity and its past activity.

While RNN‟s success was critical in speech and pattern recognition due to its ability to remember temporal dependencies, it was not without problems. RNNs were able to connect previous information to current task, only when the gap between the information was small. As the gap widened, RNNs started to perform poorly. Also, the depth and complexity of layers are increase, the vanishing gradient problem causes difficulty in training. Long Short Term Memory (LSTM) is an extension of simple RNNs, which reduce the vanishing gradient problem and can remember dependencies over larger gaps [15].

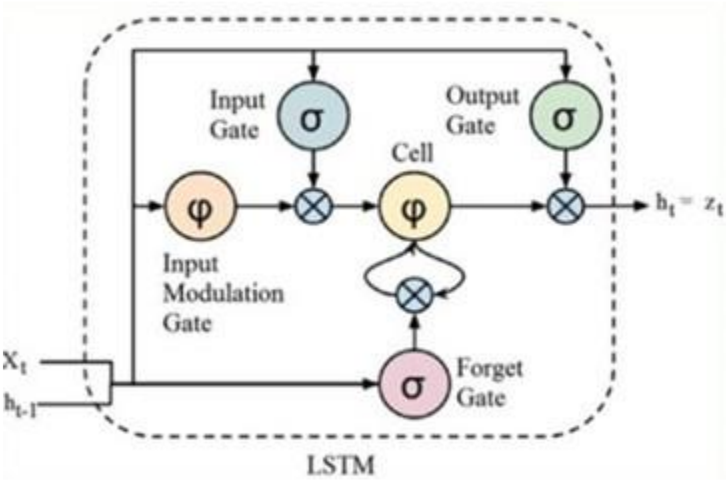

Figure 2.9 Long-Short Term Memory architecture

In the figure capture the LSTM model where σ is the logistic sigmoid function. i, f, o, and c are the input gate, forget gate, output gate, and memory cell activation vectors respectively.

$$i_t = (W^{(xi)} x_t + W^{(hi)} h_{t-1} + W^{(ci)} c_{t-1} + b^{(i)})$$

$$f_t = (W^{(xf)} x_t + W^{(hf)} h_{t-1} + W^{(cf)} c_{t-1} + b^{(f)})$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tan h \, (W^{(xc)} x_t + W^{(hc)} h_{t-1} + b^{(c)})$$

$$o_t = (W^{(xo)} x_t + W^{(ho)} h_{t-1} + W^{(co)} c_t + b^{(o)})$$

$$h_t = o_t \cdot \tan h \, (c_t)$$

Figure 2.10 Long-Short Term Memory functions and equations

Bi-directional LSTMs are an extension of traditional theoretical LSTMs that can improve model performance in the case of sequence classification problems. In all timesteps of input sequences available problems. Bidirectional LSTMs train two instead of one LSTMs on the input sequence. The first on the input sequence as-is and the second on a reversed copy of the input sequence. This can provide additional context to the network and result in faster and even fuller learning on the problem.

2.5 CHAPTER SUMMARY

This chapter encompasses the inner structural process of different methodologies and techniques used in supervised machine learning algorithms and deep learning methods. SVM, LR, NB, Random Forest, DT, and K-NN are the supervised learning classification algorithms discussed in this part. Moreover, RNN, CNN, LSTM, are the deep learning models discoursed in this chapter.

CHAPTER III

METHODOLOGY

3.1 INTRODUCTION

In this chapter a brief overview of the overall working procedure of the proposed work described in a crystal clear form. This chapter deals with the beginning of the research work from data collection to the implementation of the proposed work. This detailed explanation will help to get a crystal clear idea of the proposed work, its working procedure. The content of this chapter sequentially organized in various sections and every section deals with the description of a specific working methodology. Dataset is the main object in the world of information. Without data, there is nothing to proceed to deal with a specific problem.it is considered as the heart in both cases of machine learning and deep learning. in this part, the clear interpretation of a collection of the dataset, preprocess of the dataset, and implementation of proposed work clearly defined. Finally, the chapter concluded with the total implementation of our lodged methodology.

3.2 RESEARCH SUBJECT AND INSTRUMENTATION

The research subject is a term that traces out the systematic exploration by facts and analysis and worked with a fine outcome of a specific topic.it is dispensed with different methods and throughout the methods, new knowledge is identified to imply the proposed work. Our framework deals with sentiment analysis by automatic text classification of Bangla product reviews and classifies the reviews in positive, negative polarity. In this event, instrumentation plays a vital role to get an authentic outcome. So in terms of classifying Bangla product review, our approach is deep learning-based, we apply it with the convolutional neural network (CNN), bidirectional LSTM and a combination of both networks.

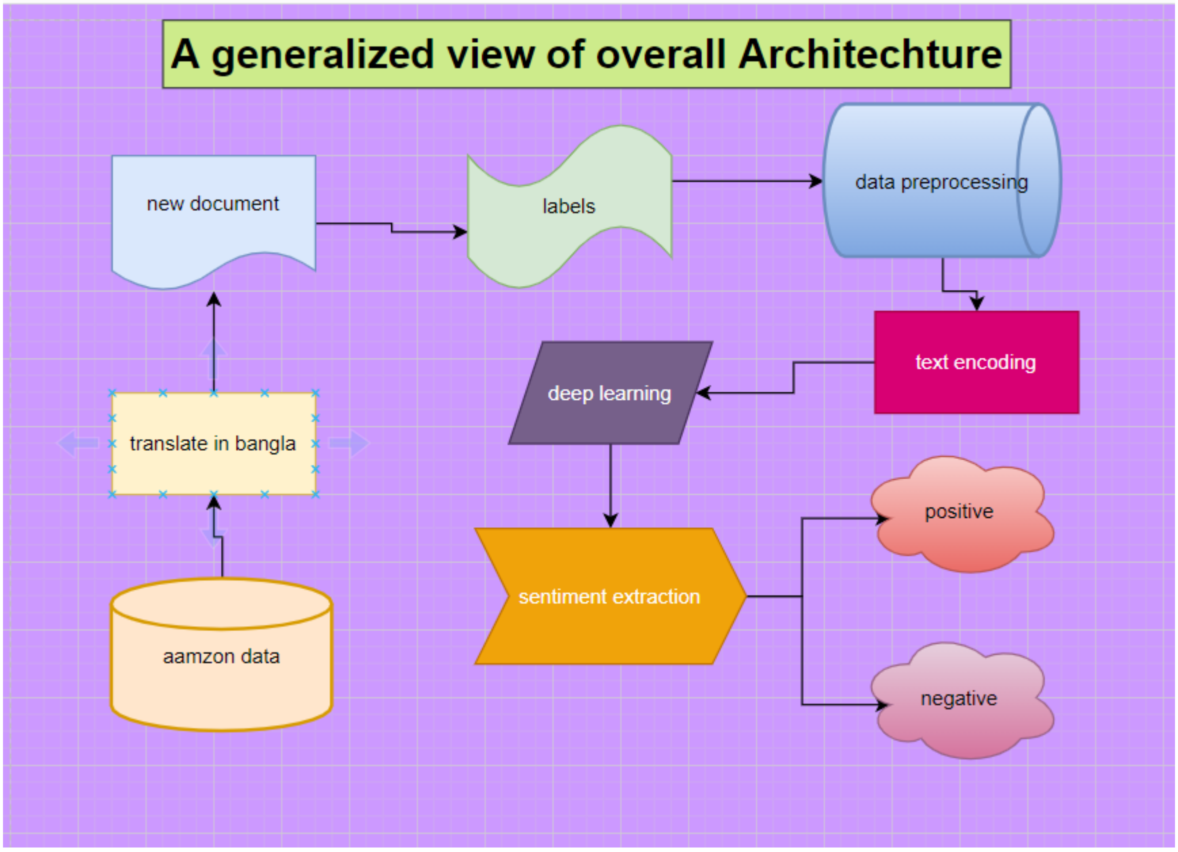

Figure 3.1 A generalized view of overall architecture

3.3 DATA COLLECTION PROCEDURE AND SOURCES

Data plays an important role in data mining. Data mining refers to extract new knowledge by analyzing the data. A dataset is a collection of related data. In the field of research, the main entity is data. Data is the most crucial part to get expected outcomes in the research. Data is processed by different tools and methods for desirable results. A dataset must be collected in a precise way so that the expected outcome furnished in a lucid manner. Our research is based on the Bangla language. There is no available dataset of the Bangla language publicly which can be utilized for our research. So that we need to build up our datasets for our experiment. We constructed our datasets from the AMAZON products review. The dataset of reviews compressed with reviews from different categories of products i.e.

- Mobile reviews
- Kindle reviews
- Book reviews

3.3.1 Data Collection

There is no available dataset in the Bangla language in terms of a product review. There are lots of E-commerce pages on social media and websites for product selling. All those sites and pages have a review section where reviews are available. There is a restriction about web scrapping so that we cannot get Bangla reviews from those pages and websites. And another problem arises about data collection, our research motive is review analysis, we have to collect the datasets in an authentic way so that we can get both results which are our expected outcomes. We thought about collect data manually, whenever we try to collect data manually from the website we get a total imbalance dataset that is if we get 1000 positive reviews there are only 10 or 20 negative reviews which are not good to get authentic results. This kind of dataset will not give a good accuracy to our experiments. So we have to construct our own dataset, we collect data from Amazon product reviews. We collect the amazon product reviews from Github and many web links. The dataset which we collect from a website and many other sources is in the English language. In term of our proposed work we need dataset in Bangla language, our main motive is to learn machine the Bangla language and build a framework which can produce a result in terms of Bangla. So for the purpose of the Bangla dataset, we translate the reviews in the Bangla language and label those datasets. We collect datasets of amazon products from different categories which are kindle review, book review, mobile review, beauty product review, electronic product review, food review, etc. Figure-3.2 illustrates the process of data collection which is given in below-

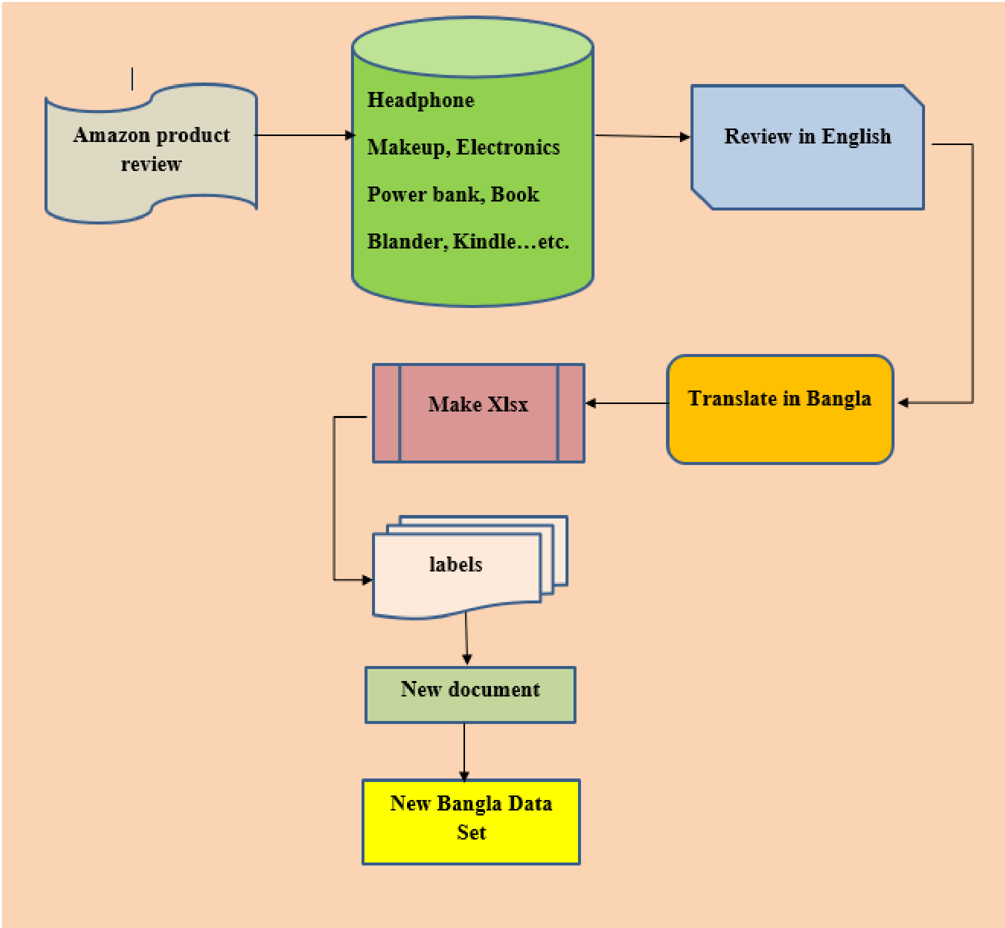

Figure 3.2 Process of data collection

3.3.2 Data Statistics and About Data

Our dataset is about amazon product review which is translated in Bangla, the reviews are from different categories of product and collected from different websites. We have a total of 15000 product reviews. We have a balanced dataset, our dataset contains 7500 positive data and 7500 negative data. Feature of the datasets are-

- Reviews (In Bangla)
- Polarity

The column Reviews contain about the content of viewpoints of products after using the accessories. The column polarity tends to positive and negative. Table-3 gives brief statics about reviews procured from different sources and different categories. We have collected datasets from AMAZON product reviews from different sites like Kaggle, GitHub, websites, etc. So the whole 15000 reviews are considered for sentiment extraction in two polarities positive and negative.

Table-3 shows statistics of our collected AMAZON reviews which is based on amazon different categories of product

Table 3 A Statistics table of collected data

Data(review)	**Quantity**
Power bank	2600
book	800
Blander machine	1000
Head-phone	2100
iPhone	1250
Kindle	510
Beauty and Personal Care	320
jewelry & watches	1042
Kitchen and Housewares	911
Outdoor living	500
Mobile	3016
Gourmet food	951
Total	15000

3.4 DATA PREPROCESSING

Data preprocessing is a technique of data mining that involved the process of making an understandable data format. Our research is about text classification, we classify the product reviews text and apply it with deep learning. The first and foremost step for process the data with deep learning is preprocessing of data. The data is the most crucial part of the process of data mining. Data preprocessing is important because the data which we collect or aggregate from various sources can contain noisy, missing values or in an inconsistent and inconvenient form. We have to preprocess data in a format so that it can be processed with a deep neural network in a convenient form. On the other side if we do not preprocess data it will make an impact on results, sometimes results are not so accurate for unprocessed data.

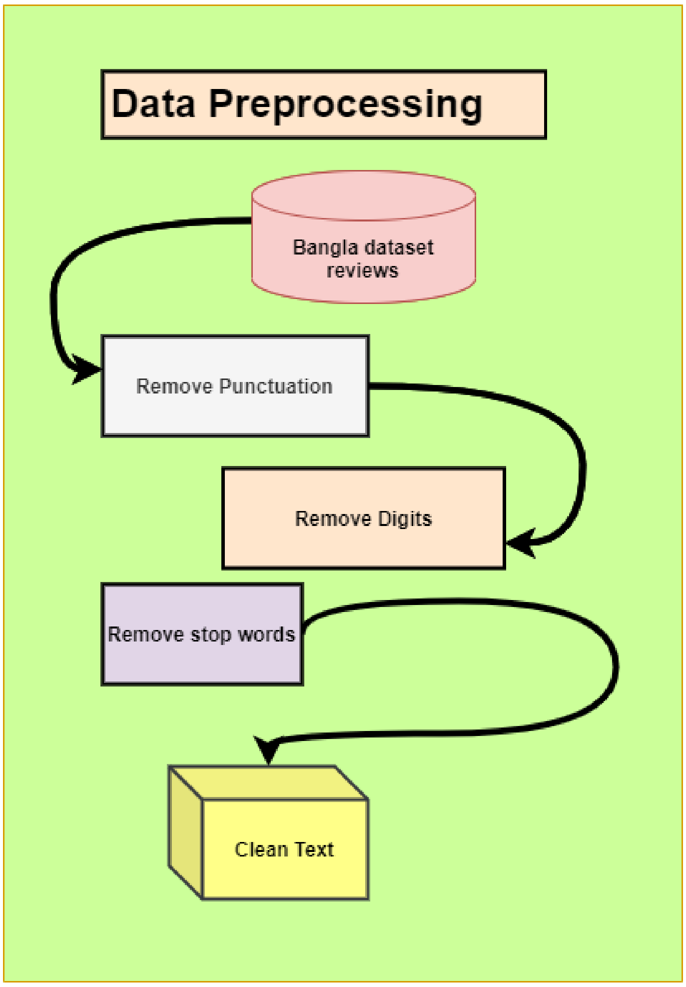

Figure 3.3 Data Preprocessing

3.4.1 Tokenization

Tokenization is a technique that splits the text into tokens. The task of segment a bunch of texts into words or linguistic tokens or segments is known as tokenization. Data preprocessing is a data mining technique that involves transforming raw data into an understandable format. there are two kinds of tokenization one is word tokenization and another is sentence tokenization. Sentence tokenization split the sentence into single words. Word tokenization split the words into single characters. Preprocessing deals with tokenization, removing delimiters, removing stop words, and so on. So the first step is tokenization. Tokenization can be done with the NLTK library Or Keras library. Here we use Tokenizer from Keras Preprocessing. text for tokenization. We use Sentence tokenization to split the sentence into tokens. Tokens retrieved after compiling these steps. Table- 4 shows a sample input of tokenization of sentence into words

Table 4 Example of sentence tokenization

Example of sample text input	**Tokenize text**
এই ঘড়ির মুখটি খুব ছোট। সাবধান হোন. আমি নিশ্চিত নই যে কেন তারা এটিকে ইউনিসেক্স ঘড়ি বলছে কারণ কোনও লোকই এর চেয়ে ছোট কিছু চাইবে না।	‘এই’ ‘ঘড়ির’ ‘মুখটি’ ‘খুব’ ‘ছোট’ ‘।‘ ‘সাবধান’ ‘হোন’. ‘আমি’ ‘নিশ্চিত’ ‘নই’ ‘যে’ ‘কেন’ ‘তারা’ ‘এটিকে’ ‘ইউনিসেক্স’ ‘ঘড়ি’ ‘বলছে’ ‘কারণ’ ‘কোনও’ ‘লোকই’ ‘এর’ ‘চেয়ে’ ‘ছোট’ ‘ কিছু’ ‘চাইবে’ ‘না’ ‘।‘
আমি যখন আমার পিএস ভিটা পেয়েছিলাম তখন আমি এটি বিনামূল্যে পেয়েছিলাম এবং এটি খুব ভাল জিনিস যেহেতু এটি অকেজো হওয়ার পাশেই রয়েছে। আমি এই বিষয়ে লোকেদের মোটেই অর্থ ব্যয় করার পরামর্শ দেব না।	‘আমি’ ‘যখন’ ‘আমার’ ‘পিএস’ ‘ভিটা’ ‘পেয়েছিলাম’ ‘তখন’ ‘আমি’ ‘এটি’ ‘বিনামূল্যে’ ‘পেয়েছিলাম’ ‘এবং’ ‘এটি’ ‘খুব’ ‘ভাল’ ‘জিনিস’ ‘যেহেতু’ ‘এটি’ ‘অকেজো’ ‘হওয়ার’ ‘পাশেই’ ‘রয়েছে’’।‘ ‘আমি’ ‘এই’ ‘বিষয়ে’ ‘লোকেদের’ ‘মোটেই’ ‘অর্থ’ ‘ব্যয়’ ‘করার’ ‘পরামর্শ’ ‘দেব’ ‘না’ ‘।‘

3.4.2 Punctuation Removal

In the data preprocessing after segmentation of text, removal of delimiter is processed. Punctuation marks make a document in a readable format, those punctuation having no importance on text classification. A clean text document is easy to process and help to get a more authentic result. So punctuation removal is an important task in data preprocessing. In the punctuation removal step, we use sklearn library for removing. We can add extra literals If we want to remove them. In the phase of punctuation removing (!"#$%&'()*+,-./:;<=>?@[\]^_`{|}~) this m symbol is removed for better data preprocessing. Table- 5 shows a sample input of punctuation removing—

Table 5 Example of punctuation remove

Sample input of text	After remove punctuation
এই ঘড়ির মুখটি খুব ছোট। সাবধান হোন. আমি নিশ্চিত নই যে কেন তারা এটিকে ইউনিসেক্স ঘড়ি বলছে কারণ কোনও লোকই এর চেয়ে ছোট কিছু চাইবে না।	▪ 'এই' 'ঘড়ির' 'মুখটি' 'খুব' 'ছোট' 'সাবধান' 'হোন'. 'আমি' 'নিশ্চিত' 'নই' 'যে' 'কেন' 'তারা' 'এটিকে' 'ইউনিসেক্স' 'ঘড়ি' 'বলছে' 'কারণ' 'কোনও' 'লোকই' 'এর' 'চেয়ে' 'ছোট' ' কিছু' 'চাইবে' 'না'

3.4.3 Digit removal

Our datasets are based on Bangla product review, these reviews contain some English Or Bangla numeric symbol which has no significant meaning in the text. This symbol can make an effect on data redundancy which can make an impact on the negative results, it has to be removed.

3.4.4 Stop Word removal

Stop words are a list of those words which are frequently used in a sentence. In-text mining we often worked with removing stop words. Removing stop words is vital for data preprocessing. Because if we remove those words which frequently used in a sentence we can focus on the important words instead.it will help to easily process data throughout the neural networks and produce better outcomes. In-text classification, texts are classified in various categories. So without stop word text can be processed with focused words. In the Bangla language, there are so many stop words, we have listed down 335 Stop words, and we worked throughout it.

Table 6 shows example of some Bangla Stop Words

Categories	**Words Example**
Postpositions	[„ডেছক‟ (dike) 'towards' ‘সহ’ (shôho) 'with', 'including']
Conjunctions	[„এবং‟ (ebang) „and‟, ‘ডকন্তু’ (kintu) „but‟]
Interjections	▪ [„বাহ‟ (bah) „well‟, ▪ „সাবাে!‟ (shabash) „bravo‟]
▪ Pronouns	[‘আডম’ (Āmi) „i‟, ‘েু ডম’ (Tumi) „you‟]
Some adjectives	[„েোছল্া‟ (bhalo) „good‟, „কৃ ষ্ণ‟ (krishno) „black‟]
Some adverbs	[„খুব’ (khub) „very‟, „কাছল্েছে’ (kālēbhadrē) „rarely‟]
All articles	[„একটি‟ (ekti) „one‟]
Proper nouns	[„আগ্রাবাে‟ (agrabad) „Agrabad‟]

Table 7 shows example of after removing Bangla stop words

Bangla Stop Word	**After Removing Stopword**
এই ঘড়ির মুখটি খুব ছোট। সাবধান হোন. আমি নিশ্চিত নই যে কেন তারা এটিকে ইউনিসেক্স ঘড়ি বলছে কারণ কোনও লোকই এর চেয়ে ছোট কিছু চাইবে না।	‘ঘড়ির’ ‘মুখটি’ ‘ছোট’ ‘সাবধান’ ‘হোন’. ‘নিশ্চিত’ ‘এটিকে’ ‘ইউনিসেক্স’ ‘ঘড়ি’ ‘কোনও’ ‘লোকই’ ‘ছোট’ ‘চাইবে’

3.5 IMPLEMENT WITH DEEP LEARNING

Our dataset is about Bangla product reviews, After preprocessing data we have to implement our dataset with deep learning. In deep learning, we use CNN, Bidirectional LSTM, and CNN with LSTM. In-text classification first of all text should be passed by Embedding layer then all other neural network layers. After feeding with neural networks we will get a final outcome or our expected result.

3.5.1 Text to Word Sequence

Natural language processing is a technique of processing the human language in a convenient way to the machine So that it can understand by machine and can produce the expected outcome. In deep learning, when we work with text classification, deep learning and machine learning both are worked with numeric values, not with texts. In deep learning, inputs are represented in vectorize form, in the class of text to sequence all the unique token get an integer value in term of that word. All the integer value represented as dimensional vector in the embedding layer.

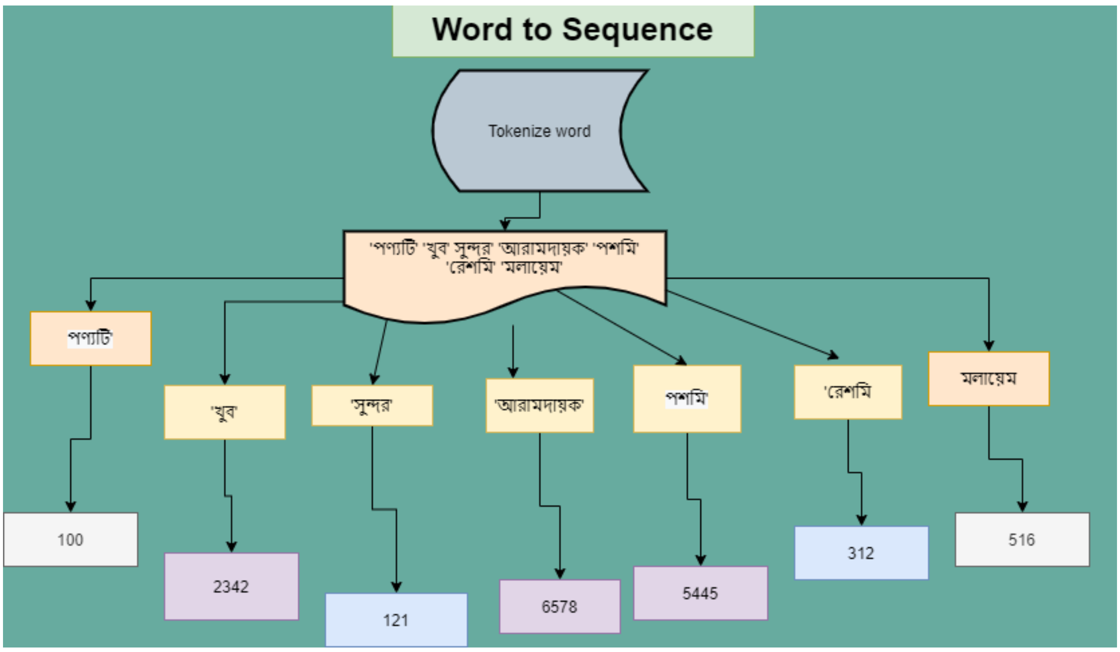

Figure 3.4 Process of Text to sequence

3.5.2 Pad sequence

Pad sequence is used to make the all sequences in the same length. The term of pad sequences is called Padding. Padding can be done in two ways one is pre and another post. Padding is done by adding random 0 until the sequence is the as same length as the longest length. In-text classification Padding is important in neural network input. All the neural networks required have the same shape input and size. In our dataset, the text of reviews, the length of all sentences are not the same. Some are longer and some are shorter. When we convert all the text in a relevant sequence number, the length of all sequences is not the same length. So for feeding the inputs to the neural networks padding is necessary. By padding, we can make all sequences to the same length and get a fine input for neural networks. Figure-3.5 demonstrates the process of padding..

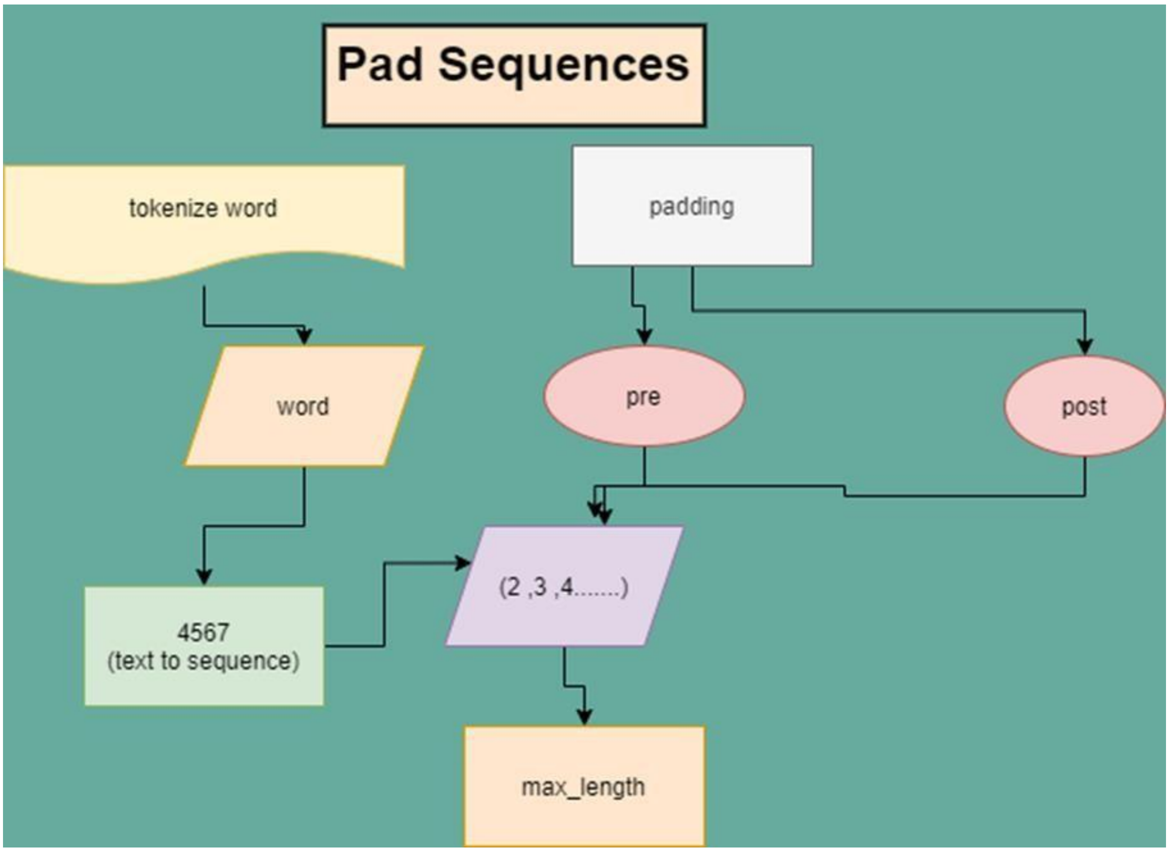

Figure 3.5 Pad sequence figure

3.5.3 Embedding layer

The embedding layer is an important layer used in deep neural networks when it implements with text classification. After preprocessing the dataset, the data column which belongs to the text. That text is given some sequence number and worked with padding for the same input shape. In CNN and LSTM those sequences are worked as input. The vocabulary, length of the sentence, and embedding dim are the content of this embedding layer. Here we use a different number of embedding dim for our proposed methodology in the embedding layer. The embedding layer is used to compress the input feature into a smaller section. Embedding layers are trained like other layers in the model, the used to utilize the loss function in term of the optimizers. Embedding layer used to index a table that belongs to the embedding vector. Fig-3.6 shows a graphical view of the embedding layer.

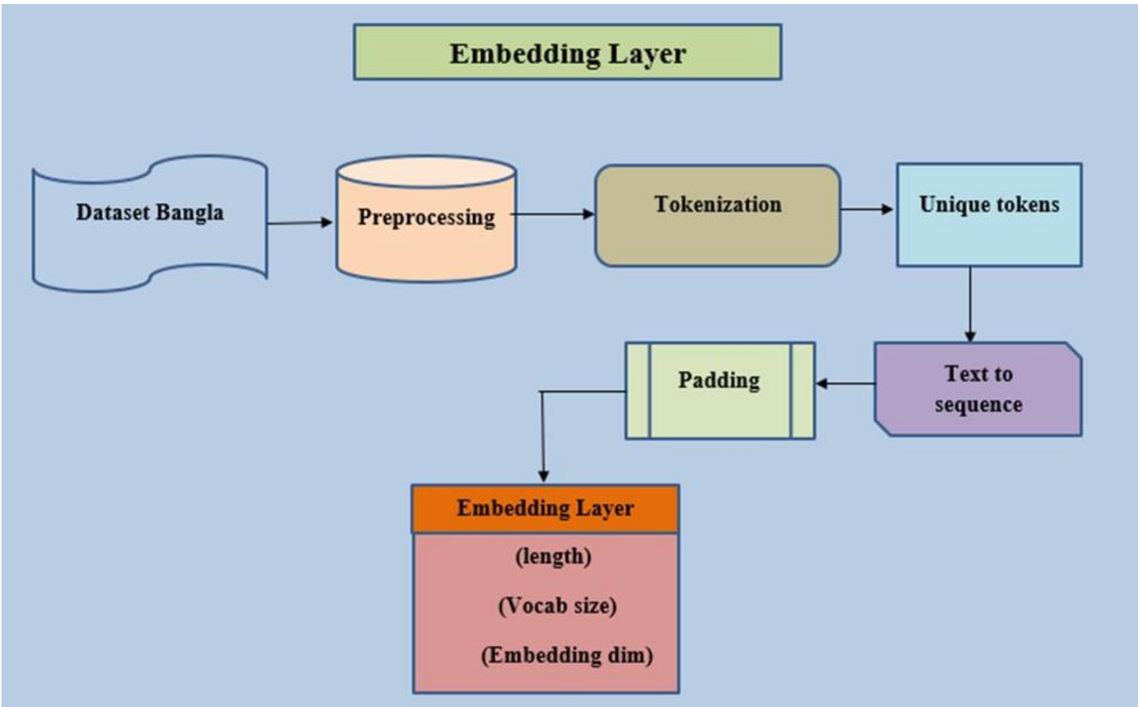

Figure 3.6 A graphical view of embedding layer

3.6 METHOD IMPLEMENTATION

3.6.1 Implement With Convolutional Neural Network

In CNN it worked with 4 steps – Convolution Layer, Maxpooling1D, Flatten, and Fully connected.in NLP when applied with CNN 1-D vector is represented the texts. When convNet is used in text classification it classifying the sentence in terms of text to sequence or pre-trained model. A sequence of words where each sequence is correlated with an embedding vector of dimension. A convolution filter or kernel will be applied in the convolution layer. After processing the input in the embedding layer, it feed-forward to the convolution layer and so on. In the final layer, it classifies our dataset in the positive and negative categories, here we use the Adam optimizer and sigmoid function to classify the outputs. Here we use 20 epochs for the neural network. In CNN we use a three-layer of convolution layer, three max pool layer and three flatten layer. We also used three embedding layers with different embedding dim. For the purpose of reducing overfitting we have. Used several regularizer like l1, l2 in term of kernel and bias regularizer Figure-3.7 illustrate the architecture of CNN.

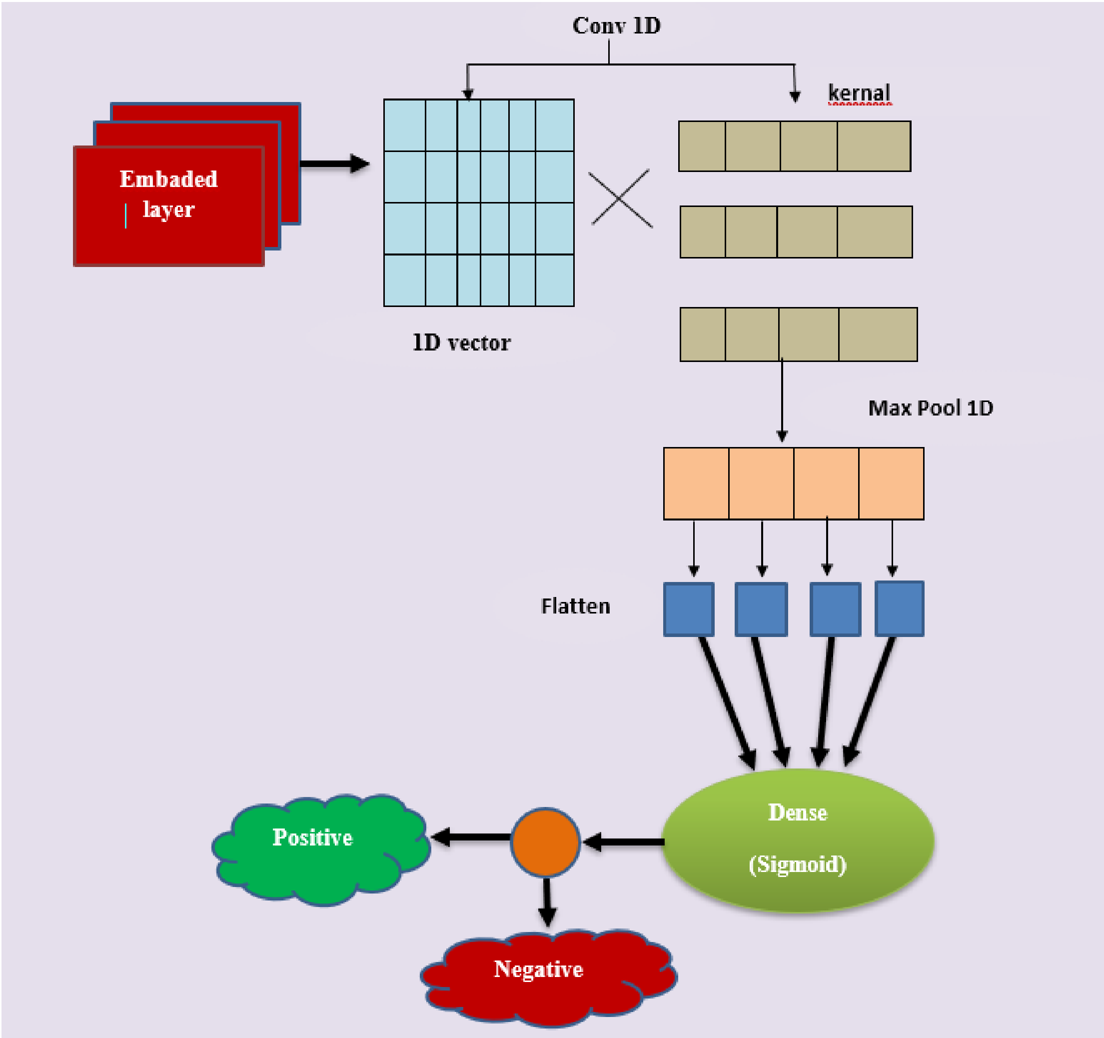

Figure 3.7 A view of CNN

3.6.2 Implement with Bidirectional Long Short Term Memory

LSTM is a special type of recurrent neural network. In LSTM for text classification. LSTM is a kind of sequence to sequence learning. The addition of traditional LSTM is known as Bidirectional LSTM.

In this kind of network, training is accomplished on two sides, where available timesteps of the input sequence are found. After embedding layer input, we feed the network with a Bi-LSTM layer with a different number of nodes and we also use Spatial dropout for better output. We have also used extra layer with nodes for optimal output. Different kinds of regularizers such as L1, L2 in terms of bias regularizer, kernel regularizer, recurrent regularizer also used for reducing overfitting. In Bi-LSTM here we use the Adam optimizer, Relu activation function. We have also used the dropout layer after each layer of feed-forward for better output we use the sigmoid activation function to classify the output for review analysis because it classifies the output in two different categories which positive and negative review. Figure-3.8 illustrate the institution of Bi- LSTM

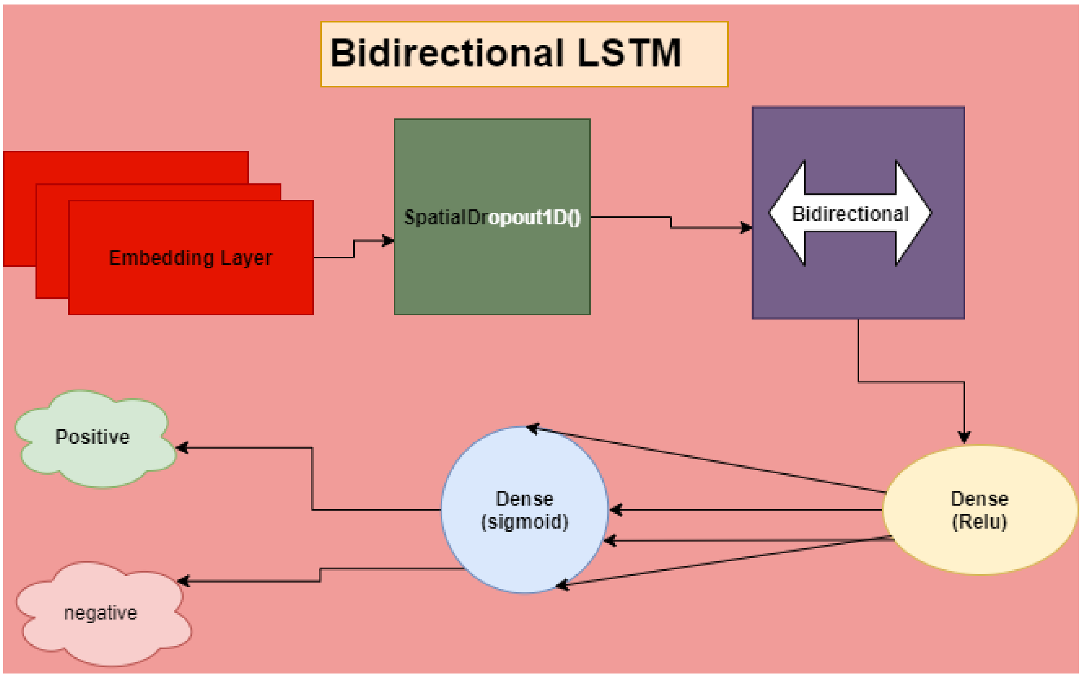

Figure 3.8 Different layer of Bi LSTM

3.7 SYSTEM ARCHITECTURE

3.7.1 A overall Architecture of Deep Learning

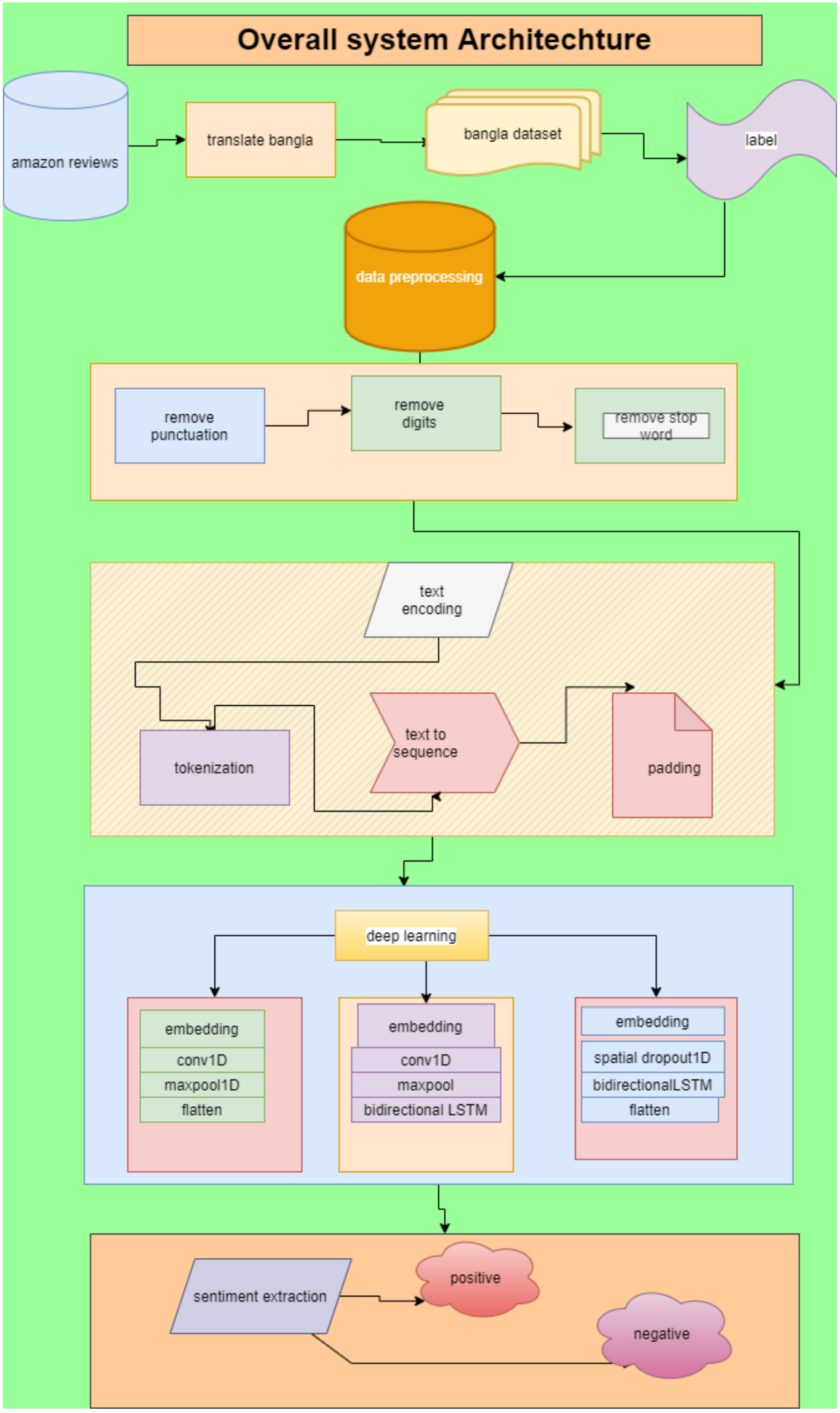

Figure 3.9 A overall architecture of deep learning

3.7.2 A overall Architecture of CNN

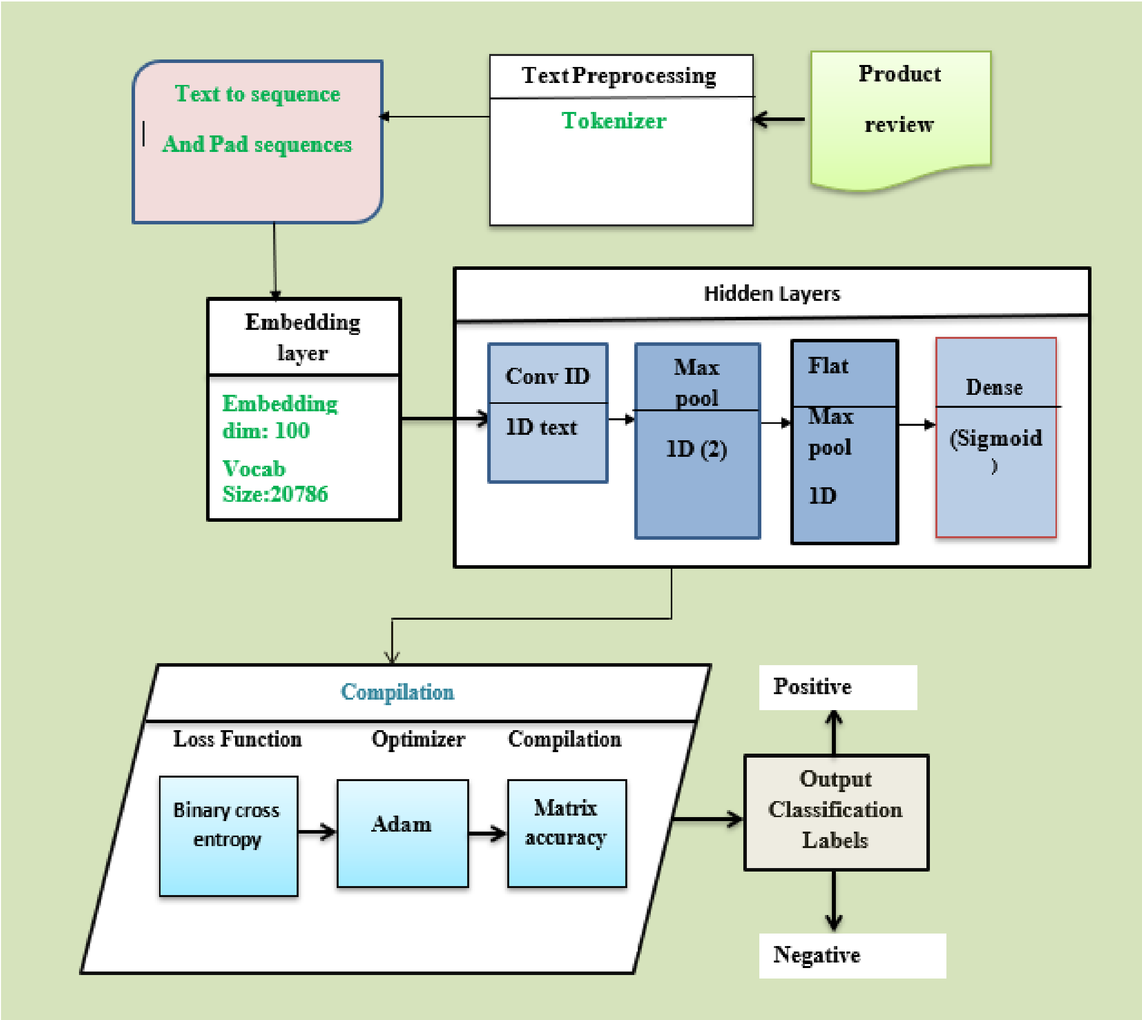

Figure 3.10 A overall architecture of CNN

3.7.3 A overall Architecture of Bi LSTM

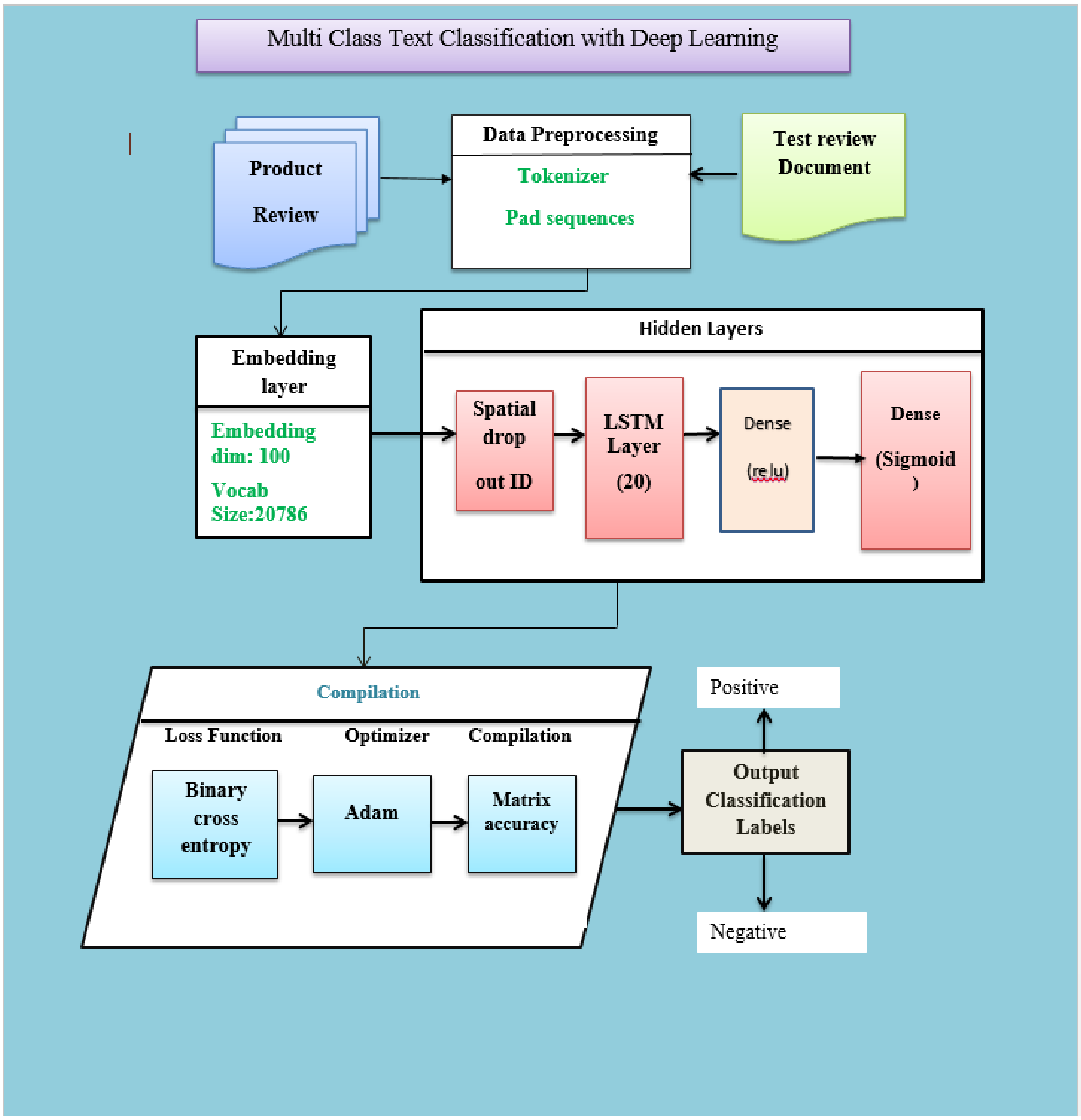

Figure 3.11 A overall architecture of Bi LSTM

3.7.4 Architecture of CNN with LSTM

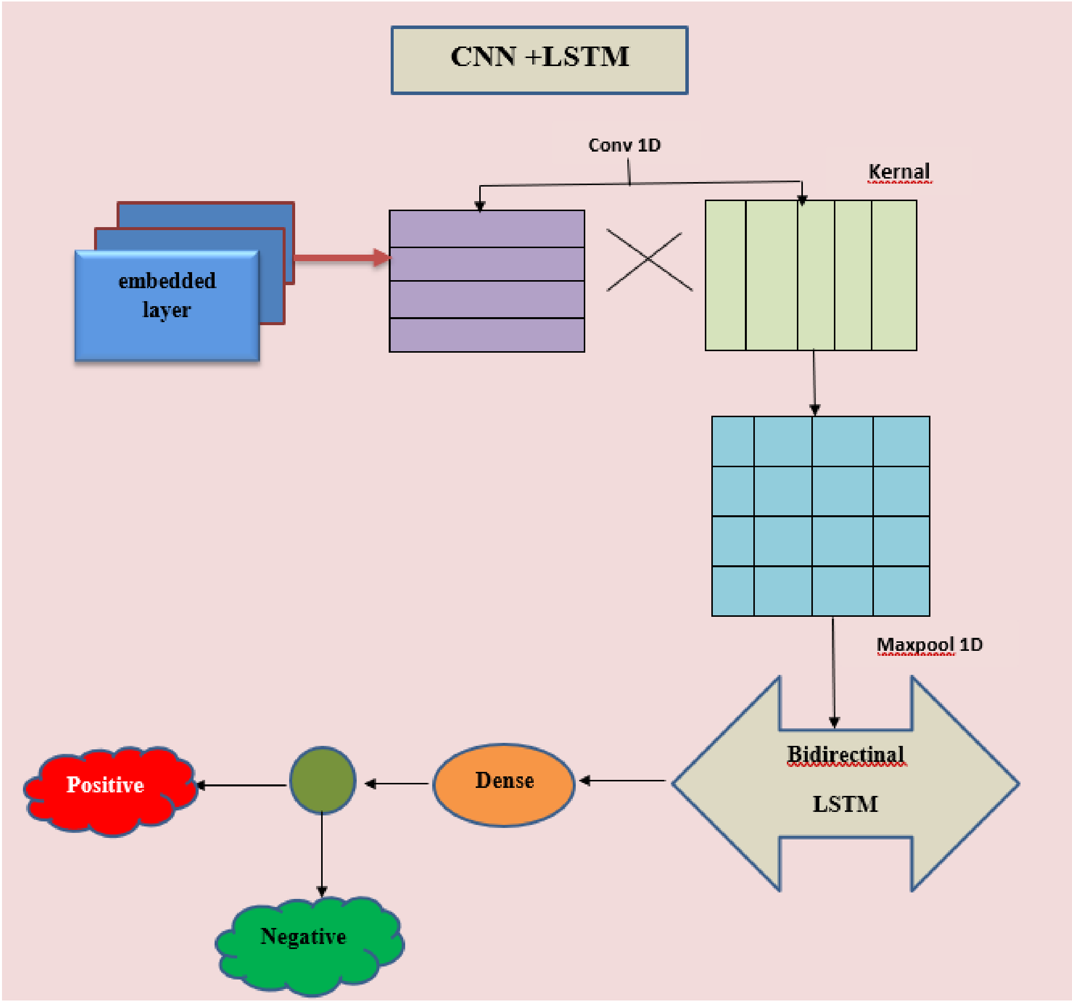

Figure 3.12 Architecture of CNN with LSTM

3.8 CHAPTER SUMMARY

In this chapter, the overall process of data collection from sources, construction of the Bangla dataset, and preprocessing of the dataset are discussed elaborately. Also, the implementation of methods with datasets are mentioned in a crystal clear way.

CHAPTER IV

RESULTS AND DISCUSSION

4.1 INTRODUCTION

The content of this chapter is about the result of our proposed framework and data analysis. For accomplishing this implementation, a popular platform “Google Colab’’ which is an online platform by using this program in the popular PYTHON language. Sentiment Analysis based on text classification and the outcomes of final results is elaborately discussed with Accuracy in this section. The performance of the framework measure by accuracy, precision, recall, and f1 are mentioned. Different kind of graphs which constructed in term of different parameters also shown in this particular section. On the other hand demonstration of implement with different deep learning algorithm and their relative comparison results are given also in this section.

4.2 TRAINING & TESTING

Training and testing are some of the biggest parts of the implementation. Training refers to the part of the task by which a portion of the dataset is applied to the framework to learn the machine. In the text classification training portion classify the text and extract the features and gain the characteristics for the test. Testing tends to the part of that task by which a portion of the dataset is being tested. After learning the characteristics from the training portion, the machine applies the learning at the time of testing. In the training and testing part, we split our dataset into the training and testing part. Two popular techniques of training and testing are-

- Percentage split validation Approach
- Cross Validation approach.

4.2.1 Percentage Split validation Approach

In this approach dataset are applied in training and testing by divided it in two parts.one part is used for training and another part is used for testing. The model acquire knowledge from train set and the knowledge is used for validation part. In this context here , we use 70% of our dataset for training and 30% for testing. We have 15000 data in our dataset so we use 10500 data for our training and 4500 data for testing. For preparing the model on the basis of data we use a fine portion of data set for training. We use 3:1 ratio for acquire the knowledge from learning and validation. Fig:4.1 shows an pi chart of train test split--

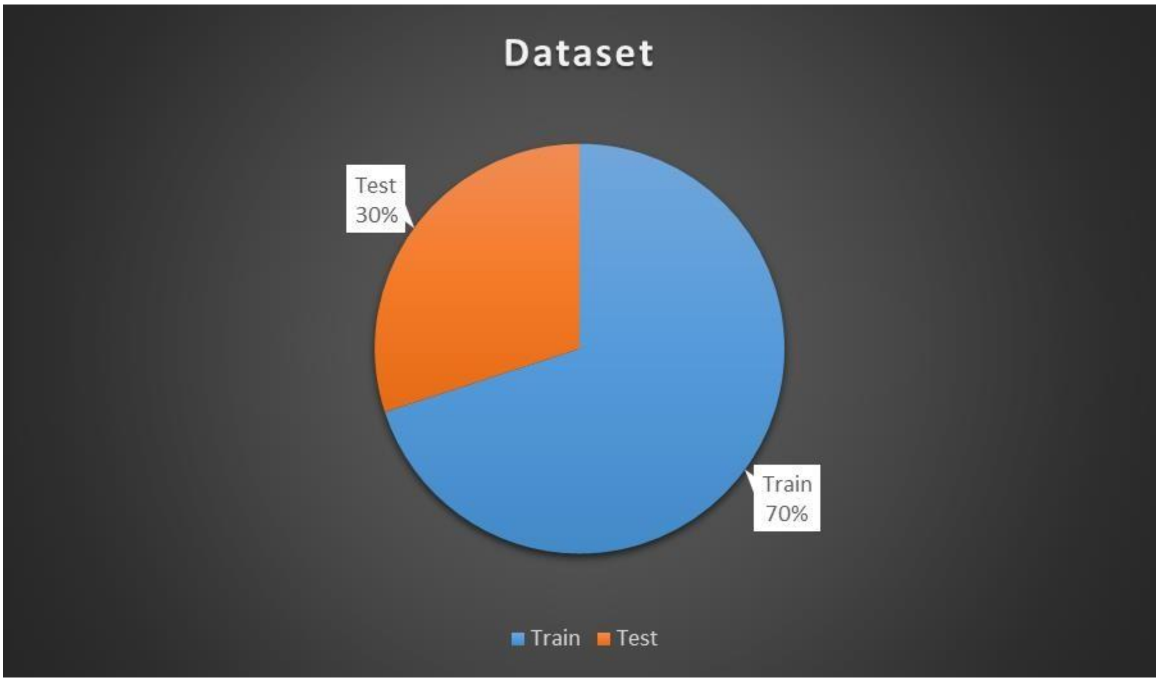

Figure 4.1 A chart of percentage validation approach

4.2.2 Cross Validation Approach

We evaluate the performance of the proposed work determined by accuracy. We split our Dataset in a margin, with regards to that margin some are used to train and some are used to testing. Evaluate the performance of the proposed methodology based on an error matrix to determine accuracy. However, the method is not so reliable to gain accuracy from a specific test set. The accuracy obtained from different test sets can make a difference in the accuracy obtained from a specific test set. K-fold cross-validation is an approach to test the dataset in a shuffling way.in this approach, datasets are split in k number of the fold, every time a fold is picked up for the test. K fold cross-validation (CV) ensures that every fold is used in a test in the same manner.

K Fold CV split the dataset in K Fold, every time fold are used as testing at the same point. Assuming the scenario of 5 fold CV, in the 5 fold CV datasets are split in 5 fold. At the time of the first iteration, 1st Fold is used as a testing and the rest of the data are used for training. At the time 2nd iteration, 2nd fold is used as testing, and the rest of the data is used as training. Data are split in k fold/section and the process is worked in this way until every fold is used as testing. We have used 5 fold cross-validation approach for our dataset so that the output of the dataset can be utilized in an optimal way.figure-4.2 shows a glance of 5 fold CV

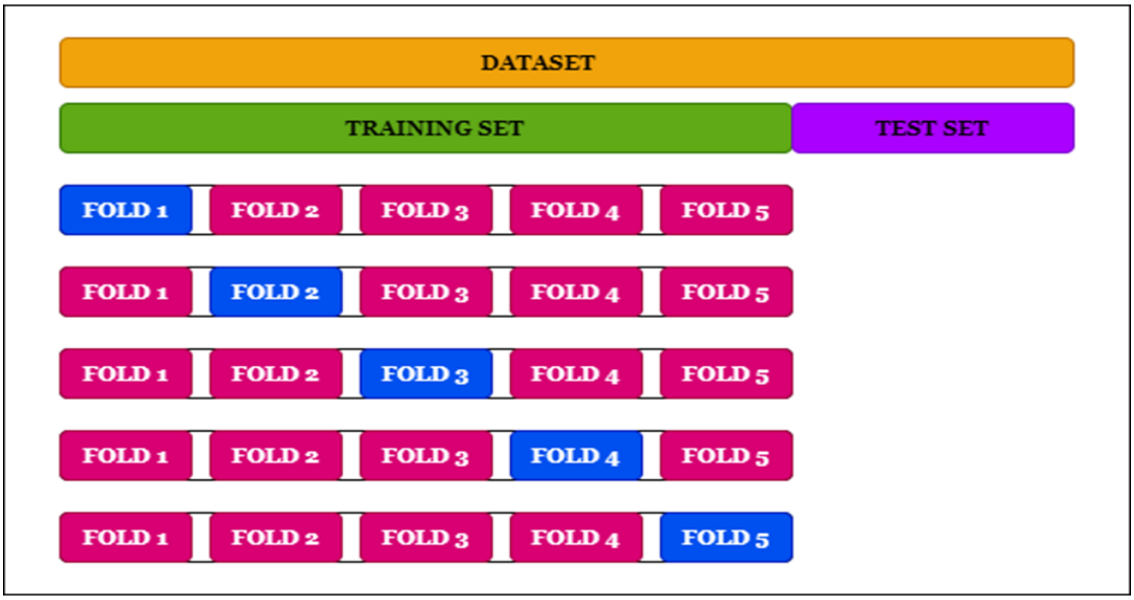

Figure 4.2 Fold cross validation approach

4.3 PERFORMANCE MEASURE

Different algorithms work with their defined procedure and produce different results. For our proposed methodology we have used different Deep Learning Algorithms and obtained outcomes. For simplifying our experiments popular technique of performance measure we adopted for accomplishing the measurements. We have used Precision, Recall, F-measure (F1) with an accuracy score.

Accuracy is a term which is deals with the correct output throughout the total input. Precision tends to the percentage of result which is relevant. On the other side, Recall is the measure of percentage relevant result acquire correctly classified by the algorithm. F- measure is basically the harmonic mean of the precision and recall.

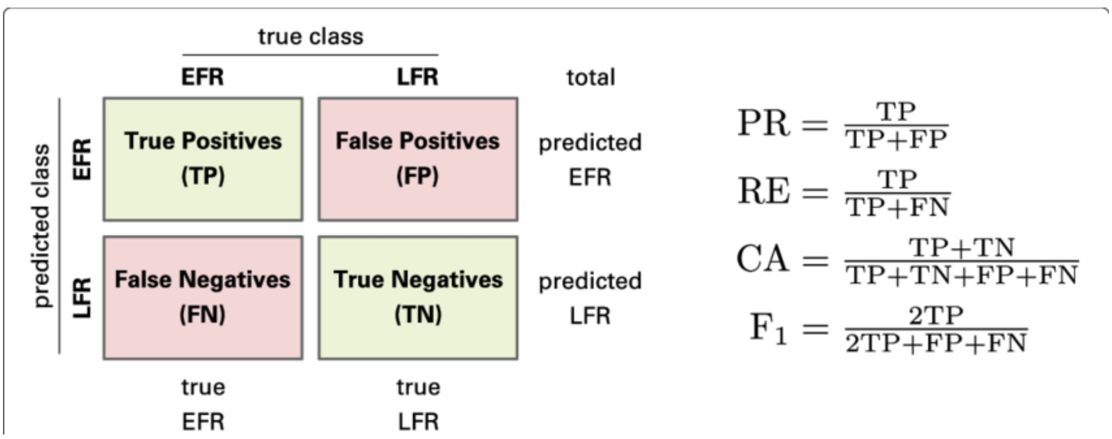

Figure 4.3 Equation of Precision, Recall and F1-measure

Precision measures the quality, whether Recall measures the quality. Higher precision is referred that the algorithm return more relevant results rather than the irrelevant one and Higher relevant is mean that the algorithm returns most of the relevant results.

F1 measure is most useful than Accuracy, especially when we have an asymmetric distribution of data.F1 score reaches the best value t 1 and worst at 0, high F1 means that the algorithm works properly and provides relevant results rather irrelevant one.

4.4 RESULTS FOR TEXT CLASSIFICATION

Sentiment analysis on the basis of text classification done with different deep learning algorithms. Text classification implemented with CNN, Special type of LSTM named Bi-LSTM, CNN with Bi-LSTM By using all these algorithms different types of results are obtained. So, those results are discussed elaborately with performance measurement

Text classification with CNN

Datasets are implemented with CNN by percentage validation. In CNN, all necessary Keras layers, pre-processing techniques, and tools are imported at the time of implementation. Tokenizing and pad sequences are applied before the process in the input layer. Text is processed through the first input layer is embedding layer than one by one hidden layer those are convolution layer and all other layers. It goes through by activation function called Relu. Finally, the model compiled with three different functions named binary cross-entropy, an optimizer named by adam. The loss function depicts to define the loss of model about fitting. In the output layer, it goes through the sigmoid function. To evaluate the model type of metrics named Accuracy is used.By Applying CNN we getting a result of accuracy that is 91.49%So here the graph of loss and validation loss is conveyed and the graph of accuracy and validation accuracy also portrayed. Here we implement text classification with 70% training data which is 10500 data and 30% for testing which is 4500 data. We apply CNN with 20 epochs with batch size 32. Figure:4.4 shows the graph of loss and validation loss function of CNN

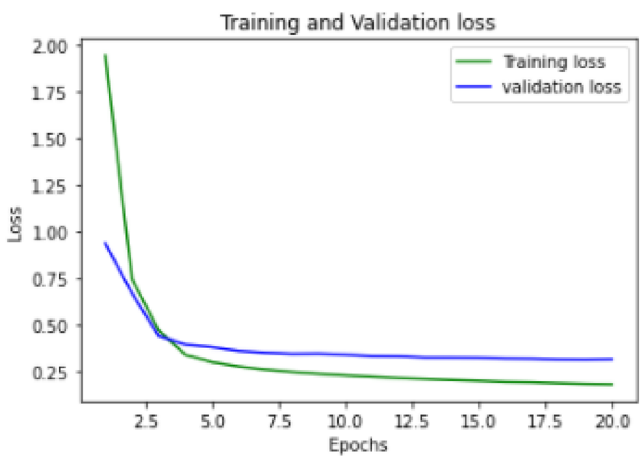

Figure 4.4 Loss and validation loss curve of CNN

From the graph, by analyzing the loss curve it can be described as the loss of training and validation data from the starting to the end decrease gradually.

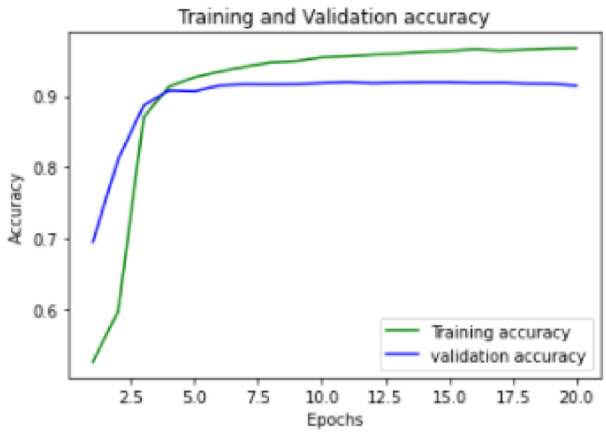

Figure 4.5 Training and validation accuracy graph of CNN

On the other side, the accuracy of testing increases gradually per step. In the end, training data achieve accuracy of 97% and the accuracy of testing data is.91.49%

Confusion metrics are used to evaluate the classification model. Confusion metrics portrayed a view of true values and predicted values. We use a confusion matrix to evaluate our model.

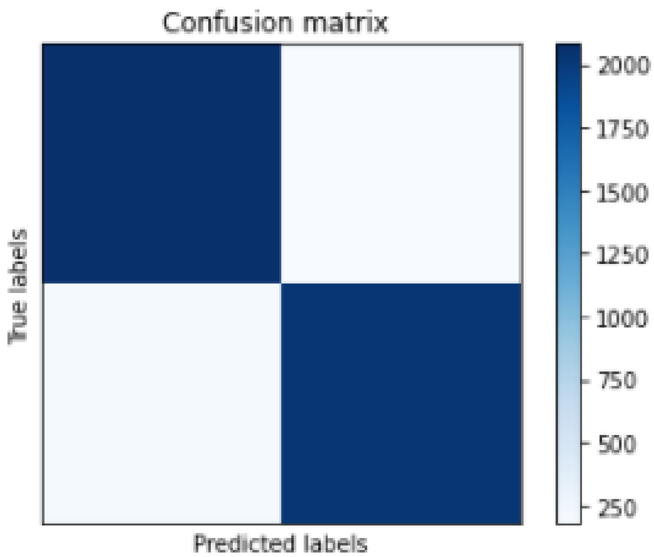

Figure 4.6 Confusion metrics of CNN

From the confusion metrics

[[2081 179]

[204 2036]]

It can be said that the CNN model correctly classifies the 4117 data out of 4500 testing data. We make a classification report of the CNN model, which describes the precision, recall, and f1 measure of the model.

```
              precision    recall  f1-score   support

           0       0.91      0.92      0.92      2260
           1       0.92      0.91      0.91      2240

    accuracy                           0.91      4500
   macro avg       0.91      0.91      0.91      4500
weighted avg       0.91      0.91      0.91      4500
```

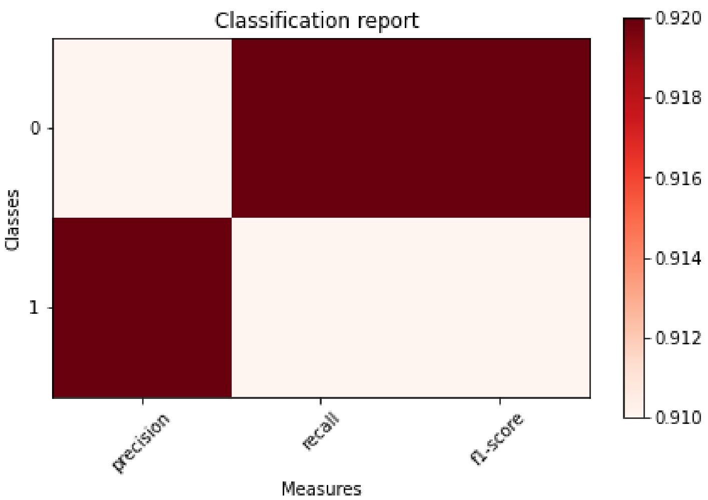

Figure 4.7 Classification report of CNN

It can be observed that Negative and positive polarity obtain different F1, precision and, recall score.

Table 8 Showing accuracy of CNN classifiers using 5- fold cross validation form

Split Method on Dataset	**Classifiers of DL**	**Accuracy in different steps**	**Average Accuracy**
5-fold cross validation	Convolutional neural network	0.937 0.9596666666666667 0.98 0.8806666666666667 0.914971657219073	93.45%

4.4.1 Text classification with Bi LSTM

We implement our dataset with bidirectional LSTM. From the very beginning, the input is passed through the first input layer which is the embedding layer. Then its transits to all the hidden layers one by one beginning with spatial dropout, then bidirectional LSTM layer. Then the model compiles with different functions like loss function as binary cross-entropy, Adam as an optimizer, and sigmoid active function. At the output layer, its transits through the sigmoid function and different kinds of dropout. further, accuracy metrics accomplish the task of experiments. By applying Bidirectional LSTM 92.13% accuracy obtained. From the graph, by analyzing the loss curve it can be described as the loss of training and validation data from the starting to the end decrease gradually.

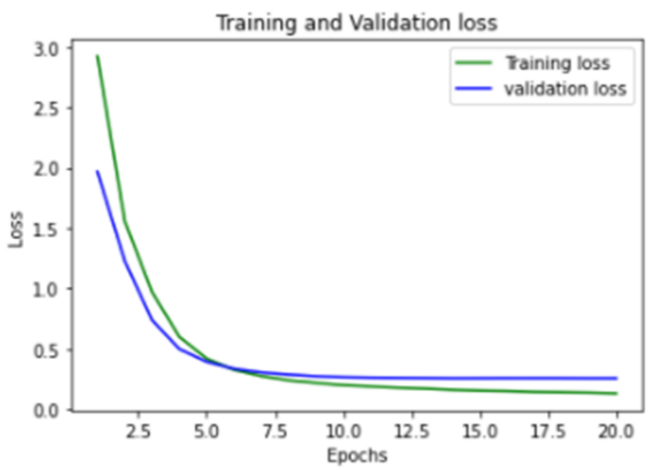

Figure 4.8 Loss and validation loss curve of Bi LSTM model

In fig 4.9 the accuracy of training and testing data implemented with Bi LSTM is plotted. From the curve, we can say that the accuracy of training data from the starting is relatively less than testing data. afterward, per step, the accuracy of training data increases gradually.

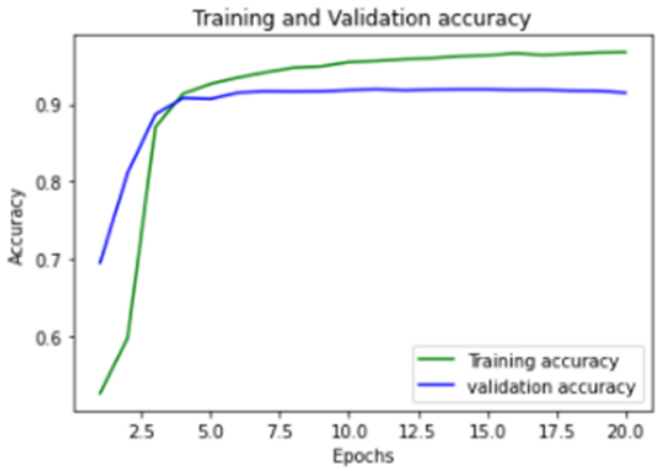

Figure 4.9 Training & validation accuracy graph of Bi LSTM model

In the following figure 4.10 show the confusion metrics for Bi LSTM model. This metrics portrayed a view of true values and predicted values.

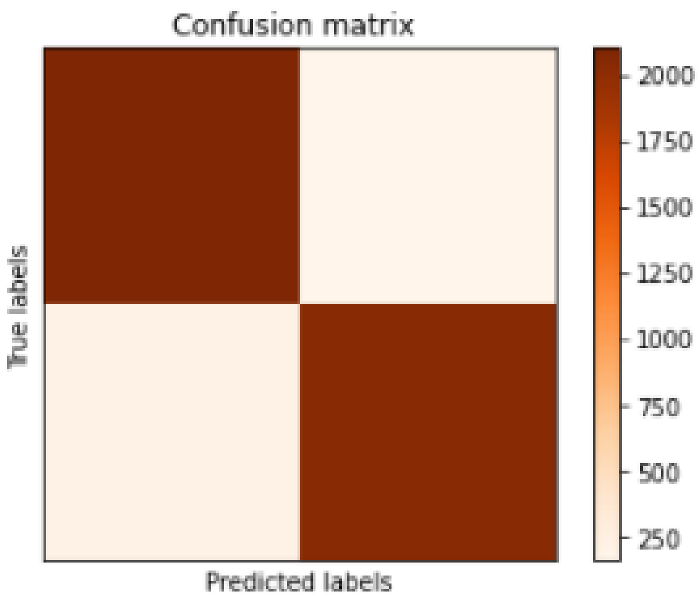

Figure 4.10 Confusion metrics of Bi LSTM

From the confusion metrics

[[2100 160

201 2039]]

It can be said that the bidirectional LSTM model correctly classifies the 4139 data out of 4500 testing data. A classification report is plotted for bidirectional LSTM which produce precision, recall, and f1 score for this bidirectional LSTM model.

	precision	recall	f1-score	support
0	0.93	0.92	0.92	2260
1	0.92	0.93	0.92	2240
accuracy			0.92	4500
macro avg	0.92	0.92	0.92	4500
weighted avg	0.92	0.92	0.92	4500

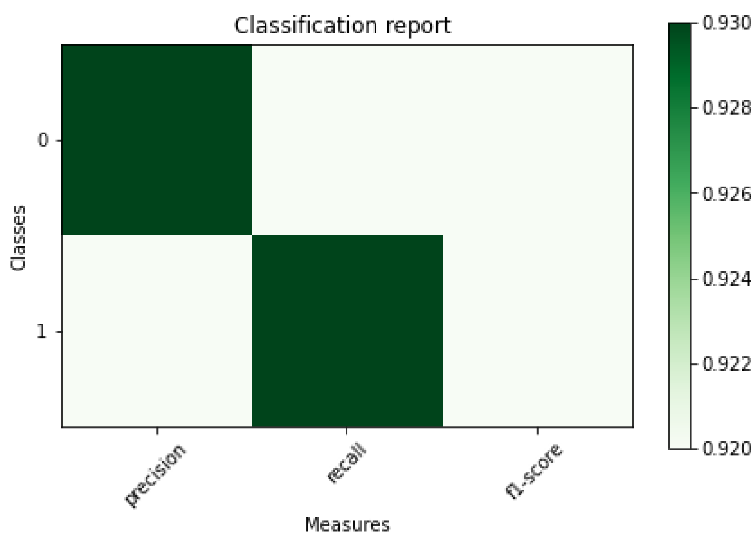

Figure 4.11 Classification report of Bi LSTM model

It can be described that negative and positive polarity obtain the same F1, precision, and recall different.

Table 9 Showing accuracy of LSTM classifiers using 5- fold cross validation form

Split Method on Dataset	**Classifiers of DL**	**Accuracy in different steps**	**Average Accuracy**
5-fold cross validation	Bi Long Short Term Memory	0.9413333333333334 0.9553333333333334 0.9866666666666667 0.932 0.9213	94.75%

4.4.2 Text classification with CNN with LSTM

CNN with LSTM is a combination of the Convolution layer and LSTM layer. LSTM is a special type of recurrent neural network. After the pre-processing of the data inconvenient manner, the data are processed with the first layer which is the embedding layer, embedding layer the inputs are passed through the convolution layer, max pool. Afterward, the input passes the layer of Bidirectional LSTM.it goes through an activation function called Relu. Moreover, three types of functions also used. One of them is Binary cross-entropy, an optimizer name Adam also processed. in the final layer, the inputs are passed by the sigmoid function. Accuracy metrics are used to evaluate the rate of the classification model.By implementing CNN with LSTM 91.01%.Here the graph of the loss function of training and validation. The curve of training loss portrays that the loss of training decrease gradually. On the other hand, validation loss conveyed a zig-zag curve, sometimes it decreases with epoch or sometimes it increases with epochs.

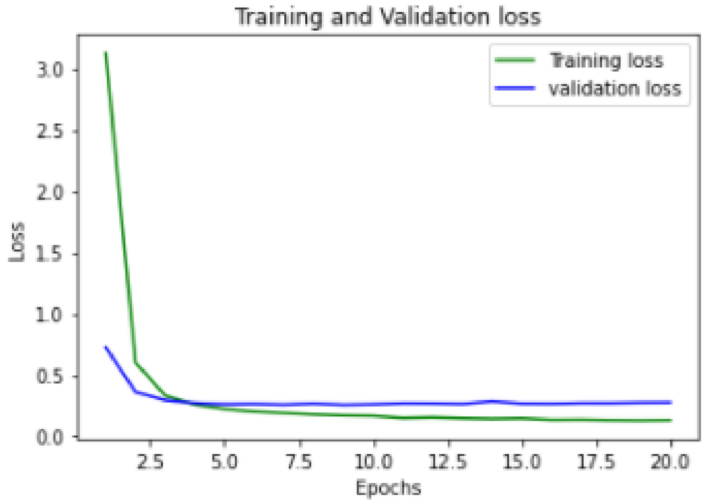

Figure 4.12 Loss and validation loss curve of CNN with LSTM

From the graph of training and testing accuracy, it can be observed that the training accuracy increase in a good margin at the second epoch after that certain point the training accuracy increases gradually.

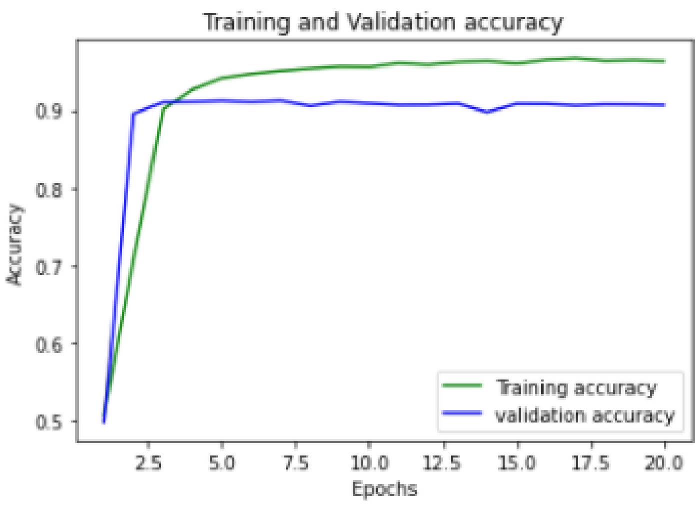

Figure 4.13 Training & validation accuracy graph of CNN with LSTM

The validation accuracy is in a zigzag form at the second epoch the validation was increased then afterward it decreases. It just works with increase and decreases form. In the end, the accuracy of training data was 97% and the accuracy of validation data was 91.01%.

In the following figure 4.14 show the confusion metrics for CNN with LSTM mode. This metrics portrayed a view of true values and predicted values.

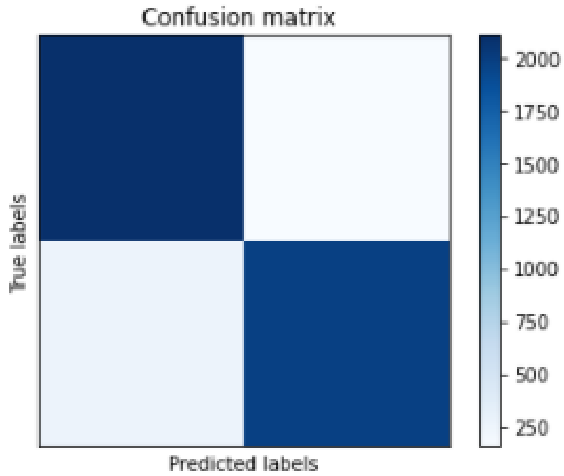

Figure 4.14 Confusion metrics of CNN with LSTM

From the confusion metrics

[[2073 187]

[209 2031]]

It can be said that the model CNN with LSTM can correctly classify 4104 data out of 4500 testing data.

We make a classification report on this model, which makes a result about precision, recall, and f1.

```
              precision    recall  f1-score   support

           0       0.91      0.92      0.91      2260
           1       0.91      0.91      0.91      2240

   micro avg       0.91      0.91      0.91      4500
   macro avg       0.91      0.91      0.91      4500
weighted avg       0.91      0.91      0.91      4500
 samples avg       0.91      0.91      0.91      4500
```

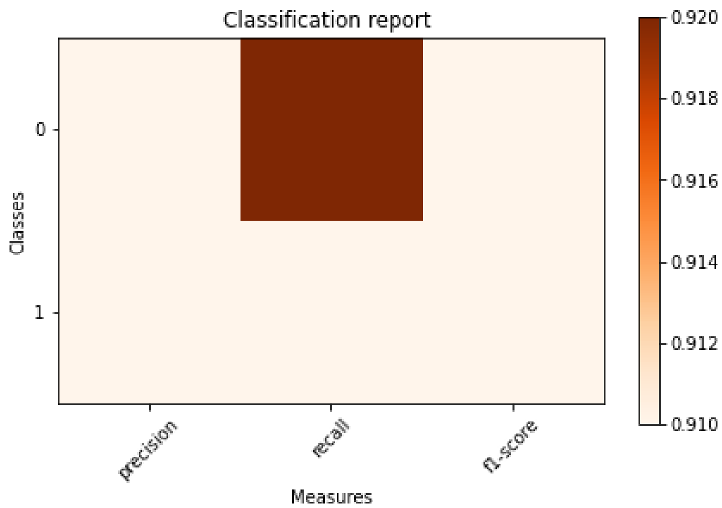

Figure 4.15 Classification report of CNN with LSTM

It can be observed that negative and positive polarity obtain same F1, precision, and recall different.

Table 10 Showing accuracy of CNN with LSTM classifiers using 5- fold cross validation form

Split Method on Dataset	**Classifiers of DL**	**Accuracy in different steps**	**Average Accuracy**
5-fold cross validation	CNN with LSTM	0.955 0.9293333333333333 0.9826666666666667 0.8243333333333334 0.9064	91.95%

4.5 COMPERATIVE RESULT OF THREE MODEL

Product review analysis in term of Bangla language is by using a deep learning approach is our proposed methodology. We configure three different models for our classification problem. The models are CNN, Bi LSTM, and CNN with LSTM. Implement our dataset with these three different types of models we get different types of accuracy because the working procedure of these models varies from each other. the specific model working with their own layer. So that the accuracy quite varies from each of them. A comparative graph of these three models are given below-

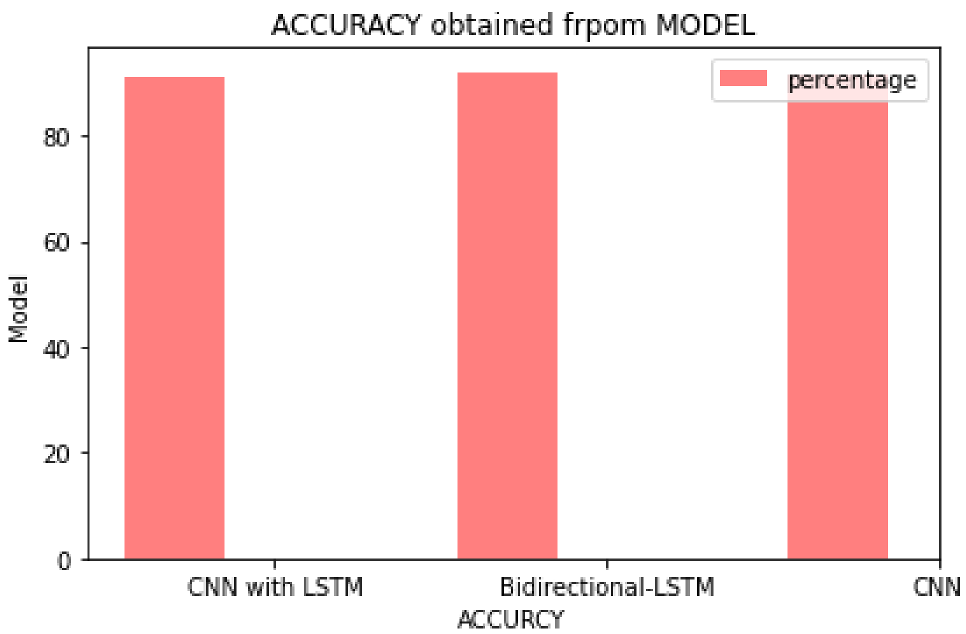

Figure 4.16 Comparative graph of three model

From the graph, it can be observed that by Implementing CNN model with our dataset which is containing 15000 Bangla data obtain 91.49% accuracy in term of 30% testing data.By implementing the Bi LSTM model with the Bangla dataset, it obtains 92.13% accuracy. By implementing CNN with LSTM it obtains 91.01% accuracy.

From these comparison graphical views, it can be said that the model bidirectional LSTM obtain the highest accuracy among these three models.

.

4.6 ROC CURVE OF THREE MODEL

ROC curve is a curve that summarizes the tradeoff between true positive rate and false-positive rate with different thresholds. by the curve the measure of the specific model rate of classification

AUC is a score under the ROC curve, which provides an aggregate measure of performance of all possible values beyond the thresholds. AUC score of 0.5 to 1 means it is an excellent model and less than 0.5 means the model is not working properly

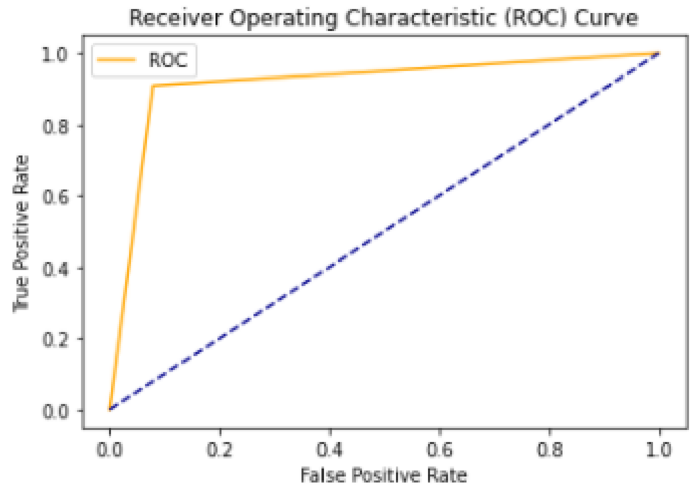

Figure 4.17 ROC curve of CNN

The AUC score of the CNN is 0.91 which is in the range of 0.5 to 1. That means the model work Pretty well.

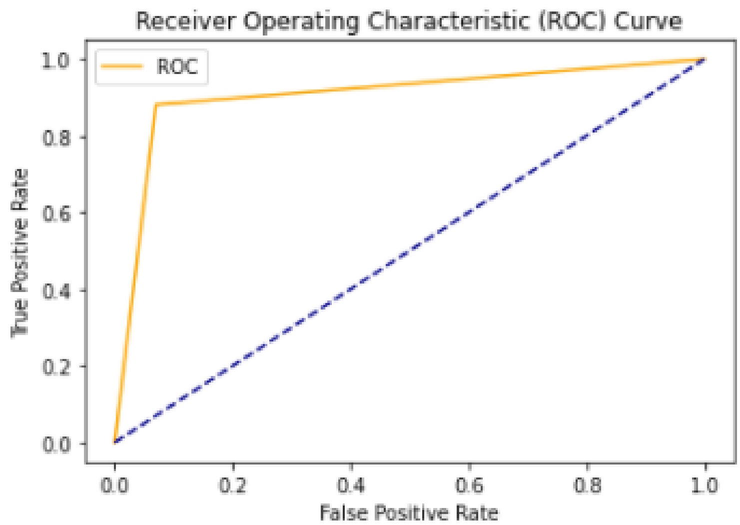

Figure 4.18 ROC curve of CNN with LSTM

The AUC score of the CNN with LSTM model is 0.91 which is pretty good in term of the range of AUC.

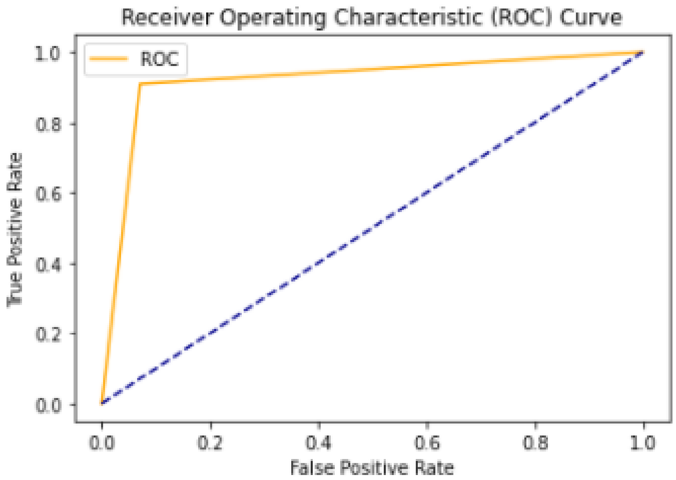

Figure 4.19 ROC curve of bidirectional Bi LSTM

The AUC score of the Bi LSTM model is 0.92 which is pretty good in term of the range of AUC.

4.7 CHAPTER SUMMARY

In this chapter, the content has been described in a crystal clear manner. The type of splitting dataset in training and testing, performance measure with different parameters, evaluate the model with different metrics also mentioned. The different kind of graph which illustrate the figure in term of different parameters also plotted in this section. The performance of individual model which is measure by accuracy also described, also a comparative view of used models also demonstrated. Overall in this section, sentiment analysis implemented with different deep learning approaches portrayed with accuracy, performance, and graphical representation.

CHAPTER V

CONCLUSION AND FUTURE WORKS

5.1 CONCLUSION

In the summary of the study, there is no doubt that there is a lot of research work in natural language processing, especially in the English language. Due to textual complexity, review analysis from the Bangla language is very hard. Extensive research has been done on social media to analyze the content of English texts, but less one has done this work with Bangla data yet. While the results of this type of work are revolutionizing change in our computing life, recently, such kind of research is being increased this time. With the blessings of such research work, we get some outstanding real-life applications. However, it is very unfortunate that there is less research work on the Bengali language. But the hope for us is that many researchers from different countries have started researching in this field. In our research work, we do some approaches to classify product reviews in the Bangla language. The key findings of our experiments are summarized as follows:

- With the deep learning approach, a good accuracy of 91.49% has been achieved with the CNN approach in the case of sentiment analysis.

- Implementation of sentiment analysis yields an accuracy of 91.01% with the deep learning approach convolutional neural network with LSTM.

- In the case of sentiment analysis using the Bidirectional LSTM approach a fair accuracy of 92.13% has been achieved.

5.2 SUMMARY OF FINDINGS

Developing efficient Bangla text classification is a crucial issue, Bengali is a rich language with many variations as well as electronic documents based on Bengali have been published very fast both online and offline. We can say without any doubt that the Bangla language is rich enough to work with and apply to various NLP tasks.

5.3 LIMITATIONS

The following is a list of limitations of our present work:

- In the case of sentiment analysis all models cannot show satisfactory performance and in the hybrid model Combination of CNN and Bi-LSTM has shown an accuracy less than the highest achieved one in terms of the deep learning approach.

- Dataset for sentiment analysis is comparatively less authentic.

5.4 FUTURE SCOPE

The future scope of our research work is to enlarging our dataset. Making an approach that can detect fake reviews. Working with neural layer. Adding neutral polarity in sentiment analysis and Implement with other deep learning methods including HAN .

5.5 CHAPTER SUMMARY

In this chapter, the contents have been arranged in such a way that narrates the procedure starting from explaining and figuring out the overall findings from our experimental results This also summarizes and gives a comparative analysis about the result. The later sections unbraid the remarks that are needed to be worked out in the future and the scope of future work at the borderline of this chapter.

REFERENCES

[1] Turney, Peter D., “Thumbs up or thumbs down? Semantic Orientation Applied to Unsupervised Classification of Reviews”, Proceedings of Association for Computational Linguistics, Philadelphia, PA. July 2002, pp. 417-424.

[2] Ji fang, Bi Chen, “Incorporating Lexicon Knowledge into SVM Learning to Improve Sentiment Classification”, Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP), IJCNLP 2011, pp. 94–100.

[3] Xing Fang* and Justin Zhan “Sentiment analysis using product review data” Fang and Zhan Journal of Big Data (2015) 2:5.

[4] Hu M, Liu B (2004) Mining and summarizing customer reviews. In: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, New York, NY, USA. pp 168–177.

[5] R. Moraes, J. F. Valiati, and W. P. G. Neto, “Document-level sentiment classification: An empirical comparison between svm and ann,” Expert Systems with Applications, vol. 40, no. 2, 2013, pp. 621–633.

[6] Han, J. and Kamber, M. ”Data Mining: Concepts and Techniques.” 2nd Edition. Morgan Kaufmann Publishers, San Francisco, USA. (ISBN-55860-901-6), 2006.

[7] R. Duda, P. Hart, and D. stork, ”Pattern Classification.” 2nd Edition, Wiley Interscience, 2001.

[8] E. Frank, and I. Witten,”Data Mining: Practical Machine Learning Tools and Techniques.” 2nd Edition, Morgan Kaufmann, San Francisco, 2005.

[9] Galathiya, A. S., A. P. Ganatra, and C. K. Bhensdadia. ”Improved Decision Tree Induction Algorithm with Feature Selection, Cross Validation, Model Complexity and Reduced Error Pruning.” International Journal of Computer Science and Information Technologies 3.2 (2012): 3427-3431.

[10] C. Z. Jiajun Zhang, Neural Networks in Machine Translation: An Overview, IEEE Intell. Syst., pp. 17241734, 2015.

[11] Bullinaria, J.A., Recurrent neural networks. Neural Computation: Lecture, 2013.12

[12] Olah, C., Understanding LSTM Networks. 2016.

[13] Hochreiter, S. and J. Schmidhuber, Long short-term memory. Neural computation, 1997. 9(8): p. 1735-1780

[14] Gers, F.A., J. Schmidhuber, and F. Cummins, Learning to forget: Continual prediction with LSTM. Neural computation, 2000. 12(10): p. 2451-2471.

[15] Elman, J.L., Finding structure in time. Cognitive science, 1990. 14(2): p. 179-211

[16] Rohit Narayan, Review Spam Detection Using Machine Learning Techniques, 2016 5:p. 769-008.

[17] M.N. Istiaq Ahsan, Tamzid Nahian, Abdullah All Kafi, Md. Ismail Hossain, and Faisal Muhammad Shah. An ensemble approach to detect review spam using hybrid machine learning technique. 19th International Conference on Computer and Information Technology, Dhaka, December 18-20, 2016.

[18] H. M. Ahmed, M. Javed Awan, N. S. Khan, A. Yasin, and H. M. Faisal Shehzad, “Sentiment analysis of online food reviews using Big Data analytics,” 2021.

[19] S. Al Mostakim, F. Ehsan, S. Mahdiea Hasan, S. Islam and S. Shatabda, "Bangla Content Categorization Using Text Based Supervised Learning Methods," 2018 International Conference on Bangla Speech and Language Processing (ICBSLP), 2018, pp. 1-6.

[20] A. Dhar, H. Mukherjee, N. Sekhar Dash, and K. Roy, “Performance of classifiers in Bangla text categorization,” in 2018 International Conference on Innovations in Science, Engineering and Technology (ICISET), 2018, pp. 168–173.

[21] M. Mansur, “Analysis of N-Gram based text categorization for Bangla in a newspaper corpus,” BRAC University, 2006.

[22] A. K. Mandal and R. Sen, “Supervised learning Methods for Bangla Web Document Categorization,” arXiv [cs.CL], 2014.

[23] R. A. Rabbi, M. W. A. Khan, Shaown, and M. S. Alam, “Product rating generation based on public opinion using sentiment analysis,” BRAC University, 2017.

[24] R. Safrin, K. R. Sharmila, T. S. Shri Subangi, and E. A. Vimal, “SENTIMENT ANALYSIS ON ONLINE PRODUCT REVIEW,” Cloudfront.net, 2008.

[25] X. Fang and J. Zhan, “Sentiment analysis using product review data,” J. Big Data, vol. 2, no. 1, 2015.

[26] X. Li, L. Xie, F. Zhang, and H. Wang, “Online deceptive product review detection leveraging word embedding,” in 2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech), 2017, pp. 867–870.

[27] A. Tripathy, A. Agrawal, and S. K. Rath, “Classification of sentiment reviews using n-gram machine learning approach,” Expert Syst. Appl., vol. 57, pp. 117–126, 2016.

[28] D. Zhang, H. Xu, Z. Su, and Y. Xu, “Chinese comments sentiment classification based on word2vec and SVMperf,” Expert Syst. Appl., vol. 42, no. 4, pp. 1857–1863, 2015.

[29] A. Tripathy, A. Agrawal, and S. K. Rath, “Classification of sentiment reviews using n-gram machine learning approach,” Expert Syst. Appl., vol. 57, pp. 117–126, 2016.

[30] Lee, S. Yi, S. Hyun, and C. Kim, “Review on the recent welding research with application of CNN-based deep learning part I: Models and applications,” J. Weld. Join., vol. 39, no. 1, pp. 10–19, 2021.

REBUTTAL LETTER
AFTER PRE-DEFENSE EXAMINATIONS

Submission ID:137
Submission Title: Product Review Analysis by using deep learning approaches.

Examiner Name:

CHAPTER	EXAMINER'S SUGGESTION	PAGE	CORRECTION CONTENTS	PAGE
3	Reducing the number of Dataset	26-27	Data Statistics and About data	26-27
3	Reduce data pre-processing steps	28	Data pre-processing	28
3	To change optimizer for Bidirectional LSTM model	36-37	Implementing with Bi-LSTM	36-37
4	Overfitting problem	48-49	Text classification with CNN	48-49
4	Overfitting problem	52-53	Text classification with Bi LSTM	52-53
4	Overfitting problem	57-58	Text classification with CNN with LSTM	57-58

STUDENT NAME (ID): AYESHA JULEKHA(C163237)

PROMA MUTSUDDI(C163215)

Ayesha Proma

SIGNATURE

SUPERVISOR NAME: DR. ABDUL KADAR MUHAMMAD MASUM

SUPERVISOR SIGNATURE: