

**SENTIMENT ANALYSIS ON BANGLADESHI E-COMMERCE
PRODUCT REVIEW DATA**

BY

SHRABONI SARKER

ID: 173-15-10324

MAHBUBA YESMIN

ID: 173-15-10291

AND

A.K.M EAMIN ASRU

ID: 173-15-10265

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

Supervised By

Mr. Md. Azizul Hakim

Lecturer (Senior Scale)

Department of CSE

Daffodil International University

Co-Supervised By

Md. Sanzidul Islam

Lecturer

Department of CSE

Daffodil International University



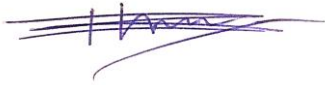
DAFFODIL INTERNATIONAL UNIVERSITY DHAKA, BANGLADESH
September 2021

APPROVAL

This Project/internship titled “Sentiment Analysis on Bangladeshi E- commerce Product Review Data”, submitted by Shraboni Sarker, ID No: 173- 15-10324, Mahbuba Yesmin, ID No: 173-15-10291 and A.K.M. Eamin Asru , ID No: 173-15-10265 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 9 September 2021.

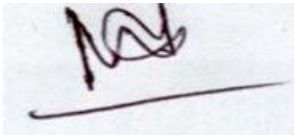
BOARD OF EXAMINERS

Chairman



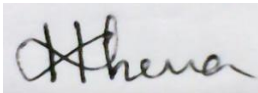
Dr. Touhid Bhuiyan
Professor and Head

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Dr. Md. Ismail Jabiullah
Professor

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Most. Hasna Hena
Assistant Professor

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Internal Examiner

Internal Examiner

External Examiner

Dr. Dewan Md. Farid
Associate Professor
Department of Computer Science and Engineering
United International University

Declaration

We hereby declare that, this project has been done by us under the supervision of **Mr. Azizul Hakim, Lecturer (Senior Scale), Department of CSE** 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:



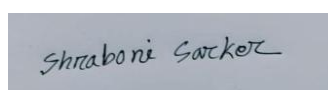
Mr. Md. Azizul Hakim
Lecturer (Senior Scale)
Department of CSE
Daffodil International University

CO-SUPERVISED BY:



Md. Sanzidul Islam
Lecturer
Department of CSE
Daffodil International University

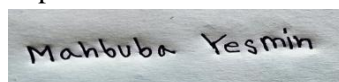
SUBMITTED BY:



SHRABONI SARKER

ID: 173-15-10324

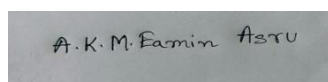
Department of CSE



MAHBUBA YESMIN

ID: 173-15-10291

Department of CSE



A.K.M EAMIN ASRU

ID: 173-15-10265

Department of CSE

Acknowledgment

First, we express our heartiest thanks and gratefulness to almighty Allah for His divine blessing makes us possible to complete the final year Thesis successfully, Alhamdulillah.

We really grateful and wish our profound our indebtedness to Supervisor **Mr. Md. Azizul Hakim, Senior Lecturer** , Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of “*Artificial Intelligence*” to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stage have made it possible to complete this project. We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

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

ABSTRACT

In this paper, we propose a sentiment analysis process for online product reviews written in Bengali. In recent years, with the rapid development of Internet technology, online shopping has become a dominant method of user purchasing and consumption. Analyzing the emotions of multiple numbers of user reviews on e-commerce platforms can help determine whether a user's attitude is positive, neutral or negative, it captures the opinions, beliefs and feelings of each user about the respective product. This work presents a machine learning-based technique to identify sentiment polarities (positive, negative or neutral category) from online product reviews in Bengali. To evaluate the effectiveness of the proposed method, a corpus containing 2228 reviews was extracted using the Daraz API of Bengali online product reviews is being developed. Comparative analysis with various approaches LR, DT, RF, MNB, KNN, SVM, SGD etc. Consider the characteristics of unigram, bigram, and trigram features, respectively. Experimental results reveal that the stochastic gradient descent (SGD) with the Tri-gram feature outperforms the other techniques with 88% accuracy.

CONTENTS	PAGE NO
Approval	iii
Declaration	iii
Acknowledgment	iv
Abstract	v
CHAPTERS1:INTRODUCTION	1-5
1.1 Introduction:	1
1.2 Motivation	3
1.3Rationale of the Study	3
1.4Research Question	3
1.5 Expected Outcome	4
1.6 Report Layout	5
CHAPTER2:BACKGROUND	5-10
2.1 Introduction	5
2.2Related works	5
2.3 Research Summary	7
2.4 Scope of the Problem	9
2.5 Challenges	9
CHAPTER 3 RESEARCH METHODOLOGY	10-17
3.1 Introduction	10
3.2 Research Subject and Instrumentation	10
3.3 Data Collection Procedure	11
3.4 Statistical Analysis	13
3.5 Implementation Requirements	13
	vi

3.5.1 Data collection	15
	16
3.5.2 Data Pre-processing	16
3.5.3 Tokenization and Punctuation Removal	16
3.5.4 Remove Stopword	17
3.5.5 Stemming	17
3.5.6 Applying Algorithm	18
3.5.7 Accuracy	18
CHAPTER 4: EXPERIMENTAL RESULTS AND DISCUSSION	19-23
4.1 Introduction	19
4.2 Experimental Results	19
4.3 Descriptive Analysis	21
CHAPTER 5: SUMMARY, CONCLUSION, RECOMMENDATION AND IMPLICATION FOR FUTURE RESEARCH	24-26
5.1 Summary of the Study	24
5.2 conclusions	25
5.3 Recommendations	25
5.4 Implication For Further Study	26
REFERENCES	27

LIST OF FIGURES

FIGURES	PAGE NO
Figure 3.1: All Required Library Functions are Imported	11
Figure 3.2: Dataset	12
Figure 3.3: Data statistics	13
Figure 3.4: Workflow	14
Figure 3.5: Length Distribution	15
Figure 4.1: Classifiers performance on Uni-gram features	21
Figure 4.2: Classifiers performance on Bi-gram features	21
Figure 4.3: Classifiers performance on Trigram features	22

LIST OF TABLES

TABLES	PAGE NO
Table 2.1 : Research summary	8
Table 3.1: Sample reviews	12
Table 4.1: Performance measurement accuracy of dataset using different algorithms	20

LIST OF ABBREVIATION

DIU – Daffodil

International University

CSE – Computer Science

and Engineering

ML – Machine Learning

NLP – Natural

Language Processing

NLTK – Natural

Language Toolkit

SVM – Support

Vector Machine

SGD- Stochastic

Gradient Descent.

TF-IDF-Term

Frequency — Inverse

Document Frequency

CHAPTER:1

INTRODUCTION

1.1 Introduction

E-Commerce is a significant part of the economy that establishes buying and selling of services using the internet between merchants and customers beyond the barricades of geographical areas. The ongoing pandemic has locked the entire world at home, but the demand has not left the market. As more and more customers shop online rather than in traditional retail stores, it is becoming a rapidly growing business unit. Shopping online is on the increase. The government has designated product delivery as an urgent service in this circumstance. Internet connections have improved in recent years, and the number of people with internet access has expanded, allowing the e-commerce industry to flourish. Manual analysis of textual reviews is required a too long time. For a vast amount of textual data analysis, it is quite impossible. For actual knowledge extraction from this vast amount of textual data, we need to create a system that can be done through Sentiment Analysis, which plays a significant part as a Natural Language Processing tool to detect the polarity of the textual data.

E-commerce platforms also allow the opportunity for customers to post reviews as positive, neutral, or negative about the product. This becomes very influential on the future customers who want to purchase the same product. It helps consumers to appreciate the information such as the quality of the product but also helps manufacturers to upgrade product quality, and merchants can adjust sales strategies and increase service quality based on product reviews, which are of large significance for e-commerce, manufacturers, and potential consumers.

So it's important to be able to search the sentiment of customers to step up the status according to customer's expectations. Trained machines by labeled data would give the satisfaction of future analysis. Many works in this state to be specific Bangla text reviews of the online marketplace have been done in the past few years. Works like Combined CNN-LSTM model for Sentiment Analysis of Restaurant Reviews [5], The naive Bayes (NB) classifier is using to extract the sentiment orientation (positive or negative) from the Amazon product reviews[7]. Recurrent Neural Network(RNN)

using for sentiment classification of Bengali text[11].

In this paper, we have developed a polarity detection system on Bengali textual reviews/opinions of products, services, using Python and approach where we have collected 2228 opinions/reviews by using Daraz API from Daraz Online Shopping Mall. Daraz is one of the most popular e-commerce sites in Bangladesh, specializing in electronics, home appliances, skincare, clothing and more. Our contribution is the first thing that comes out that's our dataset of 2228 reviews to check the performance power model which taken from sources owned by Bangladesh. Initially, reviews are unstructured and unlabeled data. so for that, we will at starting make our data into a structure and set the label as positive, negative, and neutral. Data is preprocessed for getting clean and in an appropriate format. Tokenization, punctuation removing, stop words removing & stemming are used in preprocessing. Then after some preprocessing stuff, we will nourish our data into our applied model. We at last, reached sentiment results. We also have estimated the system efficiency and performance by contrasting with well-known machine learning classifiers. The evaluation of our developed system has presented with stronger accuracy, precision, recall & F1-score.

1.2 Motivation

With the rapid development and popularity of e-commerce technology, more and more users prefer to shop on different e-commerce platforms. Analyzing sentiment trends in consumer rating not only provides a reference for other consumers but also helps businesses improve service quality and consumer satisfaction on e-commerce platforms. There is a large amount of research on product review data. In any case, most importantly, there is no good work on product review data from the point of view of information on Bangladeshi online business websites. This facility gives us the impetus to develop it. The second important thing that made it clear to us is that this is no good work done on Bangladesh online business dataset.

Sentiment analysis is a process of automatically excluding sentiments or opinions from online product review data. Our preparation model will help potential customers make informed decisions about their purchases and organizations to improve their items or services.

1.3 Rationale of the Study

Our research will support the owner of the organization or company to contribute to the improvement of products and services. Automated sentiment polarity detection techniques can be used to make better decisions, efficiently manipulate and express customer feelings/opinions about a particular product. In addition, customers can get information formed before buying or taking products and services. In our work, we use statistical methods. The most important thing that has inspired us is that there is no good study has been conducted to date on sentiment analysis of Bangladesh e-commerce trade revision dataset in the Bengali language. We choose for this reason. We think it will be a unique thing for others.

1.4 Research Question

We have selected several questions which are addressed step by step in our research work.

- Which data should be collected?
- How to recognize customer sentiment from their reviews?
- What is sentiment analysis and who benefits from sentiment analysis?

- Which algorithm works best for predicting customer emotions?

1.5 Expected Outcome

The model's performance will be tried by applying various sorts of machine learning algorithms to our dataset. The work would test how precise calculations respond to our dataset. The research work will convey a magnificent after-effect of the review of desired online business item review data by utilizing the Model. Regardless of whether it is positive, negative, or neutral. Clients and proprietors of the online business will be profited by expecting the key purposes of the average assessment which have been given by the past clients of that website.

1.6 Report Layout

Chapter 1 discusses the motivation of our thesis, the rationale of the study, the expected results, and the research question.

Chapter 2 Introduction to the background of our study. It also provides us with the facts of the work involved. Difficulties are also mentioned here.

Chapter 3 Discussion of our research technique. Details about data collection, data processing, machine learning and NLTK technique. The data collection process is also covered here.

Chapter 4 Discuss the details of the results and discuss our project with experience and results.

Chapter 5 Discusses our research with potential future scopes and conducts research work

CHAPTER 2

BACKGROUND

2.1 Introduction

Sentiment analysis is one of the broad research areas for modern technical service monitoring and online audience response visualization. So far, many studies have been conducted basically on international language preferences. Since it is the most popular language and used as an international language, there is a large amount of research work that uses various data mining algorithms applied to the English language to analyze emotions from around the world. Since we have not done much work to focus on Bengali data, our project focuses on adding Bengali and Bengali text data to the national e-commerce platform. For this, we use Comparative Analysis and various methods such as LR, DT, RF, MNB, KNN, SVM, SGD etc. to help AI recognize more human responses and order them to classify different sentiments from different services.

2.2 Related works

The work-related to sentiment analysis is more specific than before. Since this is an era of extensive Internet connectivity, the possibilities are increased through research. Next, we focus on some surveys used to obtain typical data or characterize sentiment analysis.

Shaozhang Xiao et. al [1] proposed a deep learning-based sentiment analysis solution for product reviews. Product reviews are collected from the Jingdong Mall website with at least 100,000 reviews already collected as training data set. After data preprocessing operations such as word segmentation and removal of stop words for product reviews to remove useless features, Feature vectors are built based on the bag of word model and word2vec method, then three classification algorithms, namely LSTM, Naive Bayes and logistic regression are used to model the assessments and the accuracy. In the dataset related to product reviews, there are around 80%. Omar et. al [2] presented a machine learning-based sentiment classification framework that can identify sentiment into positive and negative categories from Bengali book reviews by exploiting different feature extraction techniques. They developed their dataset on

sentiment polarity by collecting data from the available web resources such as Facebook and e-commerce sites. Among these classifiers, the MNB with uni-gram characteristics provides the highest accuracy rates of 87% and 84% for the validation and test data sets. Eftekhari et al [3] worked on a deep learning-based program for the sentiment analysis of restaurant reviews in Bengali. Word2vec embedding technique is used to consider the semantics of Bengali comments. BiLSTM network tuned to find out the optimal hyperparameter combination. To assess the efficacy of the suggested method, a corpus of 8435 Bengali restaurant evaluations was created. The BiLSTM technique produced the highest accuracy of 91.35 percent on the test dataset, according to the results of the evaluation. Ganesh et al [4] used sentiment analysis on Amazon product reviews for products sold under the category of Electronics. They explained in detail the work of sentiment analysis performed by stacking LSTM-GRU. The performance of LSTM, GRU, and stacked LSTMGRU is also compared to understand neural network dynamics in sentiment analysis tasks. They don't give importance to score. Md. et al [5] worked on a combined CNNLSTM architecture to implement sentiment analysis in Bengali online restaurant reviews. They collected data from FoodPanda and Shohoz Food. They made a dataset of 1000 reviews with 500. A combined CNN-LSTM architecture was used in their dataset and got an accuracy of 94.22%. Sheikh et al [6] presented the very first research work on the sentiment toward covid-19 using the Bangla language data set. They've got higher accuracy in CNN. Dataset collected from comments of Facebook, news pages. Applied various deep learning algorithms and found the highest accuracy in CNN 97.24% and LSTM 95.33%. Jian et al [7] discussed methods based on linear regression and multiple linear regression models to analyze data from three products representing ratings and reviews provided by customers. A naive Bayes classifier is applied to extract positive or negative affective orientation. Emotional orientation is quantified into 11 levels. Md. et al [8] contribute to the study of the determinants of customer adaptation to e-commerce by categorizing variables into influencing factors and problems. Key data were collected from 120 e-commerce users. They revealed that around 85.8% of the total respondents believe that the e-commerce market in Bangladesh will grow significantly soon, where only 14.2% of the respondents disagree. Rumman et al [9] proposes a process for analyzing the feelings of film reviews written in the Bengali language. The dataset used in this test was manually collected and tagged from comments and publicly available posts on social media

sites. Using the Support Vector Machine algorithm, this model achieved an accuracy of 88.90 % on the test set, and using a long-term memory network, this model achieved an accuracy of 82.42%. Susanti et. al [10] presented sentiment analysis provided by. U uses TFIDF and SVM methods in Google Play store reviews. TFIDF + SVM with 5-fold verification produces satisfactory accuracy, with an average accuracy of 84.7%, an accuracy of 84.9%, a recovery rate of 84.7%, and a measurement of 84.8%. The best accuracy of Fold 2 is 86.1%. Abdullah et.al [11] introduced a sentiment classification of Bengali text using Recurrent Neural Network is presented. Using a deep recurrent neural network with BiLSTM, they achieved an accuracy of 85.67%. Fabliha et. al [12] presented a review analysis system on Bangla and Phonetic Bangla. They used machine learning-based approaches namely Support Machine, Decision Tree, and Logistic Regression applied to rank the reviews where SVM provided better accuracy with 75.58%. Asif et.al [13] applied sentiment analysis on romanized Bangla and Bangla text obtained from various social media platforms. They trained their model using deep RNN and achieved an accuracy of 78 percent with categorical cross-entropy loss. Amitava et.al [14] used A hybrid method that was proposed to identify the sentiment from the sentence. The authors first determined if a sentence was a subject, and they designed a word model from a mixture of different Part of Speech (POS) characteristics gathered from sentence-level similarity, then they used syntax modeling to perform sentiment analysis. By doing so they achieved an overall 63% recall rate using SVM on the news data.

2.3 Research Summary

After this work, we found a workspace for the characterization of specific data based on regional user interference. Analysis of sentiments of Bengali categories and expressions. Sufficient collection of Bengali word data, keyword sentiment ranking, and getting these results can view the text, helping to

understand product results and information about the human behavior of Bengali people.

TABLE 2.1:RESEARCH SUMMARY

No	Author	Year	Language	Domain	Algorithms	Accuracy
1	Shaozhang Xiao et. al	2020	English	E-Commerce	LSTM, Naive Bayes logistic regression	80%
2	Omar et al.	2021	Bangla	Book Review	Multinomial naive Bayes	87% 84%
3	Eftekhar et.al	2021	Bangla	Restaurant reviews.	. BiLSTM network	91.35%.
4	Ganesh et. al	2020	English	E-Commerce	LSTM-GRU	
5	Md. et.al	2020	Bangla	Restaurant reviews	Combined CNN-LSTM	94.22%.
6	Sheikh et.al	2020	Bangla	Sentiment toward covid-19	Combined CNN-LSTM	97.24% 95.33%
7	Jian et.al	2020	English	E-Commerce	Naive Bayes (NB)	
8	Md. et.al	2020	Bangla	E-Commerce	SVM	85.8%
9	Rumman et.al	2019	Bangla	Movie reviews	Support Vector Machine (SVM),LSTM	88.90% 82.42%
10	Susanti et.al	2020	English	Google Play Store Reviews	TF-IDF SVM	86.1%
11	Abdullah et.al	2018	Bangla	Sentiment classification	BiLSTM RNN	85.67% .
12	Fabliha et. al	2019	Bangla	Restaurant reviews	SVM DT LR	75.58%.
13	Asif et.al	2016	Bangla	Social media	LSTM	78%
14	Amitava et.al	2010	Bangla	Newspaper Text	SVM	63%

2.4 Scope of the Problem:

There are a lot of problems in this work when sorting keywords. Users can use slang or unusual typing, such as shorter or misspelled words. Since there is no meaningful concrete model yet, it can be a bit difficult to understand the sentiment behind it.

2.5 Challenges:

New and changed words and emoticons need to be updated regularly. It's also important to try and sort out patterns of slang or unusual words for more accurate results. There will be various difficulties in our research:

- Gathering a large volume of data.
- Manage the data after it has been collected.
- The data procedure is our main problem in this study.
- There is also a problem between the experimental results and the outcome results.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

Since reviews are textual and differing an intuitive portrayal would be steady to users through summing up opinions and giving criticism. As text mining is scarcely selective from numeric data as text needs to data pre-processing and we additionally use Bangla data so finding the classifications is another level challenge for completing our research work.

As a matter of first importance, we have pre-processed our dataset, TF-IDF used for feature extraction and a comparative analysis with different approaches was also performed by examining characteristics of the unigram, bigram and trigram respectively to decide if the review is positive, negative or neutral.

3.2 Research Subject and Instrumentation

We are working based on the sentiments and opinions of Bangladeshi e-commerce site users. For this, we used to review data from a Bangladeshi e-commerce website. We have collected our required data using the Daraz API. We work on the sentiments and reviews of Bangladeshi e-commerce customers. For this, we collected review data from a Bangladeshi business website. From that point on, we used NLTK (Natural Language Toolkit) to remove stop words and additionally, to remove punctuation. Here we have used the sci-kit learning library implemented with the python programming language.

It is now becoming a fully common and valuable tool for textual data analysis. The advantage of learning sci-kit is that we can import many different types of libraries that have built-in algorithm visualization tools. To finish this analysis, we imported the necessary library functions (Figure: 1).

Import Libraries

```
import numpy as np
import pandas as pd
import pickle
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns
import re,json,nltk
from sklearn.preprocessing import LabelEncoder
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.feature_extraction.text import TfidfTransformer
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report,accuracy_score,precision_score,recall_score,f1_score
stopwords_list = 'stopwords-bn.txt'
```

Figure 3.1: All Required Library Functions are Imported

3.3 Data Collection Procedure

Data collection is the process of gathering information from all sources related to data collection, finding answers to research questions, testing hypotheses, and evaluating results. The efficiency of the machine depends on the quality and quantity of the data. The more data we can provide, the better the machine can predict and analyze perfectly. Bangladesh has a large number of web-based business organizations of various shapes and sizes, but for our information assortment, we choose "Daraz Bangladesh" because they have a large number of clients, as well as a large number of Bengali reviews. "Daraz Bangladesh" is an e-commerce company located in Bangladesh, whose customers and vendors are from Bangladesh. The data was extracted using the Daraz API and downloaded in JSON format from the website and converted to an xlsx file for additional ease of work. The data was divided into a training set (80%) and a test set (20%) which is in the Bangla language. We collected 2228 product reviews of Bangladeshi products in different product categories such as electronics, household machines, skincare products, fashion, etc. We divide these data into three categories based on the type of review. They are positive, negative and neutral. Subsequently, our data set contains two-column reviews and categories. Some samples in the data set are:

TABLE 3.1:SAMPLE REVIEWS

Sample Reviews	Sentiment
সব দেখি চিটিং এর দুনিয়া।আমি অর্ডার করছি একটা দিছে অন্য কালার।	Negative
এসকল ফাইজলামির কোন মানে হয় না এত নিম্নমানের প্রডাক্ট কোয়ালিটি কিভাবে মার্কেট তারা ধরে রাখবে মানুষকে ঠকানো ছাড়া আর কিছুই না	Negative
দাম হিসেবে প্রোডাক্ট যথেষ্ট ভালো ছিল	Positive
আলহামদুলিল্লাহ পন্যগুলো ঠিকঠাক মতো হাতে পেলাম।।প্যাকেজিং ভালো দ্রুত ডেলিভারি সব পারফেক্ট,প্রত্যেকটাই জোশ	Positive
প্রত্যেক আইটেমের প্রাইস গুলো এড করলে ভালো হবে	Neutral
স্কুলে থাকাকালীন বাটা জুতা আমার প্রথম পছন্দ ছিলো	Neutral

Table-2 is a overview of sampling data. Among all of those Product review we've found 948 under positive review , 920 for negative review and 360 in Neutral review section. All of the classes pi-charts are shown in Figure-2.

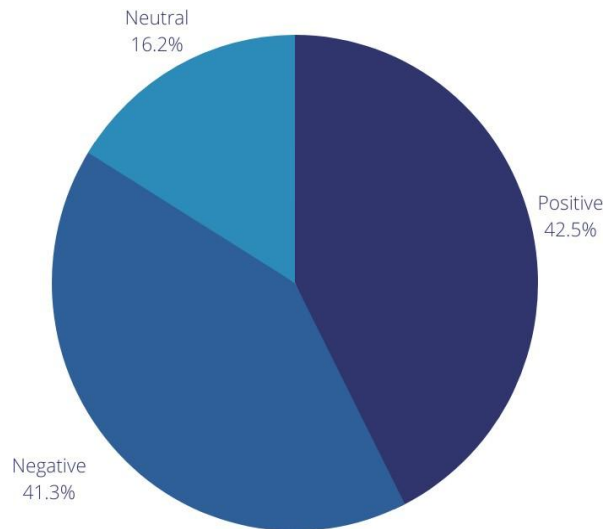


Figure 3.2: Dataset

3.4 Statistical Analysis

There are countless comment data on the "Daraz Bangladesh" site. From there, we selected 1656 Bengali data for our system training data set. Sorting multiple reviews will provide better results and finally, we perform a legitimate data review on the full-text training dataset. Figure -3 is the overview of data statistics of our dataset.

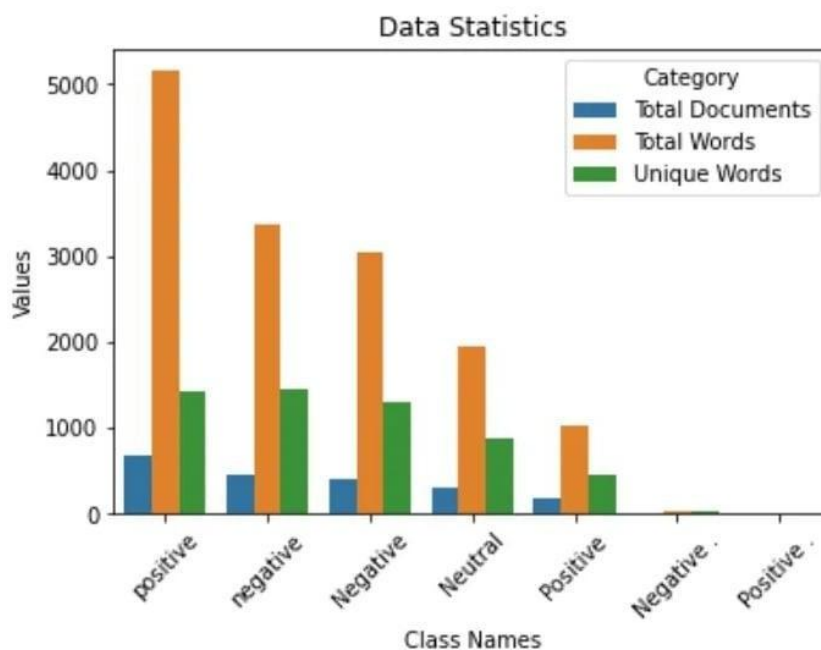


Figure 3.3: Data statistics

3.5 Implementation Requirements

To complete our research work, we followed several steps. They are all interrelated. Since we use text data, we need to sort the data before fitting it to our machine learning algorithm. Our exploration work has also made some more significant progress. Figure demonstrating the entirety of the means of our research procedure.

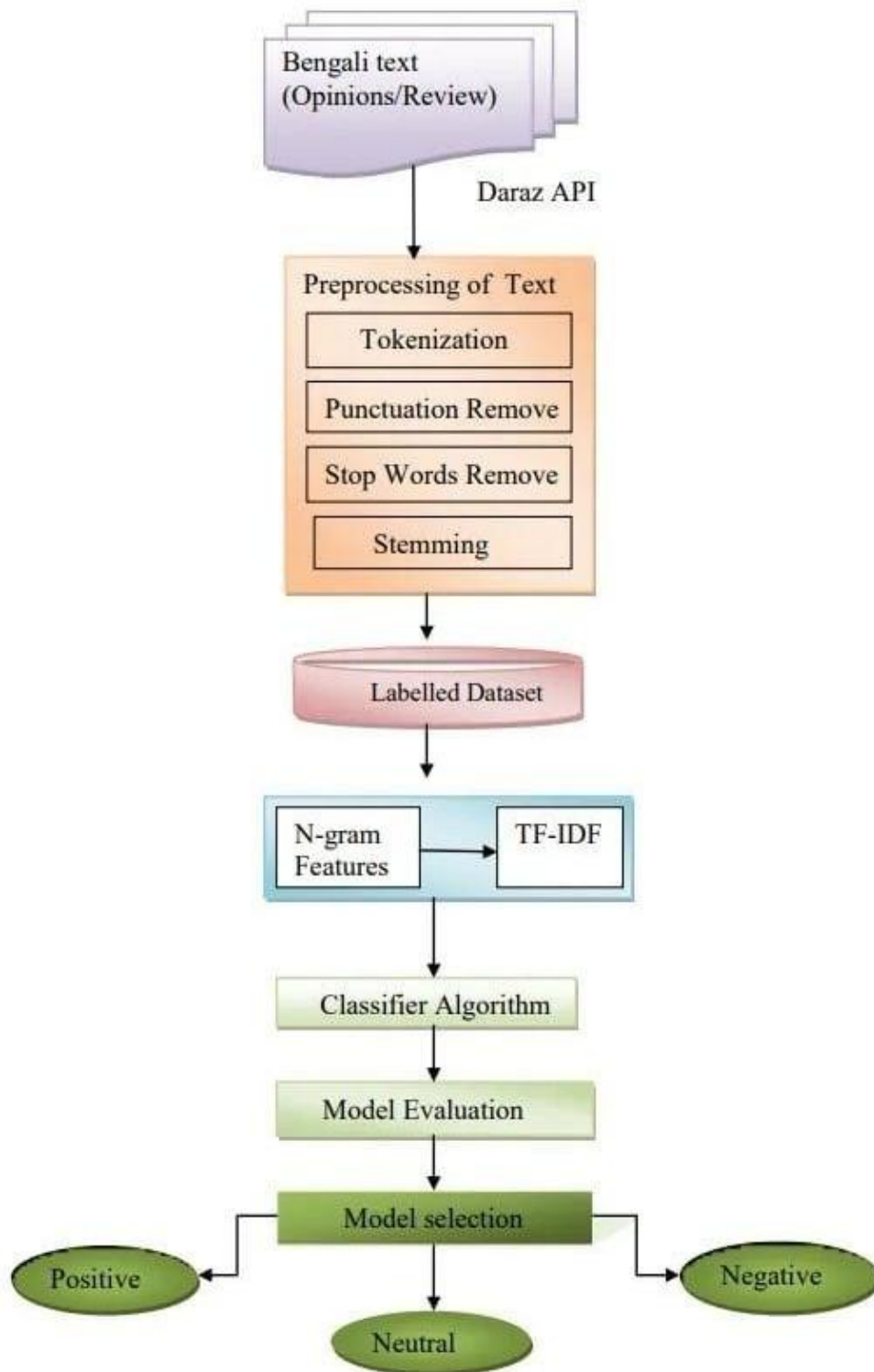


Figure 3.4: Workflow

3.5.1 Data collection:

For our research work data assortment was another level challenge as there are no open-source databases accessible on e-commerce product review dependent on the Bangladeshi marketplace. We have to collect data in solitude. We collect data from a Bangladeshi site called "Daraz Bangladesh". The data was extracted using the Daraz API and downloaded in JSON format from the website and converted to an xlsx file. Therefore, we must complete additional data collection work. Figure-5 is a overview of length distribution of product review.

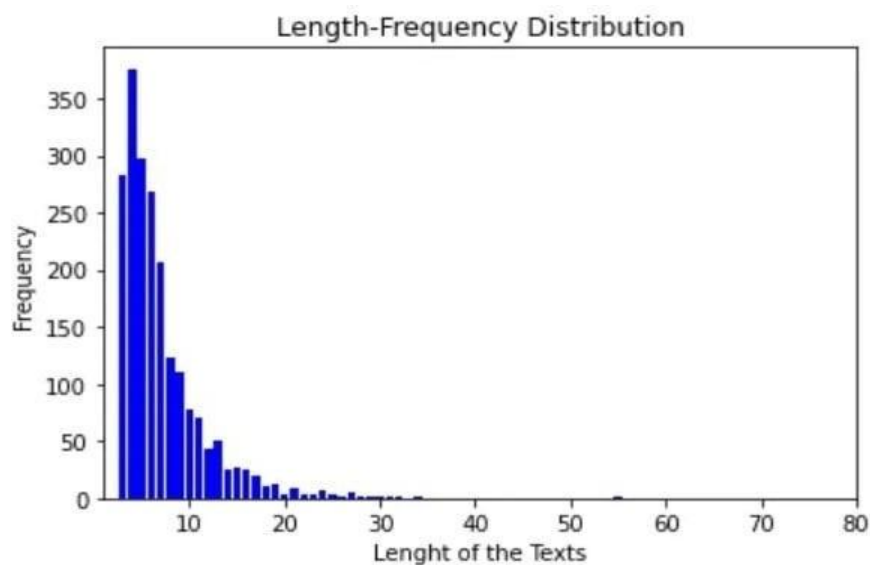


Figure 3.5: Length Distribution

Maximum Length of a review: 55

Minimum Length of a review: 3

Average Length of a reviews: 7.0

3.5.1 Data Pre-processing:

Data preprocessing is a method of data mining that involves converting raw data into an understandable format. It is too important to clean up data that helps us optimize test errors. The raw data collected does not meet the criteria for direct classification. It contains a lot of punctuation marks, emoticons, etc. that have nothing to do with the emotional analysis process. The data set should be preprocessed before the start of the classification process for higher accuracy. Depending on the language of the data set, many preprocessing steps are important to the data set. Pre-processing is a key step before the sorting process begins. Some notable methods used are tokenization, emoji punctuation and removal, stemming, stop word removal, etc.

3.5.2 Tokenization and Punctuation Removal:

Tokenization means breaking certain text into sentences called meaningful words, sentences, numbers, or tokens. In this experiment, words were separated by segmenting sentences according to space. Unnecessary items from data such as punctuation marks, alphabets, emoticons, etc. were also deleted when tokenizing was performed on each data sample. An array was then created carrying the sub-array of tokenized data. Other arrangements containing labels have been published.

3.5.3 Remove Stopword:

Words of negligible importance from the text corpus are called stopwords. This word has no meaning when classifying documents. In English, words such as "a", "of", "the", "for", and "my" are stop words. Similarly in Bangla, the words. ["অতএব", "অথচ", "অথবা", "অনুযায়ী", "এটা", "এটাই"] etc. are considered as stopwords. The list of stopwords in Bangla was collected from [15]. For instance, from the sentence["দাম কম হলেও অনেক ভাল "], After removing the stop words, the following token is obtained: [দাম , কম, হলেও , ভাল] Here " অনেক " was excluded because it was considered a stopword.

3.5.4 Stemming:

The term stemming specifies to reduce word transformations to the basic form. Words come in many forms, depending on the context in which they are used. For instance, ["করা", "করিছ", "করিছলাম"]etc. for all these words, "কর" is the original word. : The main purpose of stemming is to simplify the combined form of words into a common basic form. In this way, the total number of words that the classifier must use can be greatly reduced. To perform this process, common prefixes and suffixes used in Bengali words are stored in an array. Use Python regular expression library to detect prefixes and suffixes in words, and add shortened versions of words to the newly processed corpus. In this sentence[" দাম কম হলেও অনেক ভাল"] next adding the stem, the word is (no stop word): [দাম , কম, হলেও , ভাল] Here, "হলেও " Change its basic form "হল".

3.5.5 Applying Algorithm

TF-IDF refers to the inverse frequency of the term document. It is one of the most significant techniques for information retrieval, used to express the importance of a particular word or phrase to a particular document. The importance increases in proportion to the frequency of a word in the document, and the higher the frequency of the word in the corpus, the lower the importance. The sci-kit learn function Tf-IDF Vectorizer creates a TF-IDF feature array from the original document collection. The feature extraction process determines the success rate of any system based on machine learning. The combination of the N-gram feature(unigram, bigram, and trigram) and TF-IDF value is used as the feature of the proposed system.

There are various types of machine learning models. Our proposed work aims to develop a sentiment classifier that classifies online product reviews into positive, negative, or neutral categories. The N-gram functions with TF-IDF are mainly used for model preparation. The extracted characteristics are divided into three random distributions, and each distribution is applied to a specific stage of model preparation,

such as classification.

Classification:

At this stage, a series of classifiers have been developed by applying different learning algorithms to the training set. These algorithms are logistic regression (LR), decision tree (DT), RF, multinomial naive Bayes (MNB), KNN, SVM and stochastic gradient descent (SGD) respectively.

3.5.6 Accuracy

We used logistic regression (LR), decision tree (DT), RF, multinomial naive Bayes (MNB), KNN, SVM and stochastic gradient descent (SGD) respectively for our research work. And tried to find out accuracy with different algorithms.

- **Hardware and Software**

- **Python 3.7:-** Python 3.7 is a version of Python. It is a high level programming language It is used by the majority of researchers to conduct their research.
- **Google Colab:-** Google Collab is a free and open-source Python programming language distributor.
- **Operating System:-** Windows (7, or 10).
- **Browser:-** Google Chrome.
- **RAM & ROM:-** - Hard Disk (at least 4 GB), ROM (Minimum 4 GB)

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Introduction

To increase production results, our research project is based on the selection of suitable data and ideal models. To use the text content dataset to build a model and find the best possible result, the language preparation part is particularly important.

4.2 Experimental Results

In this task, we use some machine learning models of which best match this sentiment analysis on considered data for emotion analysis, accuracy, precision, recall, F1 score.

A) Accuracy: The accuracy is the fraction of the correct predictions made by the model.

B) Precision: Precision is understood as to how often models are correct when it classifies sentiment as positive.

C) Recall: Recall calculated the actual number of positives that the model got by labeling as positives.

D) F1 score: The F1 score is an overall measure of the accuracy of the sample, combining both the accuracy of and recall. A good F1 score is a score of low false positives and low false negatives. When the F1 score is 1, the model is deemed perfect, but when it is 0, the model is considered a complete failure. The evaluation of the performance of the proposed technology is carried out through the use of various graphs and statistical measures such as f1 score, recall, precision, ROC, and precision versus recall curve. For the preliminary development of the sentiment classification model, eight classifiers were selected for training. In Table represents the accuracy of

Model performance for our dataset

TABLE 4.1: PERFORMANCE MEASUREMENT ACCURACY OF DATASET USING DIFFERENT ALGORITHMS

Features	Classifier	Accuracy(%)	Precision(%)	Recall(%) f1	Score(%)
Uni-gram	LR	85	82	96	89
	DT	73	84	71	77
	RF	86	88	90	89
	MNB	86	89	88	89
	KNN	86	87	91	89
	Linear SVM	69	67	98	80
	RBF SVM	79	76	96	85
	SGD	85	87	90	88
Bi-gram	LR	78	75	96	84
	DT	76	82	78	80
	RF	83	83	91	87
	MNB	84	89	85	87
	KNN	80	90	76	82
	Linear SVM	66	65	100	78
	RBF SVM	66	65	100	78
	SGD	88	87	95	91
Tri-gram	LR	72	70	96	81
	DT	69	79	70	74
	RF	86	85	95	89
	MNB	84	89	85	87
	KNN	77	89	71	79
	Linear SVM	63	63	100	77
	RBF SVM	66	65	100	78
	SGD	86	85	95	89

Table 3 shows the performance of the eight classifiers in terms of accuracy, precision, recall, and f1 measures. It observed that in the case of the Uni-gram feature, random forest (RF), multinomial naive Bayes (MNB) and KNN achieved the highest accuracy of about 86%. For bi-gram features, stochastic gradient descent (SGD) achieved the highest accuracy of about 88%, whereas SVM provided the lowest accuracy (66%). On the contrary, for tri-gram features, RF and SGD both achieved the highest accuracy of about 86%. The investigation found that for Tri-gram features, stochastic gradient descent (SGD) outperformed the other classifiers and feature extraction methods. Result reveals that the stochastic gradient descent (SGD) with

the Tri-gram feature outperforms the other techniques with 88% accuracy.

4.3 Descriptive Analysis:

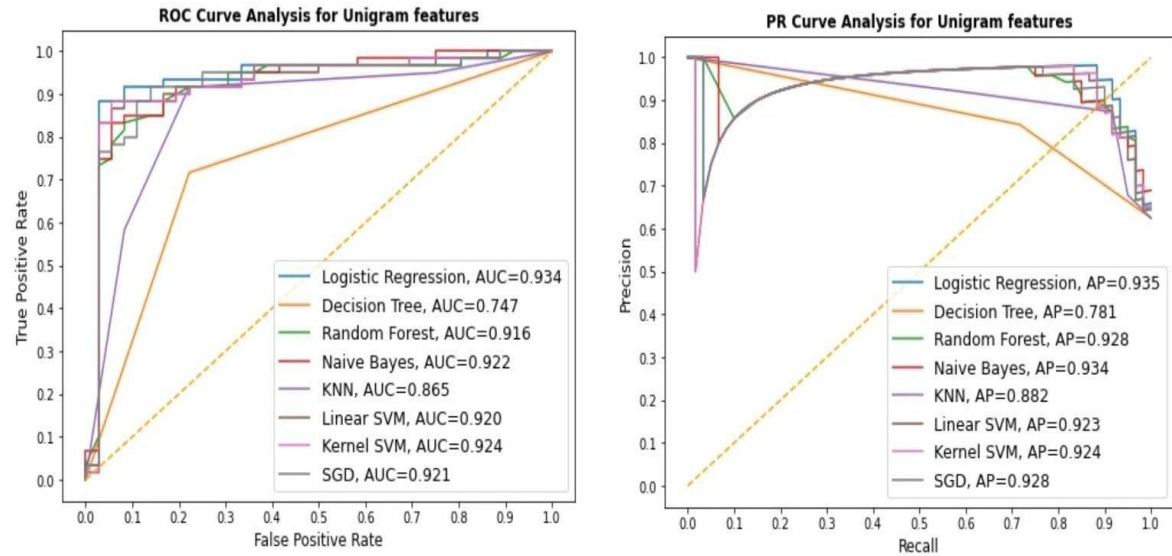


Figure 4.1: Classifiers performance on Uni-gram features

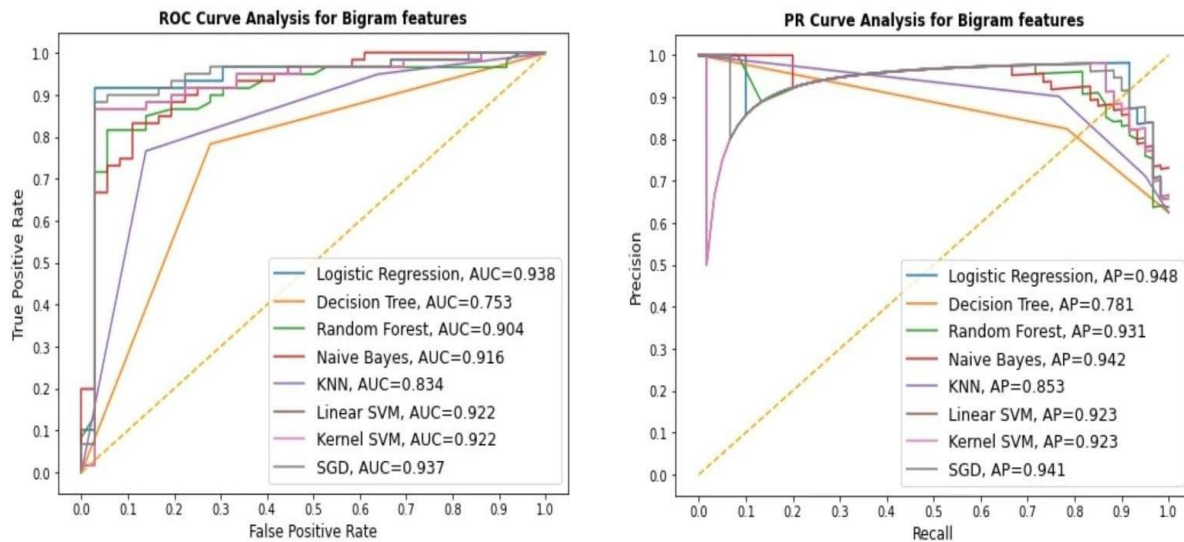


Figure 4.2: Classifiers performance on Bi-gram features.

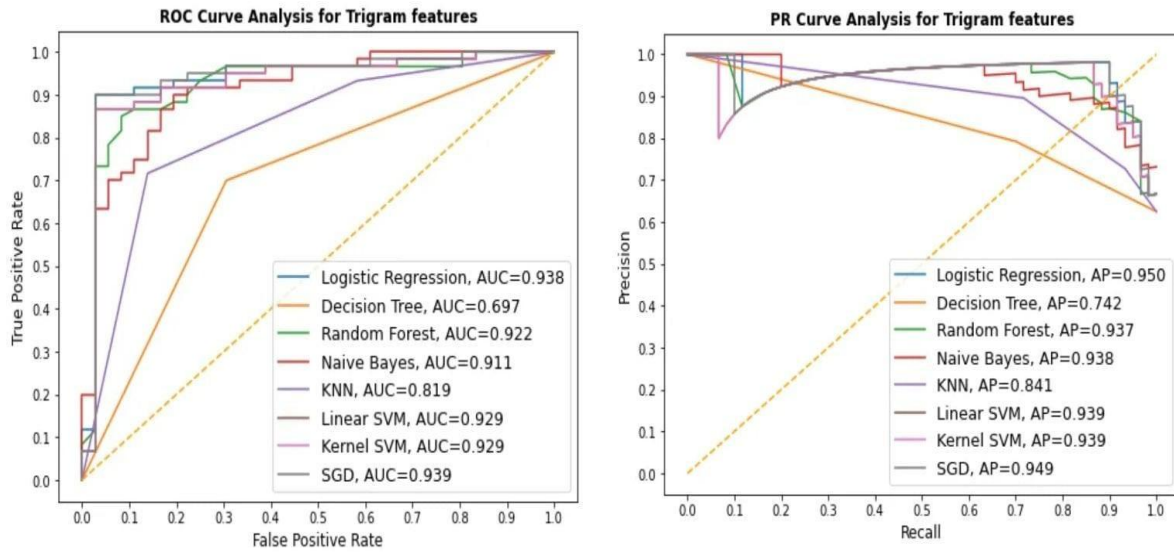


Figure 4.3: Classifiers performance on Tri-gram features.

The experiment is run again for graph analysis for all classifiers. Figures 6 to 8 show the ROC and PR curves of the eight selected classifier models with unigram, bigram, and trigram features respectively. In the case of unigram features, the AUC and AP scores show the lowest scores for all classification models. The majority of the strong sentiment carrying words and phrases misrepeated in most of the review .

On the other hand, for bigram features, the classifiers jointly give good AUC and AP score for ROC and PR curves, respectively. In the case of trigram features, SGD provides the highest AUC at 93.9% and average precision at 94.9%. Detailed analysis of results provides the rating that SGD with Trigrams features for highest accuracy, AUC and AP scores over all other gram feature and classifier models. Thus, the stochastic gradient descent (SGD) classifier was selected as the final model.

CHAPTER 5

SUMMARY, CONCLUSION, RECOMMENDATION AND IMPLICATION FOR FUTURE RESEARCH

In this part of our report, we describe some important parts of our research, namely a summary of the research, suggestions, significance of the research work, and finally the future scope of this research. Part of the entire market segment is as follows:

5.1 Summary of the Study

To complete our research work, we focus on how to use machine learning programs to obtain solutions to different types of artificial intelligence problems. There are two forms of machine learning: supervised machine learning and unsupervised machine learning. This research made us realize that our research is about supervised machine learning, and it is a classification model problem. To conduct research based on machine learning, we need a considerable amount of data. In any case, there is no reliance on the Bangladeshi online market or Bengali open source databases. So we must collect our database separately. To obtain the required database, we use the Daraz API. The data is extracted using Daraz API and downloaded from the website in JSON format, and is converted to xlsx files to facilitate work. We also learned machine learning, model classification and related algorithms. Comparative analysis of and various methods LR, DT, RF, MNB, KNN, SVM, SGD, etc. Consider the characteristics of a unigram, bigram, and trigram, respectively. There have many tools and techniques suitable for this field. We have also discovered which program is best suited to complete our research work. We use various types of tools to complete our research work. We have used various tools and techniques to achieve our goals. We have deep knowledge of Pandas, JSON format, Numpy, Sci-kit Learn, NLTK and Sci-pack learning library and Python programming language.

5.2 Conclusions:

When shopping online, analyzing consumer feedback is just as important as any other strategic goal of any business. In addition to branding, it provides feedback that businesses need to test and make improvements in areas where it is lacking. It also provides a general concept of the competitive market and the company's expansion efforts. This paper presents a machine learning-based sentiment classification framework that can identify sentiment into positive, negative and neutral categories from online product reviews by exploiting extraction techniques different features. The combination of the TF-IDF values with N-gram features is considered to be the best feature of the proposed framework. These extracted characteristics were applied to classify the sentiment inherent in Bengali textual reviews using different ML techniques. Among these classifiers, stochastic gradient descent (SGD) with Tri-gram features provided the highest accuracy of 88%. In the future, the dataset will need to be improved by adding more data samples. It will be interesting to see how the model works when working with a large dataset. Deep learning techniques and sophisticated feature extraction methods such as word2Vec or Glove can be applied to improve accuracy.

5.3 Recommendations

Although we have mentioned some related work, the amount of work is not large. In addition, there is no study like ours based on data from Bangla or assessment data from Bangladesh. We understood all of their research processes and working styles and then began to define our own research goals. After a period of grueling effort in completing all the final work, we are at a point where we can say that this is our expected research goal. With that, to do this kind of research work, it takes a tremendous amount of work to guide us in the right direction of research. We ran into different problems at the beginning of our research. We also stuck with data collection

5.4 Implication for Further Study

From this research work, latent Dirichlet Allocation (LDA) topic modeling can be performed using organizations and e-commerce business owners who can work on improving their products or services. Additionally, customers can gain knowledge before purchasing or using a product or service. We will be able to rate a seller or e-commerce site based on their review data. For the implementation, we created a website using Bootstrap and Flask Framework. The Heroku Cloud Platform is used to deploy the model. Using the Daraz API, we can also collect a review dataset very easily. After collecting this data, it will be stored in the database. Then the next procedure is to preprocess this data and classify all the data using our proposed model. Then compare it with data from other e-commerce sites to generate a sorted ranking. In addition, we hope that users can use this system using the website and Android mobile app.

REFERENCES

- [1] S. Xiao, H. Wang, Z. Ling, L. Wang, and Z. Tang, "Sentiment Analysis for Product Reviews Based on Deep Learning," *J. Phys. Conf. Ser.*, vol. 1651, no. 1, 2020
- [2] E. Hossain, O. Sharif, and M. Moshuiul Hoque, "Sentiment Polarity Detection on Bengali Book Reviews Using Multinomial Naïve Bayes," *Adv. Intell. Syst. Comput.*, vol. 1299 AISC, no. July, pp. 281–292, 2021
- [3] E. Hossain, O. Sharif, M. M. Hoque, and I. H. Sarker, "SentiLSTM: A Deep Learning Approach for Sentiment Analysis of Restaurant Reviews," *Adv. Intell. Syst. Comput.*, vol. 1375 AIST, pp. 193–203, 2021
- [4] A. M. Suresha, "Sentiment Analysis on Amazon Product Reviews with Stacked Neural Networks,"
https://www.researchgate.net/publication/344518660_Sentiment_Analysis_on_Amazon_Product_Reviews_with_Stacked_Neural_Networks, no. October, 2020
- [5] N. Hossain, M. R. Bhuiyan, Z. N. Tumpa, and S. A. Hossain, "Sentiment Analysis of Restaurant Reviews using Combined CNN-LSTM," *2020 11th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2020*, 2020
- [6] M. S. A. Pran, M. R. Bhuiyan, S. A. Hossain, and S. Abujar, "Analysis of Bangladeshi People's Emotion during Covid-19 in Social Media Using Deep Learning," *2020 11th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2020*, no. July, 2020
- [7] J. Dong *et al.*, "Potential Trend for Online Shopping Data Based on the Linear Regression and Sentiment Analysis," *Math. Probl. Eng.*, vol. 2020, 2020
- [8] N. Science and N. Science, "Investigating the factors influencing customers towards e-commerce shopping in Bangladesh," no. December, 2020
- [9] R. R. Chowdhury, M. S. Hossain, S. Hossain, and K. Andersson, "Analyzing Sentiment of Movie Reviews in Bangla by Applying Machine Learning Techniques," *2019 Int. Conf. Bangla Speech Lang. Process. ICBSLP 2019*, pp. 27–28, 2019
- [10] S. Fransiska and A. Irham Gufroni, "Sentiment Analysis Provider by.U on Google Play Store Reviews with TF-IDF and Support Vector Machine (SVM) Method," *Sci. J. Informatics*, vol. 7, no. 2, pp. 2407–7658, 2020
- [11] A. Aziz Sharfuddin, M. Nafis Tihami, and M. Saiful Islam, "A Deep Recurrent Neural Network with BiLSTM model for Sentiment Classification," *2018 Int. Conf. Bangla Speech Lang. Process. ICBSLP 2018*, no. September, 2018
- [12] F. Haque, M. M. H. Manik, and M. M. A. Hashem, "Opinion Mining from Bangla and Phonetic Bangla Reviews Using Vectorization Methods," *2019 4th Int. Conf. Electr. Inf. Commun. Technol. EICT 2019*, no. July, 2019
- [13] M. T. Aziz, "4th International Conference on Electrical Engineering and Information and Communication Technology, iCEEiCT 2018," *4th Int. Conf. Electr. Eng. Inf. Commun. Technol. iCEEiCT 2018*, pp. 8–11, 2019.

- [14] A. Das and S. Bandyopadhyay, "Phrase-level Polarity Identification for Bangla," vol. 1, no. 1, pp. 169–182, 2010.
- [15] Stopwords-Iso, "stopwords-iso/stopwords-bn," Available: <https://github.com/stopwords-iso/stopwordsbn>

Similarity Index 6%	Similarity by Source Internet Sources: N/A Publications: N/A Student Papers: 6%
exclude quoted exclude bibliography exclude small matches mode:	
quickview (classic) report <input type="button" value="Change mode"/> print refresh download	
1% match (student papers from 16-Aug-2020) Submitted to National College of Ireland on 2020-08-16	✖
1% match (student papers from 07-Apr-2018) Submitted to Daffodil International University on 2018-04-07	✖
<1% match (student papers from 09-Aug-2020) Submitted to National College of Ireland on 2020-08-09	✖
<1% match (student papers from 22-May-2021) Submitted to North South University on 2021-05-22	✖
<1% match (student papers from 08-Mar-2021) Submitted to Bangladesh University of Professionals on 2021-03-08	✖
<1% match (student papers from 29-Oct-2020) Submitted to CSU, San Jose State University on 2020-10-29	✖
<1% match (student papers from 08-Feb-2021) Submitted to BPP College of Professional Studies Limited on 2021-02-08	✖
<1% match (student papers from 06-Jun-2021) Submitted to Özyegin Üniversitesi on 2021-06-06	✖
<1% match (student papers from 22-Nov-2020) Submitted to Cornell University on 2020-11-22	✖
<1% match (student papers from 02-Nov-2020) Submitted to University of Strathclyde on 2020-11-02	✖

t.turnitin.com/newreport_classic.asp?lang=en_us&old=1627351068&ft=1&bypass_cv=1

Turnitin

<1% match (student papers from 23-May-2021) Submitted to Victorian Institute of Technology on 2021-05-23	✖
<1% match (student papers from 17-Nov-2020) Submitted to Darton State College on 2020-11-17	✖
<1% match (student papers from 31-Mar-2021) Submitted to Kaunas University of Technology on 2021-03-31	✖
<1% match (student papers from 02-May-2017) Submitted to The University of Manchester on 2017-05-02	✖