ABUSIVE LANGUAGE DETECTION FROM REAL-TIME RADIO MESSAGE GATEWAY

BY

SUMAIYA SALIM RITU
ID: 172-15-9727

JOYSURYA MONDAL
ID: 172-15-10207

AND

MD. MOJNU MIA
ID: 172-15-9625

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

Supervised By

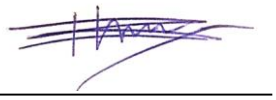
Mr. Ahmed Al Marouf
Lecturer
Department of CSE
Daffodil International University

Co-Supervised By
Shah Md Tanvir Siddiquee
Assistant Professor
Department of CSE
Daffodil International University

APPROVAL

This Research Based Project titled "Abusive Language Detection from Real-Time Radio message Gateway", submitted by Sumaiya Salim Ritu 172-15-927, Joysurya Mondal 172-15-10207, MD. Mojnu Mia 172-15-9625 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.

BOARD OF EXAMINERS



Dr. Touhid Bhuiyan **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Subhenur Latif **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Md. Abbas Ali Khan **Internal Examiner**
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Md Arshad Ali **External Examiner**
Associate Professor
Department of Computer Science and Engineering
Hajee Mohammad Danesh Science and Technology University

DECLARATION

This is to certify that this project titled “Abusive Language Detection from Real-Time Radio message Gateway "has been done by us under the supervision of Mr. Ahmed Al Marouf, Senior Lecturer, Department of CSE, Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

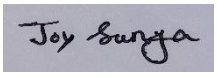
Mr. Ahmed Al Marouf
Senior Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

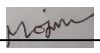
Shah Md Tanvir Siddiquee
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:

Sumaiya Salim Ritu
ID: 172-15-9727
Department of CSE
Daffodil International University



Joysurya Mondal
ID: 172-15-10207
Department of CSE
Daffodil International University



MD. Mojnu Mia
ID: 172-15-9625
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, we express our heartiest thanks and gratefulness to Almighty God for His divine blessing that makes us possible to complete the final year project/internship successfully.

We grateful and wish our profound indebtedness to **Mr. Ahmed Al Marouf**, **Lecturer**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "*Big Data Analysis*" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

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

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

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

ABSTRACT

Most of the research work on identifying objectionable texts is in English and some of them can detect objectionable texts but there are some works available in the Bengali language. People in Bangladesh usually feel comfortable expressing their views in Bengali on media platforms. The identification of offensive text in the Bengali language will be effective in preventing cybercrime, which has become a major concern of Bangladesh nowadays in preventing online language harassment, blackmailing, and cyberbullying. It is also challenging to include them in different groups depending on the text. In this project, we present our work on detecting abusive language from real-time radio message gateway. Usually what happens on radio stations is that users give messages but some messages are good and some messages are offensive. They follow all messages except filtered from their server. We have created a data set of about 45,000 messages, including good messages and offensive messages. We have leveled the good and offensive messages. We have leveled the good messages as 1 and the offensive messages as 0. Our algorithm will check the dataset when a live radio program is broadcast and listeners will try to communicate by sending a direct message and will only show positive comments. In real-time, all abusive messages will come to the spam folder here to save time.

TABLE OF CONTENTS

LIST OF FIGURES **PAGE NO**

CHAPTER 1

INTRODUCTION

1.1 Introduction

Social media is a popular medium for discussion, communication of views, sharing of content, and advancement of ideas and products. Especially in Asian countries like Bangladesh where popular use social media more than anything. So, to reach with radio programs also the stations have to use social media. Now, people have freedom of speech and so, they write and tell whatever they want and actions are not trending as this is their citizen right. If you go on social media like Facebook, Instagram in Bangladesh you will see people using in both English and Bangla writing comments as good and bad. Radio is the original social medium. Nowadays radio stations create websites /pages on social platforms and host programs to directly interact with their listeners so that they can have a visual treat as well as listen. And there is no better proof than hearing directly from a radio listener. So, while hosting a live radio show the host or the RJ connects with the listeners through the comments of that program. Now sometimes some people comment abusive types of comments and in a live program to filter comments and read is very tough so sometimes the abusive comments are also read and it becomes very uncomfortable for the host. So our project aims to filter the abusive comments by detecting their type and show positive comments to the host while on live radio programs.

As in Bangladesh, people have a different regional way of speaking, right now we have worked on the standard format of the Bangla language, and in the future, we are planning to include the regional languages a well.

1.2 Motivation

Bangla has been positioned at seventh situation among the 100 most spoken languages and 228 million native speakers in the world in 2019 [1]. Now, people have freedom of speech and so, they write and tell whatever they want. We suffer day-to-day problems from different spam messages which are sometimes not filtered. In our mobile and mail and other social sites, this is available but on the live radio station, it is not available. Now sometimes some people comment abusive types of comments and in a live program to filter comments and read is very tough so sometimes the abusive comments are also read and it becomes very uncomfortable for the host. So, this project is going to save a lot of time searching for good messages and also will make the work of radio stations more efficient and so the radio jockeys can talk to the viewers more identically.

1.3 Objective

The title of our proposed project abusive language detection from real-time radio message gateway. We are developing this project in a way that it can detect any abusive message in conversations. This project is to serve people by helping them from wasting their energy and time on abusive, problematic, and not important things and people and their messages. All the abusive messages in real-time will get to the spam folder here to save time.

1.4 Expected Outcome

Radio is the original social medium. Nowadays radio stations create websites /pages on social platforms and host programs to directly interact with their listeners so that they can have a visual treat as well as listen. Now sometimes some people comment abusive types of comments and in a live program to filter comments and read is very tough so sometimes the abusive comments are also read and it becomes very uncomfortable for the host. So our project aims to the abusive message were detected automatically. And detected messages sent to the spam folder directly and were not visible to the reader.

1.5 Report Layout

Chapter 1: Introduction

In chapter one we examined objective, motivation, expected outcome in our project work, and the report layout.

Chapter 2: Background

Chapter two is background. We talk about the foundation conditions of our task. In this chapter, we also talk about related works. Scope of problems, comparatives studies, and challenges in our project work etc.

Chapter 3: Research Methodology

In this section we discuss data collection procedure, data preprocessing, proposed methodology, representation test raw comments, and implementation requirements.

Chapter 4: Experimental Result and Discussion

In this chapter, we talk about the result of our study and analysis of the result of what we found by experiment.

Chapter 5: Impact on Society, Environment and Sustainability

In this section, we talk about the Impact on Society, Environment and Psychological state, and ethical aspect.

Chapter 6: Summary, Conclusion, and Implication for Future Research

Chapter six is Summery, Conclusion, and Implication for Future Research. Here we discuss the summary of our project, conclusions, and implication for future study.

CHAPTER 2

BACKGROUND

2.1 Preliminaries/Terminologies

2.1.1 Natural Language Processing

Computer don't understand our language. In need to convert our language in some numeric values. Then it can work through it. This is the place where we can use the natural language processing. The Fundamental of Natural Language Processing endeavors consolidate parsing and tokenization, stemming/lemmatization, recognizing confirmation of semantic association and syntactic component marking language disclosure.

2.1.2 Remove Bangla Punctuation

Remove bangla punctuation is the way where we can mark all bangla punctuation and remove it as there are no need of them. Bangla punctuations are Full-stop(.), Semi-colon, Comma(,), Question Mark(?), Exclamatory Mark(!), Inverted comma(“”), Colon(:), Colon dash(:-), Dash(-), Hyphen(-). Here we just remove all of these bangla punctuation from our comments.

2.1.3 Contractions

Contraction is the way where we convert short form word in details.

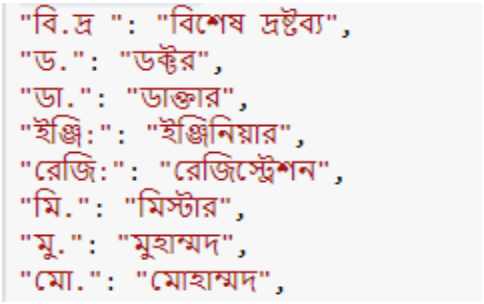


Fig 2.1: Contraction

Here we take the short form □□.□□□ and then converted into □□□□□ □□□□□□□□□□.And we did the same for all of short forms.

2.1.4 Format words and remove unwanted characters

This is the way where we can format the words by removing the unwanted characters. Unwanted characters like (), {}, [], #, $, %, ^, *, ~, ` etc. There are no need of these characters in our work. So what we have done here we just remove the unwanted characters.

2.1.5 Tokenization and CountVectorizer

Tokenization is breaking down raw material into little alleys. Tokenization breaks raw content in words, sentences are called tokens. For example, a token in a word the sentence, and a sign in a passage. These tokens help to understand Creating models of specific situations or natural language processing (NLP). Tokenization helps to understand the significance of the content.

2.1.6 Natural Language Toolkit (NLTK)

It is a set-up of open-source devices initially made in 2001 at the University of Pennsylvania to make building NLP measures in Python simpler. This bundle has been extended through the broad commitments of open-source clients in the years since its unique turn of events. NLTK fundamentally gives a kick off to building any NLP cycle by giving the essential apparatuses. We can chain these fundamental instruments together to achieve the last objective as opposed to building each one of those apparatuses without any preparation.

2.2 Related Works

Automated hate speech detection and the problem of offensive language, They Controversial statements Controversial, hostile and categorized tweets between these two and contrasts intense speech from hostile language.

Detecting offensive tweets via topical feature discovery over a large scale twitter Corpus, here they suggest designing a lexical syntactic feature for recognition Distinguish between unfavorable substances and expected unfavorable clients in online media.A survey on

different text categorization techniques for text filtration [24], here they.The study is led on content substance filtration utilizing diverse content order strategies.

Comparison of Text Classification Algorithm, on paper, they compare SMV, NB algorithm for text order where SMV is better than NB algorithm. Lots of Instructions are provided on the type of test content grouping and its applications.Detecting sentiment from Bangla text using machine learning technique and feature.

Analysis, they proposed a flat plane Count to discover the tyranny of a Bengali comment. They have worked with 300 Bengali comments have not yet found any accuracy using that information.

An approach to detect abusive Bangla text, they proposed a root leveled calculation to discover the oppressiveness of a Bangla remark. They worked with 300 Bangla remarks yet didn't found any precision utilizing that information.

An application of machine learning to detect abusive Bengali text, in this paper they attempted to identify oppressive Bangla text utilized Support Vector Machine, Random Forest along with Radial Basis operations, Polynomial, graphical, and n-gram based count victimizer to recognize Bengali damaging content.

A Bangla news classifier using distinct ML algorithms is created by M. S. ET. Al. Naïve Bayes, KNN, Decision Tree, SVM, and Random Forest have been used by numerous classification methods. They employed the uncertainty matrix for calculating precision. Naïve Bayes has far more precision than other ML algorithms by using all the algorithms they have discovered. Naïve Bayes' precision was 85 percent.

2.3 Comparative Studies

In our project the Random Forest classifier(RF) gave us the best accuracy if we compare it with the other algorithms.Then the best accuracy that we got was from LogisticRegression(LR). And after that the best we got was from MultinomialNB (NB). The difference of accuracy between the LR and the NB is a little. They both gave us almost the same result.

2.4 Scope of Problem

Our work is to detect the offensive Bengali languages that are used in real-time radio message gateways. We can counter the cyber threat of social media through our model. Radio Jockey doesn't have to be embarrassed to be on Radio Live and we can make that environment safer. We have worked with the standard Bangla language. If someone writes the wrong spelling, he will not detect it.

2.5 Challenges

We face numerous challenges in managing this research project at every stage. Due to the structure of the Bengali language, the Bengali language is much more complex than the English language. It is difficult to handle Bangla text. The main challenge of this work was collecting dataset and working with them. We collected The comments for the dataset almost for five months. And that was so difficult to collect the Bangla comments.

CHAPTER 3

Research Methodology

3.1 Data Collection Procedure

We first needed a data set to move this project forward. Since we will detect Bangla's abusive language, we have to collect Bangla data. But we did a lot of searching but could not find the Bangla data set online. So we created it by collecting data from various radio live programs and other sources. We have collected these from the social media accounts of various celebrities and programs on the social media platforms of radio stations. I put all our data in a CSV file. And to see the results of this project we planned to link to the website that we will create later. And that's how we plan to do it. And while collecting these we just realized that this project is worth making.

3.2 Preprocessing

We have some good and some bad data in this dataset. After the dataset is processed and data is all collected in sheet 1 we differentiated its type and started labeling it. Computers do not understand our language so computers need to be trained to understand language. We have been leveling the entire data set to 0 and 1 sections. We defined Section 0 as a good comment and Section 1 as an offensive comment. We have trained the data set using a machine learning algorithm. Now if the comments that users make will match the level of comment 0 then it will be displayed in the interface but when the comment

matches the level of comment 1 then that comment is offensive so it will automatically go to spam and will not be displayed in the user interface.

3.3 Proposed Methodology

The architecture we proposed for this is given below.

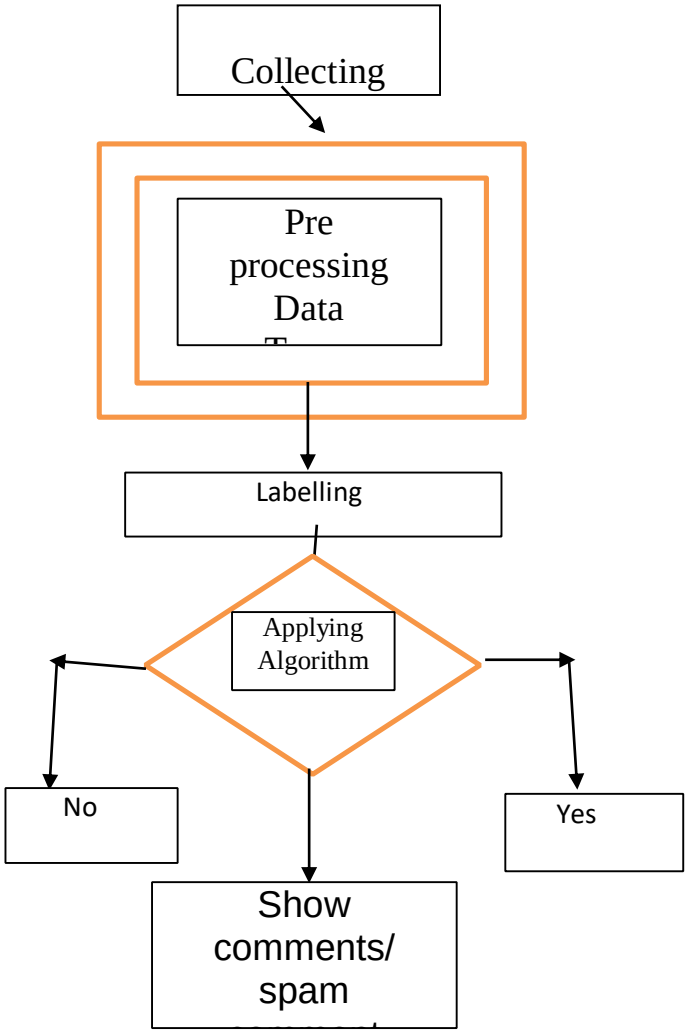


Fig 3.1: System Architecture

3.4 Representation Test Raw Comments

Below are some examples of the stage asset illustrations of our proposed algorithm Comments.

comments	label
ভাল্লাগছে	0
দারুন লাগছে সবাইকে	0
তোমাকে শুভ নববর্ষের সুবেচ্ছা।	0
তুমি অনেক সুন্দর	0
ঈদ মোবারক	0
বেয়াদব একটা	1
তোকে গরু চোরের মতো লাগছে	1
বাজারে মেয়ে এই গুলো	1
হিরু পাগল	1
বেটা পাগল	1

Figure 3.2: Representation Test Raw Comments

Here, if there is an abusive word in a sentence then 1, otherwise 0.

3.5 Implementation Requirements

3.5.1 Local System

• CPU: Intel® Core™ i7

• RAM: 8 GB

• Clock Speed: 2600 MHz

• Video Memory: 2 GB

• L3 Cache: 2 MB

• Operating System: Windows 10

3.5.2 Implementation Requirements

• Language: Python (3.9.1+)

• Code Executer: Google Colab

CHAPTER 4

Experimental Result and Discussion

4.1 Experimental Result and Analysis

We applied three separate algorithms to the pre-processed data to calculate the accuracy of our work. Applying the multinomial algorithm results in an accuracy of 68.93% which is quite successful. And since it has been applied to any Bangla data set as identification of our project, .68.93% of the success is quite good. From Figure 8 we see that the logistic regression algorithm gives slightly better results than the multinomial algorithm. The accuracy rate of this algorithm is 69.52% which is 0.59% higher than the first one. From Figure 9 we have applied a random forest classifier algorithm with an accuracy rate of 75.5% which is better than the accuracy rate of the above two algorithms.

Multinomial Algorithm:

Here is a figure of different scores regarding the value of recall, precision, accuracy, F1 score, and some code parts. It illustrates the comparison of each score for the multinomial algorithm.

```
Confusion Matrix for Multinomial Naive Bayes:
[[2331 1091]
 [ 960 2219]]
Score: 68.93
Classification Report:              precision    recall  f1-score   support

           0       0.71      0.68      0.69      3422
           1       0.67      0.70      0.68      3179

    accuracy                           0.69      6601
   macro avg       0.69      0.69      0.69      6601
weighted avg       0.69      0.69      0.69      6601
```

Fig 4.1: Multinomial algorithm with some code part

Logistic regression:

Here is a figure of different scores regarding the value of recall, precision, accuracy, F1 score, and some code parts. It illustrates the comparison of each score for the logistic regression algorithm.

```
Confusion Matrix for LogisticRegression Classifier:
[[2573  849]
 [1163 2016]]
Score:  69.52
Classification Report:
              precision    recall  f1-score   support

           0       0.69      0.75      0.72      3422
           1       0.70      0.63      0.67      3179

    accuracy                           0.70      6601
   macro avg       0.70      0.69      0.69      6601
weighted avg       0.70      0.70      0.69      6601
```

Fig 4.2:Logistic Regression algorithm with some code part

Random forest classifier:

Here is a figure of different scores regarding the value of recall, precision, accuracy, F1 score, and some code parts. It illustrates the comparison of each score for the random forest classifier algorithm.

```
Confusion Matrix for Random Forest Classifier:
[[2696  726]
 [ 891 2288]]
Score: 75.5
Classification Report:              precision    recall  f1-score   support

           0       0.75      0.79      0.77      3422
           1       0.76      0.72      0.74      3179

    accuracy                           0.76      6601
   macro avg       0.76      0.75      0.75      6601
weighted avg       0.76      0.76      0.75      6601
```

Fig 4.3:Random Forest Classifier with some code part

4.2 Sentence Tokenization

We use CountVectorizer to convert a sentence into word. Like the sentence □□□ □□□□□□ and here what it will do? The CountVectorizer will tokenize it as word (□□□) and (□□□□□□). And it will also generate the vector representation by enabling the pre-processing of text data prior.

4.3 Confusion Matrix Test Results and Classification Report:

Confusion Matrix score and Classification Report shown in the figure 4.4

```
Confusion Matrix for Support Vector Machines:
[[2537  885]
 [1001 2178]]
Score: 71.43
Classification Report:              precision    recall  f1-score   support

           0       0.72      0.74      0.73      3422
           1       0.71      0.69      0.70      3179

    accuracy                           0.71      6601
   macro avg       0.71      0.71      0.71      6601
weighted avg       0.71      0.71      0.71      6601
```

Figure 4.4 :Confusion Matrix score and Classification Report.

4.4 Discussion

After all the logical examination of the results given in Section 4.1. We can say that

Our natural language processing and machine learning algorithm is given an acceptable result.

CHAPTER 5

Impact on Society, Environment and Sustainability

5.1 Impact on Society, Environment and Psychological state

In the current times, social media has a lot of impact on society and our environment. The Internet is changing the expression of communication and values. the Internet has considerably changed the way we get our news, talk to our friends and usually live our lives. Its distributed environment makes it a perfect place for amateurs and professionals comparable to create and share ideas, data, pictures, videos, art, music, and more.Despite, or perhaps because of, its self-governing environment, the Internet is also occupied with Web sites committed to provoking hate contrary to particular ethnic, spiritual, racial, or sexually concerned with groups such as women, Jews, African, Americans, Muslims, and members of the lesbian, gay, bisexual and transgendered (LGBT) community. Occurrences have been reported on almost every continent about how a considerable amount of people of the world now communicates on social media, with approximately a third of the world's population active on Facebook only Social platforms also compromise passionate actors the chance to expose their acts.

Relationships play massive roles. Teens who were elevated in a family where harsh inquiry and pitiless abuses were the values, that can reveal that words can wound much more than body wounds. because they drive in themselves in youthful personalities, combined with the suffering that somebody who should love and appreciate, a mother, a dad, clearly doesn't have a promising opinion by any spring of the dreams. Those sorts of unbearable damages repeat themselves for quite a long time to come. Not many kids

growing up can release themselves from the scimitar named on them by a brutal parent or another powerful figure.

In 2011 a survey by Jeffrey Bowers declares application boundary line has to sway and bend human behavior. This inspection led to consider phonetic comparativeness concerning adjuration terms and dual speak. In A piece of study, 24 members among the ages of 18-26 using the age of 21 reliant on upon 20-minute experimental counting their reactions adjuration terms verbally expressed resoundingly and their reactions marked. Furthermore, their electro dermal action valued utilizing firm gadgets that deliberate changes in skin hindrance because of the adjuration terms. One more examination at Stanford in 2016 confirmed an instant connection between impudence and genuineness. In view of this investigations of 307 members, it worked out that the best two US states on impudence were additionally the most noteworthy on respectability. Past ideas about youngsters picking swearing from grown-up behavior have been discovered wrongly while facts view that they figure out how adjuration so the behavior of concern.

There are a lot of cultures that view the state of mind of grown-ups which were fetched by paying attention only to challenging and cruel fittings so they are on the upsurge. Several young folks counting stranded rational oneself-opinions take part in awful practices that hurt them much more terrible: sexual depravity, lingering drug habit, even offenses like burglary. These young grown-ups tend to lash out, and this is usual for young people who are on the surge down like this masses to turn out to be more spiteful than children brought up in a more sure and supporting environment. Ruthlessness raises violence. A youngster who knows nothing else except for puns, vocally abusing, and sarcasm will turn into the harasser of his society. It is the foremost pathway he understands how to sustain. The impacts brought about through adverse times, in case that is each of the youngster knows, are tough and extensive. They will do to others as they've had done to them, and that is can capably affect folks who never knew the victim as a youngster, however, who may meet him in a vague back way as an adult.

Remarkable-word can impact the depletion of qualities that manage corporeal and enthusiastic pressure. Absolute words, for example, "serenity" and "amour," can adjust the outflow of qualities, stimulating regions in our forward forecasts, and progressing the mind's psychologically occupied. They drive the stimulating focuses of the mind without hesitancy, as indicated by the makers, and concept strong point. Then again, the ill-disposed language can upset obvious potentials that have a serious influence on the creation of neurochemicals that armor us from pressure. People are intended to stress some percentage of our basic minds protecting us from dangers to our durability so our reflections normally go here damsel.

In several situations, an unsocial contradictory term can figure the action in our amygdala. It carries many burdens creating hormones and synapses, which thus delays our cerebrums' work. (That is mainly as to foundation, cause, and language.) Energetic terms communication attention information via the mind, and they centrally untie the validation and thoughtful emphases situated in the front projections, compose Newberg and Waldman.

It's everything all through media TV, web-based media, music, and so on. Awful arguments seem to be inevitable. The web has taken into explanation offensiveness to show up all over the place and to seem classless, accordingly causing these words to lose their head-turning regard. To preference reflection nowadays, an individual must apply a pile of words suspended organized inventively to cause an impact of utmost crudeness.

Abusive words increasingly tiptoe into our waffle without us considerate until they become a distinguishing response in a variety of conditions. Cursing can have somebody with an awful effect, sign an lack of control, and show an awful nature of enthusiasm. This is much added honest concerning our age involving with individuals from more experienced ages.

5.2 Ethical Aspect

Bearing in mind these topics, abusive or indecent, or cursed words that hinder the society and connection among colonizes we must prevent these from being used in our lives. For this purpose, we have built a program that will detect and classifies cursed words that are used in social media or internet-based social platforms. This not will detect but also will avoid being assembled in these platforms. Abusive words are proliferative that's why we must prevent them from being used and our program will be a step in the direction of a harmless and tranquil atmosphere.

CHAPTER 6

Summary, Conclusion, and Implication for Future Research

6.1 Summary of the Study

The title of our proposed project is detecting Bangla abusive Language from a real-time radio message gateway. We create a dataset for our projects that are collected from social media accounts of various celebrities and programs on social media platforms of radio stations. We have leveled the entire data set using 0 and 1 sections. We have created a website using Django. It will be used as the personal website of the radio stations and we have connected our dataset to the website using machine learning algorithms. Our algorithm will check the dataset when a live radio program is broadcast and listeners will try to communicate by sending a direct message and will only show positive comments. . All the abusive messages in real-time will get to the spam folder here to save time.

6.2 Conclusions

Since we have done this project to detect the Bangla abusive language, we have used three algorithms for best results as Bangladesh is a regional language. In the future, we have added an empty dataset to collect more data where new comments will enter directly. We have decided to work for 99% accuracy in regional languages. We plan to study using different levels of syntactic analysis and characters, and hope to tokenize these in word and region-based language, this processing can lead to improved recovery.

6.3 Implication for Future Research

Since there are obstacles in our project, the proposed strategy can be expanded in the future.

- We will break down every word and tokenize it so that more accurate answers come.
- We, Will, work for more accuracy with regional languages.