FINAL YEAR PROJECT REPORT

TOXIC COMMENT CLASSIFICATION

In fulfillment of the requirement
For degree of
BS (COMPUTER SCIENCES)

By

BILAL AHMED	51893 BSCS
AHMED ABDULLAH	51480 BSCS
MAAZ AMIR	51848 BSCS

SUPERVISED

BY

AMNA IFTIKHAR

BAHRIA UNIVERSITY (KARACHI CAMPUS)

SPRING-2021

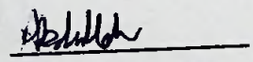
DECLARATION

We declare that this project report is based on our original work except for citations and quotations which have been duly acknowledged. We also declare that it has not been previously and concurrently submitted for any other degree or award at Bahria University or other institutions.

Signature :

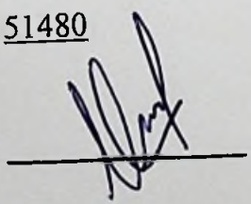
Name : BILAL AHMED

Reg No. : 51893

Signature : 

Name : AHMED ABDULLAH

Reg No. : 51480

Signature : 

Name : MAAZ AMIR

Reg No. : 51848

The copyright of this report belongs to Bahria University according to the Intellectual Property Policy of Bahria University BUORIC-P15 amended on April 2019. Due acknowledgement shall always be made of the use of any material contained in, or derived from, this report.

© 2019 Bahria University. All right reserved.

ACKNOWLEDGEMENTS

We would like to thank everyone who had contributed to the successful completion of this project. We would like to express my gratitude to my research supervisor, Miss Amna Iftikhar for her invaluable advice, guidance and her enormous patience throughout the development of the research.

In addition, we would also like to express our gratitude to our loving parent and friends who had helped and given me encouragement.

TOXIC COMMENT CLASSIFICATION

ABSTRACT

We live in a technological age where most of us have easy access to the internet. With the evolution of the Internet, the use of social media and communication forums for communication has increased significantly in recent years. But this progress also opens the door to trolls poisoning these social media and forums by behaving rudely towards others. Detecting toxic comments online has become an important issue in recent years. Toxic comments are defined as obscene, inappropriate or abusive comments that leave you speechless. Current methods of dealing with online poisoning often rely heavily on manual moderation and are not measurable enough to handle the growing number of users on a daily basis. Our project Toxic Comment Classification model through which we can identify toxicity, identity-hate, threats, severe-toxic, obscene etc and through this we can automate the process of identification of toxic material on online forums and other communication platforms. Different techniques and models which we used for the identification and different stages involving like data analysis, data processing, TF/IDF, logistic regression and other neural network and machine learning models which will be studied and discussed.

TABLE OF CONTENTS

CHAPTER 1

1 INTRODUCTION 16

1.1 Background 16

1.2 Problem Statements 16

1.3 Aims and Objectives 17

1.4 Scope of Project 17

CHAPTER 2

2 LITERATURE REVIEW 18

CHAPTER 3

3 DESIGN AND METHODOLOGY 21

3.1 Website Input Design 21

3.2 Flow Diagram 22

3.3 Data Gathering 23

3.4 Data Analysis 23

3.4.1 Data Processing 24

3.4.1.1 Tokenization 24

3.4.1.2 Removal of Stop words, punctuations, unwanted tokens 24

3.4.1.3 Lemmatization/Stemming 24

3.4.2 TF/IDF 25

3.4.3 Machine learning Models 25

3.4.3.1 Logistic Regression 25

3.4.3.2 SVM (Support Vector Machine) 26

3.4.3.3 Naïve Bayes 26

3.5 DEEP LEARNING APPROACH 27

3.5.1 Data Processing 27

3.5.1.1	Removal of Stop words, punctuations, unwanted tokens	27
3.5.1.2	Tokenization	27
3.5.1.3	Glove Embedding	27
	3.5.2 Deep learning Models	28
3.5.2.1	CNN (Convolutional Neural Network)	29
3.5.2.2	LSTM (Long Short-Term Memory)	29

CHAPTER 4

4	**IMPLEMENTATION AND TESTING**	**30**
4.1	MACHINE LEARNING	30
	4.1.1 Data Analysis	30
	1. Data Processing	34
4.1.1.1	Tokenization	34
4.1.1.2	Removal of Stop words, punctuations, unwanted tokens	34
4.1.1.3	Lemmatization/Stemming	35
	4.1.2 TF/IDF	36
	4.1.3 Models	36
4.1.3.1	Logistic Regression	36
4.1.3.2	SVM (Support Vector Machine)	41
4.1.3.3	Naïve Bayes	45
4.2	DEEP LEARNING	49
	4.2.1 Data Processing	49
4.2.1.1	Removal of unwanted words and Tokenization	51
	4.2.2 Glove Embedding	54
	4.2.3 Deep learning Models	55
4.2.3.1	CNN (Convolutional Neural Network)	55
4.2.3.2	LSTM (Long Short-Term Memory)	58
4.3	WEBSITE	63

CHAPTER 5

5.1	DISCUSSION / CONCLUSION	70