



FINAL YEAR PROJECT REPORT

CYBER BULLING DETECTOR FROM MIXED-CODE TEXT

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

By

**FURQAN AHMED
MUQADUS**

**79012 (BSCS)
79215 (BSCS)**

SUPERVISED

BY

MS. ASIA SAMREEN

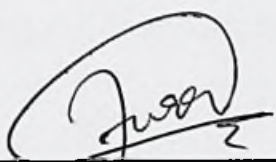
BAHRIA UNIVERSITY (KARACHI CAMPUS)

SPRING-2025

DECLARATION

We hereby 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.

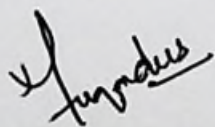
Signature :



Name : Furqan Ahmed

Reg No. : 79012

Signature :

Name : MuqadusReg No. : 79215

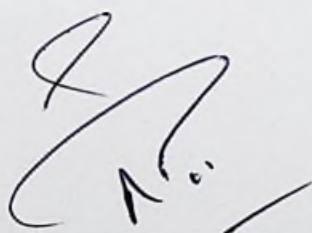
Date :

26/08/25

APPROVAL FOR SUBMISSION

We certify that this project report entitled **“CYBER BULLING DETECTOR FROM MIXED-CODE TEXT”** was prepared by **Furqan Ahmed**, and **Muqadus** has met the required standard for submission in partial fulfilment of the requirements for the award of Bachelor of Computer Science at Bahria University, Karachi.

Approved by,



Signature :

Supervisor : Miss Asia Samreen

Date :

26 | 05 | 25

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.

© 2025 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 project supervisor, Miss Asia Samreem 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 parents and friends who have helped and given us encouragement.

CYBER BULLING DETECTOR FROM MIXED-CODE TEXT

ABSTRACT

Cyberbullying has become a growing concern in Pakistan's digital landscape, where social media and online platforms have become primary means of communication. The vast linguistic diversity of Pakistan makes it harder for detection systems to identify harmful content throughout the nation. The combination of English with Roman Urdu and Roman Pashto and Roman Balochi and Roman Sindhi in user communications creates challenges for standard content moderation tools that aim to accurately determine offensive language.

Online systems designed to detect cyberbullying exist mainly for single-language environments but fail effectively when analyzing code-mixed text because they only use basic keyword filters. The combination of wrong positives with wrong negatives due to this approach diminishes the general success rate of online content moderation systems.

Our project presents an NLP-based cyberbullying detection system which specifically addresses the needs of Pakistan's digital environment that uses multiple languages. The system combines rule-based technical methods with scoring procedures to evaluate various regional language text written in Roman script. The system performs real-time detection of harmful content through automated alerts that lead to administrator intervention for proper actions after repeated violation detection.

The method provides inclusive accurate and flexible detection of cyberbullying for users across Pakistan's multilingual online landscape to establish secure digital communication environments.

TABLE OF CONTENTS

DECLARATION	ii
APPROVAL FOR SUBMISSION	iii
ACKNOWLEDGEMENTS	v
ABSTRACT	vi
TABLE OF CONTENTS	vii
LIST OF TABLES	x
LIST OF FIGURES	xi

CHAPTERS

1	INTRODUCTION	1
1.1	Background	1
1.2	Problem Statement	1
1.3	Aims and Objectives	2
1.4	Scope of Project	2
1.5	Expected Outcomes	3
1.6	Chapter Summary	3
2	LITERATURE REVIEW	4
2.1	Background	4
2.2	Related Work	4
2.2.1	Multilingual Cyberbullying Detection	4
2.2.2	Cyberbullying Detection in Roman Urdu	5
2.2.3	Machine Learning-Based Approaches	5
2.2.4	Deep Learning-Based Approaches	5
2.2.5	Hybrid Approaches for Mixed-Code Text	5

2.2.6	Social Media-Specific Features in Cyberbullying Detection	6
2.2.7	Rule-Based Scoring Approaches	6
2.3	COMPARISON TABLE WITH EXISTING STUDY	6
2.4	Chapter Summary	12
3	REQUIREMENT SPECIFICATIONS	13
3.1	Existing System	13
3.2	Proposed System	15
3.2.1	System Workflow:	15
3.3	Requirement Specification	16
3.3.1	Functional Requirements	16
3.3.2	Non-Functional Requirements	16
3.4	USE CASE	16
3.4.1	Detect Cyberbullying in Mixed-Code Text	16
3.4.2	Issue Warning for Harmful Content	17
3.4.3	Flag & Report Content to Admin	17
3.4.4	Allow Moderators to Adjust Cyberbullying Dictionary	17
3.5	Tools & Technologies	18
3.6	Chapter Summary	18
4	DESIGN AND METHODOLOGY	19
4.1	Proposed Model	19
4.2	Techniques For Cyberbullying	20
4.3	Software Development	22
4.3.1	Process Model: Agile Methodology with Scrum Framework	22
4.4	System Modules	23
4.5	Detailed Overview of the Detection Model	26
4.6	Flow Chart	28
4.7	Chapter Summary	29
5	SYSTEM IMPLEMENTATION	30
5.1	Detection Model	30

5.2	Chrome Extension	33
5.3	User Registration Handler	35
5.4	Admin Dashboard	36
5.5	User Dashboard	37
5.6	Chapter Summary	38
6	TESTING AND EVALUATION	39
6.1	Tools and Frameworks Used	40
6.2	Unit Testing	40
6.3	Endpoint Testing	41
6.4	Authentication Testing	41
6.5	Integration Testing	42
6.6	Results	42
6.7	Chapter Summary	43
7	CONCLUSION	44
	REFERENCES	46