ACKNOWLEDGEMENT

We are very thankful and grateful to our beloved guide, Professor **Mrs. K. Elakia** whose great support in valuable advices, suggestions and tremendous help enabled us in completing our project. He has been a great source of inspiration to us.

We also sincerely thank our Head of the Department, **Dr . R. G. Suresh Kumar** whose continuous encouragement and sufficient comments enabled us to complete our project report.

We thank all our **Staff members** who have been by our side always and helped us with our project. We also sincerely thank all the lab technicians for their help as in the course of our project development.

We are grateful to our Chairman, **Shri. M. K. Rajagopalan**. He has been a constant source of inspiration right from the beginning.

We would also like to extend our sincere gratitude and grateful thanks to our Principal **Dr . E. Vijayakrishna Rapaka** for having extended the research and development facilities of the department.

We wish to thank our family members and friends for their constant encouragement, constructive criticisms and suggestions that has helped us in timely completion of this project

ABSTRACT

Cyberbullying Detection uses a combination of machine learning techniques such as TF-IDF vectorization, logistic regression, multilayer perceptron, CNNs and LSTM networks to create a robust model for detecting cyberbullying. By Employing BERT model, it's able achieve higher accuracy and better performance in identifying offensive content on social media platforms.

The existing System for detecting cyberbullying in Indian Language Bengali on social media. The model uses text preprocessing, TF-IDF, and Instance Hardness Threshold (IHT) for resampling, it uses multiple Machine learning algorithms for detection of online harassment.

However, the existing System does not address the practical challenges like Real-Time Detection and the Technique used for Resampling deduce the actual size of dataset to balance the dataset which leads to lower accuracy rate. To overcome these Limitations, the proposed system uses the BERT model, known for its advanced contextual understanding and bidirectional processing capabilities, to enhance prediction accuracy.

KEYWORDS:

Cyberbully, Machine learning, Social media, NLP, Twitter, Facebook.

LIST OF TABLES

TABLE NO.	NAME OF THE TABLE	PAGE NO.
2.6	Literature Survey Comparison	15

LIST OF FIGURES

FIGURE NO	NAME OF THE FIGURE	PAGE NO
1.1	Deep learning Architecture	3
1.2	Machine Learning & Deep Learning	4
1.3	Neural Network Architecture	4
1.4	Unrolled RNN	7
1.5	Architecture of BERT	8
3.1	Bengalibullying detection architecture	18
4.1	Proposed System architecture	21
5.1	Environmental setup screenshot	22
6.1	Pre-Training & Fine-Tuning of BERT	23
7.1	Accuracy metrics Graph	27
7.2	Comparison Graph	28

LIST OF ABBREVIATIONS

1. **AI** - Artificial Intelligence
2. **ANN** - Artificial Neural Network
3. **BERT** - Bidirectional Encoder Representations from Transformers
4. **Bi-LSTM** - Bidirectional Long Short-Term Memory
5. **CNN** - Convolutional Neural Network
6. **CML** - Conventional Machine Learning
7. **DNN** - Deep Neural Network
8. **GPU** - Graphics Processing Unit
9. **IDE** - Integrated Development Environment
10. **IHT** - Instance Hardness Threshold
11. **LIWC** - Linguistic Inquiry and Word Count
12. **ML** - Machine Learning
13. **MLP** - Multi-Layer Perceptron
14. **NLP** - Natural Language Processing
15. **RNN** - Recurrent Neural Network
16. **RoBERTa** - A Robustly Optimized BERT Pretraining Approach
17. **SVM** - Support Vector Machine
18. **TF-IDF** - Term Frequency-Inverse Document Frequency
19. **VCS** - Version Control Systems

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NO
	ABSTRACT	i
	LIST OF TABLES	**ii**
	LIST OF FIGURES	**iii**
	LIST OF ABBREVIATIONS	**iv**
1.	**INTRODUCTION**	**1**
	1.1 CYBERBULLY	1
	1.2 DEEP LEARNING	2
	1.2.1 DEEP LEARNING & MACHINE LEARNING	3
	1.2.2 NEURAL NETWORKS	4
	1.2.3 TYPES OF DEEP NEURAL NETWORK	5
	1.2.4 MULTI-LAYER PRECEPTRONS	5
	1.2.5 CONVOLUTIONAL NEURAL NETWORK	5
	1.2.6 ADVANTAGES OF CNN	6
	1.2.7 RECURRENT NEURAL NETWORK	6
	1.2.8 ADVANTAGES OF RNN	7
	1.2.9 BERT	7
2.	**LITERATURE SURVEY**	**10**
	2.1 LITERATURE SURVEY - I	10
	2.2 LITERATURE SURVEY - II	11
	2.3 LITERATURE SURVEY - III	12
	2.4 LITERATURE SURVEY - IV	13
	2.5 LITERATURE SURVEY – V	14
	2.6 LITERATURE SURVEY COMPARISON	15

3.	**SYSTEM STUDY**		**17**
	3.1	EXISTING SYSTEM	17
	3.1.1	EXISTING SYSTEM ARCHITECTURE	18
	3.1.2	LIMITATIONS IN EXISTING SYSTEM	19
4.	**PROPOSED SYSTEM**		**20**
	4.1	PROPOSED SYSTEM INTRODUCTION	20
	4.2	PROPOSED SYSTEM ARCHITECTURE	21
	4.3	ADVANTAGES OF PROPOSED SYSTEM	22
5.	**SYSTEM REQUIREMENTS**		**22**
	5.1	HARDWARE REQUIREMENTS	22
	5.2	SOFTWARE REQUIREMENTS	22
	5.3	ENVIRONMENTAL SETUP	22
6.	**DESIGN AND IMPLEMENTAION**		**24**
	6.1	PROPOSED SYSTEM ARCHITECTURE	24
	6.2	IMPLEMENTATION	24
	6.3	MODULES USED	26
7.	**RESULT AND DISCUSSION**		**26**
	7.1	RESULT	26
	7.2	DISCUSSION	26
	7.3	ACCURAY METRICS	27

8. CONCLUSION AND FUTURE ENHANCEMENT 28

8.1 CONCLUSION 28

8.2 FUTURE ENHANCEMENT 28

APPENDIX 30

APPENDIX 1 30

APPENDIX 2 31

REFERENCE 37