INFORMATICS INSTITUTE OF TECHNOLOGY

INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER

CyberSuccor : Intelligence Technique for Sinhala Cyberbullying Detection on Social Media

A Project Proposal by

Mr. Vajith Chamuditha

Supervised by

Ms. Niwarthana Kariyabaduge

Submitted in partial fulfillment of the requirements for the BEng (Hons) Software Engineering degree at the University of Westminster.

November - 2022

Table of Contents

1. PROJECT PROPOSAL 1

1.1 Introduction 1

1.2 Problem domain 1

1.3 Problem definition 2

1.3.1 Problem statement 2

1.4 Research Motivation 3

1.5 Existing work 3

1.5.1 Machine Learning (ML) / Deep Learning (DL) approach 4

1.6 Research Gap 7

1.7 Contribution to the body of knowledge 8

1.7.1 Technological Contribution 8

1.7.2 Domain Contribution 8

1.8 Research Challenges 8

1.9 Research questions 8

1.10 Research Aim 9

1.11 Research Objective 10

1.12 Project Scope 11

1.12.1 In-scope 11

1.12.2 Out-scope 11

1.12.3 Features of the Prototype 13

2 METHODOLOGY 14

2.1 Research methodology 14

2.2 Development methodology 14

2.2.1 Design Paradigm 15

2.2.2 Evaluation methodology 15
2.3 Project management methodology 15
2.3.1 Resource Requirements 16
2.3.2 Gantt Chart 18
2.3.3 Deliverables 19
2.3.4 Risk Management 19
REFERENCES i

List of Tables

Table 1: Existing Works 4
Table 2: Research Objectives 9
Table 3: Research methodology 13
Table 4: Hardware Requirements 16
Table 5: Software Requirements 16
Table 6: Skills Requirements 17
Table 7: Data Requirements 17
Table 8: Deliverables 19
Table 9: Risk Management Plan 20

List of Figures

Figure 1: Gantt chart 18

Figure 2: Prototype Feature Diagram 13

Figure 3: Gantt chart 18

List of Abbreviations

Abbreviation	Definition
NLP	Natural Language Processing
ML	Machine Learning
DL	Deep Learning
CNN	Convolutional Neural Networks
SVM	Support Vector Machine
OCR	Optical character recognition

1. PROJECT PROPOSAL

1.1 Introduction

Bullying is defined as the behavior of forcing, threatening, or coercing another person to abuse, threaten, or physically dominate them. Bullying through electronic media is referred to as cyberbullying. This could manifest as the publication of gossip, threats, sexual remarks, personal details about the victim, or derogatory terms (Akhter et al., 2019). Sinhala is the native language of the Sinhalese individuals who make up the biggest ethnic group of Sri Lanka. In this research project, the author tries to find a novel approach for cyberbullying detection of Sinhala language comments in both text statements and text written images and categorization of negative comments in social media. The problem, the research gap, the research challenge, and the research approach that the author plans to use over the coming several months are all outlined in this document. Additionally, reviews are the necessary problem proofs and earlier research areas of interest. The anticipated schedule for the project's deliverables is lastly stated in the Work Plan.

1.2 Problem domain

Social media

With the development of technology, the internet has become the changing face of culture and communication. Due to vast usage of the internet, social media also makes a significant impact on how people obtain news, communicate with friends, and go about their daily lives. Although technology has many advantages for people, it also has a "black side," as some adults and even some of the youngsters themselves may misuse it to their detriment. Numerous studies have looked at various social media topics as a result of the rise in social media usage. Many offline issues now have an online counterpart as a result of the social media industry's spectacular growth over the past 10 years. Currently, both locally and globally, bullying people is a problem on social networking sites like Facebook, Twitter, Instagram, and others, and these kinds of instances are all too typical these days.

Bullying and Cyberbullying

Bullying and hatred have been issues in society for a long time. However, these bullies can now discreetly use their social media platforms to write disrespectful, abusive, or cruel texts about another person or group of people while hiding behind a computer or smartphone. Cyberbullying is a phenomenon that has harmed many people and led to depression in many

kids and teenagers (Behzadi et al., 2021). Direct cyberbullying and relationship-based cyberbullying are the two categories into which cyberbullying is divided. A victim is directly attacked when someone engages in direct cyberbullying, such as by posting unpleasant and derogatory comments about them on social media. Relationship-based cyberbullying is a novel form of cyberbullying that targets people indirectly, such as by isolating a victim by disregarding the victim's messages. It is exceedingly challenging to identify relationship-based cyberbullying because the indirect attack does not appear in text immediately (Li & Tagami, 2014). Social media posts include both text and images, and depending on the context and other background knowledge of the users, these posts may also have hidden meaning.

According to law enforcement agency CID, a study in Sri Lanka found that over 1000 cyberbullying complaints were reported in 2018. About 90% of college students suffer from cyberbullying, and almost all survey respondents said they knew someone who was bullied online. 80% of cyberbullying occurs on social media, with Facebook being the most common bullying website and the most common bullying offense among students is posting embarrassing videos and photos (65%) was. Posting personal information of victims online was his second (15%), followed by spreading gossip and lies (9%) and posting insults (2%) (Ariyadasa, n.d., 2019).

Natural Language Processing

Making computers understand statements or words written in human languages is the goal of the domain of artificial intelligence and linguistics known as "Natural language processing," or NLP. The development of natural language processing was made possible by the user's desire to interact with computers in natural language and to make their work easier. NLP serves users who lack the time to acquire new languages or to refine their existing ones because not all users will be fluent in machine specific language (Khurana et al., 2022).

1.3 Problem definition

A deliberate and hostile conduct carried out over time against a helpless victim who is unable to defend themselves is defined as cyberbullying. It can be committed by one individual or a group of individuals (NaliniPriya G & Asswini M, 2015). In today's modern world, social media platforms serve as the primary means of communication. Cyberbullying poses a real

threat to the short- and long-term prosperity of users of online media. Web 2.0 provides easy, interactive access to online communities anytime, anywhere, but also provides opportunities for cybercrime such as cyberbullying. Life-threatening cyberbullying among young people has been reported internationally. Addressing this problem in an online environment requires the ability to easily spot cyberbullying and identify the role participants play in social interactions. Additionally, young adults and teenagers are especially at risk for cyberbullying. Currently, there is no application for identifying and categorizing Sinhala cyber bullying in social media.

1.2.1 Problem statement

Cyberbullying has been an underlying problem for quite some time in Sri Lanka, but at this time there are no recognized adequate technical solutions to distinguish cyberbullying at its underlying stages by analyzing text comments and text written images.

1.4 Research Motivation

According to SLCERT, there were about 15,000 cases of online abuse recorded last year, and the prevalence of cyberbullying on social networking sites has considerably increased nationwide. Many people began using internet platforms when the COVID-19 epidemic broke out, and kids were encouraged to use them for learning as well. Approximately 8000 instances of cyberbullying on social media platforms were documented in the first half of 2021 (Sri Lanka News - Newsfirst, 2021).

1.5 Existing work

With the advancement of artificial intelligence, there has been a paradigm change toward actively combating cyberbullying through the use of computer techniques like Artificial intelligence to detect and prevent it. With time, numerous approaches were tried and developed by the researchers to distinguish best-performing algorithms, models for the purpose of detecting cyberbullying using a variety of datasets.

The majority of studies employed lexicon-based and machine learning models extensively for text mining. Additionally, text mining was explored by researchers that are exploring deep learning. The unsupervised learning to detect cyber bullying is still the subject of limited investigation (Hettiarachchi, 2020).

1.4.1 Machine Learning (ML) / Deep Learning (DL) approach

Citation	Brief	Limitation of their research	Contributions of their research
Srinandhini, B., & Sheeba, J. I. (2015). Online social network bullying detection using intelligence techniques. *Procedia Computer Science*, *45*(C), 485–492. https://doi.org/10.1016/j.procs.2015.03.085	Starts the process of recognizing cyberbullying behavior with a social network input dataset. The input data is pre-processed before being supplied to Feature Extraction. The learning algorithm receives the extracted features. The core component (algorithm) consists of a genetic algorithm for simulating exploratory and adaptive behavior. The output from the learning unit is supplied to the classifier approach, which uses the chromosomal fitness value to categorize the cyberbullying behaviors. The output is categorized according to the chromosome with the highest fitness value.	Not stated	Data pre-processing, which involves deleting stop words, excess characters, and hyperlinks.

Jikriya, S. S. (2021). Cyber bullying Detection in social media using Supervised ML & NLP Techniques. *International Journal for Research in Applied Science and Engineering Technology*, *9*(VI), 2259–2264. https://doi.org/10.22214/ijraset.2021.35483	Identify posts that are inappropriate for children or teenagers, and can successfully deal with the wrongdoings that are committed via these platforms. An algorithm that uses Supervised learning to identify and stop cyberbullying on Twitter. Naive Bayes and Support Vector Machines are the models' descendants, and the TFIDF vectorizer was also used in mining. The results demonstrate that Support Vector Machine, which is superior to Naive Bayes, offers exceptional accuracy for spotting social media bullying content. The created model will protect people from online media threats' assaults.	utilized the TFIDF vectorizer.	Provide modifications or additions to the anticipated grouping and arrangement calculations to achieve a more extensive implementation. Increase the recognition accuracy and decrease the rate of hostile tweets, a tried-and-true combination of information mining techniques, additional mixes, used.
NaliniPriya G, & Asswini M. (2015). *A DYNAMIC COGNITIVE SYSTEM FOR AUTOMATIC DETECTION AND PREVENTION OF CYBER-BULLYING ATTACKS*.	Create an automated system for the detection and prevention of cyber-bullying attacks by analyzing both textual features as well as social network features to detect cyberbullying. When the offensive message threshold level is exceeded, the system has the ability to send alert signals to the trusted contacts. The technology automatically alerts the user's trusted contacts in order to ask for their assistance in assisting the victim of bullying when the illegal posts or messages are posted on the user's page more than five times. The user's closest people could be the trusted contact. The users choose the reliable contacts on their own.	With the introduction of merging two or more cyberbullying detection elements, accuracy of bullying detection should improve. Improve image	Gathering all data types, including text, image, and social media data. Analyze both textual and social network elements. .

10(10). www.arpnjournals.com		analysis.	
Li, M., & Tagami, A. (2014). A study of contact network generation for cyber-bullying detection. *Proceedings - 2014 IEEE 28th International Conference on Advanced Information Networking and Applications Workshops, IEEE WAINA 2014*, 431–436. https://doi.org/10.1109/WAINA.2014.70	Suggesting a framework for creating a contact network. And show some methods to detect a friendship. The building phase and the extension phase make up the proposed framework. From the gathered SNS data, the building phase generates an undirected graph, and the expansion phase fills in the gaps to finish the formation of the contact network.	Should try to create a more accurate contact network. Look into cutting off the extra links. should try to utilize additional data. e.g., the dynamic anchors	Create a dataset using "DECOLOG", website.

Table 1: Existing Works

1.6 Research Gap

Numerous different types of study have been done in this area throughout the years, and the majority of the research was done to find cyber bullying systems for social networks in different countries for their native languages. Most of the research, projects and models have been implemented for the English language. Despite these results, it is obvious that Sri Lanka lacks adequate detection and classification techniques to identify Sinhala cyberbullying in text statements and text written in photos. When compared to the current systems, the proposed system's primary goal is to improve accuracy and detect cyberbullying terms, particularly for the Sinhala language as it is an area which many researchers are not focused on because it is a bit challenging. Even though there has only been a little amount of research on the Sinhala language, none of it has attempted to analyze both text comments and text written in images with categorization of the negative comments according to the bullying type. It might need proper optical character recognition techniques to identify exact text in the image and accurate classification for negative comments as many comments have hidden meanings and contain multiple bullying types such as age-based bullying with appearance-based bullying etc. The main issue of the existing research is the accuracy and none of the methods offer any recommendations for the victim of bullying. With reference to the current systems, the proposed system's primary goal is to improve accuracy and forecast cyberbullying terms for the Sinhala language.

1.7 Contribution to the body of knowledge

The suggested research fits into the applied research category, and it may be determined how much it adds to the body of knowledge after analyzing the prior work. The key contributions from this research effort can be categorized as solving the observed gaps in both the technological and domain aspects.

1.6.1 Technological Contribution

The key technological contributions of this research are development of a novel algorithm that can improve the performance and accuracy of current systems, analyzing negative statements in text comments and texts written in images using feature extraction and OCR techniques with text classifier methods and evaluating the performance of the suggested approach to the approaches already in use by applying evaluation methodologies (benchmarking and confusion matrix).

1.6.2 Domain Contribution

- Identifying Sinhala cyberbullying comments in social media platforms.
- Analyzing statements in both text comments and text written in images.
- Categorization of negative comments according to the type.
- Recommending solutions for the victims.

1.8 Research Challenges

As Sinhala is a language which has more variety and complexity than other languages, it's very challenging to adapt with existing technologies, models and algorithms which are developed focusing mostly on the English language. Detections need to be very accurate and reliable because cyberbullying can even worsen into serious conditions that affect both mental and physical health.

1.9 Research questions

- What are the existing technology frameworks/algorithms/network architectures used in the text classification for cyberbullying detection?
- What are the most suitable deep learning approaches to detect cyberbullying and categorization of negative comments?
- What are the techniques that can be used to expand the dataset?
- What data augmentation techniques can be used to enrich datasets for model training and testing?
- What are the most suitable techniques that should be used for Optical Character Recognition?

1.10 Research Aim

This research project aims to use the NLP or DL approach to identify cyberbullying in Sinhala language statements and to design, develop and evaluate a system that may achieve a high accuracy rate.

In further development, the proposed system classifies bullying statements (bullied based on appearance, status, education level, relationship status, etc.) with the use of the NLP approach. The approach will be selected after reviewing previous literature in the field to

identify the most powerful and accurate approach. Due to the lack of research and bullying in social media platforms is gradually getting high in Sri Lanka, the proposed system is targeted at social media users in Sri Lanka. The system is able to identify bullying comments in Sinhala and help users by providing solutions to those who have already been bullied.

1.11 Research Objective

Research Objectives	Explanation	Learning Outcome
Problem Identification	● Find a problem in the real world that can be solved with the current technologies. ● Understand the current state of the art in cyberbullying detection measures and their efficacy. ● Determine the parameters of the proposed solution.	LO1
Literature Review	● Determine the research gaps and limitations in the current body of knowledge for the selected domain. ● Understand about the technologies and algorithms utilized in earlier studies, as well as how well they performed and were accurate. ● Make a concept map and Identify top-performing NLP algorithms and techniques used to address this research challenge.	LO1 LO4
Data Gathering and Analysis	● Making a new, significant Sinhala multi-label dataset with instances of various types of online abuse and cyberbullying that have been annotated as well as a sizable amount of offensive content from trusted and valid sources. ● Gather the opinions of industry and domain experts on the suggested system. ● Interaction with stakeholders, including young people, parents, medical professionals, and law enforcement, to understand their perspectives on cyberbullying and its	LO2 LO3 LO6

	prevention.	
Research Design	• Create a high-level architecture based on the proposed research. • Design diagrams derived from the planned implementation strategy and the defined requirements.	LO2 LO5
Implementation	• Implement a model with the use of suitable algorithms for high accuracy and performance. • Deliver the design prototype as a web application that is free to use for the public as a tool to prevent cyberbullying.	LO3 LO5 LO7
Testing and Evaluation	• Utilize various testing methods to evaluate the model that was implemented. Including unit testing • After testing, evaluate the model's and the prototype's performance in terms of accuracy. • Evaluate the system's performance in terms of responsiveness, scalability, and acceptability by conducting assessment studies. • Showcase the prototype to domain and industry professionals to assess the research effort.	LO8

Table 2: Research Objectives

1.12 Project Scope

Considering the objectives of the research and analyzing the research which are already done, the project scope can be defined as follows. The purpose of this research is to find a most suitable algorithm to detect cyberbullying statements from given text comments and text written images and categorize the negative comments.

1.11.1 In-scope

- Identifying given text as cyberbullying or not-cyberbullying.
- Reading the texts in given images and identifying the text as cyberbullying or

not-cyberbullying.

- Classification of the negative comments according to the type.
- Web application for the cyberbullying detection system.
- Recommending solutions to victims.

1.11.2 Out-scope

- Identifying given videos as cyberbullying or not-cyberbullying.
- Analysis of negative comments from other languages such as English, Tamil, Hindi, French and Russian etc.

1.11.3 Features of the Prototype

The proposed prototype is made up of features to address the identified research gap and enhance the prototype's user experience. The primary planned prototype features are graphically shown in Figure 3.

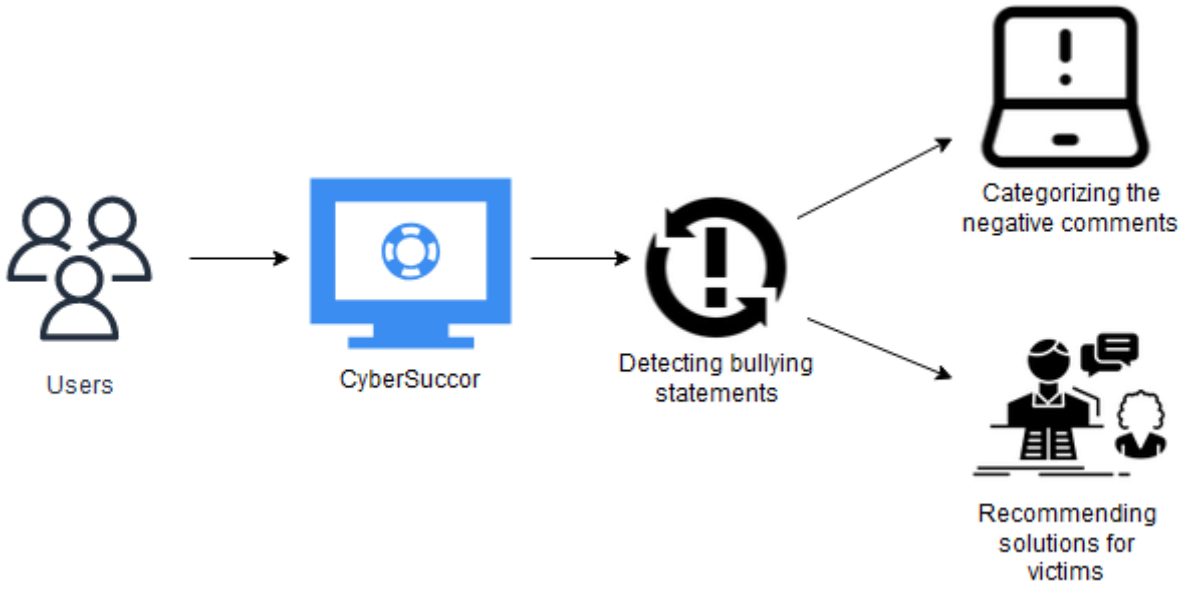

Figure 3: Prototype Feature Diagram

2 METHODOLOGY

For the research project to be completed successfully and in line with the project's goals, it must adhere to a few established procedures and deadlines. Additionally, the project's risks and mitigations are mostly mentioned here, along with the deadlines for deliverables. The research methodology, development methodology, and project management methodology discussed in this chapter all adhere to this methodology. This system is intended to be built and developed, thus in order to accomplish the goal, the author must take into account the requirements, adopt an approach that is specific to this project, and choose the finest techniques that would help to successfully finish the research project.

2.1 Research methodology

Research Philosophy	Pragmatism was chosen as the research philosophy since the suggested research is mutually exclusive between quantitative and qualitative methodologies, and the research questions identified may be answered by employing both quantitative and qualitative research strategies. A practical solution will also be helped by the study by highlighting the many approaches that may be used to solve the issue, which is similar to the pragmatism research ethos.
Research Approach	Hybrid is selected as the research strategy since the research has the features of both quantitative and qualitative research approaches.
Research Strategy	Questionnaires, interviews and surveys may all be used to acquire information for research topics, and because the methodology is hybrid any quantitative and qualitative technique can be applied.
Research Choice	As the research paradigm is pragmatism and it demonstrates both quantitative and qualitative research approaches, the mixed-method was chosen as the research option.
Time zone	The time zone will be cross-sectional since the study will acquire the necessary data for the model's training and testing throughout one project time frame.

Table 4: Research methodology

2.2 Development methodology

Following an appropriately defined methodology enables a project to produce better estimates, deliver solid systems, develop a clear picture of the task ahead, and spot difficulties earlier, giving plenty of time to make improvements. Out of the waterfall, spiral, rapid, prototype, and agile methodologies, agile will be chosen for the proposed system development because of flexibility, better control, Complete visibility of the progress.

2.2.1 Design Paradigm

The proposed model's design strategy was decided upon using the Object-Oriented Analysis and Design (OOAD) technique. OOAD is preferred over other design paradigms like the Structured System Analysis and Design Method (SSADM) because it uses an incremental or iterative methodology to enhance and update the designs as well as because applying analysis techniques has a minimal risk. Additionally, this paradigm's effectiveness in comparison to other paradigms is demonstrated by its usage of diagrams like activity diagrams, use case diagrams, sequence diagrams for the requirement engineering process.

2.2.2 Evaluation methodology

A program, practice, intervention, or initiative can be evaluated to see how effectively it meets its objectives. Evaluation offers a methodical way to do this. For the purpose of evaluating the model performance, the benchmarking and confusion matrix methodologies were suggested. Confusion matrix is helpful since it is able to gauge the recall, precision, and accuracy of the model, whereas benchmarking can be used to determine data capacity, training speed, and reaction speed, which can then be used as a scale to assess the performance and accuracy of the model.

2.3 Project management methodology

Project management ensures a consistent attention across all projects and enables the replicate effective elements and learn from mistakes, leading to a process of continual progress. For this research project, PRINCE2 Agile was selected as the project management methodology. Flexibility, reactivity, and governance are features shared by both the agile methodology and the PRINCE2 techniques in PRINCE2 Agile. This technique is ideal for the research project because it places a strong emphasis on time management, consistently meeting deadlines, scaling requirements, and having the flexibility to respond favorably to shifting project requirements.

2.3.1 Resource Requirements

1. Hardware Requirements

Core i5 processor	To model training, processing, and testing duties as well as other high-processing-power development tasks.
8GB RAM	Modern software needs at least 8gb of RAM to run smoothly.
Disk space of 40GB	To keep all the papers, documents, datasets, models and other data associated to research

Table 5: Hardware Requirements

2. Software Requirements

Operating System – Windows 10 / Linux	An OS that can manage such computing capacity is necessary because NLP processes require a moderate amount of processing power.
Google Drive / OneDrive	To keep and manage the essential backups.
Mendeley / Zotero	To organize research articles and to include citations.
MS Word, MS Excel, MS PowerPoint (Google Docs, Sheets, Forms)	For the purposes of data collection, documentation, management, and presentation.
StarUML / Draw.io	To design and draw the diagrams related to this research project.
Git / Github / Bitbucket	To use as version control tools.
Figma / Adobe XD / Zeplin	To design the interfaces of the web application.
PyCharm / Jupiter Notebook / Vs code	To use as IDEs for implementation.
Keras / Tensorflow / Pytorch	Libraries which are needed for model implementation.

Table 6: Software Requirements

3. Skills Requirements

Domain knowledge	It is necessary to have a fundamental understanding of the research project's non-technical domain in order to identify the problem.
Documentation Skills	Standards for documentation must be taken into consideration since research must be recorded and a final thesis must be submitted that includes information about prior work on the subject of studying.
Project management Skills	In order to use the project management techniques effectively, knowledge of such methodologies is necessary. Therefore, standard project management and development techniques are used in project management.
NLP skills	The suggested study is based on natural language processing, thus having strong natural language processing knowledge is necessary to implement a model that fills in the research gaps with high accuracy.
Web application development and UX / UX skills	In order to implement the user-friendly and responsive web application to detect cyberbullying, web application development skills and UI / UX skills are needed.

Table 7: Skills Requirements

4. Data Requirements

Reliable and verified dataset for train, test and evaluate the modal.	To train the cyberbullying detection model, test and evaluate the implemented model and calculate accuracy with error ratio.

Table 8: Data Requirements

2.3.2 Gantt Chart

Figure 4: Gantt chart

2.3.3 Deliverables

Deliverable	Date
Project Proposal	03rd Nov 2022
Ethics Form	03rd Nov 2022
Literature Review	27th Oct 2022
Software Requirement Specification	24th Nov 2022
Project Specifications Design and Prototype (PSDP)	30rd Jan 2023
Prototype Demo	4th Mar 2023
Test and Evaluation report	23rd Mar 2023
Final Thesis	27th Apr 2023

Table 9: Deliverables

2.3.4 Risk Management

Risk Item	Severity	Frequency	Mitigation Plan
Project scope and requirements change.	5	5	Arrange the changes in order of importance and use project management and development approaches to address the issue.
Unable to meet the deadline because of other academic work.	5	4	Follow the Gantt chart's project tasks while scheduling time for other academic tasks.
Failing to finish the work by the due date.	5	4	In accordance with the Gantt chart, follow and finish the daily tasks allocated to the project.

Not able to gather necessary datasets.	5	3	In the early stages of the research study, begin gathering the necessary datasets.
insufficient understanding in the domain	4	3	Follow tutorial and get help from domain experts.

Table 10: Risk Management Plan

REFERENCES

Akhter, A., Uzzal, K. A., & A. Polash, M. M. (2019). Cyber Bullying Detection and Classification using Multinomial Naïve Bayes and Fuzzy Logic. *International Journal of Mathematical Sciences and Computing*, *5*(4), 1–12. https://doi.org/10.5815/ijmsc.2019.04.01

Sri Lanka News - Newsfirst. (2021). *Significant rise in Cyberbullying cases in Sri Lanka: SLCERT*. [online] Available at: https://www.newsfirst.lk/2021/07/30/significant-rise-in-cyberbullying-cases-in-sri-lanka-slcert/.

Behzadi, M., Harris, I. G., & Derakhshan, A. (2021). Rapid Cyber-bullying detection method using Compact BERT Models. *Proceedings - 2021 IEEE 15th International Conference on Semantic Computing, ICSC 2021*, 199–202. https://doi.org/10.1109/ICSC50631.2021.00042.

Ariyadasa, A. (n.d.). *Legal Consultant, Senior Lecturer in Commercial Law, International Commercial Arbitration, Cyber Law, International Human Resources Management, Company Law, Criminal Law*. LLM in Commercial Law. https://dictionary.cambridge.org/dictionary/english/cyber.

Li, M., & Tagami, A. (2014). A study of contact network generation for cyber-bullying detection. *Proceedings - 2014 IEEE 28th International Conference on Advanced Information Networking and Applications Workshops, IEEE WAINA 2014*, 431–436. https://doi.org/10.1109/WAINA.2014.70.

NaliniPriya G, & Asswini M. (2015). *A DYNAMIC COGNITIVE SYSTEM FOR AUTOMATIC DETECTION AND PREVENTION OF CYBER-BULLYING ATTACKS*. *10*(10). www.arpnjournals.com

Srinandhini, B., & Sheeba, J. I. (2015). Online social network bullying detection using intelligence techniques. *Procedia Computer Science*, *45*(C), 485–492. https://doi.org/10.1016/j.procs.2015.03.085

Jikriya, S. S. (2021). Cyber bullying Detection in social media using Supervised ML & NLP Techniques. *International Journal for Research in Applied Science and Engineering Technology*, *9*(VI), 2259–2264. https://doi.org/10.22214/ijraset.2021.35483

NaliniPriya G, & Asswini M. (2015). *A DYNAMIC COGNITIVE SYSTEM FOR AUTOMATIC DETECTION AND PREVENTION OF CYBER-BULLYING ATTACKS*. *10*(10). www.arpnjournals.com

Li, M., & Tagami, A. (2014). A study of contact network generation for cyber-bullying detection. *Proceedings - 2014 IEEE 28th International Conference on Advanced Information Networking and Applications Workshops, IEEE WAINA 2014*, 431–436. https://doi.org/10.1109/WAINA.2014.70

Khurana, D., Koli, A., Khatter, K., & Singh, S. (2022). Natural language processing: state of the art, current trends and challenges. *Multimedia Tools and Applications*. https://doi.org/10.1007/s11042-022-13428-4

Hettiarachchi, N. W. (2020). *Detecting Hate Speech in Social Media Articles in Romanized Sinhala*.