

Informatics Institute of Technology
In Collaboration With
University of Westminster, UK



University of Westminster, Coat of Arms

“Safe-Social”- A Cyberbullying Detection System on Social Media using Hybrid algorithm for Singlish Language

A dissertation by
Induwara Mapalagama
W1790257 / 2019779

Supervisor: Guhanathan Poravi

Date: 10th of May 2023

Key Words: natural language processing, preventing cyberbullying, mixed code

This project proposal is submitted in partial fulfillment of the requirements for the
BEng(Hons) Software Engineering degree at the
University of Westminster.

ABSTRACT

Cyberbullying has become an important concern in today's culture due to the increasing usage of social media. With recent events such as Covid-19 and the economic crisis taking place, the number of social media users in Sri Lanka has grown to the top on a vast scale. Despite investing millions of dollars to solve this issue, prominent social media sites such as Facebook, Instagram, and Twitter remained struggling to make a significant shift on their platforms. Social media platforms have grown in popularity as a tool of communication for languages with low resources. However, the detection of cyberbullying in various transliterated forms of such languages is limited. Due to a lack of expertise of such languages, it is difficult for platforms to respond quickly during riots on social media platforms. This study utilizes an automated technique to detect cyberbullying in Romanized Sinhala, a low-resource language which will help the community to overcome this major concern.

With this in place, cyberbullying in the Singlish (Sinhala & English) mixed-code language has been highlighted as a big concern in the Sri Lankan community. In respond to this problem, a cyberbullying detection system for Singlish was proposed, which utilized a hybrid algorithm combining Naive Bayes and MLP Classifier. The suggested technology aims to identify cyberbullying automatically on social media platforms, hence lessening the detrimental impact of cyberbullying on Sri Lankan society. The goal of this study is to develop a critical tool for identifying cyberbullying on social media using the proposed hybrid algorithm for Singlish language.

This study provides a thorough examination and critique of past studies and technologies in order to support the suggested solution. Furthermore, the author goes into detail about the testing and evaluation procedures that were used. Based on the results of the assessments, the author considers that the proposed system is a very effective solution to the current problem.

Keywords: Cyberbullying Detection, Hybrid Algorithms, Machine Learning, Neural Network, Social Media, Data Science, Ensemble, Mixed-Code Language, Singlish