INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER

A Machine Learning Approach for Depression Detection on Social Media Content Written in Sinhala Unicode and Sinhala – English Code-Mixed Text

A Dissertation by

Bagya Gunawardana (W1810492/20200330)

Supervised by

Mr. Prasan Yapa

September 2023

Submitted in partial fulfilment of the requirements for the MSc in Advanced Software Engineering degree at the University of Westminster.

ABSTRACT

This research study offers a new machine learning method for detecting depression in Sinhala-language social media content. Even though one previous study had explored on the topic that approach is not realistically applicable to the real setting of Sri Lankan social media because of the numerous input patterns employed, such as Sinhala Unicode and Sinhala-English mixed code. Therefore this research study proposes a novel approach to address the aforementioned gaps and introduce a model for detecting depression on content written using both Sinhala unicode and code-mixed content while improving the accuracy and performance. This machine learning model will enable development of new tools, approaches to detect depression based on Sinhala content and help health sector to conduct easy diagnosis of the disease, patient monitoring and duly treatments.