Informatics Institute of Technology

In Collaboration with

University Of Westminster

Deep Learning Approach for Fake News Detection in Sinhala English Mixed Code Language in Social-Media

A Dissertation by

Ms Adikaram Duvini

Supervised by

Mr Lakna Gammedda

Submitted in partial fulfilment of the requirements for the BSc (Hons) Computer Science degree at the University of Westminster, UK

May 2023

ABSTRACT

In recent years, fake news has proliferated substantially more frequently and has disseminated rapidly through online social networks. These platforms' explosive expansion has created an ideal environment for the spread of false information which is motivated by both political and commercial interests. False narratives are now widespread in the world of technology and posing a serious problem for society as a whole. False narratives have become commonplace in the digital realm and presenting a grave challenge to society at large. Online viewers are easily swayed by these false news reports due to the use of deceiving language which has significant effects offline. Detecting and identifying fake news is essential for improving the veracity of information in online social networks.

The primary goal of this research endeavor is to develop a system that can effectively identify fake news written in Romanized Sinhala by utilizing a bi-directional LSTM model. By leveraging deep learning techniques and natural language processing, the project seeks to enhance the precision and effectiveness of detecting deceptive information across various online platforms. While bi-directional LSTM models have demonstrated promising outcomes in various fields, this study focuses on their application specifically to Romanized Sinhala fake news detection. By examining the potential advantages and optimizations of machine learning techniques in this area, the project seeks to contribute to the advancement of reliable information dissemination and combat the proliferation of false information.

To train the algorithm for false news identification in this study, a dataset with 10,269 was used. The model's evaluation showed a 0.57 accuracy. Although these results show some progress, more work has to be done to increase the model's performance. In order to achieve improved accuracy and more reliable identification of false information, this highlights the necessity for ongoing study and development in the field of fake news detection in Romanized Sinhala.

Keywords:

Subject Descriptors: