Sentiment Based Approach to Analyze Fake News Related to COVID-19 in Sri Lanka

K.I. Somasinghe
(Reg. No.: MS19816500)
M.Sc. in Information Systems

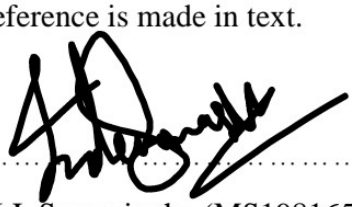
Supervisor: Mr. Prasanna S. Haddela

September 2021

Department of Computer Systems Engineering
Faculty of Computing
Sri Lanka Institute of Information Technology

Declaration

I certify that this report does not contain any material previously submitted for a degree or diploma at any university without acknowledgement, and that it does not contain any material previously published or written by another person to the best of my knowledge and belief, except where due reference is made in text.



19.09.2021

K.I. Somasinghe (MS19816500) Date

22.10.2021

Supervisor: Mr. Prasanna S. Haddela Date

Abstract

Generation and spread of fake news have drastically increased with the growth of technology and advancement of online media platforms. Today, rather than using traditional resources to get information, most people rely on the internet and it has become a part of every individual's life since this is a one of the simplest methods to acquire information on almost everything. This internet based media has become a source of sharing news and these sources are used by companies, political parties as well as social influencers etc.

Fake news changes the perception of the viewers and diverts them from the reality. By analyzing fake news in Sinhala Language related to COVID-19 which is a disastrous situation to not only Sri Lanka but the whole world it would be a great advantage to notify people regarding fake news and the resources they use to spread fake news and reduce unethical sharing of news, to protect the authenticity of news that reaches people and the authenticity of the journalism field.

This research presents an approach to analyze the effect of polarity in the sentiment of the news data and analyze how it affects towards fake news in Sinhala language using textual data. The proposed model uses natural language processing techniques such as sentiment analysis and machine learning algorithms such as Logistic Regression, Support Vector Machine, Naïve Bayes.

Acknowledgements

First of all, I would like to thank my supervisor, Mr. Prasanna S. Haddela, for providing an opportunity and continuous guidance for me, and for the patience, encouragement and expertise of my research and review of the thesis. The encouragement he gave helped me to constantly plan and write the thesis.

I want to thank Dr. Anuradha Jayakody, the coordinator of the IS program, for his guidance and the support given me during the hardest times.

Last but not least, I want to thank my family and friends for their support throughout my life.

Table of Contents

List of Figures

List of Tables