Enhancing Sinhala Hate Speech Detection in Online Platforms

W. M. R. D. Silva
(Reg. No.: MS23006966)

A THESIS
SUBMITTED TO
SRI LANKA INSTITUTE OF INFORMATION TECHNOLOGY
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE IN INFORMATION TECHNOLOGY
(ENTERPRISE APPLICATIONS DEVELOPMENT)

December 2024

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Dr. Dilshan De. Silva

Approved for MSc. Research Project:

MSc. Programme Co-ordinator, SLIIT

Approved for MSc:

Head of Graduate Studies, FoC, SLIIT

DECLARATION

This is to certify that the work is entirely my own and not of any other person, unless explicitly acknowledged (including citation of published and unpublished sources). The work has not previously been submitted in any form to the Sri Lanka Institute of Information Technology or to any other institution for assessment for any other purpose.

Sign: 

W. M. R. D. Silva

Date: 15/01/2025

ABSTRACT

Enhancing Sinhala Hate Speech Detection in Online Platforms

W. M. R. D. Silva
MSc. in Information Technology (EAD)
Supervisor: Dr. Dilshan De Silva
December 2024

The rise of deep learning methodologies has indeed revolutionized text analysis, enabling more sophisticated and nuanced understanding of language dynamics. With the proliferation of social media platforms, these advancements have been particularly crucial in navigating the vast amounts of data generated by online interactions. However, amidst the benefits of this digital age, the prevalence of hate speech has emerged as a pressing concern, transcending linguistic and cultural boundaries. In the context of Sinhala, a language rich in nuances and deeply intertwined with cultural complexities, the challenges in detecting and mitigating hate speech are further compounded. Language is not merely a tool for communication but also a reflection of societal norms, values, and power structures. In the Sinhala-speaking context, historical legacies, religious beliefs, and political tensions intertwine to shape discourse in multifaceted ways. Consequently, any hate speech detection mechanism must navigate these intricate layers of meaning, accounting for cultural sensitivities and contextual nuances to ensure accurate identification of harmful content. The integration of deep learning techniques and advanced semantic analysis holds promise in enhancing hate speech detection in Sinhala. By leveraging the power of neural networks to discern patterns and contexts within textual data, such mechanisms can offer a more nuanced understanding of language dynamics. Moreover, the evaluation of these tools on real-world social media data not only validates their effectiveness but also provides insights into the evolving nature of online discourse. Ultimately, addressing hate speech in Sinhala and similar low-resource languages requires a multifaceted approach that combines technological innovation with cultural sensitivity and community engagement to foster safer and more inclusive online spaces.

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to Dr. Dilshan De Silva and Mr. Lasitha Petthawadu for their invaluable support and guidance throughout the course of this research. Dr. De Silva's expert advice, insightful feedback, and encouragement were instrumental in shaping the direction and success of this study. His deep understanding of the subject matter and willingness to share his knowledge greatly enhanced the quality of the research.

Without their dedicated support and mentorship, this research would not have been possible. Their contributions have been invaluable, and I am deeply appreciative of their efforts to help me achieve my research goals.

TABLE OF CONTENTS

List of Figures

List of Tables