

# Automated Spelling Checker And Grammatical Error Detection And Correction Model for Sinhala Language

Mithma Goonawardena  
Department of Computer Science and  
Software Engineering  
Sri Lanka Institute of Information  
Technology, Sri Lanka  
it19135830@my.sliit.lk

Ashini Kulatunga\*  
Department of Computer Science and  
Software Engineering  
Sri Lanka Institute of Information  
Technology, Sri Lanka  
it19136134@my.sliit.lk

Raveena Wickramasinghe  
Department of Computer Science and  
Software Engineering  
Sri Lanka Institute of Information  
Technology, Sri Lanka  
it19146652@my.sliit.lk

Thisuraka Weerasekara  
Department of Computer Science and  
Software Engineering  
Sri Lanka Institute of Information  
Technology, Sri Lanka  
it19190808@my.sliit.lk

Hansi De Silva  
Department of Computer Science and  
Software Engineering  
Sri Lanka Institute of Information  
Technology, Sri Lanka  
hansi.d@ sliit.lk

Samantha Thelijjagoda  
SLIIT Business School  
Sri Lanka Institute of Information  
Technology, Sri Lanka  
samantha.t@sliit.lk

**Abstract** - Sinhala is a native language spoken by the Sinhalese people, the largest ethnic group in Sri Lanka. It is a morphologically rich language, which is a derivation of Pali and Sanskrit. The Sinhala language creates a diglossia situation, as the language's written form differs from its spoken form. With this difference, the written form requires more complex rules to be followed when in use. Manually proofreading the content of Sinhala material takes up much time and labor, and it can be a tedious task. Hence, a system is necessary which can be used by different industries such as journalism and even students. At present, there are a handful of systems and research that have automated Sinhala spelling analysis and grammar analysis. In addition, the existing systems are mainly focused on either spelling analysis or grammar analysis. However, the proposed system will cover both aspects and improve upon existing work by either optimizing or re-building the process to provide accurate outputs. The proposed system consists of a suffix list built for verbs and subjects, which helps the system stand out from the current proposed solutions. This research intends to implement a service for spell checking and grammar correctness of formal context in Sinhala. The research follows a rule-based approach with some components adopting a hybrid approach. As per the literature survey, many papers were analyzed, related to different aspects of the proposed system and complete systems. The proposed system would be able to overcome most barriers faced by previous papers whilst it takes a fresh take on providing a solution.

**Keywords** - automated Sinhala spelling analysis, grammar analysis, part of speech tagging, subject characteristics analysis