



INFORMATICS  
INSTITUTE OF  
TECHNOLOGY

UNIVERSITY OF  
WESTMINSTER<sup>®</sup>

**SINHALA GRAMMAR EVALUATION THROUGH  
NATURAL LANGUAGE PROCESSING APPROACHES**

A dissertation by  
**Kokila Widyaratna (2015307/w1608492)**

Supervised by  
**Mr. Asanjay Fernando**

**Submitted in partial fulfilment of the requirements for the  
BEng (Hons) Software Engineering degree  
Department of Computing**

**May 2019**

**© The copyright for this project and all its associated products resides  
with Informatics Institute of Technology**

## Abstract

An automated grammar checking tool for Sinhala language is a shortage so far. Efforts taken so far have been mostly considered rule-based approaches map Sinhala grammar rules. However, major limitations identified in exiting literature are, rule-based approaches being unable to understand the context of the sentences and extensive work has to be done to scale the system introducing new grammatical features. This research will focus on tackling the free word order nature and complex morphology of Sinhala language in order to implement a more reliable grammar checker.

The solution of the project is utilising a neural machine translation model to learn the probable grammar errors made by day today people and the respective corrections for those errors. The system is composed of a sequence to sequence model with attention module which is widely used architecture in machine translation applications. Grammar correcting model was trained with generated grammatical errors related to verbs of Sinhala language. Within the considered scope, the performance of the system is up to the expectations and highly promising. Also, the system is highly scalable with the data driven approach taken.

### **Key Words:**

Sinhala Grammatical Error Correction

Machine Translation

Natural Language Processing