



**INFORMATICS
INSTITUTE OF
TECHNOLOGY**

INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER

Singlish-Translator:

An approach to convert Singlish text to English text

A dissertation by

Mohamed Isthikar Imsar Ahamed

(2018262 / W1715778)

Supervised by

Ms. Krishnakripa Jayakumar

Submitted in partial fulfilment of the requirements for the BEng Software Engineering degree
at the University of Westminster.

July 2022

Singlish-Translator: An approach to convert Singlish text to English text.

Abstract

In this technologically advanced era, communication is a core aspect of maintaining mutual relationships among business and personal ones. Today Machine Learning (ML) and Artificial Intelligence (AI) have actively involved in bridging the gaps in language barriers as humans have to interact with various people worldwide.

In Sri Lanka, the primary language used is Sinhala even though English is the universally accepted language. In that case as all of the main e-commerce application and the social media application use English. So, there is a problem in communication within and out of country. In order to overcome this problem people have come up with a way of writing the Sinhala words in English by directly converting the Sinhala pronunciation. This being called as Singlish is being used widely among the users in the Sri Lanka.

In this research a, ML is used in converting the Singlish text to English text so that local users who do not understand Singlish or foreign users who try to communicate with local users can use this option to convert the languages vice versa without any miscommunication. In this approach a LSTM-based model is used together with more than one hundred thousand datasets utilized to train the model. As per the progress a 40% of accuracy has been obtained via a dataset trained using ten thousand data.

Keywords: Singlish, Sinhala, Natural Language Processing, LSTM, Machine Translation