MACHINE TRANSLATION SYSTEM FOR SINGLISH TO ENGLISH

Mohomed Haffaz

A dissertation submitted in partial fulfilment of the requirement for

Bachelor of Engineering (Honours) degree in Software Engineering

Department of Computing

Informatics Institute of Technology, Sri Lanka

in collaboration with

University of Westminster, UK

2020

3. Abstract

Technological advancement in the communication field has laid the groundwork to reach audiences that are a thousand miles away with relative ease through the use of effective translation. With the growing popularity of social media during the past few decades, the amount of unfiltered data used for various purposes is rising exponentially. Switching back and forth between multiple languages also known as code-mixing is a frequently encountered phenomena are users in countries with a significant number of multilingual/ bilingual speakers. In Sri Lanka this phenomenon takes the form of Singlish, where Sinhala serves as the host. Singlish is often left untouched by non-Singlish due to lack of proper translation mechanism.

This research presents a Machine translation system that will enable the ability of translating Singlish into English. The system is able to derive the intended word from varying spelling variations of a Singlish word caused due to lack of a standardization. The outcomes of the research have been satisfied by industry and domain experts. The target audience of this research is general internet users.

Keywords: Natural Language Processing, Machine Translation, Transliteration, Code-mixing