



INFORMATICS
INSTITUTE OF
TECHNOLOGY

UNIVERSITY OF
WESTMINSTER^ෆ

Informatics Institute of Technology

In Collaboration With

UNIVERSITY OF WESTMINSTER, UK

Reviewly

Applying Sentiment Analysis for Romanized Sinhala-English Mixed Reviews

A Dissertation by

Mr. Sashminda Iranga Withanage

w1790117 / 2019586

Supervised by

Mr. Deshan Sumanathilaka

April 2023

Submitted in partial fulfilment of the requirements for the
BSc (Hons) Software Engineering degree at the University of Westminster.

Abstract

The e-commerce world is improving at a rapid phase. Sellers and buyers on major platforms like Amazon, eBay and AliExpress use customer reviews to make decisions. NLP and its various techniques are being used to analyse reviews. Although review analysis systems are available to analyse English reviews, there is no distinct system to handle Sinhala-English code mixed reviews. A dedicated system which can evaluate and output a sentiment score with a summary would be helpful for monolingual users when making decisions based on user reviews.

The proposed system Reviewly is capable of analysing and providing the user with a sentiment score and a summary of a particular code mixed review. When a review is given, the proposed model will transliterate into Sinhala first and then translate it into English. After translating the review will be analysed using a Sentiment analysis model. This type of approach would outperform a rule-based traditional approach which is common in romanized transliteration.

A product that is able to analyse romanized Sinhala hasn't been attempted before because Sinhala can be identified as a low-resource language. The proposed solution was developed using current trends in sentiment analysis and natural language processing. The BERT model was fine-tuned and used for the final minimum viable product. The author was able to achieve over 90 percent accuracy with the used dataset.

Keywords: Transliteration, Translation, Romanized Sinhala, Sentiment analysis

Subject descriptors:

- Computing methodologies → Artificial intelligence → Natural language processing → Machine translation