Enhancing Conversational AI Model Performance and Explainability for Sinhala-English Bilingual Speakers

Ishara Dissanayake[1], Shamikh Hameed[1], Akalanka Sakalasooriya[1], Dinushi Jayasinghe[1], Lakmini Abeywardhana[1], and Dinuka Wijendra[1]

[1]Department of Information Technology, Sri Lanka Institute of Information Technology, Sri Lanka.
isharadissanayake@icloud.com, shamikhhameed@gmail.com, himesha@outlook.com, jayasinghedt@gmail.com, lakmini.d@sliit.lk, dinuka.w@sliit.lk

Abstract

Natural language processing has become an essential part of modern conversational AIs. However, applying natural language processing to low-resource languages is challenging due to their lack of digital presence. Sinhala is the native language of approximately nineteen million people in Sri Lanka and is one of many low-resource languages. Moreover, the increase in using code-switching: alternating two or more languages within the same conversation, and code-mixing: the practice of representing words of a language using characters of another language, has become another major issue when processing natural languages. Apart from natural language processing, the explainability of opaque machine learning models utilized in conversational AIs has become another prominent concern. None of the existing modern conversational AI platforms supports explainability out of the box and relies on just a performance score such as accuracy or f1-score. This paper proposes a no-code conversational AI platform with a series of built-in novel natural language processing, model evaluation, and explainability tools to tackle the problems of processing Sinhala-English code-switching and code-mixing natural language data and model evaluation in modern conversational AI platforms.

Keywords

Conversational Ai, Natural Language Processing, Code-Switching, Code-Mixing, Explainable Ai