

UNSEEN MULTI-LABEL INTENT DETECTION

Akalahandi Malitha Wijewardhane

A dissertation submitted in partial fulfilment of the requirement for
Bachelor of Science (Honors) degree in Software Engineering

Department of Computing
Informatics Institute of Technology, Sri Lanka
in collaboration with
University of Westminster, UK

2021

Abstract

With the recent advancement in the field of Natural Language Processing, conversational agents are becoming important components in many applications such as mobile apps, companion bots and virtual assistants. These conversational agents are mainly developed with the combination of language understanding, dialogue management and dialogue generation modules. Intent detection is a crucial sub task carried out by the language understanding module with the objective of to understand user's intentions. By analyzing the literature on conversational domain, it is evident that Conversational Agents faces unknown/unseen user intents and furthermore these user utterances could contain more than one intention. This dissertation dissects forementioned dilemma and proposes an approach to detect unseen intents in scenarios where utterances contain many intentions.

Based on the analysis on the literature, relevant existing works has been identified to support the research. After exploring many approaches, multi-label classification along with local-outlier detection was selected to solve the research problem. Using BiLSTM model to address the multi-label classification and LOF algorithm to tackle outlier detection, this research proposes an intent detection module which is capable of detecting unseen intent without letting user utterance be misclassified.

The proposed model was tested and evaluated with **MixATIS** and **MixSNIPS** which are standard benchmarking datasets in intent detection domain. Our model was able to provide solid performance on both datasets scoring better f1-scores on MixATIS and MixSNIPS datasets under 25% unseen intent concentration. Furthermore, the model was evaluated against baseline models and benchmarking models and has shown state-of-the art performance in MixATIS dataset and competitively equal performance in MixSNIPS dataset. This model was successfully implemented in intent detection module in the domain of flight booking. Key importance of the contribution is that this module can be integrated in any conversation AI and repurpose it in any domain.

Keywords: Intent Detection; Multi-Label Classification; Conversational AI; Local Outlier Detection; Unseen Multi-Intent Detection.