STACK TAGGER
AUTOMATED QUESTION TAGGING FOR STACK OVERFLOW

Kurukulasuriya Yukthika Nuwantha Fernando

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

Abstract

Tagging is being carried out everywhere from tagging being used by the websites to help better identify the information in defining math's equations. Online communities have become the next frontier of the problem solvers ranging from life skills to all fields of study. StackOverflow is one such online community for developers. This not only has a userbase that provide answers to questions. It has also transformed to the ideal platform for Developers as it also includes job posting. This has a userbase over 12 million users and increasing over a million annually. The topics covered are categorized into appropriate tags to represent the questions. Currently the tags are entered manually apart from the code block analysis. Due to the immense number of tags present, it is useful to adapt an auto tagging system. The current system analyses the code blocks added to the body of the question and suggests suitable tags to classify the questions. In most cases this can classify the problem to suggest a few tags. But isn't the title the most important area of the question. The proposed solution is to provide a multi label classification for the question titles with the use of Bert. This will help the responders of the question to easily identify a question they are able to provide the help, which the poster is looking for from the community.

Keywords: - **Stack Overflow**, **Multi Labeled Classification**, **Tagging** & **BERT**