MSc Project Report

Classification of Tender Notices through Deep Learning Concepts

Dissertation by:

Damian Perrin Vanniasinkam (1714925)

Supervised by:

Suhothayan Sriskandarajah

August 2019

A report submitted as part of the requirements for the degree of MSc Big Data Analytics at Robert Gordon University, Aberdeen, Scotland

Abstract

With the continuous growth of the tender industry in Sri Lanka, the visibility of tenders to the suppliers is a cause for concern. Tender alert service companies have minimised this complexity but due to the cumbersome task of sourcing papers daily and categorising tenders, they are not always efficient or accurate.

Keeping in mind the escalated growth of the machine learning industry, this thesis takes the tender alert service industry in to consideration and looks at multi-label classification methods to automatically classify tenders. The goal is to identify what categories a tender belongs to by just considering its heading. A dataset with 36 tender categories was used for this purpose.

A thorough literature review session was conducted to identify the best methods to classify the tenders and it was identified that converting the multi-label problem to binary problems was the best solution to ensure high accuracy. As such, the data was processed through 3 classifiers, namely, Linear SVM, Random Forest and Neural Networks, to identify suitable classifiers for each category.

Apart from a single category that didn't have many supporting records, all other categories were able to produce classifiers with a f1-score of above 85%. Although the hamming loss could have been better, most of them were below 10%. Based on this, we can conclude that a satisfactory classification model for tenders was achieved.