

ABSTRACT

Al-Qur'an is an eternal miracle form for characterizing linguistic, correct, and validating perfection of the latest scientific findings. Muslim must understand and practice the command and stay away from the prohibitions mentioned in the Qur'an. Each verse of the Qur'an has a different meaning, and can have more than one topic that can be studied. To understanding the verses of the Qur'an needs to be classified against the verses of the Qur'an. In this Final Project research, a classification model was developed which was able to identify the classes in each verse of Al-Qur'an by multi-label.

In the previous study, the probabilistic approach yielded the best performance compared to other approaches. The probabilistic approach is considered to be common in completing text classification. Thus, in this study classification model was built using the method of Tree Augmented Naïve Bayes (TAN). In improving system performance, used Mutual Information (MI) to select the dependent variables used. Performance calculations that are considered appropriate for the multi-label classification are hamming loss.

The results of the tests have shown that the classification model built with MI resulted in the best performance with hamming loss value of 0.1121, while the classification model built without MI resulted in hamming loss value of 0.1208.

Keywords: quranic verses, *multilabel classification, mutual information, tree augmented naïve bayes, hamming loss*