Abstract

Hadith is the second source of law for Muslims after the Qur'an which comes from various forms of the words, actions and stipulations of the Prophet Muhammad or referred to as his sunnah. In order to make it easier for Muslims to apply the teachings of the hadiths, a classification system is needed that can categorize a hadith into a class or a combination of two of the three classes which called a multi-label classification. In building a text classification system, there are various classification techniques, one of which is k-Nearest Neighbor (KNN). KNN is a simple and effective classification method for text classification, but has a weakness in processing data with high vector dimensions so that the computation time is higher and the efficiency of text classification is very low. Mutual Information (MI) is used as a feature selection method to reduce vector dimensions because it has the ability to show how strong a feature is in making a correct prediction of a class. In this study Problem Transformation Method with the Binary Relevance (BR) approach is used so that the multi label classification process can be accomplished. The optimum results obtained in this study shows the value of hamming loss is 0.0886 or about 91.14% of data were correctly classified and computational time for 595 seconds by using MI as a feature selection, but without stemming.