

Abstract

Hadith is used as the source of Islamic law other than Qur'an, Ijma and Qiyas, hadith is the second of Islamic law after the Qur'an. This study attempted to build a system than can classify shahih hadith of Bukhari in Indonesian Translation. This topic was chosen to help Muslims who want to understand from each hadith is in the form of informations, prohibitions or suggestions. Support Vector Machine was chosen because it can perform classification by providing good performance for dataset with a large number of features. Latent Semantic Indexing as a feature selection method was chosen because it can reduce features by eliminating unimportant features (noise term). This study also using Bootstrap Aggregating (Bagging) method to improve accuracy. The accuracy results show that by using Latent Semantic Indexing and Bagging on Support Vector Machine classification is 84.67% of single label hadith data.

Keywords: hadith, support vector machine, feature selection, latent semantic indexing, bootstrap aggregating, text classification, single label