Abstract

Short messages or more common known as SMS (Short Message Service) is a service for exchanging messages among users. Because of the increasing number of SMS's user, some of the parties exploit to gain advantages for themselves by either spreading or sending junk SMSes, which is known as SMS spam. Hence, this final project is trying to classifyed the SMSes with spam (junk) or ham (legitimate) class. This classification is done by using collaborative naïve-bayes method which is oriented of recommendation from some users and content-based naïve bayes by analyzing the content of a SMS. Recommendation's data are obtained by spreading 300 SMSes to some users. Preprocessing is needed for content-based naïve bayes in order to get uniform content, have useful information, and to expedite the computation. Slang handling, stopword removal, and stemming are used for Preprocessing. SMSes are divided into training set and testing set according to 5-fold and 10-fold data's selection method. From this experiment, the average result for 5-fold is 97.12% and for 10-fold is 97.28% .

Keyword: *Classification, Collaborative, Content-based, Naïve bayes, Preprocessing.*