Abstract

SMS (Short Message Service) is still the primary choice as a communication media although nowadays mobile phone is growing with a variety of communication media messenger application. SMS is considered as one of the communication media which is simple because cheap, easy to use for various groups of mobile users, and can be documented. Along with the development of various other communication media, some carriers in some countries decrease the SMS rates to keep attracting the mobile phone users. However, this tariff reduction led to increased SMS spam, as used by some people as an alternative to advertisement even fraud. It becomes an important issue because it can annoy and harm to the user.

Naive Bayes is considered as one of the learning algorithm which is very effective and important for machine learning in information retrieval. Naive Bayes proved to have a good performance in the classification of text and SMS spam detection [2, 10] by showing high accuracy. With the collaboration algorithm that is able to determine frequent itemset well then able to produce a better accuracy rate [2]. Because not only considered each and every word as independent and mutually exclusive but also frequent words as a single, independent and mutually exclusive [2], so it increase probabilty value and lead system more accurate to classification. In this case, FP-growth is used for mining frequent pattern that has good performance and efficient because it does not require frequent candidate generation [4]. Results of using collaboration of Naive Bayes and FP-Growth showed the highest average accuracy is 98,506 % and 0,025% better than without using FP-Growth for dataset SMS Spam Collection v.1, and improve the precision value so the classicication results more accurate.

Keywords : SMS, *Naive Bayes, FP-Growth, frequent itemset, spam*