Abstract

Twitter is one of the most popular social media in the world. Almost everyone takes advantage of the tweets feature on this social media to exchange information with people around them. The feature is broadcast so that it can facilitate users to disseminate information to other users. Along with the growing needs of the community, today many people use the tweet feature for business needs such as advertising, product promotion, and others. Although the information in a single tweet is limited to 140 characters, usually those users who want to promote something on Twitter attach the source link to their tweets to provide further information about the contents of the promotional tweet. Of course the information is very useful for those other users who want the information, but otherwise this would be a disadvantage and can even be considered as spam if the content in the promotional tweet is not relevant to the link. By facing the problem, therefore in this final task will be created a system using Naive Bayes method and URL Analysis to classify a set of tweets that contain links into two categories of spam and non spam. Before entering the classification process, the system preprocessing the dataset used, it aims to improve the performance of the two methods used. In this study Naïve Bayes and URL Analysis yielded a fairly high accuracy value of 91% and 93%.

Keywords : *Twitter, Tweet, Spam, Naive Bayes, URL Analysis.*