

Abstract

Communications media that is currently commonly used by people to interact is a mobile phone (mobile phone). One of the most widely used services for messaging is Short Message Service (SMS). Sms was chosen because of its cheap tariff, but now sms is also used by pilk who is not responsible for disseminating unsolicited commercial advertisement (Sms Spam). For some people consider this trivial problem, but not a few also feel annoyed by the existence of this spam sms. In this final project the authors do an algorithmic analysis to detect sms spam using sms data that have been collected by the previous author. The methods used in this case are Vector Space Model and K-means clustering. Where Vector Space Model can recognize the pattern of sms by converting the word into the word weight and k-means clustering that serves as a method that classifies the weight of the word so that after done various test scenarios produced 60% accuracy means that both of these methods are able to detect and classify between sms spam and sms not spam (sms ham) but not too optimal due to the lack of sms as training data.

Keywords: *SMS, Spam Sms, Vector Space Model, K-Means Clustering, classification, Ham Sms.*