

Abstract

SMS is still one of the most important services in communication media. However due to its cheap cost, many people abuse function of SMS which is called SMS *spam*. This SMS *spam* is being used to promote and fraud. Not many users believe SMS directly to the content of the SMS. But there are also who believe in the content of the SMS.

Therefore, the application was built to filter the contents of decent SMS received by users. In this thesis research on SMS *Spam* Filtering. Which aims to classify into two classes, namely SMS *spam* and *ham*. To Conduct this research, the writer will pas through some processes. The first process is *preprocessing* which generally aims to tidy up the content of the SMS. In *preprocessing* there is *Case Folding* process which is used to make all character become uppercase or lowercase.

After all the data through *preprocessing* the next algorithm is Dynamic Markov Compression. This algorithm is use for text compression. The next stage is classification by Minimum Cross Entropy. The data will be used is Indonesian language. The desired result of this research is the value accuracy above 70%.

Key words : SMS *Spam* Filtering, *Preprocessing*, Dynamic Markov Compression Algorithm, Classification