Abstract

The need of information is increasing year to year in line with technological developments and the needs of the public to know the latest information. The information currently available does not only come from newspaper, but also from the internet that has become one of the main sources of information. According to the website http://direktori.perusahaan.web.id/index.php?search=berita there are about 136 sites that are listed as indonesian language site that contains a variety of information such as music, technology, national and international. Some of the sites are http://www.liputan6.com/ , http://www.metrotvnews.com/,http://www.wartajazz.com/, http://www.musikindo.com/ ,and others. The emergence of online news website is in line with the indonesian internet users in indonesia that have reached 63 miliion people.

Due to amount of information in circulation, the classification of information into the right category becomes important. In order to avoid confusion among readerss, the proper categorazation becomes useful for the reader as he can get what he wants. One of solutions to solve this problem is to perform automatic classification of existing information. Classification is the process of grouping documents according to the categories of documents that are contained in the document contents.

In this final project,the classification is using support vector machine with decission tree architecture(SVM-DTA) method. By using the method, system can automatically classify news.

The best classification are obtained when using RBF kernel with parameter C=10 and gamma=1 and in polynomial kernel with parameter C=10 and degree=2, value of error rate is smaller then the other test that is 4%.

Keywords: classification, text, SVM-DTA, tree,