ABSTRACT

News is an important role in human life. Because people always require information, person can access the news through various media, including: print media, electronic media, and Internet media. Most people want a piece of news can be accessed anywhere and anytime, then arose an idea how to make the news becomes portable. In this final project, will develop a server application responsible for crawling and manage news, News was addressed to the mobile user such as a cell phone. The aim of this project which provides convenience for mobile phone users to get the latest news from the site on the internet news provider. By utilizing the RSS feature of the site, the system will get the latest news from the site. Then look for the complete news by taking the content page in the link results RSS document. News that is still embedded in HTML documents will be separated by using RegEx or Regular Expression. Then the news is categorized by using text mining and correlation analysis. The calculation of the correlation, which becomes variable is the news and categories of documents that have been provided by the system. Then stored into the database which will be accessed by a WML server. The expected result is the news access to the WML server with load data faster than with direct access to the original site. And categorization of news with an error rate of 20 to 30%.

Keywords: news, rss, regex, text mining, correlation analysis, wml.