

## ABSTRACT

As the increasing of the development of the internet, the number of online news web available in the net is also increasing. Basically, the online news web provides important information. Various techniques in text mining can be applied to gain more advantages from the available information, such as using page-based clustering or even keyword based search. But, pages on newspaper, usually consist of many kinds of news item with unrelated topic each other, so the page-based clustering gives an optimal result deficiently

In this final assignment, an approximation is applied by extracting the news items on web pages individually and by mining it separately using pattern based strategy. This approximation is using the pattern URL text and anchor text in order to extract news link and also using crawler for news link search to seek the full story from each news item.

The analysis and testing phase results that the pattern based strategy approach build, proved can extract the full story of the news web page although not all full stories from each news item is extractable. The extracted news item will reach the optimal value if the links on the input web is homogeny

**Kata Kunci :** pattern based strategy , URL text , anchor text, newspaper web, full story.