# Abstract

Technological developments and the need for information led to the accumulation of information and data in text documents in the form of softcopy and hardcopy. This condition causes difficulty in finding the appropriate documents with the needs. For ease of finding documents as needed, take the keywords that accompany each document. Keyword is a word that represents the contents of a document as a whole. With keywords, not only make it easier to find documents as needed, but also helped to find other documents that have relevance to the topic of a document.

In this final task, the method implemented Artificial Immune System (AIS) are based on information theory to extract the keywords from a document. This method adopted the Natural Immune System's ability to recognize the substances that enter the body's harmful. With this method the content of each word is calculated to determine the information value of keywords.The process of learning is done by storing information every word every time, or the extraction of training conducted for use in the next extraction.

Tests were conducted to obtain the minimum value of the percentage content of the best information on a word using AIS method based on information theory parameters precision, recall, and f-measure in predicting the 10 terms with the highest tf-idf value. In addition the test was also performed to determine the influence of the addition of a process of elimination of stopwords extracted. Based on a minimum percentage of the test obtained the best information on the content of a word that is at least 90% of the total entropy of the document sets and about 50% to 90% of the entropy of the category in which the word originated. The lower the minimum percentage of the value of the information content in a word, the bigger the influence of stopwords elimination process increased the value of precision results of extraction. Conversely, the higher the minimum percentage of the value of information content in a word, the smaller the effect of stopwords elimination process increased the value of precision results of extraction. But the addition of stopwords elimination process can also be lowered if there is a recall value among the 10 terms with highest tf-idf values found in the list of stopwords which ultimately reduce the value of f-measure.


**Keywords**: *Artificial Immune System* (AIS), keyword extraction, keyword, Information Theory.