

## **Abstract**

Document Summarization is aimed to get the core information of a document. In this final project research MMR is used to reduce the redundancy in document. The data set is derived from some official news website. The input's data are news and query, also the process that consist of text preprocessing, tfidf calculation, cosine similarity calculation, MMR calculation and the output's data is the rank of sentence will be extracted from this document. By building this system, will gained a summary that describes the news and relevant to a query that is entered, and contains no redundancy sentence.

**Keyword:** document summarization, cosine similarity, maximum marginal relevance, redundancy in sentence