

ABSTRACT

COMPARISON OF VECTOR SPACE MODEL ALGORITHM AND RABIN KARP ALGORITHM FOR MEASURING SIMILARITY DOCUMENTS IN THE PROPOSAL SKRIPSI

Plagiarism is rampant in Indonesia. So many ways can used to identify pralgiarism, one of them is comparing documents submitted with other documents. If the degree of similarity between documents is high, then plagiarism is suspected in the document, so we need application can measure similarities between documents.

In this research, we compare the vector space model algorithm and rabin karp algorithm to see ore efficient algorithm. This application is designed using the waterfall method, created with PHP language and MYSQL database. Data to be used is from proposal document at Faculty of Science Technology Catholic Musi Charitas University informatics study program. To measure similarity, two approaches are used cosine similarity for vector space model algorithm and dice similarity for Rabin Karp algorithm.

The final result is vector space model algorithm is more efficient than Rabin karp

Keyword: Information Retrieval, Vector Space Model Algorithm, Rabin Karp Algorithm