IMPLEMENTATION OF RABIN-KARP ALGORITHM ON INDONESIAN TEXT DOCUMENT PLAGIARISM DETECTION APPLICATION

ABSTRACT

Nowadays, internet is a common source of information. This paradigm causes a tendency to duplicate text documents from the Internet and claim it as his/her own works. This action is called plagiarism. One effort to prevent plagiarism is to make an application that detects plagiarism in text documents, especially in Indonesian language.

This thesis implements a plagiarism detection application using Rabin-Karp algorithm. This string matching algorithm is used because it is capable to match multiple patterns. Pattern matching is performed to determine which same patterns are in two documents so that the percentage of similarity between those documents can be calculated. In addition, there is a modified Rabin-Karp algorithm. Modified Rabin-Karp algorithm requires shorter detection time than Rabin-Karp algorithm without modification. Beside that, the resulting percentage of similarity does not differ much.

Keywords: plagiarism, text document, Indonesian language, string matching, Rabin-Karp algorithm, modified Rabin-Karp algorithm