

Abstract

Technological developments are increasingly sophisticated with the internet. The internet that can be easily accessed to find information and documents can lead to acts of plagiarism. Every document that carries out plagiarism will be difficult to recognize without a system that can recognize the similarity of documents. A system that can detect plagiarism by looking for similarities to the required documents. In this study Smith-Waterman algorithm was used to identify the most significant similarity (local alignment) of two sequences of nucleotide sequences of processes (sequences) or protein sequences so that the similarities between the two documents can be seen. The Nazief & Andriani algorithm is a stemming algorithm on text-preprocessing as a supporting algorithm in the process of determining the similarity of text documents. The results obtained in this study are the comparison of two sequences with the help of preprocessing has a greater level of similarity calculation in detecting document similarities. From the average test on the original document and the test document, the level of similarity produced is more than 50%, which means that the results can be stated as a plagiarism action.

Keywords: *Document, Plagiatrism, Algoritma Smith-Waterman, Algorithm Nazief & Andriani*