

ABSTRACT

Plagiarism, or better known as plagiarism is not foreign to our ears. Moreover, among the students with so many tasks, it was inevitable plagiarism will surely coloring educational word. Actually, plagiarism can be solved by building a system that capable to detecting the similarity among documents. Many algorithms that can be used to building these systems, such as Rabin Karp, Smith Waterman, Edit Distance, etc.. All these algorithms have advantages each other. But Edit Distance algorithm offers the comparison process more quickly and accurately. To increase the accuracy of the system required preprocessing process. Preprocessing process contains case folding, tokenizing, filtering, and stemming.

Keywords : Plagiarism, Edit Distance, Preprocessing, String Matching