

CHAPTER I

INTRODUCTION

1.1 Background of Study

Plagiarism is presenting someone else's work or ideas as your own, with or without their consent, by incorporating it into the work without full acknowledgment [1]. Plagiarism has happened in academic and non-academic environment. Plagiarism is used to happen in the academic environment. Since students' awareness in making paper or essay is too average, it triggers a lot of the occurrence of plagiarism. Plagiarism committed by students is often denied by the reason they don't copy the other people's work, but they only get inspiration from them.

In era of globalization the internet is more advanced. [2]. Prevention and detection are some ways that can be done in order to reduce plagiarism. Prevention means to hindering the emergence of plagiarism which is concerned with the moral community and the education system. This will provide remarkable long-term effects. The detection means a way to reveal plagiarism.

Some software which designed for detecting plagiarism documents are Turnitin, Plagium, Dupli Checker, iThenticate, Plagiarism Checker, and etc [3]. Plagiarism also accomplished by the students of Information Engineering Bachelors in Dian Nuswantoro University (UDINUS). Because of this, the writer made a plagiarism detection and this system will be applied in information retrieval and the process will use vector space

model. Methods of Vector Space Model is one of the simplest methods used to search for documents in common with other documents. This method also has several stages in the search for common document, which looks at the frequency of occurrence of the word in the document preprocessing stage, then calculate the similarity of a document with a document that is compared by calculating the term frequency – inverse document frequency for terming documents and the cosine similarity for similarity of the documents.

1.2 Problem Statement

From the background of study above, the problem statement of this study is how Vector Space Model method can be used for plagiarism detection document effectively.

1.3 Scope Of The Study

The scope of problem in this study are:

1. The data will be taken from undergraduate student's thesis of informatics engineering in Dian Nuswantoro University.
2. The chapter 1 of the thesis document that will be used.
3. The method use in this research is Vector Space Model Approach with Term Frequency – Inverse Document Frequency (TF-IDF) and the similarity was using Cosine Similarity Algorithm.
4. The documents that be used are: *.docx.

1.4 Objective of Study

The objective of this study is to implement Vector Space Model Method which is used to detect possible plagiarism in the text document, so it can maximize for originality in writing for undergraduate thesis in Dian Nuswantoro University.

1.5 Benefit

1.5.1. Benefit for author

Doing the research of plagiarism using Vector Space Model can improve the writer's knowledge in plagiarism and Vector Space Model itself.

1.5.2. Benefit for academic

- a. Increasing to the academic literature, also can provide information to the readers as their references for those who are interested in developing a plagiarism detection system on a text document.
- b. Providing access to students who want to check whether the tasks that will be made have in common with the final task-a task that has previously been made.