

**IMPLEMENTASI TEXT MINING UNTUK MENDETEKSI  
KEMIRIPAN DOKUMEN TEKS YANG BERBENTUK SKRIPSI  
MENGGUNAKAN METODE LEVENSHTEIN DISTANCE**  
**(Studi Kasus: Universitas Satya Negara Indonesia)**



**FAKULTAS TEKNIK  
UNIVERSITAS SATYA NEGARA INDONESIA  
JAKARTA  
2018**

**IMPLEMENTATION OF TEXT MINING FOR DETECTION  
SIMILARITY OF TEXT DOCUMENT IN THE FORM OF THESIS  
USING LEVENSSTEIN DISTANCE METHOD  
(Case Study: Satya Negara Indonesia University)**



**FACULTY OF ENGINEERING  
SATYA NEGARA INDONESIA UNIVERSITY  
JAKARTA  
2018**

**IMPLEMENTASI TEXT MINING UNTUK MENDETEKSI  
KEMIRIPAN DOKUMEN TEKS YANG BERBENTUK SKRIPSI  
MENGGUNAKAN METODE LEVENSHTEIN DISTANCE**  
**(Studi Kasus : Universitas Satya Negara Indonesia)**



**FAKULTAS TEKNIK  
UNIVERSITAS SATYA NEGARA INDONESIA  
JAKARTA  
2018**

## ABSTRAK

Skripsi merupakan karya tulis ilmiah berdasarkan hasil penelitian yang dilakukan oleh seorang mahasiswa sebagai tugas akhir untuk memperoleh gelar sarjana. Akan tetapi mahasiswa mengalami kesulitan dalam penyusunan skripsi, sehingga mahasiswa banyak melakukan tindakan plagiarisme. Salah satu cara yang dapat dilakukan untuk mencegah terjadinya tindak plagiarisme yaitu dengan cara mendeteksi, maka diperlukan sebuah sistem yang dapat membandingkan isi yang ada di dalam skripsi.

Untuk membangun sistem tersebut, maka diperlukannya sebuah metode yang dapat membandingkan isi dokumen skripsi. metode *levenshtein distance* merupakan metode pencocokan antara dua *string* sehingga cocok digunakan dalam pembuatan sistem pendeteksi kemiripan dokumen skripsi. Hasil dari penerapan metode *levenshtein distance* yaitu berupa informasi persentase kemiripan antar dokumen skripsi.

Setelah dilakukan pengujian sistem dengan menggunakan dokumen skripsi mahasiswa Teknik Informatika menghasilkan kemiripan dokumen yang terendah yaitu 19.7% dengan klasifikasi plagiarisme ringan. Dan hasil kemiripan dokumen yang tertinggi yaitu 33.27% dengan klasifikasi plagiarisme sedang.

Kata kunci: *Text Mining, Levenshtein Distance, Skripsi*

## **ABSTRACT**

*Thesis is a scientific paper based on the results of research undertaken by students as the final task to obtain a degree. However, students have difficulty in preparing the thesis, so that many students do plagiarism. One way that can be done to prevent the occurrence of plagiarism is to detect, required a system that can compare the content in the thesis.*

*To build the system, need a method that can compare the content of thesis documents. The levenshtein distance method is a matching method between two strings making it suitable for use in making the system resemble the same thesis document. The result of applying levenshtein distance method is in the form of percentage of information similarity between thesis document.*

*After testing the system by using the thesis document of Informatics Engineering students yielded the lowest document similarity that is 19.7% with the classification of mild plagiarism. And the highest document similarity result is 33,27% with the classification of medium plagiarism.*

*Keywords:* Text Mining, Levenshtein Distance, Thesis