

## SISTEM PENDETEKSI PLAGIARISME DOKUMEN SKRIPSI DENGAN MENGUNAKAN PERBANDINGAN *N-GRAMS*

### ABSTRAK

Plagiat bukan menjadi hal baru dalam dunia pendidikan. Dengan adanya kemudahan bagi setiap orang untuk dapat bertukar informasi. Sehingga memberikan kemudahan dalam tindakan plagiarisme terutama dikalangan mahasiswa. Yang menjadi salah satu keprihatinan dalam dunia pendidikan yaitu adanya tindakan plagiat pada naskah dokumen skripsi yang banyak dilakukan oleh para mahasiswa. Oleh karena itu diperlukan suatu sistem pendeteksian plagiarisme pada dokumen teks yang dilakukan secara terkomputerisasi.

Dalam hal ini sistem pendeteksi plagiat dokumen skripsi yang akan digunakan yaitu dengan menggunakan perbandingan *n-grams*. Perbandingan *n-grams* merupakan sebuah metode sederhana dengan membandingkan potongan teks berdasarkan *n-grams* kata.

Pengujian dari sistem ini dilakukan dengan melakukan perhitungan *recall* dan *precision* dari hasil testing program dengan membandingkan dengan hasil deteksi secara manual. Dari hasil pengujian yang dilakukan terhadap dokumen skripsi dengan menggunakan perbandingan *n-grams* dengan nilai  $n=3$ ,  $n=4$ , dan  $n=5$  yang memiliki tingkat ketepatan terbaik adalah yang nilai  $n=5$  dengan nilai *precision* = 90.12% dibandingkn dengan nilai  $n =3$  dengan hasil *precision* 10.61% dan  $n=4$  dengan hasil *precision* 78.12%.

Kata kunci : sistem pendeteksi plagiarisme, plagiarisme dokumen skripsi, perbandingan *n-grams*

**SYSTEM PLAGIARISM DETECTION DOCUMENT THESIS  
BASED ON N-GRAMS COMPARISON  
(A CASE STUDY AT DEPARTMENT OF INFORMATICS  
ENGINEERING , SANATA DHARMA UNIVERSITY )**

**ABSTRACT**

Plagiarism is not a new thing in the world of education. With the convenience for everyone to be able to exchange information. So much easier and acts of plagiarism, especially among students. That was one of concern in education is the act of plagiarism in the thesis text document is mostly done by the students. Therefore we need a system for plagiarism detection in text documents is done in computerized.

In this system plagiarism detection document thesis based on n-grams comparison. N-grams comparison is a simple method by comparing the bits of text based on word n-grams.

Testing performance of the system with comparing of recall and precision of the testing program and the detection results manually. From the results of tests performed on thesis document based on n-grams comparison with a value of  $n = 3$ ,  $n = 4$  and  $n = 5$  which has the best accuracy rate is the value of  $n = 5$  with the precision = 90.12% compared with a value of  $n = 3$  with the precision 10.61% and  $n = 4$  with 78.12% precision results.

Key words : plagiarism detection systems, document plagiarism thesis, n-grams comparison