

Pengembangan sistem pendeteksi plagiarisme karya tulis digital dwi bahasa (Indonesia-Inggris) menggunakan algoritma winnowing dan googletrans api dengan bahasa pemrograman python = Development of plagiarism detector system to bilingual digital essay (Indonesian-English) using winnowing algorithm and googletrans api with python programming language

Addi Ryan, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20472934&lokasi=lokal>

Abstrak

Pengembangan sistem pendeteksi plagiarisme dibuat untuk mengatasi masalah plagiarisme yang kerap terjadi pada dunia akademis. Pada skripsi ini dibuat sistem pendeteksi plagiarisme otomatis pada karya tulis digital dwi bahasa Indonesia-Inggris dengan Bahasa Indonesia digunakan sebagai karya tulis yang akan diuji dan Bahasa Inggris sebagai karya tulis referensinya. Sistem menerapkan algoritma winnowing yang dilengkapi dengan metode penerjemah bahasa Googletrans API dan similar words. Algoritma winnowing merupakan algoritma yang dapat mendeteksi kesamaan antar teks dengan menggunakan fingerprint yang didapat dari proses hashing karakter teks. Penelitian dilakukan untuk meningkatkan akurasi sistem dan mengetahui metode penilai kesamaan teks yang akurat.

Dari hasil penelitian, didapatkan bahwa parameter terbaik algoritma winnowing secara umum terdapat saat nilai k-window = 5 dan nilai basis bilangan prima bernilai 3. Parameter n-gram bernilai kecil akan lebih akurat pada teks yang memiliki jumlah kata lebih sedikit dan/atau tingkat plagiarisme tinggi dan sebaliknya. Tingkat akurasi sistem pendeteksi plagiarisme otomatis dwi bahasa yang dikembangkan berkisar antara 75.02 hingga 99.51.

Metode Cosine Similarity menjadi metode penilai kesamaan teks terbaik dari hasil penelitian ini. Selain itu, metode penerjemahan Googletrans API juga memberikan kelebihan dalam hal akurasi dan kelengkapan data kamus dibandingkan dengan metode kamus terjemahan database.

The development of plagiarism detection system is made to overcome the problem of plagiarism that often occurs in the academic world. In this thesis, an automatic plagiarism detection system on bilingual digital paper Indonesian English is created with Indonesian is used as the tested paper and English as the reference paper. The system implements the winnowing algorithm that comes with the Googletrans API language translator method and similar words. Winnowing algorithm is an algorithm that can detect similarity between text by using fingerprint obtained from hashing process of text character. The study was conducted to improve system accuracy and to know accurate method of text equality assessment.

From the study result, it is found that the best parameter of winnowing algorithm is generally occurred when the value of k window 5 and the base value of the prime number is 3. The smaller value of n gram parameter will be more accurate in text that has fewer word counts and or high plagiarism levels and vice versa. The accuracy level of the automatic plagiarism detection system in the developed language ranged from 75.02 to 99.51 .

The Cosine Similarity method is the best method of text equality assessment according to results of this study. In addition, the Googletrans API translation method also provides advantages in terms of accuracy and completeness of dictionary data as compared to database translation dictionary method.