SISTEM DETEKSI PLAGIASI MENGGUNAKAN ALGORITME FREQUENCY BASED HASHING-S PADA FILE PDF

SKRIPSI

Oleh
Thamrin Auliya
180210117

FAKULTAS TEKNIK
PROGRAM STUDI TEKNIK INFORMATIKA
UNIVERSITAS PUTERA BATAM
TAHUN 2021/ 2022

SISTEM DETEKSI PLAGIASI MENGGUNAKAN ALGORITME FREQUENCY BASED HASHING-S PADA FILE PDF

SKRIPSI

Untuk memenuhi salah satu syarat
memperoleh gelar sarjana

Oleh
Thamrin Auliya
180210117

FAKULTAS TEKNIK
PROGRAM STUDI TEKNIK INFORMATIKA
UNIVERSITAS PUTERA BATAM

TAHUN 2021/ 2022

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini penulis:

Nama : Thamrin Auliya

NPM : 180210117

Fakultas : Teknik dan Komputer

Program studi : Teknik Informatika

Menyatakan bahwa **"Skripsi"** yang dibuat dengan judul:

"**SISTEM DETEKSI PLAGIASI MENGGUNAKAN ALGORITME FREQUENCY BASED HASHING-S PADA FILE PDF"**

Ini adalah karya sendiri dan bukan "duplikasi" dari karya orang lain. Sejauh yang penulis tahu dalam teks skripsi ini tidak ada karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang disebutkan dalam teks ini dan disebutkan dalam sumber referensi kutipan.Apabila terdapat didalam naskah Skripsi ini dapat dibuktikan terdapat unsur unsur PLAGIASI, saya bersedia naskah Sikripsi ini digugurkan dan gelar akademik yang saya peroleh dibatalkan, serta diproses sesuai dengan peraturan perundangundang yang berlaku.
Demikian penyataan ini saya buat dengan sebenarnya tanpa ada paksaan dari siapapun.

Batam, 8 Agustus 2022

Thamrin Auliya
180210117

SISTEM DETEKSI PLAGIASI MENGGUNAKAN ALGORITME FREQUENCY BASED HASHING-S PADAFILE PDF

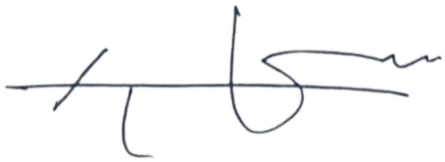
SKRIPSI

Untuk memenuhi salah satu syarat
memperoleh gelar sarjana

Oleh
Thamrin Auliya
180210117

Telah disetujui pembimbing pada
tanggalseperti yang tertera dibawah
ini

Batam, 8 agustus 2022



Cosmas Eko Suharyanto, S.Kom., M.MSI

Pembimbing

ABSTRAK

Di era komputasi digital sekarang ini, telah terjadi pertumbuhan produksi data digital yang sangat tinggi. Dikutip dari laman forbes.com, ada 2,5 quintillion bytes data yang diproduksi setiap hari. Salah satu data digital tersebut adalah tugas mahasiswa yang dikumpulkan melalui media Google Classroom. Dari data mahasiswa yang dikumpulkan tersebut, banyak tugas yang terindikasi plagiasi dengan tugas mahasiswa yang lainnya. Jika dilakukan anilisis secara manual, maka akan memakan waktu yang lama dan sangat melelahkan. Untuk efisiensi pemanfaatan waktu dan sumber daya, dibutuhkan proses penyaringan yang memiliki kemampuan untuk menghitung tingkat plagiasi dari setiap tugas mahasiswa. Metode Approximate Matching merupakan metode yang paling sering digunakan untuk menemukan kesamaan diantara data yang dibandingkan dengan menetapkan skor kesamaan. Pada penelitian ini algoritme Approximate Matching yang digunakan adalah algoritme Frequency Based Hashing-S. Kelebihan algoritme ini adalah aman terhadap serangan aktif dan memiliki tingkat akurasi 98% untuk format data terkompresi. Aplikasi ini bermanfaat bagi dosen yang akan memeriksa tugas mahasiswa yang banyak. Dengan adanya aplikasi ini, dosen hanya perlu memeriksa tugas mahasiswa yang tidak plagiasi saja sehingga mengurangi jumlah tugas mahasiswa yang akan diperiksa secara manual.

Kata Kunci: Frequency Based Hashing-S, Sistem Plagiasi, Sistm Penyaringan Data

ABSTRACT

In today's digital computing era, there has been a very high growth of digital data production. Quoted from the forbes.com page, there are 2.5 quintillion bytes of data created every day. One of the digital data is student assignments collected through Google Classroom media. From the student data collected, many assignments indicated plagiarism with other student assignments. If the analysis is done manually, it will take a long time and be very tiring. For efficient use of time and resources, a screening process is needed that has the ability to calculate the plagiarism level of each student's assignment. The Approximate Matching method is the method most often used to find similarities between the data being compared by setting a similarity score. In this study, the Approximate Matching algorithm used is the Frequency Based Hashing-S algorithm. The advantages of this algorithm are that it is safe against active attacks and has a 98% accuracy rate for compressed data formats. This application is useful for lecturers who will check a lot of student assignments. With this application, lecturers only need to check student assignments that are not plagiarized, thereby reducing the number of student assignments that will be checked manually.

Keywords: Frequency Based Hahsing-S, Plagiarism System, Data Filtering System

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan Kehadirat Tuhan Yang Maha Esa atas kuasa dan limpahan rahmat-Nya sehingga penulis dapat menyelesaikan proposal yang merupakan salah satu syarat untuk menyelesaikan program studi strata satu (S1) pada program studi Teknik Informatika Universitas Putera Batam.

Dengan segala keterbatasan, penulis menyadari pula bahwa proposal ini tidak akan terwujud tanpa bantuan, bimbingan, dan dorongan dari berbagai pihak. Untuk itu dengan segala kerendahan hati, penulis menyampaikan ucapan terimakasih kepada:

1. Ibu Dr. Nur Elfi Husda, S.Kom., M.SI. selaku Rektor Universitas Putera Batam.
2. Bapak Welly Sugianto, S.T., M.M. selaku Dekan Fakultas Teknik dan Komputer.
3. Ketua program studi teknik informatika Bapak Andi Maslan, S.T.,M.SI
4. Bapak Cosmas Suharyanto, S.Kom., M.Kom, selaku pembimbing skripsi pada program Stusi Teknik Informatika Universitas Putera Batam.
5. Kepada orang tua tercinta atas curahan kasih sayang, doa, nasihat, serta pesan yang disampaikan kepada penulis sehingga penulis tetap memiliki semangat juang dalam penyelesaian proposal ini.
6. Serta semua pihak yang tidak dapat saya dapat penulis sebutkan satu persatu yang telah membantu dalam penyusunan skripsi ini. yang telah membantu penyelesaian skripsi ini yang tidak dapat penulis sebutkan satu persatu

Batam, 8 Agustus 2022

Thamrin Auliya

DAFTAR ISI

DAPTAR TABEL

DAPTAR GAMBAR