

Penerapan Algoritma Knuth-Morris-Pratt dalam Mendeteksi Tingkat Kemiripan Judul Skripsi Berbasis Web

Dina Suci Islamiyati, Ahmad Fikri*

Fakultas Teknik, Teknik Informatika, Universitas Pamulang, Tangerang
Jl. Surya Kencana No.1, Pamulang Bar., Kec. Pamulang, Kota Tangerang Selatan, Banten, Indonesia

Email: ¹dinasuciislamiyati21@gmail.com, ²*dosen041033@unpam.ac.id

Email Penulis Korespondensi: dosen041033@unpam.ac.id

Submitted: 13/01/2022; Accepted: 31/01/2022; Published: 31/01/2022

Abstrak—Dalam hal pencarian topik permasalahan untuk judul Skripsi seringkali mahasiswa mengalami kesulitan dalam menentukan judul Skripsi, selain itu dimungkinkan terjadi kesamaan antara mahasiswa. Pengecekan judul secara manual perlu upaya lebih, untuk itu diperlukan sebuah sistem untuk mendeteksi tingkat kemiripan pada judul Skripsi. Untuk merancang dan membangun sistem pendeteksian tingkat kemiripan judul skripsi dengan menggunakan algoritma *Knuth-Morris-Pratt*. Untuk mengetahui persentase kemiripan pada judul, algoritma *Knuth-Morris-Pratt* pada saat pencocokkan string adalah Masukkan Query kata yang akan dicari, mencocokkan *pattern* susunan kata yang dijadikan sebagai contoh pada awal teks. Dari kiri ke kanan, algoritma ini akan mencocokkan karakter per karakter *pattern* dengan karakter di teks yang bersesuaian, Algoritma kemudian menggeser *pattern* susunan kata sampai *pattern* susunan kata yang dijadikan sebagai contoh berada di ujung teks. Berdasarkan analisis, perancangan dan implementasi pada sistem pendeteksi tingkat kemiripan dua atau lebih judul skripsi dengan menggunakan algoritma *Knuth Morris Pratt* dapat disimpulkan bahwa: Sistem pendeteksi telah berhasil dibuat dengan menggunakan algoritma *Knuth Morris Pratt (KMP)*. Telah didapatkan hasil analisa dengan proses *string matching* antara *string* dengan *pattern* yang mana sistem menampilkan hasil analisa berupa jumlah kemiripan kata dan persentasenya.

Kata Kunci: Algoritma Knuth-Morris-Pratt; Judul Skripsi; String Matching

Abstract—In terms of finding the problem topic for thesis titles, students often have difficulty in determining thesis titles, in addition, there may be similarities between students. Manually checking the title requires more effort, for it takes a system to detect the level of similarity in the thesis title. To design and build a thesis title-level detection system using the Knuth-Morris-Pratt algorithm. To find out the percentage of similarities in the title, the Knuth-Morris-Pratt algorithm at the time of matching the string is Enter query of the word to be searched, matching the word arrangement pattern used as an example at the beginning of the text. From left to right, this algorithm will match character per character pattern with characters in the corresponding text, the algorithm then shifts the wording pattern until the word arrangement pattern used as an example is at the end of the text. Based on the analysis, design and implementation of the similarity level detection system of two or more thesis titles using the Knuth Morris Pratt algorithm it can be concluded that: The detection system has been successfully created using the Knuth Morris Pratt algorithm (KMP). It has been obtained the results of analysis by the process of string matching between strings with patterns where the system postulates the results of analysis in the form of the number of word similarities and percentages.

Keywords: Knuth-Morris-Pratt Algorithm; Thesis Title; String Matching

1. PENDAHULUAN

Salah satu upaya yang dapat membantu dalam hal pengecekan kemiripan dua judul skripsi adalah *by system* dengan pencocokan string (*string matching*) pada teks yang terdapat pada judul Skripsi/TA. *String matching* adalah mencari sebuah *string* yang terdiri dari beberapa karakter yang biasa disebut (*pattern*) dalam jumlah besar [1]. Terdapat beberapa algoritma pada *string matching* ini diantaranya *Brute Force*, *Winnowing*, *Knuth-Morris-Pratt*, *Booyer-Moore*, *Rabin-Karp* dan sebagainya.

Penelitian-penelitian mengenai *string matching* atau deteksi tingkat kemiripan judul Skripsi/TA telah banyak dilakukan, seperti penelitian yang dilakukan oleh [1]. Hasil dari penelitian ini adalah dengan menggunakan algoritma *String Matching* pada makna rambu lalu lintas dapat lebih mudah. Dengan adanya aplikasi ini dapat membantu pengguna *smartphone android* dalam melakukan pencarian gambar yang memiliki kesamaan kata jika diinput dengan kata yang telah di simpan sebelumnya di dalam *database* pencarian rambu lalu lintas. Penelitian lain dilakukan oleh [2]. Hasil dari penelitian ini adalah dalam penerapan algoritma *Winnowing* dengan teknik pemotongan *string* menggunakan *k-gram* memiliki tingkat nilai *similarity* yang tinggi namun ketika nilai jumlah *k* semakin besar akan mengurangi tingkat nilai *similarity* dengan hasil rata-rata pada *k* = 2 sebesar 0.5299, *k* = 3 sebesar 0.1689, *k* = 5 sebesar 0.0283 dan *k* = 7 sebesar 0.0095. Penerapan pemotongan *string* berdasarkan kata menggunakan *n-gram* pada penerapan unigram memiliki rata-rata tingkat *similarity* sebesar 0.0683 sedangkan bigram memiliki tingkat nilai *similarity* 0.003, serta pada *trigram* dan *fourgram* memiliki tingkat kesamaan sebesar 0.000. Penelitian lainnya dilakukan oleh [3]. Hasil dari penelitian ini dapat ditarik kesimpulan bahwa algortima *Boyer-Moore* mampu melakukan pencocokan *string* dengan waktu yang cepat.

Pada penelitian ini peneliti akan mencoba menerapkan algoritma *Knuth-Morris-Pratt* untuk mendeteksi tingkat kemiripan pada dua judul skripsi. Algoritma *Knuth-Morris-Pratt* merupakan salah satu algoritma pencarian *string*, dikembangkan secara terpisah oleh Donald E. Knuth pada tahun 1967 dan James H. Morris bersama Vaughan R. Pratt pada tahun 1966, namun keduanya mempublikasikannya secara bersamaan pada tahun 1977. Algoritma *Knuth-Morris-Pratt* merupakan algoritma pencarian *string* dengan melakukan perbandingan karakter mulai dari karakter paling kanan dari *string* yang dicari [4]

2. METODOLOGI PENELITIAN

2.1 Algoritma Knuth-Morris-Pratt

Secara sistematis, langkah-langkah yang dilakukan algoritma *Knuth-Morris-Pratt* pada saat pencocokkan *string* adalah sebagai berikut [5]:

1. Masukkan *Query* kata yang akan dicari. Dengan permisalan $P = \text{Pattern}$ atau pola susunan kata yang dijadikan sebagai contoh atau pola teks yang akan dicari $T = \text{Teks}$ atau judul dokumen.
2. Algoritma *Knuth Morris Pratt* mulai mencocokkan *pattern* atau pola susunan kata yang dijadikan sebagai contoh pada awal teks.
3. Dari kiri ke kanan, algoritma ini akan mencocokkan karakter per karakter *pattern* dengan karakter di teks yang bersesuaian, sampai salah satu kondisi berikut dipenuhi:
 - a. Karakter di *pattern* dan di teks yang dibandingkan tidak cocok (*mismatch*).
 - b. Semua karakter di *pattern* cocok. Kemudian algoritma akan memberitahukan penemuan di posisi ini.
4. Algoritma kemudian menggeser *pattern* atau pola susunan kata yang dijadikan sebagai contoh berdasarkan tabel *next*, lalu mengulangi langkah no. 2 sampai *pattern* atau pola susunan kata yang dijadikan sebagai contoh berada di ujung teks.

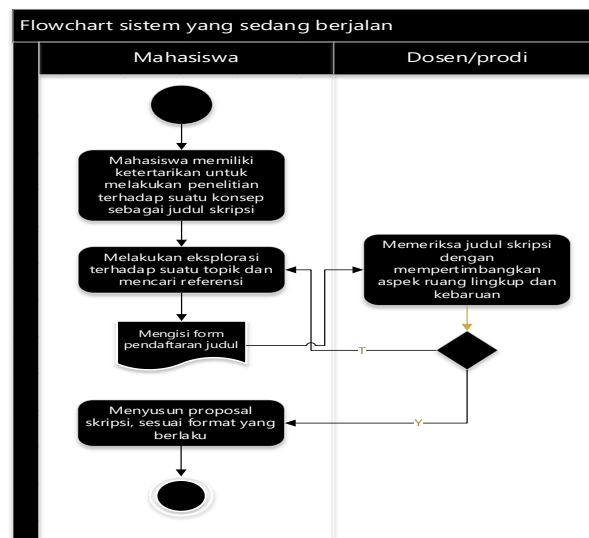
3. HASIL DAN PEMBAHASAN

3.1 Implementasi

Implementasi merupakan kelanjutan dari kegiatan perancangan sistem dan dapat dilihat sebagai upaya untuk merealisasikan sistem yang telah dirancang. Pada tahap ini, proses yang diimplementasikan adalah implementasi *hardware* dan *software* untuk mengimplementasikan dan menguji sistem yang dirancang. Hasil dari fase ini adalah sistem pemrosesan data dan informasi yang berfungsi dengan baik.

3.2 Analisa Sistem Berjalan

Analisa sistem yang sedang berjalan bertujuan untuk mendapatkan pemahaman yang lebih baik tentang cara kerja sistem dan masalah yang dihadapinya. Analisis sistem ini dapat digunakan sebagai dasar untuk rekomendasi perancangan sistem yang sedang berjalan berdasarkan urutan kejadian yang ada dan dari urutan kejadian yang mungkin dari (*Flowchart*).

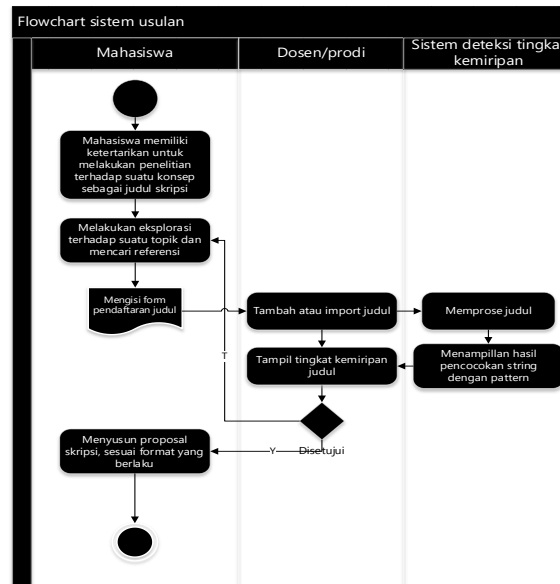


Gambar 1. Flowchart sistem yang sedang berjalan

3.3 Evaluasi Sistem Yang sedang berjalan

Setelah menganalisa sistem yang sedang berjalan dan mengidentifikasi permasalahan yang ada, maka sistem yang diusulkan adalah sebagai berikut. Setelah mahasiswa mendalami topik yang akan diangkat dengan mencari referensi dari berbagai sumber kemudian mengisi form pengajuan judul. Pihak prodi mengkaji topik yang akan dibahas dengan memperhatikan aspek ruang lingkup, kebaruan juga melakukan pengecekan tingkat kemiripan dengan judul yang sudah terdapat pada database yang ada dengan cara menambahkan atau mengimport judul ke sistem deteksi tingkat kemiripan dua atau lebih judul skripsi guna membandingkannya untuk mengecek indikasi

tingkat kemiripan diantara judul tersebut. Setelah sistem berhasil melakukan proses string matching, maka hasil akan ditampilkan bentuk popup web pada menu analisa judul. Apabila terdeteksi adanya kemiripan, maka judul yang diajukan mahasiswa tidak akan disetujui dan mahasiswa disarankan untuk mengganti judul skripsinya dan mengajukan kembali. Namun jika tidak ditemukan tingkat kemiripan maka pihak prodi akan menyetujui topik atau judul yang dimaksud maka mahasiswa dapat menyusun proposal skripsi. Gambar 2 evaluasi sitem yang sedang berjalan.

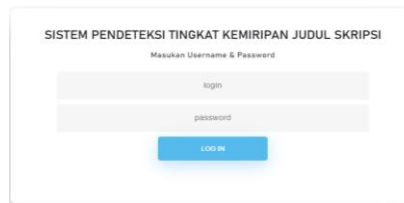


Gambar 2. Evaluasi sistem yang sedang berjalan

3.4 Sistem Interface

1. Tampilan Halaman Login

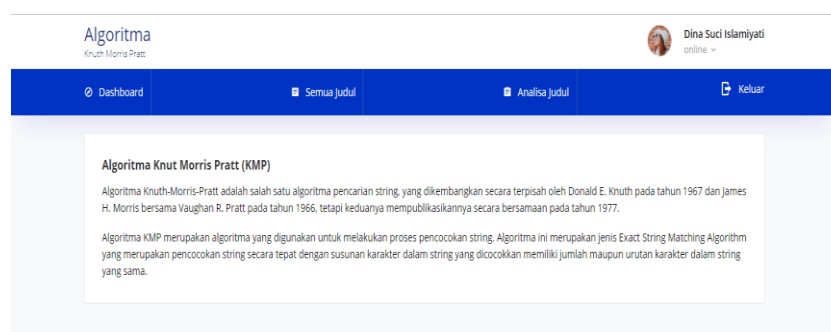
Halaman *login* merupakan halaman pertama saat aplikasi ini dijalankan. Pada halaman ini, pengguna dapat memasukkan nama pengguna dan kata sandi. Kemudian langkah selanjutnya adalah klik tombol *Sign in*. Jika pengguna benar memasukan *username* dan *password* maka sistem akan menampilkan halaman *dashboard*, namun jika salah sistem akan menampilkan pesan *error* berupa “*username/password tidak sesuai*” Form halaman *login* terlihat seperti pada gambar.



Gambar 3. Tampilan halaman login

2. Tampilan Halaman Dashboard

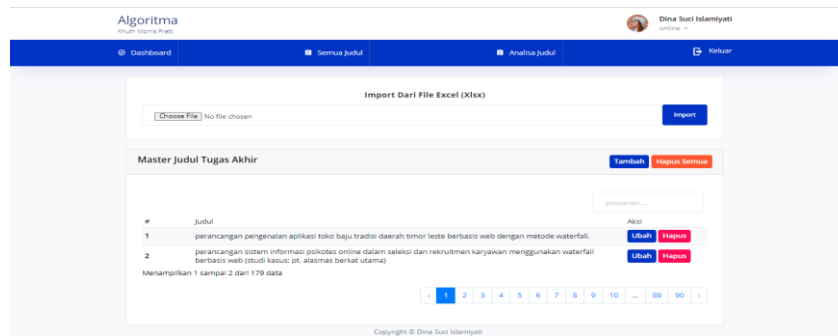
Halaman *dashboard* merupakan halaman yang muncul setelah admin berhasil melakukan proses *login*. Pada halaman ini juga terdapat info terkait algoritma *knuth morris pratt*. Halaman *dashboard* admin terlihat seperti pada gambar.



Gambar 4. Tampilan halaman dashboard

3. Tampilan Halaman Kelola Judul

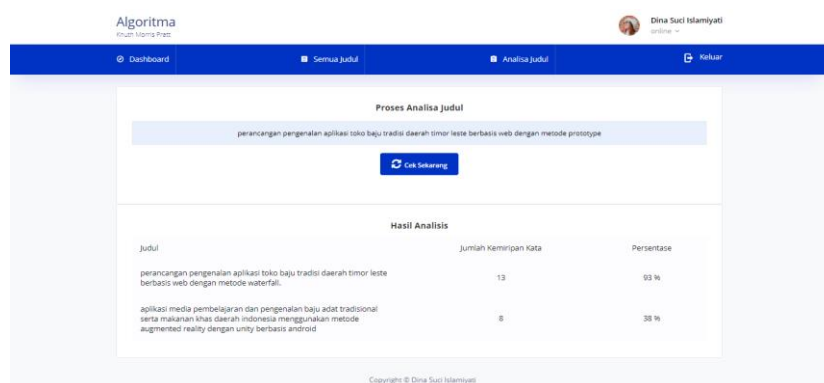
Halaman kelola judul merupakan halaman yang digunakan admin untuk mengelola data judul seperti mengimport, menambah, mengubah, menghapus juga mencari data judul pada sistem deteksi tingkat kemiripan dua atau lebih judul skripsi menggunakan algoritma *knuth morris pratt* ini. Halaman kelola judul terlihat seperti pada gambar.



Gambar 5. Tampilan halaman kelola judul

4. Tampilan Halaman Analisa Judul

Halaman analisa judul merupakan halaman yang digunakan admin untuk menganalisa judul, dimana admin memasukkan judul yang akan dianalisa kemudian sistem akan melakukan proses string matching menggunakan algoritma *knuth morris pratt*. Setelah pengecekan selesai sistem akan menampilkan hasil pengecekan tersebut yang mana admin dapat mengetahui seberapa mirip judul yang dianalisa dengan database judul yang sudah ada dalam sistem ini. Halaman analisa judul terlihat seperti pada gambar.



Gambar 6. Analisa judul

3.5 Pengujian Black Box

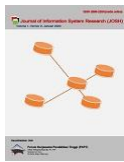
Pada tahap ini dilakukan pengujian terhadap sistem yang sudah dibuat. Dengan uji *black box* ini diharapkan jika terdapat kesalahan atau kekurangan pada aplikasi, peneliti dapat segera menemukan hal tersebut secepatnya. Uji *black box* terlihat seperti pada tabel 1.

Tabel 1. Hasil pengujian Black Box

Skenario Pengujian	Kasus Pengujian	Kasus Yang Diharapkan	Hasil Pengujian	Kesimpulan
Login	Input Username & Password	Sistem Dapat Melakukan Login	Sesuai	Berhasil
Kelola Judul	Input, Tambah, Ubah, Hapus dan Cari Judul	Data Judul Dapat Diperbaharui dengan Cara Diimport, Ditambah, Diubah, atau Diapus. Selain Itu Juga Dapat Mencari Data Judul	Sesuai	Berhasil
Analisa Judul	Input Judul yang akan Dianalisa	Dapat Menampilkan Hasil Pengecekan String Matching dengan Menggunakan Algoritma Knuth Morris Pratt	Sesuai	Berhasil

4. KESIMPULAN

Berdasarkan analisis, desain/perancangan dan implementasi pada sistem pendeteksi tingkat kemiripan dua atau lebih judul skripsi dengan menggunakan algoritma *Knuth Morris Pratt* dapat disimpulkan bahwa: Sistem



pendeteksi tingkat kemiripan dua atau lebih judul skripsi telah berhasil dibuat dengan menggunakan algoritma *Knuth Morris Pratt (KMP)*. Dan Telah didapatkan hasil analisa dari judul yang telah diajukan dengan proses *string matching* antara *string* dengan *pattern* yang mana sistem menampilkan hasil analisa berupa jumlah kemiripan kata dan presentasinya.

REFERENCES

- [1] Z. A. Matondang, “Implementasi Algoritma String Matching Pencarian Kata Dari Makna Rambu Lalu Lintas Berbasis Android,” *Jurnal Sistem Informasi Kaputama (JSIK)*, vol. II, no. 1, pp. 101-106, 2018.
- [2] W. Hidayat, E. Utami dan A. D. Hartanto, “Pemilihan Parameter Terbaik pada Algoritma Winnowing dalam Mendeteksi Tingkat Kesamaan Dokumen Bahasa Indonesia,” *Citec Journal*, vol. VII, no. 2, pp. 119-132, 2020.
- [3] Ahmad, I., Borman, R. I., Caksana, G. G., & Fakhruroz, J. (2021). Implementasi String Matching Dengan Algoritma Boyer-Moore Untuk Menentukan Tingkat Kemiripan Pada Pengajuan Judul Skripsi/TA Mahasiswa (Studi Kasus: Universitas XYZ). *SINTECH Journal*, IV(1), 53-58.
- [4] Nursobah, & Pahrudin, P. (2019). Penerapan Algoritma Pencarian Knuth-Morris-Pratt (KMP) Dalam Sistem Informasi Perpustakaan SMK TI Pratama. *SEBATIK*, XXIII(1), 112-115.
- [5] D. Ramadhani, “Perancangan aplikasi pencarian Buku Pada Perpustakaan Islamic Internasional School Darul Ilmi Murni Dengan Algoritma Knuth Morris Pratt,” *Pelita Informatika Budi Darma*, vol. XVI, no. 1, pp. 25-27, 2017.