

# IMPLEMENTASI ALGORITMA WINNOWER SEBAGAI PEMERIKSAAN KEMIRIPAN JUDUL SKRIPSI

Irman Hariman<sup>1</sup>, M Ichsan Ammarula A<sup>2</sup>

Program Studi Teknik Informatika – Fakultas Ilmu Komputer dan Sistem Informasi  
Universitas Kebangsaan Republik Indonesia

Jl. Terusan Halimun no. 37. Lingkar Selatan, Bandung, Jawa Barat 40263  
[irmanhariman@gmail.com](mailto:irmanhariman@gmail.com)<sup>1</sup>, [ichsanammarulaazwardi@gmail.com](mailto:ichsanammarulaazwardi@gmail.com)<sup>2</sup>

## Abstrak

Pengajuan judul skripsi merupakan proses yang mudah namun terkadang menjadi sulit. Banyak mahasiswa yang mengalami kesulitan dalam menentukan topik yang akan dibahas pada skripsinya. Hal ini karena adanya batas waktu untuk pengajuan judul skripsi dan penentuan tema skripsi. Berdasarkan dari permasalahan yang ada, maka diperlukan aplikasi yang dapat melakukan checker kemiripan judul skripsi sehingga dapat memberikan kemudahan bagi mahasiswa di Universitas Kebangsaan Republik Indonesia untuk mengajukan judul dan memeriksa tingkat kemiripan pada judul skripsi. Penelitian ini juga bertujuan untuk memberikan kemudahan bagi pihak fakultas dalam mengelola data judul skripsi yang ada dan yang akan diajukan oleh mahasiswa. Deteksi kemiripan judul skripsi dilakukan menggunakan Algoritma Winnower yang merupakan ekstensi dari Rabin-karp dengan penambahan metode window untuk optimalisasi hasil yang lebih baik. Serta mampu memberikan validasi checker kemiripan dokumen judul skripsi pada Fakultas Ilmu Komputer dan Sistem Informasi di Universitas Kebangsaan Republik Indonesia menggunakan Algoritma Winnower.

**Kata Kunci** : Aplikasi, checker, winnower

---

## 1. Pendahuluan

Pengajuan judul skripsi merupakan proses yang tergolong mudah namun dapat juga dikatakan sulit. Hal ini disebabkan adanya internet yang dapat mempermudah mahasiswa untuk mencari referensi judul skripsi yang ingin diajukan. Namun karena adanya batas waktu untuk pengajuan judul skripsi hal ini menyebabkan mahasiswa terjebak dengan plagiarisme sebab tidak adanya perangkat yang dapat melakukan pemeriksaan kemiripan judul skripsi mengingat jumlah data judul yang dibandingkan cukup banyak. Plagiarisme merupakan tindakan disengaja maupun tidak disengaja mengakui karya orang lain sebagai karya ilmiah diri sendiri.

Berdasarkan berita yang dimuat oleh Mashabi (2021) dalam kompas.com, kasus plagiarisme di Indonesia pernah terjadi pada saat Komisi III DPR melakukan uji kelayakan dan kepatutan (Fit and proper test) terhadap Calon Hakim Agung Yohanes Priyana. Pada kasus tersebut, Yohanes Priyana terindikasi plagiarisme setelah tidak memberikan catatan kaki pada kutipan yang seharusnya.

Kasus lainnya pernah terjadi pada seorang penyanyi Hong Jin Young asal Korea Selatan yang terbukti melakukan plagiarisme pada tesis yang dibuat. Sehingga komite program pascasarjana Chosun University di Korea Selatan tersebut mengadakan rapat dan memutuskan bahwa tesis Hong Jin Young hasil jiplak dan gelar magisternya pun dicabut.

Saat ini teknologi informasi telah berevolusi menjadi semakin canggih yang dapat mempermudah mahasiswa dalam penulisan skripsi. Disisi lain kemudahan tersebut dapat menyebabkan dampak negatif seperti terjadinya praktik *copy-paste* sehingga pendapat orang lain dijadikan seolah-olah menjadi karangan sendiri. Hal tersebut dapat didefinisikan sebagai tindakan plagiarisme. Berdasarkan dari permasalahan yang ada, maka dibutuhkan suatu aplikasi yang dapat melakukan proses pemeriksaan kemiripan terhadap judul skripsi sehingga dapat memberikan kemudahan bagi program studi Teknik Informatika di Universitas Kebangsaan Republik Indonesia untuk melakukan pemeriksaan tingkat kemiripan pada

judul skripsi. Aplikasi yang dibangun dapat digunakan secara online sehingga mahasiswa dapat melakukan pengajuan dimana saja dan kapan saja. Aplikasi yang dibangun dapat menunjukkan presentase kemiripan judul skripsi dengan penelitian terdahulu sehingga pengajuan judul skripsi secara online dapat mempermudah mahasiswa dalam mengetahui tingkat kemiripan skripsi yang dibuatnya. Salah satu metode yang dapat digunakan adalah *text mining*. Pada *text mining* algoritma dapat diterapkan pada pendeteksian kemiripan skripsi menggunakan *Winnowing*. Wibowo dan Hastuti (2016) menyatakan bahwa “Algoritma *Winnowing* merupakan implementasi dari algoritma rabin-karp dengan penambahan metode *window* yang digunakan untuk mendeteksi plagiarisme dengan melakukan *preprocessing* meliputi *case folding*, *filtering*, *stemming*, dan *tokenizing*”.

Algoritma *Winnowing* dianggap layak diterapkan karena beberapa penelitian sebelumnya menyatakan bahwa algoritma *winnowing* memiliki keakuratan persentase yang lebih laik. “Pendekatan Algoritma *Winnowing* lebih baik daripada pendekatan algoritma Rabin Karp karena menghasilkan tingkat presentase yang lebih kecil dan waktu proses yang lebih cepat, Berdasarkan hasil Pengujian terhadap perbandingan pendekatan algoritma *winnowing* dengan algoritma rabin karp dapat dilihat kemungkinan kemiripan teks judul skripsi yang terkecil adalah dengan menggunakan pendekatan algoritma *winnowing* yaitu pada ujicoba ke 8 dengan nilai  $n\text{-gram} = 9$  dan  $window = 3$ , proses waktu 0.0257 dengan tingkat kemiripan terkecil yaitu 32.6 %” (Alamsyah, 2017). Hasil penelitian tersebut menjadi dasar pemilihan algoritma *winnowing* sebagai metode yang akan diterapkan pada aplikasi yang akan dibangun untuk mendeteksi tingkat kemiripan skripsi.

## 2. Definisi Aplikasi

Aplikasi secara umum adalah alat terapan yang difungsikan secara khusus dan terpadu sesuai kemampuan yang dimilikinya aplikasi merupakan suatu perangkat komputer yang siap pakai bagi user (Siregar dkk., 2018). Menurut Nurhayati dkk., (2018) Aplikasi adalah program siap pakai untuk melayani kebutuhan pengguna dalam berbagai aktifitas untuk pengolahan. Menurut Eka Noviansyah (2008: 4) dalam (Haerulah & Ismiyati, 2017) Aplikasi adalah penggunaan dan penerapan suatu konsep yang menjadi suatu pokok pembahasan. Aplikasi dapat diartikan juga sebagai program komputer

yang dibuat untuk menolong manusia dalam melakukan tugas tertentu.

Pada jurnal penelitian ini aplikasi akan dirancang dengan pendekatan objek dengan menggunakan UML (*Unified Modelling Language*) yang menurut Rosa A. S dan Shalahuddin (2015) “UML (*Unified Modelling Language*) adalah salah satu standar bahasa yang banyak digunakan didunia industri untuk mendefinisikan requirement, membuat analisis dan desain, serta menggambarkan arsitektur dalam pemrograman berorientasikan objek”.

Sedangkan algoritma yang digunakan adalah *winnowing* yang menurut Hidayat (2016) adalah algoritma yang digunakan untuk melakukan proses pemeriksaan kesamaan kata (*document fingerprinting*) yang tentunya akan digunakan untuk melakukan pengidentifikasian plagiarisme (penjiplakan). Secara teknis *winnowing* adalah ekstensi dari implementasi algoritma rabin-karp fingerprint dengan penambahan metode *window* (R. K. Wibowo & Hastuti, 2016). Jadi dapat disimpulkan jika Algoritma *Winnowing* adalah suatu proses pengecekan kemiripan kata pada dokumen (*document printing*) menggunakan penerapan algoritma rabin-karp dengan penambahan metode *window*. Metode ini diawali dengan proses *Preprocessing*, yaitu proses manipulasi teks dengan melakukan *case folding*, *filtering*, *stemming*, dan *tokenizing*. *Case folding* merupakan proses manipulasi *case-sensitive*, semua input teks data akan diubah menjadi huruf kecil/*lower-case*. Sedangkan *filtering* atau sering dikenal dengan istilah *stopword removal* adalah proses penghapusan kata yang tidak relevan dalam teks. *Stemming* ialah proses pemisahan kata menjadi kata dasar. Tahap berikutnya adalah *Tokenizing* yang merupakan proses pemisahan kata berdasarkan susunan kata (R. K. Wibowo & Hastuti, 2016). Setelah ditemukan rangkaian string maka langkah berikutnya adalah membentuk rangkaian hash (string menjadi integer). Jika rangkaian hash telah ditemukan maka dilanjutkan dengan proses *winnowing* yang akan menghasilkan fingerprint untuk digunakan dalam pencocokan kata.

Sedangkan K-Gram merupakan serangkaian token yang memisahkan teks menjadi string mulai dari posisi tertentu dalam teks dengan panjang k. Dalam konteks komputasi secara linguistik, token ini bisa berupa kata-kata, meskipun bisa berupa karakter atau subset karakter. Jumlah token mempengaruhi nilai k untuk mengambil potongan-potongan karakter huruf secara

kontinuitas dan membentuk substring sepanjang k karakter dari sebuah string.

Selanjutnya diperlukan satu proses berikutnya yang dinamakan *Rolling Hash* yang merupakan cara untuk mengubah string menjadi nilai unik dengan panjang tertentu dengan memberikan kemampuan untuk menghitung nilai hasil tanpa mengulang seluruh string yang berfungsi sebagai penanda string. Dari proses rolling hash akan menghasilkan nilai hash yang digunakan untuk proses *window*.

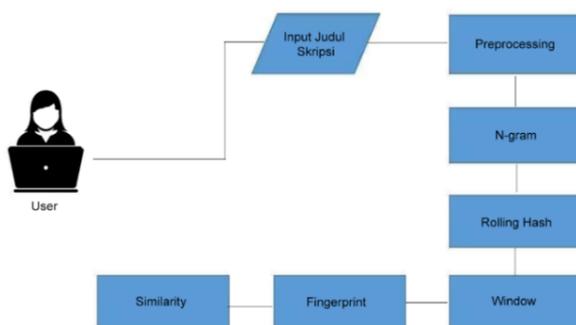
Selanjutnya *Jaccard Similarity Coefficient* merupakan persamaan yang digunakan untuk mengukur prosentase kemiripan dokumen dengan cara menghitung atau mengukur nilai persamaan atau pertidaksamaan dalam suatu dokumen. *Jaccard Similarity Coefficient* menggunakan kumpulan kata-kata yang telah dihitung nilai hash-nya dalam menentukan persentase kemiripan suatu berkas atau pada suatu dokumen.

### 3. Metodologi Penelitian

Pada jurnal ini diterapkan beberapa tahapan dalam melakukan implementasi algoritma winnowing sebagai pemeriksaan kemiripan judul skripsi adalah sebagai berikut :

- Pengumpulan Data
- Pengumpulan data-data dokumen skripsi yang ada pada Fakultas Ilmu Komputer dan Sistem Informasi Universitas Kebangsaan.
- Pemrosesan Data Menggunakan Algoritma Winnowing

Berikut langkah-langkah penerapan algoritma winnowing:



Gambar 1. Proses Algoritma Winnowing

Pada Gambar 1 digambarkan alur proses yang terjadi pada Algoritma Winnowing. Berikut penjelasan dari tahapan yang terjadi pada Algoritma Winnowing :

- Pembuangan karakter yang tidak relevan

- Langkah ini merupakan tahap penghapusan tanda baca, spasi dan simbol-simbol khusus pada judul skripsi hanya menghilangkan spasi

- Pembentukan rangkaian n-gram

Pembentukan rangkaian n-gram dilakukan dengan cara membentuk rangkaian karakter sepanjang n dari hasil pembuangan karakter yang tidak relevan dengan ukuran n.

- Perhitungan fungsi hash pada untuk setiap n-gram

Yaitu melakukan perhitungan nilai-nilai hash dari setiap gram, fungsi yang digunakan untuk menghasilkan nilai hash dari rangkaian gram dalam algoritma Winnowing adalah rolling hash. Rolling hash merupakan suatu cara untuk mentransformasi sebuah string menjadi suatu nilai yang unik dengan panjang tertentu (*fixed-length*) yang berfungsi sebagai penanda string tersebut. Fungsi hash  $H(c_1..c_k)$  didefinisikan sebagai berikut:

$$H_{(ck)} = c_1 * b^{(k-1)} + c_2 * b^{(k-2)} + \dots + c_k * b^{(k-k)}$$

Keterangan :

c = nilai ascii karakter

b = basis (bilangan prima)

k = banyak karakter

- Pembentukan *window* dari nilai hash

- Pemilihan *fingerprint* dari setiap window

Menentukan nilai *fingerprint* teks. Nilai *fingerprint* ditentukan dengan memilih nilai hash terkecil dari setiap window.

Persamaan *Jaccard Coeficient Similarity*

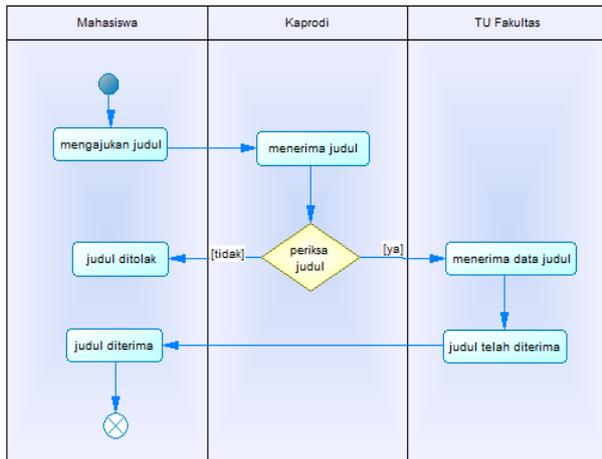
Nilai *fingerprint* yang dibentuk dari algoritma winnowing digunakan untuk mengukur presentase kemiripan teks menggunakan persamaan *Jaccard Coeficient*.

### 4. Pembahasan

#### 4.1. Sistem Yang Sedang Berjalan

Berikut gambaran dari proses yang sedang berjalan pada pengajuan judul skripsi di Universitas Kebangsaan Republik Indonesia.

Proses awal adalah mahasiswa mengajukan judul ke Kaprodi, lalu Kaprodi akan memeriksa kelayakan judul yang diajukan oleh mahasiswa. Jika judul tersebut masih belum layak maka judul akan ditolak. Namun jika judul diterima maka akan diteruskan ke TU Fakultas untuk pengarsipan dan pembuatan SK. Untuk lebih jelasnya berikut tergambarakan melalui diagram aktivitas dibawah ini :

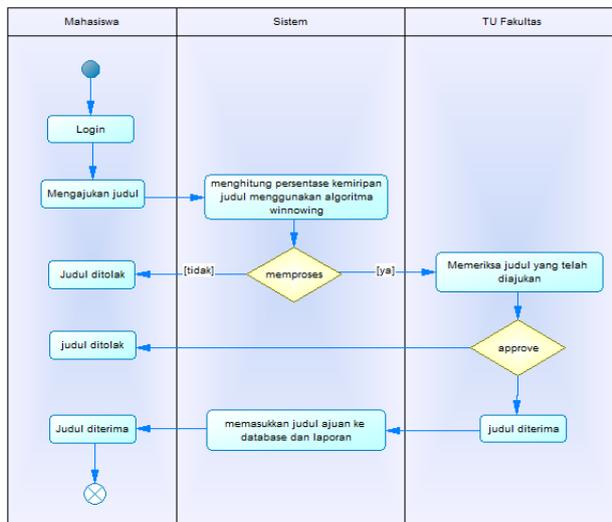


Gambar 1 Activity Diagram Sistem Yang Berjalan

Pada sistem diatas terdapat kelemahan karena tidak ada proses pemeriksaan terhadap usulan judul yang diusulkan mahasiswa. Untuk itu perlu dibuat prosedur untuk memperbaiki sistem yang berjalan dengan menambahkan aktivitas pemeriksaan usulan judul skripsi mahasiswa.

#### 4.2. Rancangan Pengembangan Sistem

Berikut rancangan sistem usulan pada pengajuan judul skripsi di Universitas Kebangsaan Republik Indonesia.

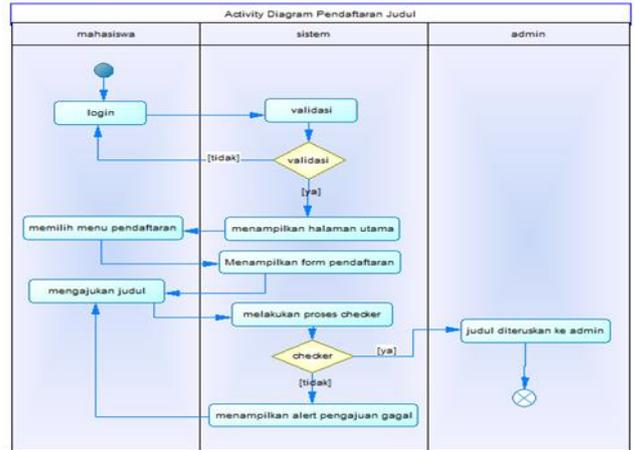


Gambar 3. Activity Diagram Sistem Usulan

Mahasiswa melakukan login ke sistem dan mengajukan judul. Sistem akan memproses judul ajuan mahasiswa dengan melakukan penghitungan persentase kemiripan judul skripsi menggunakan algoritma winnowing. Jika judul tersebut mempunyai persentase yang melebihi batas toleran maka judul tersebut akan ditolak secara otomatis oleh sistem. Namun, jika judul yang diajukan memiliki persentase kemiripan dibawah batas toleran maka judul tersebut akan diteruskan ke admin untuk proses approve. Pada

proses approve ini, admin akan memeriksa judul yang diajukan oleh mahasiswa apakah telah sesuai atau tidak. Jika tidak sesuai maka di interface mahasiswa akan menampilkan informasi judul telah ditolak. Apabila judul diterima maka judul tersebut akan diteruskan oleh sistem untuk dimasukkan ke database serta laporan dan menampilkan informasi pada interface mahasiswa bahwa judul telah diterima.

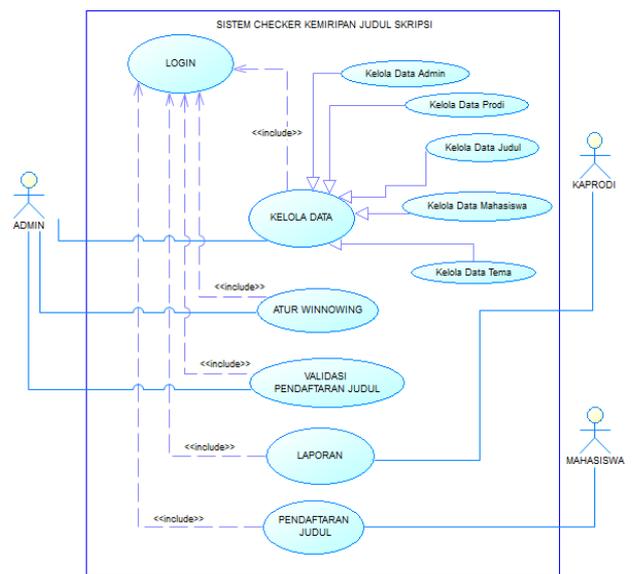
Activity Diagram Pendaftaran Judul merupakan diagram yang menjelaskan alur yang terjadi pada saat mahasiswa ingin melakukan pendaftaran/pengajuan judul.



Gambar 4. Activity Diagram Pendaftaran Judul

#### 4.3. Use Case Diagram

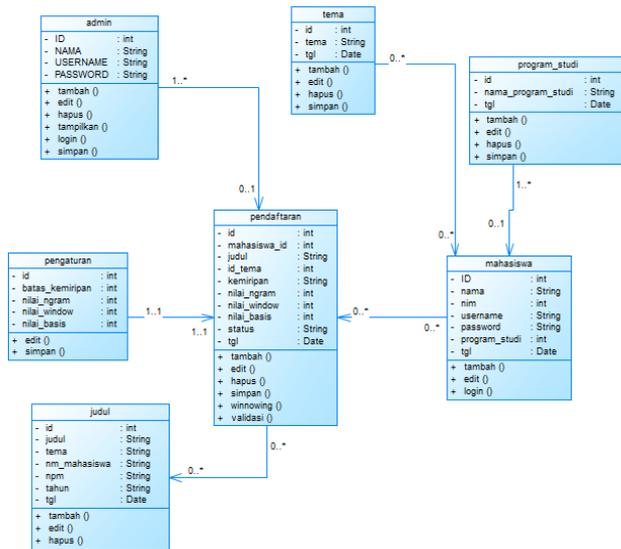
Untuk memberikan gambaran yang lebih jelas mengenai sistem tersebut maka dibawah ini akan digambarkan fungsionalitas sistem dengan menggunakan diagram use case. Berikut Use case diagram yang dirancang untuk membangun aplikasi tersebut



Gambar 5. Diagram Use Case Fungsionalitas Sistem Usulan

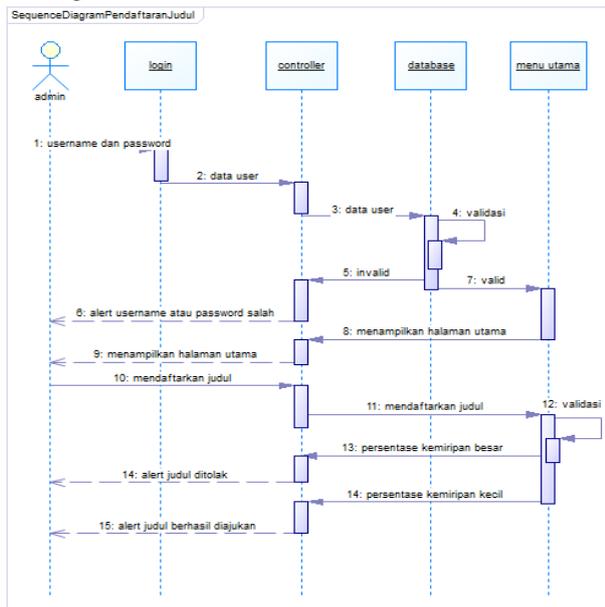
#### 4.4. Diagram Class

Untuk menunjukkan interaksi antar objek pada sistem tersebut, maka perlu digambarkan dalam diagram class, sebagai berikut :



Gambar 6. Diagram Class Sistem Usulan

Kemudian dilengkapi dengan diagram sequence untuk menunjukkan interaksi antar objek dengan ditunjukkan melalui perilaku sistem (antar objek), sebagai berikut :

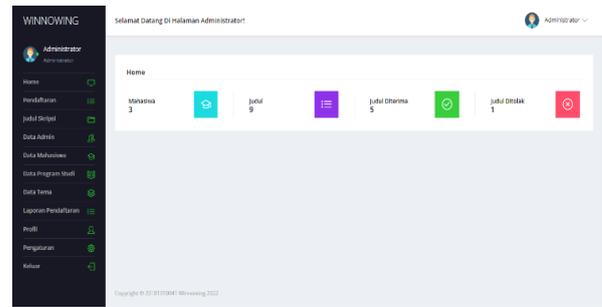


Gambar 7. Sequence Diagram Pendaftaran Judul

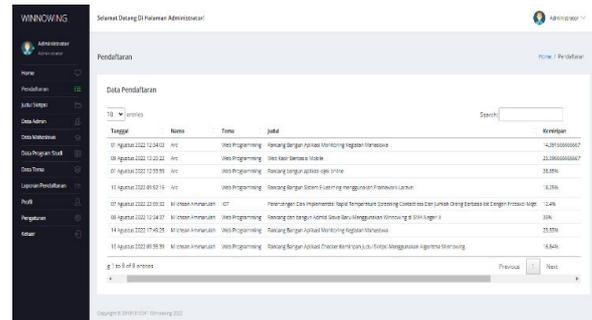
#### 4.5. Implementasi

Sistem usulan ini diimplementasikan dengan perangkat lunak berbasis Web dengan bahasa yang digunakan berbasis PHP dengan penggunaan framework Ci 4.

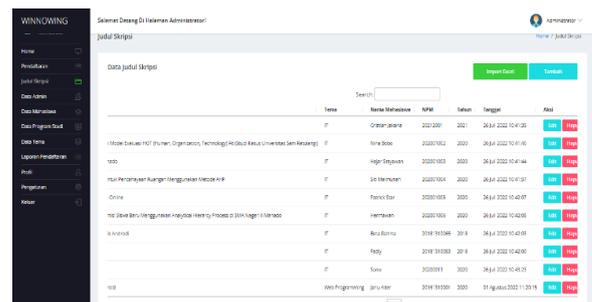
Berikut merupakan implementasi dari rancangan antarmuka sistem yang dibangun dengan urutan sebagai berikut :



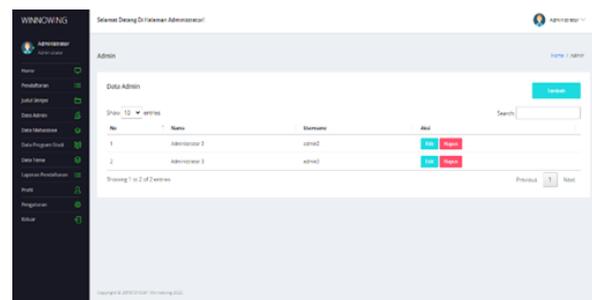
Gambar 8. Halaman Utama (Admin)



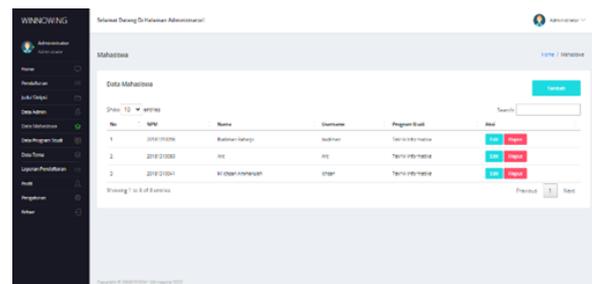
Gambar 9. Halaman Pendaftaran (Admin)



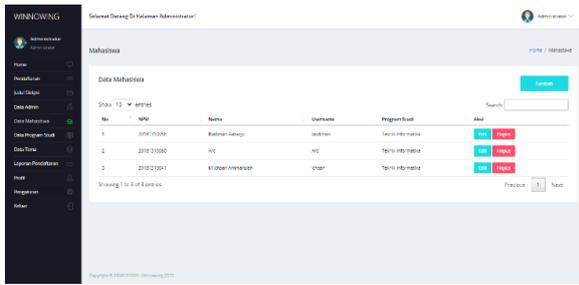
Gambar 10. Halaman Pendaftaran (Admin)



Gambar 11. Halaman Judul (Admin)



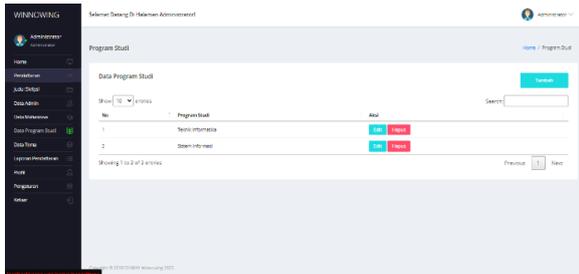
Gambar 12 Halaman Data Admin (Admin)



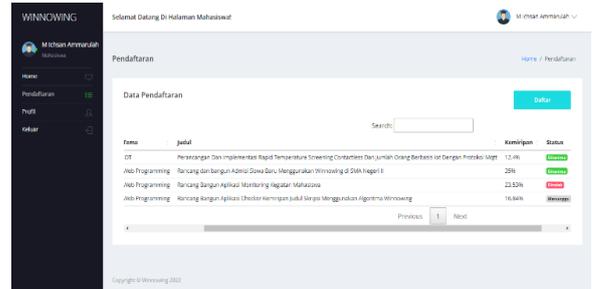
Gambar 14. Halaman Data Mahasiswa (Admin)



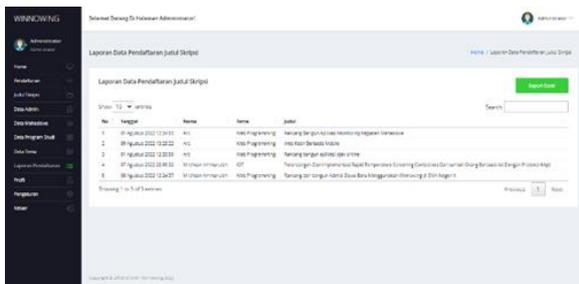
Gambar 18. Halaman Utama (Mahasiswa)



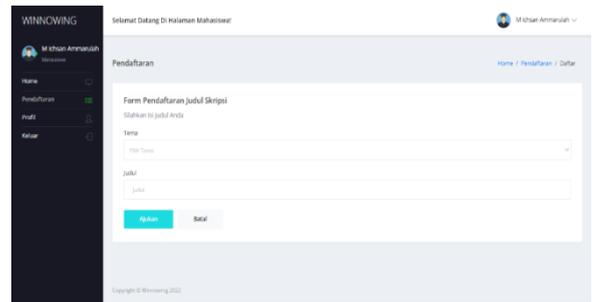
Gambar 15. Halaman Data Prodi (Admin)



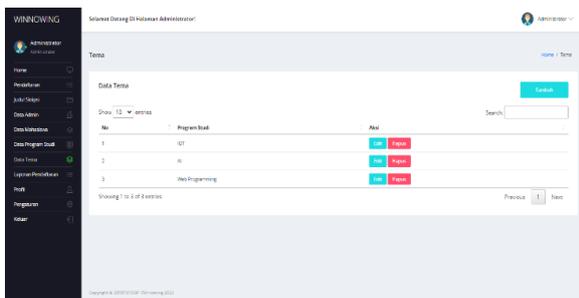
Gambar 19. Halaman Data Pendaftaran (Mahasiswa)



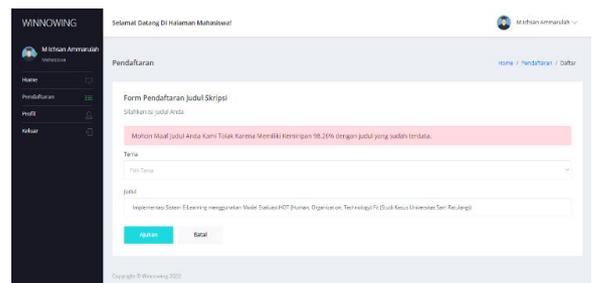
Gambar 15. Halaman Laporan Pendaftaran (Admin)



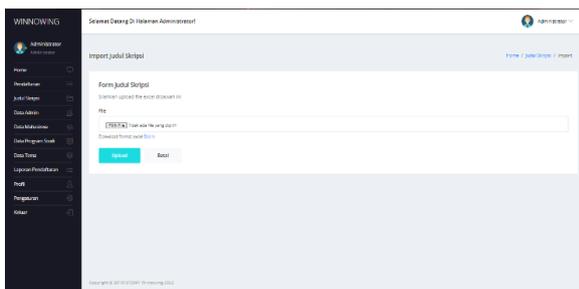
Gambar 20. Halaman Pengajuan Judul (Mahasiswa)



Gambar 16. Halaman Data Tema (Admin)



Gambar 21. Pengujian k-gram 5

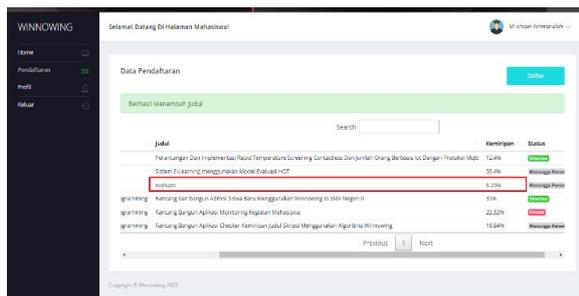


Gambar 17. Halaman Import Judul Skripsi (Admin)

#### 4.6. Pengujian Nilai K-gram

Pengujian menggunakan parameter :  
 $K\text{-gram} = 5$  ,  $W\text{-gram} = 4$  , dan nilai Basis = 3  
 Menggunakan teks pembanding yaitu :  
 “Evaluasi Implementasi Sistem E-Learning menggunakan Model Evaluasi HOT (Human, Organization, Technology) Fit (Studi Kasus Universitas Sam Ratulangi)” dan hasilnya adalah persentase menjadi tergolong stabil, hal ini ditunjukkan jika melakukan pengurangan 1 kata hasil persentase tidak melebihi 100%.

Jika hanya menggunakan satu kata, tetap menunjukkan perhitungan besaran kemiripan sebesar 6.25%



Gambar 22. Pengujian k-gram 5

Hal tersebut menunjukkan bahwa nilai k-gram 5 memiliki tingkat keakuratan yang lebih baik dibandingkan nilai k-gram yang diuji sebelumnya.

## 5. Kesimpulan

Berdasarkan hasil penelitian dari permasalahan diatas, maka didapatkan kesimpulan meliputi :

- a. Aplikasi yang dibangun telah sejalan dengan tujuan awal yaitu membangun suatu aplikasi yang dapat digunakan mahasiswa untuk melakukan pengajuan judul secara online dan fungsi checker kemiripan menggunakan algoritma winnowing.
- b. Aplikasi yang dibangun dapat digunakan oleh pihak fakultas untuk melakukan kegiatan kelola data judul skripsi, serta dengan adanya checker kemiripan didalam aplikasi membuat data judul yang diajukan oleh mahasiswa memiliki validitas persentase kemiripan yang kecil sehingga terhindar dari plagiarisme pada perpustakaan Fakultas Ilmu Komputer dan Sistem Informasi di Universitas Kebangsaan.
- c. Pemilihan nilai n-gram berpengaruh besar dalam perhitungan persentase tingkat kemiripan.
- d. Algoritma winnowing dapat diterapkan pada aplikasi berbasis web dan digunakan sebagai checker kemiripan judul skripsi serta memiliki waktu perhitungan yang cepat. Selain itu, skala penelitian yang terbatas pada judul skripsi membuat algoritma ini cocok digunakan karena proses preprocessing memiliki hasil akurasi yang tinggi jika konten yang dideteksi sangat spesifik dan terhindar dari typo.

## Daftar Pustaka

- [1] Alamsyah, N. (2017). Perbandingan Algoritma Winnowing Dengan Algoritma Rabin Karp Untuk Mendeteksi. 8(3), 124–134.
- [2] Haerulah, E., & Ismiyati, S. (2017). Aplikasi e-commerce penjualan souvenir pernikahan pada toko “XYZ.” *Jurnal Prosisko*, 4(1), 43–47.  
<https://ejournal.lppmunsera.org/index.php/PROSISKO/article/download/146/208>
- [3] Hidayat, A. (2016, April 14). Algoritma Winnowing.  
<https://arfianhidayat.com/algoritma-winnowing>
- [4] Mashabi, S. (2021, September 20). Uji Kelayakan di Komisi III, Calon Hakim Agung Ini Dianggap Lakukan Plagiat.  
<https://nasional.kompas.com/read/2021/09/20/13394501/uji-kelayakan-di-komisi-iii-calon-hakim-agung-ini-dianggap-lakukan-plagiat>
- [5] Nurdin, & Munthoha, A. (2017). Sistem Pendeteksian Kemiripan Judul Skripsi Menggunakan Algoritma winnowing. *Jurnal Nasional Informatika dan Teknologi Jaringan*, Vol 2, No, 90–97.
- [6] Nurhayati, A. N., Josi, A., & Hutagalung, N. A. (2018). Rancang Bangun Aplikasi Penjualan Dan Pembelian Barang Pada Koperasi Kartika Samara Grawira Prabumulih. *Jurnal Teknologi dan Informasi*, 7(2), 13–23.  
<https://doi.org/10.34010/jati.v7i2.490>
- [7] Rosa A. S, & Shalahuddin, M. (2015). *Rekayasa Perangkat Lunak Terstruktur dan Berorientasi Objek*. Informatika.
- [8] Siregar, H. F., Siregar, Y. H., & Melani. (2018). Perancangan Aplikasi Komik Hadist Berbasis Multimedia. *JurTI (Jurnal Teknologi Informasi)*, 2(2), 113–121. *JurTI (Jurnal Teknologi Informasi)*, 2(2), 113–121.  
<http://www.jurnal.una.ac.id/index.php/jurti/article/view/425>