

PERANCANGAN APLIKASI PENDETEKSI TINGKAT KESAMAAN ANTAR DOKUMEN DENGAN ALGORITMA WINNOWING

Arnold Pramudita Tjiawi, Dyah E. Herwindiati, Lely Hiryanto
Teknik Informatika Universitas Tarumanagara
Jl. Letjen S.Parman No.1, Jakarta 11440 Indonesia
email : tjiawi.rush@gmail.com

Abstrak

Perancangan aplikasi pendeteksian tingkat kesamaan antar dokumen ini dimaksudkan untuk menyimpan data tesis mahasiswa dan aplikasi dapat digunakan untuk membandingkan data skripsi mahasiswa dengan data baru yang ada. Hasil dari proses deteksi adalah persentase tingkat kesamaan. Algoritma yang digunakan dalam merancang aplikasi ini adalah Algoritma Winnowing. Algoritma ini termasuk dalam salah satu metode fingerprinting dokumen. Algoritma ini menggunakan rolling hashing untuk melakukan proses hashing dan menggunakan koefisien jaccard untuk menghitung tingkat kemiripan. Dalam pengujian aplikasi ini dengan data dummy diketahui bahwa aplikasi ini sangat bergantung pada urutan masing-masing lokasi substring. Aplikasi ini juga telah diuji untuk membandingkan sejumlah data skripsi yang ada dengan persentase tingkat kesamaan di bawah 30%. Aplikasi ini tidak bisa menentukan apakah sebuah dokumen plagiat atau tidak, namun aplikasi ini bisa memberikan informasi berupa persentase tingkat kesamaan dan persentase dapat digunakan oleh pihak yang memiliki wewenang untuk menentukan.

Kata kunci—Algoritma Winnowing, plagiarisme, rolling hashing, tingkat kesamaan antar dokumen, jaccard coefficient

Abstract

The design of detection application for similarity level between documents is intended to store student thesis data and applications can be used to compare a student thesis data with existing with new data. The result of the detection process is the percentage of similarity level. The algorithm used in designing this application is winnowing algorithm. This algorithm is included in one of the document fingerprinting methods. This algorithm uses rolling hashing to do the hashing process and use jaccard coefficient to calculate the similarity level. In testing this application with dummy data it is found that this application is very dependent on the order of each substring location. This application has also been tested to compare a number of existing thesis data with a percentage of similarity rate below 30%. This application can not specify a plagiarism document or not, but this application can provide information in the form of percentage level of similarity and percentage can be used by who has the authority to determine.

Keywords—Document Fingerprinting, Jaccard Coefficient, Plagiarism, Rolling Hashing, Similarity, Winnowing algorithm

1. PENDAHULUAN

Penulisan karya ilmiah dalam lingkup pendidikan tinggi merupakan suatu hal yang wajib ada. Pada jenjang Strata satu (S1) dikenal makalah dan skripsi (tugas akhir). Dalam penulisan skripsi biasa diperlukan data ataupun informasi-informasi pendukung. Penulisan karya ilmiah tidak terlepas dari sumber-sumber pustaka yang dijadikan acuan penulisan. Sumber-sumber pustaka didapat dari buku, koran, majalan jurnal dan media tulis lainnya. Data dan informasi ini dapat dengan mudah diakses dengan adanya kemajuan teknologi informasi yang ada sekarang. Tetapi kemudahan ini memiliki efek negatif, seperti penjiplakan terhadap karya ilmiah orang lain. Hal ini berbahaya dan merugikan bagi pemilik karya ilmiah maupun bagi institusi tempat penjiplak berada.

Kegiatan penjiplakan ini biasa disebut sebagai plagiarisme. Plagiarisme ini sulit untuk di deteksi jika dengan kasat mata, karena banyaknya informasi-informasi yang biasa terdapat dalam sebuah karya ilmiah. Untuk itu diperlukan sebuah perancangan aplikasi pendeteksi tingkat kesamaan antar dokumen. Aplikasi yang dirancang ini akan menggunakan algoritma winnowing. Algoritma ini menggunakan *rolling hashing* untuk melakukan *hashing* dan menggunakan *jaccard coefficient* untuk menghitung tingkat kesamaan. Hasil yang diharapkan dari aplikasi ini adalah sebuah prosentase tingkat kesamaan antar dokumen satu dengan dokumen lainnya, sehingga prosentase ini dapat digunakan untuk mengambil keputusan oleh pemilik otoritas.

2. DASAR TEORI

2.1 Plagiarisme

Plagiarisme merupakan sebuah tindakan plagiat atau melakukan plagiat. Plagiat atau penjiplakan adalah pengambilan karangan, pendapat, dan sebagainya orang lain dan menjadikannya seolah karangan dan pendapat sendiri. Misalnya menerbitkan karya tulis orang lain yang belum sempat diterbitkan oleh penulisnya atas nama diri sendiri[1]. Metode untuk mendeteksi plagiarisme dapat dibagi menjadi tiga bagian, yaitu metode perbandingan teks lengkap, metode dokumen fingerprinting, dan metode kesamaan kata kunci[2]. Metode yang digunakan dalam perancangan ini adalah metode dokumen *fingerprinting*. Metode ini menggunakan teknik hashing, berikut adalah beberapa algoritmanya yaitu *rabin-karp*, *winnowing*, dan *manber*. Pada perancangan ini digunakan algoritma winnowing.

2.2 Algoritma Winnowing

Algoritma winnowing merupakan salah satu metode dalam dokumen fingerprinting. Input dari algoritma ini adalah dokumen teks yang diproses sehingga mendapatkan output nilai hash. Algoritma yang digunakan untuk mendapatkan nilai hash pada winnowing adalah *rolling hashing*. Berikut ini tahapan proses dalam algoritma winnowing :

1. Text Pre-processing

Dalam proses ini dokumen teks diolah agar menjadi data siap proses. Pengolahan ini meliputi spasi, jenis huruf (kapital atau normal), tanda baca dan sebagainya (*whitespace insensitivity*).

Contoh teks:

Indonesia tanah air Beta.

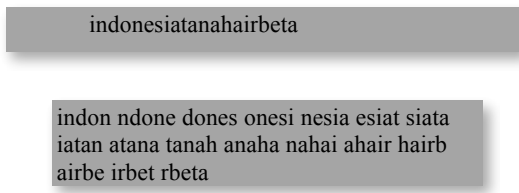
Indonesia tanah air Beta.

indonesiatanahairbeta



2. Membentuk rangkaian k-grams dari teks

Percobaan menggunakan nilai k-grams = 5, berikut ini contohnya :

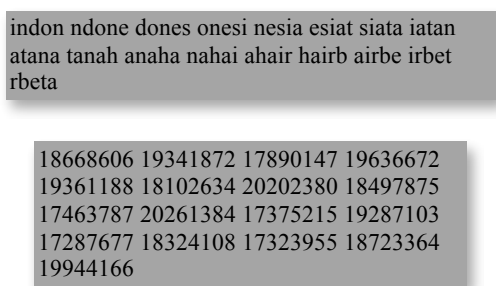


3. Melakukan rolling hashing untuk setiap k-grams

Persamaan (1) adalah perhitungan fungsi *hash* dari algoritma *winning* dengan *c* sebagai nilai ASCII, *b* sebagai nilai *basis* bilangan prima, dan banyak karakter *k* [3] :

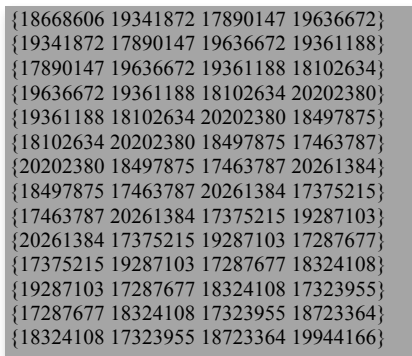
$$H_{(c_1...c_k)} = c_1 * b^{(k-1)} + c_2 * b^{(k-2)} + \dots + c_{n-1} * b + c_k \tag{1}$$

Berikut ini contoh pernggunaan *rolling hashing* pada dokumen teks yang sudah melewati proses nomor 2. Dengan menggunakan nilai *b* = 11.



4. Memilih nilai *fingerprint* dari hasil *hashing* dengan pembagian hasil *hash* berdasarkan satu nilai *window* merupakan hasil dari perhitungan $w = t-k+1$ [3].

Misalkan nilai *t* = 8, sehingga $w = 8-5+1 = 4$. Dalam satu ukuran *window* dapat terisi dengan 4 nilai *hash*.



Setelah *window* terbentuk sebanyak 14 buah, maka dicarilah nilai *fingerprint* dengan melihat nilai terkecil dari setiap *window*. Maka didapati nilai *fingerprint* : 17890147 18102634 17463787 17375215 17287677 17323955. Dari 14 *window* hanya terdapat 6 *fingerprint*, hal ini disebabkan karena ada beberapa *window* yang memiliki nilai terendah yang sama sehingga tidak perlu dituliskan kembali.

5. Menghitung tingkat kemiripan

Untuk menghitung tingkat kemiripan diperlukan minimal dua buah dokumen teks untuk dibandingkan. Carany dengan mengikuti dari proses ke-1 hingga mendapatkan *fingerprint* dari setiap dokumen teks, lalu akan digunakan *jaccard coefficient* untuk menghitung kesamaannya.

$$\text{Similarity}(d_i, d_j) = \frac{|w(d_i) \cap w(d_j)|}{|w(d_i) \cup w(d_j)|} \times 100\% \quad (2)$$

Misalkan file teks D_1 dan D_2 dengan *fingerprint* $D_1 = 17890147, 18102634, 17463787, 17375215, 17287677, 17323955$ dan *fingerprint* $D_2 = 17890147, 18102524, 17318708$ maka sesuai dengan persamaan (2) diperoleh nilai kesamaan sebesar 12.5%

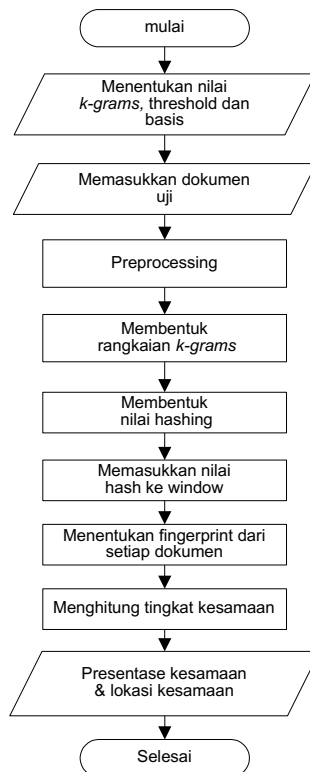
$$\text{Similarity}(d_i, d_j) = \frac{|1|}{|8|} \times 100\% = 12.5\%$$

2.3 Penentuan Nilai Parameter

Penentuan nilai parameter untuk pengujian sudah dilakukan berdasarkan pengujian yang sudah pernah dilakukan oleh Milani Winangga, Drs. Marji, M.T., dan Drs. Achmad Ridok, M.Kom. Kinerja terbaik diperoleh pada konfigurasi parameter nilai $k\text{-gram}=5$, $\text{threshold}=25$, dan $\text{basis}=11$ dengan hasil tingkat prosentase error yang berbanding lurus dengan prosentase perubahan dokumen [4].

2.4 Sistem yang Dibuat

Sistem aplikasi pendeteksi tingkat kesamaan antar dokumen dengan algoritma *winnowing* ini memiliki alur yang dapat dilihat pada Gambar 1. Pada proses pendeteksian diperlukan untuk menentukan nilai $k\text{-grams}$, *guarantee threshold (threshold)*, dan *basis*. Setelah itu user harus memasukkan dokumen uji dengan format ekstensi *.docx* serta memilih dokumen pembanding yang sudah ada datanya dalam aplikasi. Selanjutnya proses akan menjalankan proses dengan algoritma *winnowing*. Jika aplikasi sudah menjalankan urutan proses dalam algoritma *winnowing* maka akan didapati hasil prosentase tingkat kesamaan serta menampilkan kata yang sama. Prosentase ini dapat dilihat dan digunakan sebagai data untuk menentukan apakah dokumen uji tersebut plagiat atau tidak. Penentuan ini ditentukan oleh pihak yang memiliki otoritas dalam menentukan dan aplikasi hanya menampilkan prosentase tingkat kesamaan.



Gambar 1 Alur kerja sistem yang di buat

3. HASIL DAN PEMBAHASAN

Pengujian dilakukan terhadap aplikasi pendeteksi tingkat kesamaan antar dokumen dengan algoritma *winnowing* ini. Pengujian dibagi menjadi 2 yaitu pengujian terhadap modul dan pengujian terhadap data. Pengujian terhadap modul ini dilakukan untuk memastikan bahwa setiap modul dapat berjalan dan berfungsi dengan baik. Sedangkan untuk pengujian terhadap data dilakukan untuk memastikan bahwa penerapan algoritma *winnowing* pada aplikasi ini sudah berjalan sesuai dengan yang diharapkan.

Hasil pengujian terhadap modul menghasilkan bahwa setiap modul dapat berjalan dengan baik. Berikut ini daftar modul yang ada dalam aplikasi :

- Modul Login
- Modul Home
- Modul Mahasiswa
- Modul Dosen
- Modul Archieve
- Modul Deteksi
- Modul Hasil
- Modul Unggah Skripsi
- Modul Log Out

Pengujian terhadap data dilakukan dengan 2 tahap, yaitu tahap pertama dengan data dummy dan tahap kedua dengan data skripsi asli. Pengujian dengan data dummy ini untuk memastikan bahwa algoritma dalam aplikasi berjalan sesuai yang diharapkan dan untuk mencoba kinerja aplikasi. Sedangkan untuk pengujian terhadap data skripsi asli untuk memastikan bahwa aplikasi dapat berjalan sesuai dengan yang diharapkan.

Untuk pengujian menggunakan data dummy telah dilakukan dengan sebuah dokumen teks yang berisi sebuah paragraf yang diatur komposisinya sebagai berikut :

- Data dummy 100% sama persis
- Data dummy dipotong 20% kata, sehingga tersisa 80% kata
- Data dummy dipotong 40% kata, sehingga tersisa 60% kata
- Data dummy dipotong 60% kata, sehingga tersisa 40% kata
- Data dummy dipotong 80% kata, sehingga tersisa 20% kata
- Data dummy diacak 20% kata, pengacakan dilakukan secara manual
- Data dummy diacak 60% kata, pengacakan dilakukan secara manual

Hasil dari pengujian data dummy dapat dilihat pada Gambar 2.

#	FILE 1	FILE 2 (ARCHIVE)	NILAI K	NILAI B	NILAI W	PERSENTASE
38.	dummy 100.docx	dummy 100.docx	5	11	21	100
39.	dummy 80 (potong 20).docx	dummy 100.docx	5	11	21	76.92
40.	dummy 60 (potong 40).docx	dummy 100.docx	5	11	21	61.54
41.	dummy 40 (potong 60).docx	dummy 100.docx	5	11	21	38.46
42.	dummy 20 (potong 80).docx	dummy 100.docx	5	11	21	15.38
43.	dummy acak 20.docx	dummy 100.docx	5	11	21	53.13
44.	dummy acak 60.docx	dummy 100.docx	5	11	21	23.68

Gambar 2 Hasil pengujian data dummy

Pengujian terhadap data skripsi asli menggunakan laporan skripsi dari 18 orang mahasiswa. Laporan tersebut dipecah menjadi 3 bagian yaitu BAB I, BAB III, dan BAB IV. Pengujian dilakukan dengan mengambil data BAB I seorang mahasiswa dan dibandingkan

dengan 18 data lainnya. Berikut ini hasil pengujian BAB I bisa dilihat di Gambar 3, pengujian pada BAB III dapat dilihat di Gambar 4, dan pengujian pada BAB IV dapat dilihat di Gambar 5

#	FILE 1	FILE 2 (ARCHIEVE)	NILAI K	NILAI B	NILAI W	PERSENTASE
8.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Ayu Windy Astuti).docx	5	11	21	22.56
9.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Elizabeth Erlsha).docx	5	11	21	14.81
10.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Farencio).docx	5	11	21	19.62
11.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Ferry Dharmawan).docx	5	11	21	13.41
12.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Fransisca Regina).docx	5	11	21	17.1
13.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Irawati Djajadi).docx	5	11	21	19.86
14.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Jacklin Shintia Thio).docx	5	11	21	21.32
15.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Josselyn Shintia Thio).docx	5	11	21	20.84
16.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Mariana).docx	5	11	21	21.8
17.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Mishelle Tirta Winarta).docx	5	11	21	18.53
18.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Nadia Yanitra).docx	5	11	21	18.07
19.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Renaldo Ali).docx	5	11	21	15.66
20.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Rionaldy Triasaputra).docx	5	11	21	13.44
21.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Rosalinda).docx	5	11	21	17.28
22.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Paryanka Victorius Abadi).docx	5	11	21	100
23.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Sri Whisnu A.W).docx	5	11	21	26.64
24.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Vina Tendeen).docx	5	11	21	18.12
25.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Yunita).docx	5	11	21	19.07

Gambar 3 Pengujian pada BAB I

#	FILE 1	FILE 2 (ARCHIEVE)	NILAI K	NILAI B	NILAI W	PERSENTASE
26.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Ayu Windy Astuti).docx	5	11	21	26.3
27.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Elizabeth Erlsha).docx	5	11	21	18.08
28.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Farencio).docx	5	11	21	19.6
29.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Ferry Dharmawan).docx	5	11	21	20.15
30.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Fransisca Regina).docx	5	11	21	20.55
31.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Irawati Djajadi).docx	5	11	21	18.54
32.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Mariana).docx	5	11	21	21.04
33.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Mishelle Tirta Winarta).docx	5	11	21	19.48
34.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Nadia Yanitra).docx	5	11	21	18.66
35.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Renaldo Ali).docx	5	11	21	15.63
36.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Rionaldy Triasaputra).docx	5	11	21	15.74
37.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Rosalinda).docx	5	11	21	18
38.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Paryanka Victorius Abadi).docx	5	11	21	100
39.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Jacklin Shintia Thio).docx	5	11	21	21.91
40.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Josselyn Shintia Thio).docx	5	11	21	20.52
41.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Sri Whisnu A.W).docx	5	11	21	29.11
42.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Vina Tendeen).docx	5	11	21	20.84
43.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Yunita).docx	5	11	21	20.09

Gambar 4 Pengujian pada BAB III

#	FILE 1	FILE 2 (ARCHIEVE)	NILAI K	NILAI B	NILAI W	PERSENTASE
44.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Ayu Windy Astuti).docx	5	11	21	19.52
45.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Elizabeth Erlsha).docx	5	11	21	10.83
46.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Farencio).docx	5	11	21	19.9
47.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Ferry Dharmawan).docx	5	11	21	12.97
48.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Fransisca Regina).docx	5	11	21	16.13
49.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Irawati Djajadi).docx	5	11	21	14.97
50.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Jacklin Shintia Thio).docx	5	11	21	22.23
51.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Josselyn Shintia Thio).docx	5	11	21	19.42
52.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Mariana).docx	5	11	21	23.08
53.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Mishelle Tirta Winarta).docx	5	11	21	14.91
54.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Nadia Yanitra).docx	5	11	21	12.89
55.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Renaldo Ali).docx	5	11	21	13.29
56.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Rionaldy Triasaputra).docx	5	11	21	10.06
57.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Rosalinda).docx	5	11	21	13.25
58.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Paryanka Victorius Abadi).docx	5	11	21	100
59.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Sri Whisnu A.W).docx	5	11	21	25.35
60.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Vina Tendeen).docx	5	11	21	15.73
61.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Yunita).docx	5	11	21	15.15

Gambar 5 Pengujian pada BAB IV

Pengujian selanjutnya adalah dengan membandingkan beberapa skripsi mahasiswa yang memiliki topik yang mirip, berikut ini daftarnya :

- a. Pengujian antara Skripsi Farenco dengan Paryanka Victorius Abadi dapat dilihat pada Gambar 6.

FILE 1	FILE 2 (ARCHIEVE)	NILAI K	NILAI B	NILAI W	PERSENTAS
BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Farenco).docx	5	11	21	19.62
BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Farenco).docx	5	11	21	19.6
BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Farenco).docx	5	11	21	19.9

Gambar 12 Pengujian Skripsi Farenco dengan Paryanka

- b. Pengujian antara Skripsi Fransisca Regina dengan Vina Tendean dapat dilihat pada Gambar 7.

FILE 1	FILE 2 (ARCHIEVE)	NILAI K	NILAI B	NILAI W	PERSENTASE
BAB 1 (Vina Tendean).docx	BAB 1 (Fransisca Regina).docx	5	11	21	27.97
BAB 3 (Vina Tendean).docx	BAB 3 (Fransisca Regina).docx	5	11	21	36.81
BAB 4 (Vina Tendean).docx	BAB 4 (Fransisca Regina).docx	5	11	21	22.37

Gambar 7 Pengujian Skripsi Fransisca Regina dengan Vina Tendean

- c. Pengujian antara Skripsi Ferry Dharmawan dengan Renaldo Ali dapat dilihat pada Gambar 8.

FILE 1	FILE 2 (ARCHIEVE)	NILAI K	NILAI B	NILAI W	PERSENTASE
BAB 1 (Ferry Dharmawan).docx	BAB 1 (Renaldo Ali).docx	5	11	21	30.36
BAB 3 (Ferry Dharmawan).docx	BAB 3 (Renaldo Ali).docx	5	11	21	20.6
BAB 4 (Ferry Dharmawan).docx	BAB 4 (Renaldo Ali).docx	5	11	21	27.41

Gambar 8 Pengujian Skripsi Ferry Dharmawan dengan Renaldo Ali

- d. Pengujian antara Skripsi Ayu Windy, Jacklin Shintia Thio, Josselyn Shintia Thio, Michelle Tirta, Nadia Yanitra, Sri Whisnu, dan Mariana dapat dilihat pada Gambar 9.

FILE 1	FILE 2 (ARCHIEVE)	NILAI K	NILAI B	NILAI W	PERSENTASE
BAB 1 (Ayu Windy Astuti).docx	BAB 1 (Mariana).docx	5	11	21	26.66
BAB 1 (Ayu Windy Astuti).docx	BAB 1 (Jacklin Shintia Thio).docx	5	11	21	25.49
BAB 1 (Ayu Windy Astuti).docx	BAB 1 (Josselyn Shintia Thio).docx	5	11	21	25.98
BAB 1 (Ayu Windy Astuti).docx	BAB 1 (Michelle Tirta Winarta).docx	5	11	21	20.57
BAB 1 (Ayu Windy Astuti).docx	BAB 1 (Nadia Yanitra).docx	5	11	21	20.56
BAB 1 (Ayu Windy Astuti).docx	BAB 1 (Sri Whisnu A.W).docx	5	11	21	21.82
BAB 3 (Ayu Windy Astuti).docx	BAB 3 (Josselyn Shintia Thio).docx	5	11	21	29.03
BAB 3 (Ayu Windy Astuti).docx	BAB 3 (Mariana).docx	5	11	21	27.13
BAB 3 (Ayu Windy Astuti).docx	BAB 3 (Michelle Tirta Winarta).docx	5	11	21	27.21
BAB 3 (Ayu Windy Astuti).docx	BAB 3 (Nadia Yanitra).docx	5	11	21	24.31
BAB 3 (Ayu Windy Astuti).docx	BAB 3 (Sri Whisnu A.W).docx	5	11	21	42.57
BAB 3 (Ayu Windy Astuti).docx	BAB 3 (Jacklin Shintia Thio).docx	5	11	21	26.35
BAB 4 (Ayu Windy Astuti).docx	BAB 4 (Jacklin Shintia Thio).docx	5	11	21	22.64
BAB 4 (Ayu Windy Astuti).docx	BAB 4 (Josselyn Shintia Thio).docx	5	11	21	28.17
BAB 4 (Ayu Windy Astuti).docx	BAB 4 (Mariana).docx	5	11	21	22.13
BAB 4 (Ayu Windy Astuti).docx	BAB 4 (Michelle Tirta Winarta).docx	5	11	21	15.07
BAB 4 (Ayu Windy Astuti).docx	BAB 4 (Nadia Yanitra).docx	5	11	21	15.4
BAB 4 (Ayu Windy Astuti).docx	BAB 4 (Sri Whisnu A.W).docx	5	11	21	22.94

BAB 1 (Jacklin Shintia Thio).docx	BAB 1 (Josselyn Shintia Thio).docx	5	11	21	21.76
BAB 1 (Jacklin Shintia Thio).docx	BAB 1 (Mariana).docx	5	11	21	24.53
BAB 1 (Jacklin Shintia Thio).docx	BAB 1 (Mishelle Tirta Winarta).docx	5	11	21	18.22
BAB 1 (Jacklin Shintia Thio).docx	BAB 1 (Nadia Yanitra).docx	5	11	21	20.02
BAB 1 (Jacklin Shintia Thio).docx	BAB 1 (Sri Whisnu A.W).docx	5	11	21	28
BAB 3 (Jacklin Shintia Thio).docx	BAB 3 (Josselyn Shintia Thio).docx	5	11	21	26.64
BAB 3 (Jacklin Shintia Thio).docx	BAB 3 (Mariana).docx	5	11	21	22.42
BAB 3 (Jacklin Shintia Thio).docx	BAB 3 (Mishelle Tirta Winarta).docx	5	11	21	22.65
BAB 3 (Jacklin Shintia Thio).docx	BAB 3 (Nadia Yanitra).docx	5	11	21	20.87
BAB 3 (Jacklin Shintia Thio).docx	BAB 3 (Sri Whisnu A.W).docx	5	11	21	28.59
BAB 4 (Jacklin Shintia Thio).docx	BAB 4 (Josselyn Shintia Thio).docx	5	11	21	26.26
BAB 4 (Jacklin Shintia Thio).docx	BAB 4 (Mariana).docx	5	11	21	24.83
BAB 4 (Jacklin Shintia Thio).docx	BAB 4 (Mishelle Tirta Winarta).docx	5	11	21	15
BAB 4 (Jacklin Shintia Thio).docx	BAB 4 (Nadia Yanitra).docx	5	11	21	14.3
BAB 4 (Jacklin Shintia Thio).docx	BAB 4 (Sri Whisnu A.W).docx	5	11	21	26.55
BAB 1 (Josselyn Shintia Thio).docx	BAB 1 (Mariana).docx	5	11	21	23.23
BAB 1 (Josselyn Shintia Thio).docx	BAB 1 (Mishelle Tirta Winarta).docx	5	11	21	22.08
FILE 1	FILE 2 (ARCHIVE)	NILAI K	NILAI B	NILAI W	PERSE
BAB 1 (Josselyn Shintia Thio).docx	BAB 1 (Nadia Yanitra).docx	5	11	21	20.9
BAB 1 (Josselyn Shintia Thio).docx	BAB 1 (Sri Whisnu A.W).docx	5	11	21	20.43
BAB 3 (Josselyn Shintia Thio).docx	BAB 3 (Mariana).docx	5	11	21	25.58
BAB 3 (Josselyn Shintia Thio).docx	BAB 3 (Mishelle Tirta Winarta).docx	5	11	21	28.64
BAB 3 (Josselyn Shintia Thio).docx	BAB 3 (Nadia Yanitra).docx	5	11	21	24.16
BAB 3 (Josselyn Shintia Thio).docx	BAB 3 (Sri Whisnu A.W).docx	5	11	21	29.21
BAB 4 (Josselyn Shintia Thio).docx	BAB 4 (Mariana).docx	5	11	21	22.65
BAB 4 (Josselyn Shintia Thio).docx	BAB 4 (Mishelle Tirta Winarta).docx	5	11	21	14.35
BAB 4 (Josselyn Shintia Thio).docx	BAB 4 (Nadia Yanitra).docx	5	11	21	16.08
BAB 4 (Josselyn Shintia Thio).docx	BAB 4 (Sri Whisnu A.W).docx	5	11	21	22.57
BAB 1 (Mariana).docx	BAB 1 (Mishelle Tirta Winarta).docx	5	11	21	23.61
BAB 1 (Mariana).docx	BAB 1 (Nadia Yanitra).docx	5	11	21	18.19
BAB 1 (Mariana).docx	BAB 1 (Sri Whisnu A.W).docx	5	11	21	25.18
BAB 3 (Mariana).docx	BAB 3 (Mishelle Tirta Winarta).docx	5	11	21	21.22
BAB 3 (Mariana).docx	BAB 3 (Nadia Yanitra).docx	5	11	21	19.99
BAB 3 (Mariana).docx	BAB 3 (Sri Whisnu A.W).docx	5	11	21	25.25
BAB 4 (Mariana).docx	BAB 4 (Mishelle Tirta Winarta).docx	5	11	21	16.67
BAB 3 (Mariana).docx	BAB 3 (Nadia Yanitra).docx	5	11	21	19.99
BAB 3 (Mariana).docx	BAB 3 (Sri Whisnu A.W).docx	5	11	21	25.25
BAB 4 (Mariana).docx	BAB 4 (Mishelle Tirta Winarta).docx	5	11	21	16.67
BAB 4 (Mariana).docx	BAB 4 (Nadia Yanitra).docx	5	11	21	19.21
BAB 4 (Mariana).docx	BAB 4 (Sri Whisnu A.W).docx	5	11	21	25
BAB 1 (Mishelle Tirta Winarta).docx	BAB 1 (Nadia Yanitra).docx	5	11	21	28.04
BAB 1 (Mishelle Tirta Winarta).docx	BAB 1 (Sri Whisnu A.W).docx	5	11	21	19.63
BAB 3 (Mishelle Tirta Winarta).docx	BAB 3 (Nadia Yanitra).docx	5	11	21	43.36
BAB 3 (Mishelle Tirta Winarta).docx	BAB 3 (Sri Whisnu A.W).docx	5	11	21	24.07
BAB 4 (Mishelle Tirta Winarta).docx	BAB 4 (Nadia Yanitra).docx	5	11	21	25.44
BAB 4 (Mishelle Tirta Winarta).docx	BAB 4 (Sri Whisnu A.W).docx	5	11	21	14.65
BAB 1 (Nadia Yanitra).docx	BAB 1 (Sri Whisnu A.W).docx	5	11	21	19.97
BAB 3 (Nadia Yanitra).docx	BAB 3 (Sri Whisnu A.W).docx	5	11	21	23.19
BAB 4 (Nadia Yanitra).docx	BAB 4 (Sri Whisnu A.W).docx	5	11	21	15.33

Gambar 9 Pengujian antara skripsi Ayu Windy, Jacklin Shintia Thio, Josselyn Shintia Thio, Mishelle Tirta, Nadia Yanitra, Sri Whisnu, dan Mariana

4. KESIMPULAN

Kesimpulan yang dapat diperoleh berdasarkan pembuatan dan pengujian dari aplikasi pendeteksi tingkat kesamaan antar dokumen dengan algoritma winnowing ini adalah sebagai berikut :

1. Aplikasi pendeteksi tingkat kesamaan antar dokumen ini dapat dijalankan dan mengoperasikan algoritma winnowing dengan baik serta dapat memberikan hasil yang dapat di tampilkan prosesnya.
2. Berdasarkan hasil pengujian diketahui bahwa susunan kata antar dokumen sangat menentukan nilai prosentase kesamaan. Jika urutan kata antar dokumen yang sama dirubah, maka hasilnya persentase kesamaannya akan turun.
3. Aplikasi dapat melakukan perbandingan dokumen uji dengan beberapa dokumen pembanding dan dapat ditampilkan secara bersamaan.

Saran untuk yang ingin mengembangkan aplikasi pendeteksi tingkat kesamaan antar dokumen dengan algoritma winnowing ini adalah sebagai berikut :

1. Dapat mengirimkan data hasil deteksi melalui email ketika proses deteksi sudah selesai
2. Menggunakan tambahan metode dokumen *fingerprinting* lainnya agar dapat digunakan sebagai tambahan informasi ukuran kesamaan antar dokumen.

UCAPAN TERIMA KASIH

Terima kasih saya ucapkan kepada dosen-dosen Fakultas Teknologi Informasi Universitas Tarumanagara yang sudah mau membimbing penulis dalam perancangan aplikasi ini. Terima kasih juga kepada seluruh pendahulu saya yang sudah digunakan data skripsinya untuk data pembanding dalam proses perancangan aplikasi ini.

DAFTAR PUSTAKA

- [1] Depdikbud, 1997, *Kamus Besar Bahasa Indonesia*. Jakarta.
- [2] Stein, Benno & Eissen, Sven Meyer Zu. 2006. *Fingerprint-based Similarity Search and its Applications*. Bauhaus-Universitat Weimar.
- [3] Schleimer, Saul. Wilkerson, Daniel S., dan Aiken, Alex. 2003. *Winnowing : Local Algorithms for document Fingerprinting*. San Diego.
- [4] Milani W, Marji, Achmad Ridok. 2014. *Deteksi Plagiarisme Pada Dokumen Teks Bahasa Indonesia Menggunakan Algoritma Winnowing Dengan Stemming*. Malang : Universitas Brawijaya.