



Penerapan Algoritma Wnnowing untuk Mendeteksi Kemiripan Dua Teks yang Berbeda

Jupron¹, Sutrisno²

Universitas Pamulang

dosen02664@unpam.ac.id¹, dosen02673@unpam.ac.id²

Keywords:

Customer Satisfaction,
Fuzzy Inference
System, Sugeno
Method, Tsukamoto
Method, MAPE

Abstract

Untuk menilai kemiripan (similarity) teks pada dua teks di butuhkan sebuah algoritma untuk menilai apakah kedua teks tersebut memiliki kesamaan (similarity). Oleh karena itu penerapan algoritma winnowing untuk mendeteksi kemiripan teks sangatlah penting agar dapat membedakan ke aslian suatu dokumen. Algoritma winnowing adalah salah satu algoritma yang di pakai untuk menentukan atau menilai kemiripan teks. Permasalahan yang terjadi biasanya pada perhitungan nilai similarity antara perhitungan manual tidak sama dengan perhitungan di sistem di karenakan pada perhitungan manual tidak ada penyeleksian tanda baca. Penelitian ini di lakukan dengan cara mengubah huruf pada text menjadi nilai atau angka menggunakan bilangan ascii, setelah proses tersebut terbentuk dilanjutkan dengan proses window, maka terbentuklah nilai fingerprint. Kedua teks dapat di nilai kemiripannya jika teks memiliki nilai fingerprint, fingerprint inilah yang di pakai untuk menentukan kemiripan kedua teks tersebut, algoritma ini mencari nilai fingerprint dengan cara mengubah nilai rangkaian n-gram kedalam bentuk nilai angka yang di sebut nilai hash. Hasil dari pengujian winnowing ini berupa bentuk nilai presentase similarity yang bernilai dari 0% sampai 100 %, semakin kecil presentasenya maka dokumen tersebut di katakan tidak mirip (tidak sama), tetapi jika nilai similarity pada kedua dokumen tersebut besar, maka kedua dokumen tersebut dapat dikatakan mirip(sama).

Pendahuluan

Pemanfaatan kecanggihan teknologi saat ini berkembang dengan pesat dan luas di berbagai belahan dunia dan menjadi salah satu hal yang sangat penting dikehidupan sehari hari. Mengakibatkan dengan sangat mudah sekali kita dapat mengakses dan mencari informasi di belahan dunia. Internet membuat berbagai informasi sangat mudah diakses dan penyebar luasan ilmu pengetahuan pun seakan tanpa batas. Pemanfaatan - pemanfaatan teknologi dalam kehidupan

sehari - hari adalah untuk mencari, mengolah atau dapat juga menyimpan informasi dengan menggunakan kecanggihan komputer, informasi yang di peroleh biasanya berupa dokumen teks. Kemudian mengakses informasi dalam berupa dokumen teks bisa menimbulkan sifat manusia yang negatif, yaitu dengan meniru dan mencontoh hasil karya orang lain melebihi batas normal tolerir. Jika hanya mencontoh sedikit bagian atau mengutip untuk di jadikan refrensi mungkin tidak begitu masalah. Namun banyak juga kasus tingkat kemiripan dua dokumen tersebut bisa sama hampir sepenuhnya. Kita dapat mengetahui berapa tingkat kemiripan antar dua dokumien yang di memiliki kesamaan dengan sebuah sistem (Alamsyah, 2017).

Plagiarisme berdasarkan kesamaan kata antar dokumen salah satu contohnya yaitu mahasiswa sekarang sering melakukan *copy paste* dalam mengerjakan tugas yang diberikan, baik itu dari internet atau hasil tugas orang lain. Tindakan tersebut dapat memicu mahasiswa melakukan plagiasi salah satu contohnya pada penulisan proposal skripsi, dengan mengambil kutipan atau pernyataan tanpa menyertakan sumbernya.

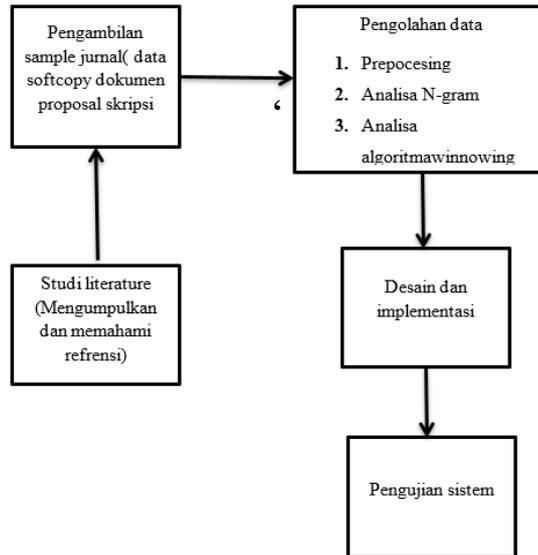
Kemiripan dokumen karya tulis ilmiah tidak di katakan plagiat jika menggunakan informasi yang berupa fakta umum, menulis kembali (dengan megubah kalimat atau parafrase) opini orang lain dengan memberikan sumber jelas, mengutip secukupnya tulisan orang lain dengan memberikan tanda batas jelas bagian kutipan dan menuliskan sumbernya. Tindakan plagiarisme dalam instansi, sektor akademis, maupun non akademis secara perlahan harus di cegah dan dihilangkan dengan melakukan pendeteksian plagiat secara manual maupun dengan memanfaatkan metode pencocokan string. Namun pendeteksian secara manual memiliki masalah yang cukup besar yaitu sangat tidak memungkinkan melakukan pendeteksian dokumen dengan membandingkannya dengan dokumen lain yang berjumlah ratusan bahkan sampai ribuan, dengan demikian melakukan pendeteksian plagiarisme secara manual sangatlah tidak efektif.

Metode

Jenis Penelitian yang digunakan adalah model eksperimen. Penelitian ini menggunakan metode Winnowing untuk memudahkan institusi dalam penilaian terhadap plagiarisme.

1. Perancangan Penelitian

Untuk mendapatkan data-data yang diperlukan dalam penelitian, penulis menggunakan beberapa tahapan yang di lakukan sebagai alur dalam me nyelesaikan penelitian tugas akhir ini, seperti gambar 3.1:



Gambar 3. 1 Tahapan penelitian

Berdasarkan Gambar 3.1 dijelaskan tahapan penelitian ini dimulai dengan studi literatur dengan mengumpulkan dan memahami referensi, pengumpulan data dengan mengumpulkan data *softcopy* dokumen jurnal, pengolahan data, desain dan implementasi, kemudian pengujian aplikasi.

2. Studi Literatur

Studi literatur dilakukan untuk mendapatkan informasi tentang permasalahan yang akan diteliti dengan mengumpulkan jurnal-jurnal penelitian sebelumnya dan buku dasar teori yang berkaitan tentang algoritma untuk mendeteksi plagiarisme. Dari studi literatur didapatkan metode yang digunakan untuk mendeteksi indikasi plagiarisme

3. Pengambilan Sampel Jurnal

Pengambilan sample jurnal untuk pengujian deteksi indikasi plagiarisme. Jurnal yang di ambil sampelnya untuk pengujian menggunakan sample jurnal mahasiswa unpad dari jurusan tehnik informatika.

4. Pengolahan Data

Pengolahan data dilakukan untuk menyelesaikan permasalahan yang akan diteliti dan diterapkan dalam sistem, berikut langkah-langkah yang dilakukan dalam mendeteksi plagiarisme menggunakan algoritma *N-gram* dan *document fingerprinting* :

a. Analisa Preprocessing

Dalam tahapan *Preprocessing* dilakukan proses pembuangan karakter yang tidak di perlukan.

- b. Analisa Algoritma N-Grams
Dalam tahapan ini dilakukan perubahan dari bentuk kata menjadi rangkaian *N-Grams* dengan urutan sesuai struktur kalimat.
- c. Analisa Algoritma Winnowing
Dalam tahapan ini dilakukan proses perubahan dari bentuk rangkaian *N-Grams* menjadi *hash* untuk menentukan *fingerprinting*. Dengan menggunakan algoritma *winnowing* dapat menentukan tingkat kesamaan kata antar dokumen
- d. Desain dan Implementasi
Desain aplikasi dilakukan untuk mendesain struktur dan tampilan *interface* dari aplikasi deteksi indikasi plagiarisme. Dalam sistem deteksi indikasi plagiarisme yang dibangun menggunakan bahasa pemrograman *PHP*, dan *database* yang digunakan yaitu *MySQL* untuk menyimpan dokumen serta *fingerprint* dari dokumen jurnal. Implementasi dilakukan dengan menerapkan hasil pengolahan data pada aplikasi deteksi plagiarisme yang sudah di rancang untuk mengetahui tingkat plagiarisme dari suatu dokumen.
- e. Pengujian
Proses pengujian menggunakan 2 skenario, scenario pertama pengujian pada judul penulis dan scenario kedua pada sample jurnal mahasiswa unpam, scenarionya sebagai berikut:
 - a) Pengujian skenario pertama yaitu menguji judul jurnal dengan cara membandingkan judul asli yang tidak diubah dalam perkataanya dengan judul asli yang di ubah 2 kata dari judul aslinya tetapi masih mengandung arti yang sama.
 - b) Pengujian skenario kedua yaitu menguji keseluruhan jurnal asli dari judul sampai daftar pustaka dengan jurnal asli yang sudah di hilangkan abstrak, bab 5 dan daftar pustaka. Pengujian menggunakan n-garm n=3 sampai n=9 (diambil dari penelitian sebelumnya)

Hasil dan Pembahasan

1. Hasil Penerapan Algoritma Winnowing Untuk Menentukan Similarity

Implementasian metode Agoritma Winnowing untuk menentukan nilai kemiripan text (similarity) dapat di hitung menggunakan aplikasi PHP MYSQL, berikut ini adalah tampilan dari aplikasinya:

a. Halaman *Login*

Berikut ini adalah tampilan dari halaman *login* dari sistem algoritma winnowing yang merupakan tampilan awal sebelum melakukan akses ke sistem, pada halaman ini terdapat *username* dan *password* sebagai syarat untuk melakukan akses ke sistem untuk tampilan dapat dilihat pada Gambar 4.1



Gambar 4. 1 Halaman Login

b. Halaman *Beranda*

Berikut ini adalah tampilan dari halaman beranda dari sistem Algoritma Winnowing, dokumen yang merupakan tampilan dari sistem setelah berhasil *login* dengan *username* dan *password* yang benar, di halama ini terdapat lima tools dalam fungsinya dapat dilihat pada Gambar 4.2



Gambar 4. 2 Halaman Beranda

c. Halaman Info

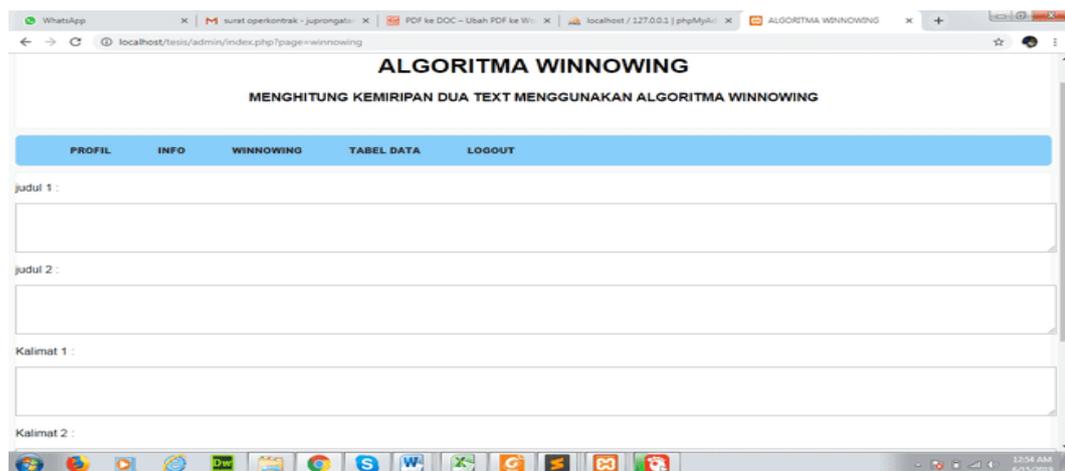
Berikut ini adalah tampilan dari halaman info dari sistem Algoritma Winnowing, berisi tentang pengertian algoritma winnowing tampilan dari sistem, dapat dilihat pada Gambar 4.3



Gambar 4. 3 Halaman Info

d. *Halaman Winnower*

Berikut ini adalah tampilan dari halaman winnowing dari sistem Algoritma Winnowing, halaman ini berisi tentang proses algoritma winnowing itu di jalankan di mana dua text saling di dibandingkan satu sama yang lain agar mendapatkan nilai similaritynya. Untuk tampilan dapat dilihat pada gambar 4.4



Gambar 4. 4 Halaman Winnower

e. *Halaman Tabel*

Halaman ini berisi tentang tabel hasil dari proses algoritma winnowing yang sudah di proses di halaman winnowing. Untuk tampilan dapat di lihat pada gambar 4.5

no	Judul1	Judul2	ngram	window	ngram2	fling1	fling2	fingers1	fingers2	union_fingers	intersect_fingers	unint	hasil
1	penerapan sistem pendukung keputusan pemilihan kepala mekarik menggunakan metode simple additive weighting (saw) pada cv surya makmur kencana	sistem pendukung keputusan pemilihan lokasi pembukaan cabang usaha vapor yang strategis menggunakan metode technique for order preference by similarity to ideal solution (topsis)	7	7	7	95000143	98169918	19205	13317	32522	5529	26993	20.483088208054
2	rekayasa perangkat lunak key performance indicator untuk penilaian kinerja sales	sistem pendukung keputusan pemilihan lokasi pembukaan cabang usaha vapor yang strategis menggunakan	7	7	7	95353405	98169918	31560	13317	44877	8211	36666	22.394043528064

Gambar 4. 5 Halaman Tabel

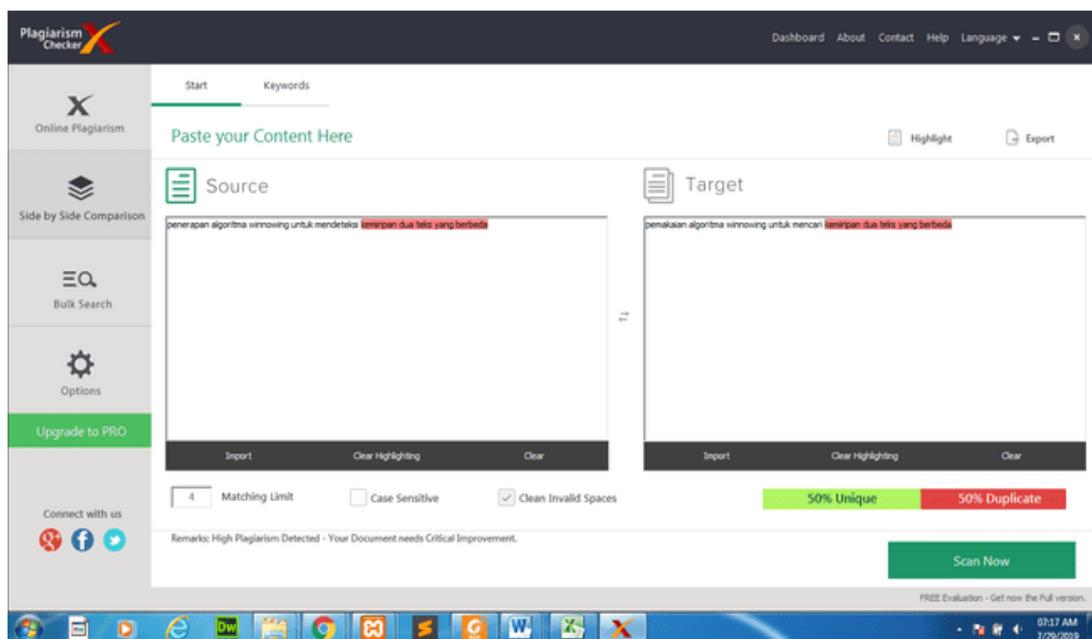
2. Pengujian Judul Menggunakan Plagiarisme Checker dan Metode Winnowing

Pada sekenario ini akan menguji judul penulis dimana dari judul tersebut di bedakan dengan cara merubah dua kata pada kalimatnya tetapi masih mengandung makna yang sama :

Teks 1 : penerapan algoritma winnowing untuk mendeteksi kemiripan dua teks yang berbeda

Teks 2 : pemakaian algoritma winnowing untuk mencari kemiripan dua teks yang berbeda

3. Pengujian Plagiarisme Checker

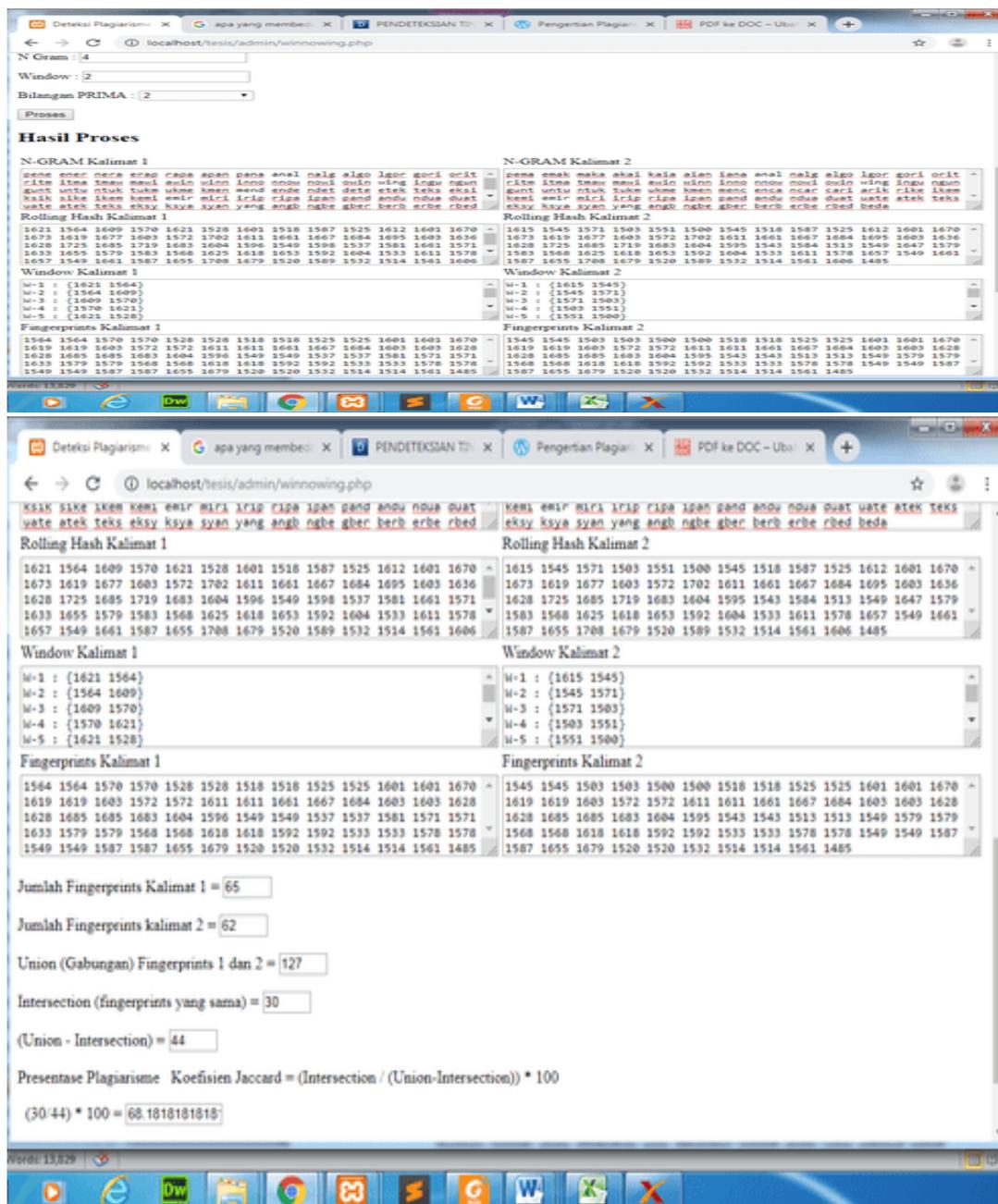


Gambar 4. 6 Pengujian Plagiarism Checker

Di lihat dari gambar 4.6 hasil dari plagiarism chaker menunjukkan bahwa nilai similarity dari pengujian tersebut adalah 50 % , kata yang di block warna merah menandakan kata tersebut memiliki kesamaan.

4. Pengujian Sistem Menggunakan Metode Wnnowing

Tahapan pengujian yang akan dilakukan yaitu pengujian metode Wnnowing. Pengujian dilakukan meliputi pengujian jumlah gram, jumlah window, basis bilangan prima, waktu proses dan jumlah langkah yang dilakukan untuk menghasilkan output. Perhitungan dengan n gram= 4, window = 2 dan bilangan prima = 2 dapat di lihat pada gambar 4.7.



Gambar 4. 7 Pengujian Sistem Algoritma Wnnowing

Dilihat dari gambar 4.7 Nilai similaritas yang terbaca adalah 68.18 %. Dari pengujian di atas dapat kita simpulkan bahwa hasil dari nilai similaritas pada sistem menghasilkan nilai yang sama pada perhitungan manual menggunakan rumus algoritma winnowing dan jika di bandingkan dengan nilai similarity plagiarism checker nilai persentasinya lebih akurat karena plagiarisme checker hanya membaca 4 kalimat yang sama sedangkan nilai similaritas perhitungan winnowing membaca seluruh nilai karakter pada teks.

5. Pengujian Jurnal Menggunakan Metode Winnowing Dan Plagiarisme Checker

Pada skenario ini penulis mengambil sampel jurnal dari mahasiswa unpad dengan judul “sistem pakar diagnosa penyakit kanker dengan metode forward chaining berbasis web” nama penulis Adrivo aziz.

Pada skenario ini jurnal di uji dengan cara membagi jurnal menjadi dua teks yang berbeda, teks pertama jurnal di uji secara keseluruhan dengan total karakter (12289) dan teks yang kedua pada judul, abstrak, kesimpulan dan daftar pustaka di hapus sehingga jumlah karakter yang di bandingkan adalah (9361)

6. Pengujian Jumlah Gram

Analisis jumlah gram dilakukan agar diketahui jumlah gram yang optimal untuk ketepatan dari perhitungan sistem. Gram yang akan dianalisis merupakan jumlah yang diinputkan oleh user. Berikut merupakan tabel hasil analisis dari jumlah gram yang telah dilakukan:

Jumlah window : 3

Basis bil.prima : 3

Dokumen 1 : Jumlah Karakter (12289)

Dokumen 2 : Jumlah karakter (9361)

Dokumen 1 memiliki total 12289 karakter dan Dokumen 2 adalah dokumen yang sama namun dihilangkan judul, abstrak, dan daftar pustaka untuk membedakan isi dari dokumen sehingga Dokumen 2 memiliki total 9361 karakter. Menurut perhitungan manual kemiripannya adalah $\frac{12289}{9361} \times 100\% = 76.17\%$

persentase kemiripan yang mendekati hasil tersebut adalah jumlah gram = 9 lebih mendekati yaitu sebesar 77.57 % dengan selisih 1,4 %. Pengujian jumlah gram dapat dilihat pada table 4.1:

Tabel 4. 1 Pengujian Jumlah Gram

ngram	jml_finger_kalimat1	jml_finger_kalimat2	fingerprint_gabungan	fingerprint_sama	unint	hasil	total_time
3	11891	9118	21009	208	222	93.69%	1.32 detik
4	11890	9117	21007	527	593	88.87%	1.39 detik
5	11889	9116	21005	1152	1343	85.78%	1.64 detik
6	11888	9115	21003	1959	2409	81.32%	1.67 detik
7	11887	9114	21001	2674	3402	78.60%	1.78 detik
8	11886	9113	20999	3180	4094	77.67%	2.18 detik
9	11885	9112	20997	3537	4560	77.57%	2.82 detik

Dari hasil analisis jumlah gram pada table 4.1 maka dapat disimpulkan bahwa semakin besar jumlah gram yang diinputkan maka persentasi kemiripan menjadi menurun. Dan presentase ke akuratan lebih akurat karena hasil mendekati hasil perhitungan manual. Perbedaan waktu proses ketika gram bertambah dapat dilihat gambar 4.8



Gambar 4. 8 Grafik Perbedaan Waktu Proses Terhadap Gram Winnowing

7. Pengujian Jumlah Window

Analisis jumlah window dilakukan agar diketahui jumlah window yang optimal untuk ketepatan dari perhitungan sistem. Jumlah window yang akan dianalisis merupakan jumlah yang diinputkan oleh user. Karena jumlah gram optimal telah ditemukan pada proses analisis sebelumnya yaitu gram = 9 maka jumlah gram yang akan dipakai untuk analisis window adalah gram = 9 Berikut merupakan tabel hasil pengujian dari jumlah window yang telah dilakukan:

Jumlah gram : 9
 Basis bil.prima : 3
 Dokumen 1 : Jumlah Karakter (12289)
 Dokumen 2 : Jumlah karakter (9361)

Dokumen 1 memiliki total 12289 karakter dan Dokumen 2 adalah dokumen yang sama namun dihilangkan judul, abstrak, dan kesimpulan daftar pustaka untuk membedakan isi dari dokumen sehingga Dokumen 2 memiliki total 9361 karakter. Dari hasil analisis dapat disimpulkan bahwa jumlah gram= 9 dan window= 3 menghasilkan 77.57 %, Persentasi tersebut adalah paling mendekati dengan jumlah persentasi perhitungan manual yaitu 76.17 % dengan selisih 1.4 %. Pengujian jumlah window dapat dilihat pada table 4.2:

Tabel 4. 2 Pengujian Jumlah Window

window	jml_finger_kalimat1	jml_finger_kalimat2	fingerprint_gabungan	fingerprint_sama	unint	hasil	total_time
2	11886	9113	20999	4704	6050	77.75%	2.41 detik
3	11885	9112	20997	3537	4560	77.57%	2.82 detik
4	11884	9111	20995	2950	3795	77.73%	3.39 detik
5	11883	9110	20993	2513	3213	78.21%	5.34 detik
6	11882	9109	20991	2172	2772	78.35%	6.12 detik
7	11881	9108	20989	1937	2472	78.36%	12.93 detik
8	11880	9107	20987	1756	2223	78.99%	15.33 detik
9	11879	9106	20985	1589	2005	79.25%	16.71 detik

Dari tabel dapat dilihat bahwa pembagian window tidak begitu mempengaruhi kemiripan, dikarenakan pembagian window hanya membagi nilai-nilai *fingerprints* saja kedalam window-window, berbeda ketika pembagian gram, pembagian gram membagi dan melakukan perhitungan pada gram yang terbagi-bagi sehingga perbedaan gram mempengaruhi nilai kemiripan. Perbedaan waktu proses ketika jumlah window bertambah dapat dilihat pada gambar 4.9



Gambar 4. 9 Grafik Perbedaan Waktu Proses Terhadap Window Winnowing

Dari gambar 4.9 di lihat pada grafik di atas, perbedaan nilai window mempengaruhi waktu proses, semakin nilai windownya besar semakin lama waktu prosesnya.

8. Pengujian Basis Bilangan Prima

Analisis bilangan prima dilakukan untuk mendapatkan informasi jumlah dari maksimal bilangan prima yang dapat digunakan oleh Algoritma Winnowing menurut rumus yang dilakukan pada proses *Rolling Hash*. Bilangan prima dalam rumus *Rolling Hash* tidak ditentukan oleh user tapi ditentukan oleh *programmer* atau pembuat sistem. Maka dilakukan analisis untuk mendapatkan nilai bilangan prima yang optimal untuk dapat diterapkan pada rumus tersebut. Karena jumlah gram dan window optimal telah ditemukan pada proses analisis sebelumnya yaitu gram=9 dan window = 3 maka jumlah gram yang akan dipakai untuk analisis bilangan prima adalah gram=9 dan window 3. Pengujian dari jumlah window yang telah dilakukan:

Jumlah gram : 9
 Window : 3
 Dokumen 1 : Jumlah Karakter (12289)
 Dokumen 2 : Jumlah karakter (9361)

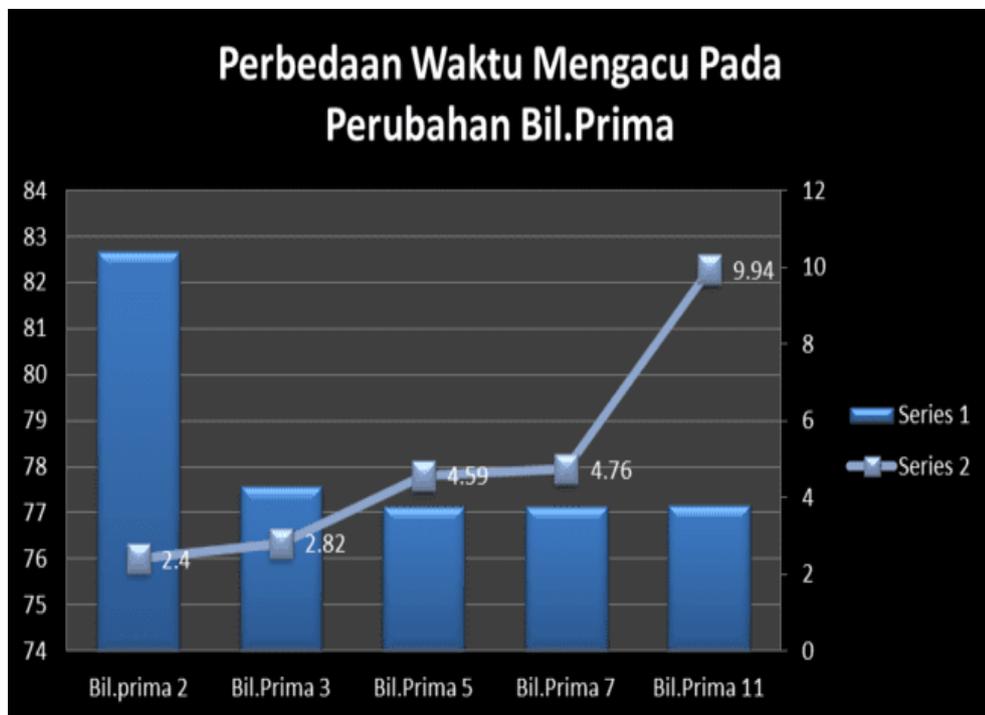
Dokumen 1 memiliki total 24785 karakter dan Dokumen 2 adalah dokumen yang sama namun dihilangkan judul, abstrak, dan daftar pustaka untuk membedakan isi dari dokumen sehingga Dokumen 2 memiliki total 21703 karakter.

Dari hasil analisis dapat disimpulkan bahwa jumlah gram=9, window=3 dan bilangan prima = 7 menghasilkan 77,13 %. Persentasi tersebut adalah paling mendekati dengan jumlah persentasi perhitungan manual yaitu 76.17 %. dapat dilihat bahwa pembagian ngram dan bilangan prima dapat mempengaruhi jumlah presentase kemiripan. Pengujian jumlah bil.prima dapat dilihat pada table 4.3:

Tabel 4. 3 Pengujian Jumlah Bilangan Prima

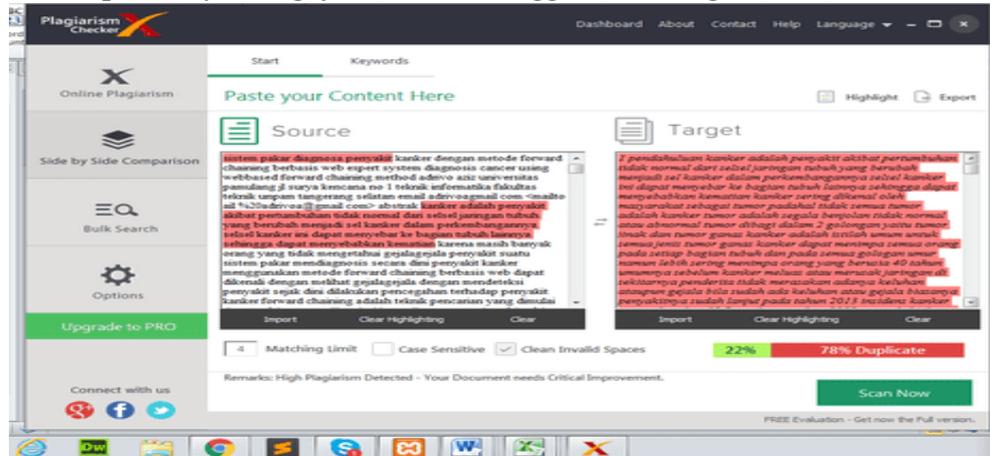
bil_prima	jml_finger_kalimat1	jml_finger_kalimat2	fingerprint_gabungan	fingerprint_sama	unint	hasil	total_time
2	11885	9112	20997	2581	3122	82.67%	2.4 detik
3	11885	9112	20997	3537	4560	77.57%	2.82 detik
5	11885	9112	20997	3533	4580	77.14%	4.59 detik
7	11885	9112	20997	3488	4522	77.13%	4.76 detik
11	11885	9112	20997	3483	4514	77.16%	9.94 detik

Dari hasil analisis jumlah bil.prima pada table 4.3 maka dapat disimpulkan bahwa semakin besar jumlah bil.prima yang diinputkan maka persentasi kemiripan menjadi menurun. Dan presentase ke akuratan lebih akurat karena hasil mendekati hasil perhitungan manual dan plagiarism checker, Perbedaan waktu proses ketika bil.prima bertambah dapat dilihat gambar 4.10



Gambar 4. 10 Grafik Perbedaan Waktu Proses Terhadap Bil.Prima Winnowing

Dari gambar 4.10 di lihat pada grafik di atas, perbedaan nilai bil. prima mempengaruhi waktu proses, semakin nilai basisnya besar, semakin lama waktu prosesnya. Pengujian Jurnal Menggunakan Plagiarism Checker



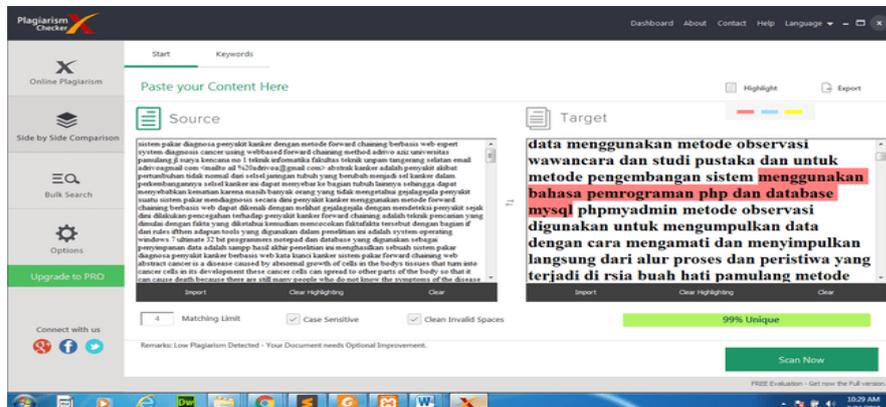
Gambar 4. 11 Pengujian Plagiarism Checker

Di lihat dari gambar 4.1 hasil dari plagiarism chaker menunjukkan bahwa nilai similarity dari pengujian tersebut adalah 78 % selisih 1.83 % dari perhitungan manual , kata yang di block warna merah menandakan kata tersebut memiliki kesamaan.

9. Pengujian Dua Jurnal Yang Berbeda Menggunakan Plagiarisme Checker dan Metode Winoing

Pada skenario ini penulis mengambil dua sampel jurnal dari mahasiswa unpm dengan judul. Pada skenario ini dua jurnal di uji dengan cara membandingkan isi keseluruhan dari kedua jurnal tersebut dari judul sampai daftar pustaka dengan menggunakan window = 3 ngram = 9 dan bilangan prima = 7 mengikuti dari hasil terbaik dalam pengujian window, ngram dan bilangan prima pda scenario kedua, teks pertama pada jurnal pertama mempunyai total karakter (12289) dan teks yang kedua pada jurnal kedua mempunyai total karakter (14709)

10. Pengujian Plagiarisme Checker Pada Kedua Jurnal

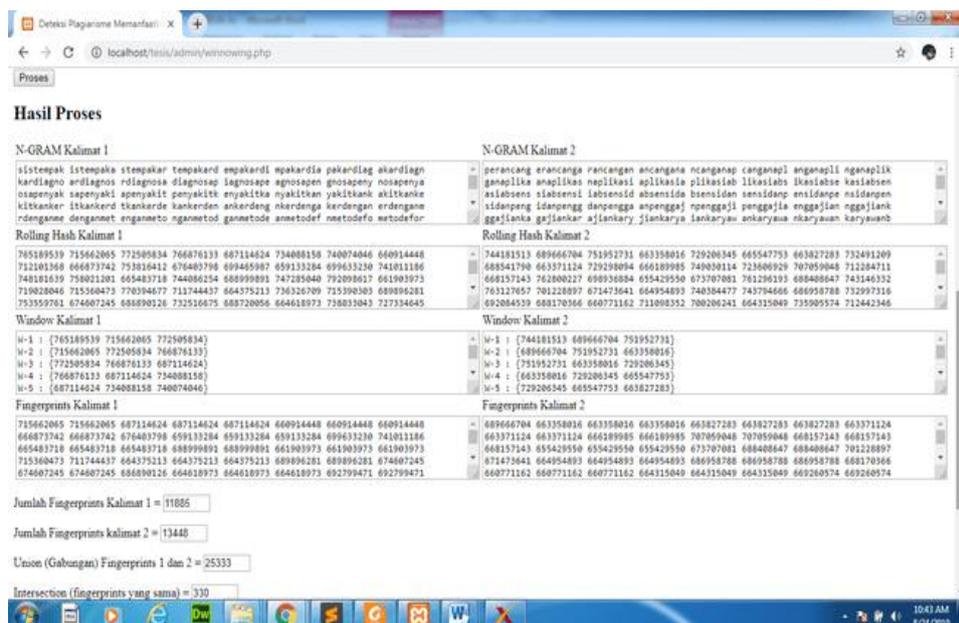
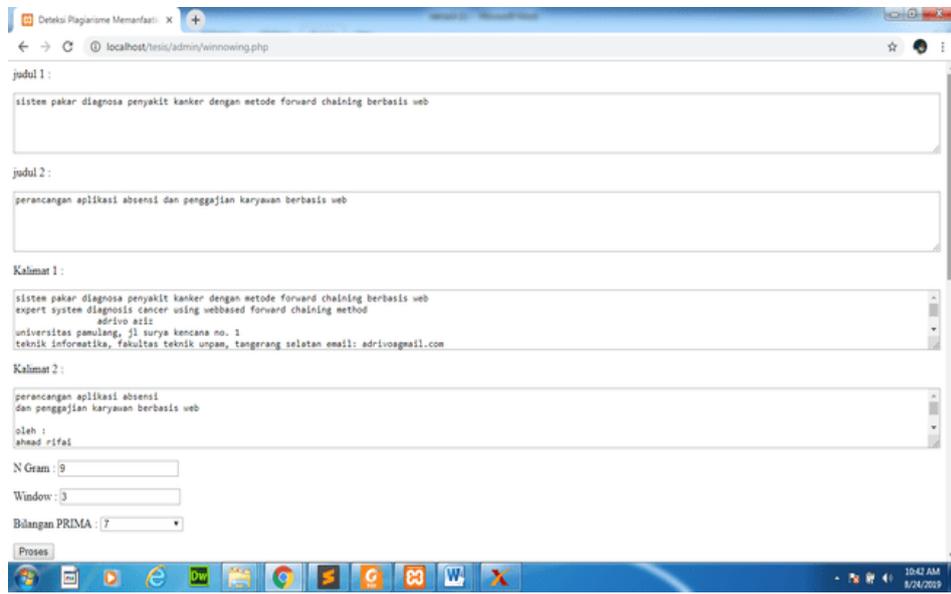


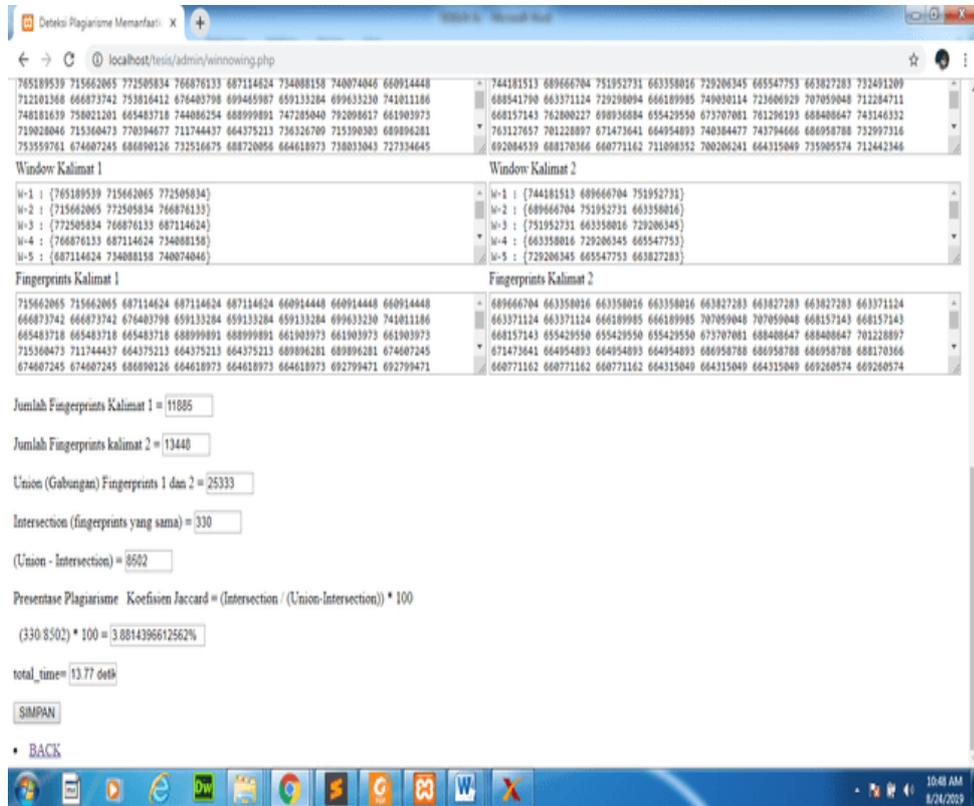
Gambar 4. 12 Pengujian Plagiarism Checker

Di lihat dari gambar 4.12 hasil dari plagiarism chaker menunjukkan bahwa nilai similarity dari pengujian tersebut adalah 1 % , kata yang di block warna merah menandakan kata tersebut memiliki kesamaan. pada pengujian kedua jurnal ini terbaca unique itu artinya kedua jurnal tersebut bisa di katakan berbeda atau tidak sama.

11. Pengujian Sistem Kedua Jurnal Menggunakan Metode Winnowing

Tahapan pengujian yang akan dilakukan yaitu pengujian metode Winnowing. Pengujian dilakukan meliputi pengujian jumlah gram, jumlah window, basis bilangan prima, waktu proses dan jumlah langkah yang dilakukan untuk menghasilkan output. Perhitungan dengan n gram= 9, window = 3 dan bilangan prima = 7 dapat di lihat pada gambar 4.13.





Gambar 4. 13 Pengujian Sistem Algoritma Winnowing

Dilihat dari gambar 4.13 Nilai similaritas yang terbaca adalah 3.88 %. jika di bandingkan dengan nilai similarity plagiarism checker nilai presentase kemiripannya lebih besar menggunakan software yang di buat oleh penulis, di karenakan pada plagiarism cheker minimal 4 kata baru terbaca kemiripannya.

Kesimpulan

Berdasarkan rumusan masalah, hasil penelitian dan pembahasan mengenai pendeteksian kemiripan dua teks yang berbeda, dapat diambil beberapa kesimpulan, yaitu:

1. Aplikasi pendeteksi kemiripan dua teks dengan metode winnowing telah selesai di bangun dengan bahasa pemrograman php dan SQL database dan dapat mendeteksi kemiripan dua teks dengan baik
2. Berdasarkan hasil perbandingan uji coba dengan sistem dengan cara manual, dapat di simpulkan bahwa sistem mampu mendeteksi kemiripan dua teks yang berbeda cukup akurat dengan selisi presentase kemiripan 1 %

Daftar Pustaka

- Alamanda, R., Suhery, C., & Brianorman, Y. (2016). Aplikasi Pendeteksi Plagiat Terhadap Karya Tulis Berbasis Web Menggunakan Natural Language Processing Dan Algoritma Knuth- Morris -Pratt. *Jurnal Coding*, 4, 33-44.
- Alamsyah, N. (2017). Perbandingan Algoritma Winnowing Dengan Algoritma Rabin Karp Untuk Mendeteksi Plagiarisme Pada kemiripan Teks Judul Skripsi. *Jurnal Ilmiah Technologia*, 8, 124-134.
- Astutik, S., Cahyani, A. D., & Sopan, M. K. (2015). Sistem Penilaian Esai Otomatis Pada E- Learning Dengan Algoritma Winnowing. *Jurnal Informatika* , 12, 47-52.
- Cahyono, D. U. (2018). Aplikasi Deteksi Dini Plagiarisme Judul Tugas Akhir Mahasiswa Sekolah Tinggi Ilmu Kesehatan Yayasan Rs. Islam Surabaya Dengan Algoritma Rabin. *Jurnal Aplied Thnology and Computing Science*, 1, 1-10.
- Faizzani, w., & Hastarita R, F. (2014). Sistem FAQ Konsultasi Dokter Gigi Menggunakan Algoritma Winnowing dan Synonym Replacement. *Jurnal Simantec*, 4, 105-114.
- Jarwati, Prihandoko, A. C., & Yulia R, W. E. (2017). Penerapan Algoritma Winnowing Pada Sistem Rekomendasi Penentuan Dosen Pembimbing Skripsi (Studi Kasus Program Studi Sistem Informasi). *Jurnal Ilmiah Berkala Saintek*, 1, 11-20.
- Magdalena, H., & Santoso, H. (2017). Studi Faktor Penilaian Plagiat Menggunakan Analytical Hierarchy Process. *Jurnal Sisfo*, 6, 183-198.
- Mulyana, I., chairunnas, A., & Maesya, A. (2015). Identifikasi Plagiasi Karya Ilmiah berbasis Temu Kembali Informasi Menggunakan Algoritam Edit Distance Melalui Perin. *Jurnal ilmiah Nasional*, 866-874.
- Putra, D. A., Sujjaini, h., & Pratiwi, H. S. (2015). Perbandingan Algoritma Winnowing dengan Algoritma Rabin Karp Untuk Mendeteksi Plagiarisme Pada Kemiripan Teks Judul Skripsi. *Jurnal Istem dan Tekhnologi Informasi*, 1.
- Radili, A., & Sanjaya, S. (2016). Penerapan Metode Winnowing Fingerprint dan Naive Bayes Untuk Pengelompokan Dokumen. *Jurnal Core IT*, 3, 69-75.
- Said, B. (2017). Klasifikasi dan Deteksi Similaritas Data Sms Center Bupati Pamekasan Menggunakan Naive Bayes dan Winnowing. *Jurnal Link*, 26, 1-5.
- Sanjaya, S., & Absar, E. A. (2015). Pengelompokan Dokumen Menggunakan Winnowing Fingerprint. *Jurnal Core IT*, 1, 50-56.