

P-ISSN : 2337 - 8344

E-ISSN : 2623 - 1247

Jurnal InformaSI dan Komputer



**Diterbitkan Oleh :
STMIK DIAN CIPTA CENDIKIA KOTABUMI**

Volume 10 Nomor 1 Tahun 2022

Penerbit

Lembaga Penelitian STMIK Dian Cipta Cendikia Kotabumi

**Hak atas naskahh/tulisan tetap berada pada penulis, isi diluar tanggung jawab
penerbit dan Dewan Penyunting**



PENGANTAR REDAKSI

Puji syukur dipanjatkan kehadirat Tuhan Yang Maha Esa, atas karunia dan limpahan rahmatNYA jualah Jurnal Informasi dan komputer (JIK) STMIK Dian Cipta Cendikia Kotabumi ini dapat terwujud. Jurnal Informasi dan Komputer (JIK) yang terbit dua (2) kali dalam setahun ini merupakan suatu wadah untuk penyebar luasan hasil-hasil penelitian, studi pustaka, karya ilmiah yang berkaitan dengan Informasi dan Komputer khususnya bagi dosen-dosen STMIK Dian Cipta Cendikia Kotabumi serta umumnya para cendekiawan, praktisi, peneliti ilmu Informatika dan Komputer.

Harapan, dengan diterbitkannya Jurnal Informasi dan Komputer (JIK) ini sebagai salah satu bentuk sumbangan pemikiran dalam pengembangan ilmu informatika dan komputer yang berkaitan dengan kajian-kajian di bidang teknologi Informatik, Komunikasi Data dan Jaringan Komputer, perancangan dan Rekayasa Perangkat Lunak, serta ilmu-ilmu yang terkait dengan bidang Informasi dan Komputer lainnya.

Berkenaan dengan harapan tersebut, kepada para peneliti, dosen dan praktisi yang memiliki hasil-hasil penelitian, kajian pustaka, karya ilmiah dalam bidang tersebut diatas, dengan bangga redaksi Jurnal Informasi dan Komputer (JIK) menerima naskah ringkasan untuk dimuat pada jurnal Informasi dan Komputer (JIK) STMIK Dian Cipta Cendikia Kotabumi dengan berpedoman pada penulisan naskah jurnal sebagaimana dilampirkan pada halaman belakang (Bagian kulit dalam) buku jurnal ini.

Mutu dari suatu jurnal ilmiah tidak hanya ditentukan oleh para pengelolanya saja, tetapi para penulis dan pembaca jualah yang mempunyai peranan besar dalam meningkatkan mutu jurnal Informatika dan Komputer ini. Merujuk pada realita ini kamu sangat mengharapkan peran aktif dari peneliti untuk bersama-sama menjaga dan memelihara keberlangsungan dari jurnal Informasi dan Komputer STMIK Dian Cipta Cendikia Kotabumi ini. Yang juga tidak kalah pentingnya dari partisipasi tersebut diatas, adalah saran dan kritik yang membangun dari pembaca yang budiman agar kiranya dapat disampaikan langsung kepada redaksi JIK. Saran dan kritik yang membangun akan dijadikan masukan dan pertimbangan yang sangat berarti guna peningkatan mutu dan kualitas Jurnal Informasi dan Komputer STMIK Dian Cipta Cendikia Kotabumi.

Tak lupa diucapkan terima kasih yang tak terhingga atas perhatian dan kerjasama dari semua pihak yang tak dapat disebutkan satu persatu hingga dapat diterbitkan nya Jurnal Informasi dan Komputer (JIK) STMIK Dian Cipta Cendikia Kotabumi. Semoga apa yang telah diperbuat untuk kebaikan akan menjadi amal ibadah, amin.

Kotabumi, 25 April, 2022



Dewan Redaksi

JURNAL INFORMASI DAN KOMPUTER

Volume 10 Nomor 2 April 2022

Jurnal Informasi dan Komputer merupakan Sarana informasi ilmu pengetahuan, Teknologi dan Komunikasi yang berupa hasil penelitian, tulisan ilmiah, Atau pun studi pustaka. Jurnal ini terbit dua kali setahun pada bulan April dan Oktober. Berisi hasil penelitian ilmiah di bidang informatika yang bertujuan untuk menghubungkan adanya kesenjangan antar kemajuan teknologi dan hasil penelitian. Jurnal ini di terbitkan pertama kali pada tahun 2013.

Penanggung Jawab:

Ketua STMIK Dian Cipta Cendikia Kotabumi

Pembina:

Ketua STMIK Dian Cipta Cendikia Kotabumi
Ketua Lembaga Penelitian STMIK Dian Cipta Cendikia Kotabumi

Pimpinan Redaksi

Dwi Marisa Efendi, S.Kom., M.Ti

Redaksi pelaksana

Rustam, S.Kom., M.Ti (STMIK Dian Cipta Cendikia Kotabumi)
Nurmayanti M.Kom (STMIK Dian Cipta Cendikia Kotabumi)
Sukatmi, S.Kom., M.Kom (AMIK DCC Bandar Lampung)
Sampurna Dadi Riskiono, M.Kom (Universitas Teknokrat Indonesia)
Ifo Wahyu Pratama, S.Kom., M.Ti (AMIK MASTER Lampung)

Mitra Bestari

Dr. RZ. ABDUL AZIZ, ST., MT (Institut Informatika dan Bisnis Darmajaya)
Dr. Dadang Sudrajat, S.Si, M.Kom (STMIK IKMI Cirebon)
Dr. Septafiansyah Dwi Putra, S.T., M.T (Politeknik Negeri Lampung)
Dr. Evi Grativiani, S.E., M.S.I (Universitas Sebelas Maret)
Rohmat Indra Borman (Universitas Teknokrat Indonesia)
Ferry Wongso, S.KOm., M.Kom (STMIK Darma Pala Riau)
Ferly Ardhy, S.Kom., M.Ti (Universitas Aisyah Pringsewu)
Firmansyah, S.E., M.Si (STMIK Darma Pala Riau)

Amarudin (Universitas Teknokrat Indonesia)
Didi Susianto, S.T., M.Kom (AMIK Dian Cipta Cendikia Bandar Lampung)
Alhibarsyah, St., M.Kom (STMIK Tunas Bangsa Bandar Lampung)
Kemal Farouq Mauladi, S.Kom .M.Kom (Universitas Islam Lamongan)
Rima Mawarni, M.Kom (STMIK Dian Cipta Cendikia Kotabumi)
Wira Jaya Hartono, S.Pd., M.Pd (STMIK Darma Pala Riau)

Penerbit : STMIK Dian Cipta Cendikia Kotabumi Bekerja Sama Dengan LPPM STMIK Dian Cipta Cendikia Kotabumi.

Alamat Redaksi/Penerbit:

Jl. Negara No. 3 Candimas Kotabumi Lampung Utara
No Telp/Fax 0724 23003
Email : lppm-stmik@dcc.ac.id



JURNAL INFORMASI DAN KOMPUTER VOL. 10 NO. 2 THN. 2022

DAFTAR ISI

	Halaman
Peningkatan Pengelolaan Arsip Surat Menyurat Melalui Aplikasi Berbasis Web Dengan Metode <i>First In First Out</i> Yuli Syafitri ¹ , Reni Astika ² , Lusia Septia Eka Esti Rahayu ³ , (AMIK Dian Cipta Cendikia ¹² , AMIK Lampung ³)	01-08
Pengembangan Sistem Informasi Perpustakaan Berbasis Web Pada Amik Dian Cipta Cendikia Bandar Lampung Sukatmi ¹ , Euis Mustika Prianganti ² , Astriyanti ³ (AMIK DCC Bandar Lampung ¹²³)	09-14
Klasifikasi Penyakit <i>Powdery Mildew</i> Pada Ceri Manis Dengan Menggunakan Algoritma <i>Convolutional Neural Network (CNN)</i> Iwansyah Edo Hendrawan ¹ , M. Ilhamsyah ² , Dadang Yusup ³ (Universitas Singaperbangsa Karawang ¹²³)	15-20
Penerapan Finite State Automata Pada Desain Vending Machine Masker Dan Hand Sanitizer Ridwan ¹ , Windu Gata ² , Hafifah Bella Novitasari ³ , Laela Kurniawati ⁴ , Sri Rahayu ⁵ (Universitas Nusa Mandiri ¹²).....	21-28
Analisis Perhitungan Muatan Sedimentasi Berdasarkan Kedalaman Air (<i>Chart Datum</i>) Pada Senipah Channel Di Kabupaten Kutai Kartanegara Berbasis Web Salmajah (Stmik Handayani Makasar)	29-43
Aplikasi Pembelajaran Ilmu Tajwid Berbasis Mobile Novita Lestari Anggreini ¹ , Ichsan Perdana Putra ² (Politeknik TEDC Bandung).....	44-49
Implementasi Algoritma Naïve Bayes Untuk Memprediksi Pengaruh Media Sosial Terhadap Semangat Belajar Mahasiswa Di Masa Pandemi Covid 19 Fiqih Satria ¹ , Hermanto ² (Universitas Raden Intan Lampung)	50-56
Klasifikasi Kinerja Pembayaran Angsuran Dengan Algoritma Naive Bayes (Studi Kasus : Data Nasabah Koperasi Simpan Pinjam Pembiayaan Syariah Bina Bersama) Dwi Marisa ¹ , Sigit Mintoro ² , Supriyanto ³ , Sani Hanika lubis ⁴ , Sri Lestari ⁵ (STMik Dian Cipta Cendikia Kotabumi)	57-61
Peningkatan Akurasi Prediksi Pengadaan Bahan Baku Produksi Dengan Menggunakan Metode <i>Neural Network</i> Mumtaz Muttakin ¹ , Sabar Hanadwiputra ²	

(STMIK Bani Saleh, Bekasi)	62-72
Penerapan Konsep Finite State Automata Pada Simulasi Vending Machine Pergantian Seragam Karyawan Ristyani Slamet ¹ , Windu Gata ² , Ketut Sakho Parthama ³ , Nita Merlina ⁴ , Eni Heni Hermaliani ⁵ (Universitas Nusa Mandiri ^{1,2,4,5} , Universitas Pramita Indonesia ³)	
	73-79
Penerapan Metode Electre Untuk Pemilihan Pengajar Terbaik Muchamad Maskhur ¹ , Wiwien Hadikurniawati ² (Universitas Stikubank, Semarang).....	
	80-88
Sistem Pendukung Keputusan Kelayakan Kenaikan Pangkat Jabatan Fungsional(Asn) Metode Topsis Nurmayanti ¹ , Merri Parida ² , M. Reka Yuansyah ³ (STMIK Dian Cipta Cendikia kotabumi)	
	89-96
Media Pembelajaran Interaktif Berbasis Articulate Storyline 3 Pada Mata Kuliah Pemrograman Berorientasi Objek Dikwan Moeis ¹ , Andi Harmin ² (STMIK Profesional Makasar ¹²)	
	97-106
Penentuan Penerima Beasiswa Di Stmik Bani Saleh Dengan Perbandingan Metode Algoritma C4.5 Dan Knearest Neighbors Siti Chodijah ¹ , Mohammad Iqbal ² (Universitas Gunadama ¹²)	
	107-114
Perancangan Sistem Informasi Kehadiran Pegawai Dan Skp (Sikap) Pada Institut Agama Islam Negeri (Iain) Metro Toto Andri Puspito (Institut Agama Islam Negeri Metro)	
	115-120
Kualitas Layanan Sistem Informasi Akademik (Siakad) Terhadap Kepuasan Mahasiswa Sebagai Pengguna Aidah Hami ¹ , Dyah Anggraini ² (Stmik Bani Saleh ¹ , Universitas Gunadarma)	
	121-129
Implementasi Metode Bag Of Visual Words Dalam Pengenalan Citra Masker Pada Wajah Komang Budiarta ¹ , I Made Budi Adnyana ² , Gede Herdian Setiawan ³ (ITB STIKOM BALI)	
	130-137
Sistem Tiket Helpdesk Pada Stmik Bani Saleh Zaenal Mutaqin Subekti ¹ , Kresno Murti Prabowo ² , Budi ³ (STMIK Bani Salih ¹²³)	
	138-144
Algoritma Naive Bayes Untuk Memprediksi Jumlah Siswa Berpotensi Drop Out Sidik Rahmatullah ¹ , Ngajiyanto ² , Pakarti Riswanto ³ , Arief Hendriawan ⁴ (STMIK Dian Cipta Cendikian Kotabumi ¹²³)	
	145-153
Pengklasteran Risiko Covid-19 Di Riau Menggunakan Teknik <i>One Hot Encoding</i> Dan Algoritma <i>K-Means Clustering</i> Silviana ¹ , Rahmad Kurniawan ² , Alwis Nazir ³ , Elvia Budianita ⁴ ,	

Fadhillah Syafria ⁵ , Siska Kurnia Gusti ⁶ (Universitas Riau ² , Universitas Islam Negeri Sultan Syarif Kasim Riau ^{1,3,4,5,6})	154-163
Aplikasi Pengelolaan <i>E-Document</i> Sistem Penjaminan Mutu Internal Menggunakan Metode <i>User Centered Design</i> Andi Harmin ¹ , Rosnani ² (STMIK Profesional Makassar ¹²)	164-173
Game Edukasi Mengenal Kepulauan Indonesia Menggunakan <i>Unity 3d</i> Untuk Meningkatkan Minat Belajar Siswa Sekolah Dasar Tri Aditama ¹ , Ade Irma Purnamasari ² , Tati Suprapti ³ (STMIK IKMI Cirebon)	174-179
Alat Pemantau Bilik Desinfektan Untuk Pencegahan Penularan Covid 19 Dengan Internet Of Things (I.O.T) Berbasis Microcontroller Yusup Supriadi (Universitas Panca Sakti Bekasi)	180-193
Penerapan Metode <i>Fuzzy Ahp (Analytical Hierarchy Process)</i> Sebagai Sistem Pendukung Keputusan Dosen Terbaik (Studi Kasus : Stmik Pringsewu) Afrizal Martin ¹ , Bambang Suprpto ² , Sulasminarti ³ , Akni Widiyastuti ⁴ , Deny Firmansyah Kurniawan ⁵ , Henry Simanjuntak ⁶ (STMIK Pringsewu ¹ , AMIK Dian Cipta Cendikia Pringsewu ²³⁴⁵⁶)	194-207
Game Edukasi Pembelajaran Anak Usia Dini Berbasis Android Ferly Ardhy ¹ Gusnaedi Adam ² Agustinus Eko Setiawan ³ Anti Aisyah ⁴ (unversitas aisyah pring sewu, STMIK Dian Cipta Cendikia Kotabumi)	208-213
Sistem Informasi Penjualan Alat-Alat Pancing Berbasis Web Pada Toko Yoggi Bersaudara Di Talang Padang Kabupaten Tanggamus (Studi Kasus Toko Yoggi Bersaudara) Rima Mawarni ¹ , Dewi Triyanti ² , Dodi Afriansyah ³ , Yoggi Kurniawan ⁴ (STMIK Dian Cipta Cendikia Kotabumi ¹⁴ AMIK Dian Cipta Cendikia Pringsewu ²³)..	214-219
Implementasi Algoritma <i>Winnowing</i> Dalam Mendeteksi Plagiarisme Pada Tugas Mahasiswa Ida Bagus Ketut Surya Arnawa (ITB STIKOM BALI)	220-230
Sistem Pendukung Keputusan Pemilihan Guru Teladan Menggunakan Metode <i>Composite Performance Index (Cpi)</i> Pada Smk Negeri 1 Kotabumi Rustam ¹ , Pakarti Riswanto ² , Dwi Marisa Efendi ³ , Asep Afandi ⁴ , Supriyanto ⁵ , Desri Arisandi ⁶ (STMIK Dian Cipta Cendikia Kotabumi ¹²³⁴)	231-238

IMPLEMENTASI ALGORITMA WINNOWER DALAM MENDETEKSI PLAGIARISME PADA TUGAS MAHASISWA

Ida Bagus Ketut Surya Arnawa
ITB STIKOM Bali
Jl. Raya Puputan No 86 Renon, Denpasar
E-mail : arnawa@stikom-bali.ac.id

ABSTRAK

Pendidikan merupakan suatu proses mengubah sikap dan perilaku seseorang atau kelompok orang dalam upaya meningkatkan kualitas hidup. Banyak perguruan tinggi baik negeri maupun swasta menyelenggarakan Pendidikan untuk mendukung meningkatkan kualitas Pendidikan dan daya saing bangsa. Institut Teknologi dan Bisnis ITB STIKOM Bali merupakan salah satu perguruan tinggi swasta yang ada di provinsi Bali. ITB STIKOM Bali merupakan perguruan tinggi IT yang dalam proses belajar mengajar semuanya menggunakan teknologi informasi. Dari berbagai kelebihan yang ditawarkan menggunakan teknologi informasi terdapat kekurangan atau kendala terutama pada sistem E-Learning, dimana E-Learning belum mampu untuk mengcover plagiarisme pada dokumen tugas mahasiswa yang dikirimkan. Dosen mengalami kendala dalam proses koreksi tugas mahasiswa yang terindikasi melakukan plagiarisme. Dalam mencegah permasalahan plagiarisme diperlukan sebuah sosialisasi mengenai pentingnya menghindari tindakan plagiarisme dan undang – undang yang mengatur tentang tindakan plagiarisme. Selain itu juga perlu dilakukan sebuah tindakan untuk mendeteksi terjadinya tindakan plagiarisme sehingga tindakan tersebut dapat dicegah. Dalam mendeteksi terjadinya plagiarisme pada tugas mahasiswa perlu sebuah tools yang mampu mencocokkan tugas mahasiswa guna mengetahui tingkat kemiripan tugas mahasiswa dengan menggunakan algoritma winnower. Hasil penelitian menunjukkan algoritma winnower mendapatkan hasil terbaik pada uji coba dengan nilai n-gram (n) = 3 dan window (w) = 3.

Kata kunci : Text Mining, Winnower, Tugas mahasiswa

ABSTRACTS

Education is a process of changing the attitudes and behavior of a person or group of people in an effort to improve the quality of life. Many universities, both public and private, organize education to support improving the quality of education and the nation's competitiveness. The ITB STIKOM Bali Institute of Technology and Business is one of the private universities in the province of Bali. ITB STIKOM Bali is an IT college which in the teaching and learning process all use information technology. Of the various advantages offered by using information technology, there are shortcomings or obstacles, especially in the E-Learning system, where E-Learning has not been able to cover plagiarism in the submitted student

assignment documents. Lecturers experience problems in the process of correcting student assignments which are indicated to be plagiarism. In preventing plagiarism problems, a socialization is needed about the importance of avoiding plagiarism and the laws governing plagiarism. In addition, it is also necessary to take an action to detect the occurrence of plagiarism so that these actions can be prevented. In detecting the occurrence of plagiarism in student assignments, a tool is needed that is able to match student assignments in order to determine the level of similarity in student assignments using the winnowing algorithm. The results showed that the winnowing algorithm got the best results in trials with a value of n-gram (n) = 3 and window (w) = 3.

Keywords : Text Mining, Winnowing, student work

1.PENDAHULUAN

Pendidikan merupakan suatu proses mengubah sikap dan perilaku seseorang atau kelompok orang dalam upaya meningkatkan kualitas hidup. Dalam menempuh pendidikan seseorang dapat menempuhnya melalui pendidikan formal atau pendidikan informal. Pemerintah Indonesia mewajibkan setiap warga menempuh pendidikan dasar. Hal ini merupakan amanah undang-undang dasar 1945. Setelah menempuh Pendidikan dasar diharapkan dapat melanjutkan sampai jenjang perguruan tinggi guna untuk meningkatkan kualitas dan daya saing bangsa. Banyak perguruan tinggi baik negeri maupun swasta menyelenggarakan Pendidikan untuk mendukung meningkatkan kualitas Pendidikan dan daya saing bangsa.

Institut Teknologi dan Bisnis ITB STIKOM Bali merupakan salah satu perguruan tinggi swasta yang ada di provinsi Bali. ITB STIKOM Bali merupakan perguruan tinggi IT yang dalam proses belajar mengajar semuanya menggunakan teknologi informasi. Teknologi yang dipergunakan dalam proses belajar mengajar yaitu teknologi berbasis website seperti Sistem Informasi Online, E-Learning, Ms Teams, Sistem Informasi Akademik, Sistem Informasi Dosen. Untuk menunjang penelitian dan pengabdian juga dikembangkan website E-Research dan untuk perpustakaan menggunakan E-Library. Dalam proses belajar mengajar E-Learning merupakan yang

paling dominan digunakan. E-Learning dapat mengcover kegiatan seperti dosen upload materi, tugas, quiz dan diskusi sedangkan mahasiswa dapat mendownload materi, mengupload tugas mengerjakan quiz dan berdiskusi di forum kelas.

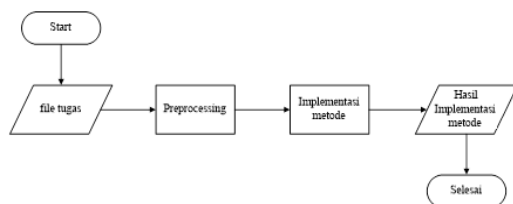
Dari berbagai kelebihan yang ditawarkan menggunakan teknologi informasi terdapat kekurangan atau kendala terutama pada sistem E-Learning, dimana E-Learning belum mampu untuk mengcover plagiarisme pada dokumen tugas mahasiswa yang dikirimkan. Dosen masih mengoreksi tugas secara manual, sehingga dosen memerlukan waktu yang cukup lama untuk mengoreksi semua tugas mahasiswa. Selain itu juga mengalami kendala dalam proses koreksi tugas mahasiswa yang terindikasi melakukan plagiarisme. Plagiarisme merupakan salah satu tindakan mengambil atau mengakui ide / hasil karya orang lain tanpa mencantumkan atau tanpa ijin dari yang punya ide/ hasil karya. Banyak faktor yang menyebabkan tindakan plagiarisme terjadi khususnya dikalangan mahasiswa, seperti kurangnya kemampuan mahasiswa dalam mengerjakan tugas yang diberikan oleh dosen. Selain itu juga beberapa mahasiswa kurang memahami tentang makna plagiarisme serta undang-undang yang mengatur tentang tindakan plagiarisme.

Dalam mencegah permasalahan plagiarisme diperlukan sebuah sosialisasi mengenai pentingnya menghindari tindakan plagiarisme dan undang – undang yang mengatur tentang tindakan plagiarisme. Selain itu juga perlu

dilakukan sebuah tindakan untuk mendeteksi terjadinya tindakan plagiarisme sehingga tindakan tersebut dapat dicegah. Dalam mendeteksi terjadinya plagiarisme pada tugas mahasiswa perlu sebuah tools yang mampu mencocokkan tugas mahasiswa guna mengetahui tingkat kemiripan tugas mahasiswa dengan menggunakan algoritma text mining. Ada banyak algoritma yang dapat digunakan untuk menganalisis kemiripan tugas mahasiswa salah satunya adalah algoritma winnowing. Dalam penelitian ini penulis akan mengimplementasikan algoritma winnowing dalam mendeteksi kemiripan tugas mahasiswa [1].

2. METODE PENELITIAN

Adapun alur penelitian yang digunakan dalam penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitian

Gambar 1. merupakan alur penelitian yang digunakan dalam penelitian ini. File tugas mahasiswa akan melalui proses preprocessing terlebih dahulu seperti Case Folding, Filtering, Tokenizing dan Stemming. Setelah proses preprocessing dilakukan maka tahap selanjutnya dilakukan implementasi algoritma winnowing dalam mengukur tingkat kemiripan tugas mahasiswa.

2.1. Algoritma Winnowing

Algoritma winnowing merupakan salah satu algoritma yang dapat digunakan untuk mengukur tingkat kesamaan kata (document fingerprinting)

untuk mendeteksi plagiarisme. Rolling hash merupakan algoritma yang digunakan mencari nilai hash dalam winnowing. Nilai hash merupakan nilai numerik yang terbentuk dari perhitungan ASCII setiap karakter. Langkah-langkah dalam penerapan Algoritma Winnowing adalah sebagai berikut [2]:

1. Langkah Pertama

Penghapusan karakter yang tidak relevan yaitu berupa spasi, tanda baca serta spesial karakter seperti !, @, #, \$, %, ^, &, *, (,), _ , - ?

Contoh : Belajar Pemrograman C# belajarpemrogramanc

Akan dirubah menjadi

2. Langkah Kedua

Pembentukan Rangkaian n-gram yaitu dengan cara membentuk rangkaian karakter sepanjang n dari hasil pembuangan karakter yang tidak relevan pada langkah pertama. Dari text diatas telah dibersihkan dengan ukuran k=5 belaj elaja lajar ajarp jarpe arpem rpemr pemro emrog mrogr rogra ogram grama raman amanc

3. Langkah Ketiga

Menghitung fungsi hash untuk setiap n-gram yaitu dengan cara menghitung nilai - nilai hash dari setiap gram, algoritma holling hash digunakan untuk menghitung nilai hash dalam algoritma winnowing. Rolling hash merupakan suatu teknik untuk mentransformasikan sebuah string menjadi nilai yang unik dengan panjang tertentu yang berfungsi sebagai penanda string tersebut. Fungsi tersebut menghasilkan nilai ini disebut fungsi hash, sedangkan nilai yang dihasilkan disebut nilai hash. Fungsi hash $H(c1...ck)$ didefinisikan sebagai berikut :

$$H(ck) = c1 * b^{(k-1)} + c2 * b^{(k-2)} + \dots + ck * b^{(k-k)}$$

Keterangan :

1247

c = nilai ascii karakter

b = basis (bilangan prima)

k = banyak karakter

hasil rolling hash dari kalimat diatas yaitu

1583490 1635489 1724342 1574366

1696180 1586683 1831680 1788777

1638938 1762281 1830629 1777214

1672790 1812547 1578302

4. Langkah Keempat

Pembentukan window dari nilai hash dari window dengan ukuran 3 yaitu sebagai berikut :

{1583490 1635489 1724342}
 {1635489 1724342 1574366}
 {1724342 1574366 1696180}
 {1574366 1696180 1586683}
 {1696180 1586683 1831680}
 {1586683 1831680 1788777}
 {1831680 1788777 1638938}
 {1788777 1638938 1762281}
 {1638938 1762281 1830629}
 {1762281 1830629 1777214}
 {1830629 1777214 1672790}
 {1777214 1672790 1812547}
 {1672790 1812547 1578302 }

5. Langkah Kelima

Langkah terakhir yaitu memilih nilai terkecil dari setiap window untuk dijadikan fingerprint, hasil dari nilai fingerprintnya sebagai berikut:

[1583490,0] [1574366,3] [1586683,5]
 1638938,8] [1762281,9] [1672790,12]
 [1578302,14]

Nilai fingerprint yang dibentuk dari algoritma winnowing digunakan untuk mengukur prosentase kemiripan teks pada persamaan Jaccard Coefficient. Persamaan Jaccard

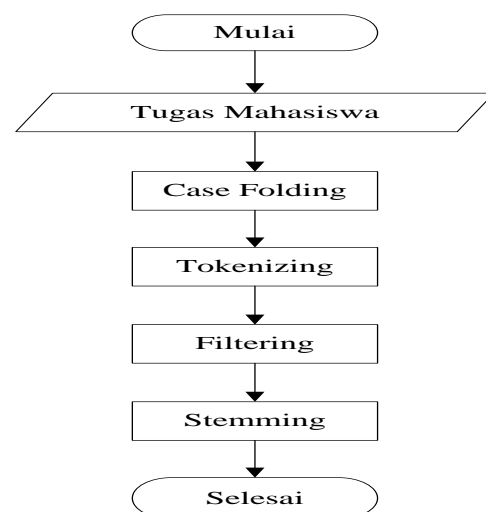
Coefficient digunakan untuk menghitung kemiripan (similarity) dari kumpulan kata-kata yang telah dihitung nilai hash nya. Berikut ini rumus persamaan Jaccard Coefficient.

$$\text{Similarity} = \frac{\text{Jumlah_fingerprint_sama}}{\text{Total_seluruh_fingerprint}} \times 100$$

3. HASIL DAN PEMBAHASAN

3.1. Flowchart Preprocessing

Pada tahapan preprocessing dilakukan beberapa tahapan yaitu diawali dengan case folding yaitu mengkonversi seluruh huruf dalam tugas mahasiswa menjadi huruf kecil. Kemudian dilanjutkan dengan tahap tokenizing yaitu proses memisahkan setiap kata yang menyusun dokumen tugas mahasiswa dengan memanfaatkan karakter spasi. Tahap selanjutnya yaitu filtering dimana tahap ini akan mengambil kata yang penting dari tahap tokenizing. Tahap terakhir yaitu stemming dimana pada tahap ini mengubah sebuah kata menjadi kata dasar dengan menghilangkan imbuhan. Gambar 3 merupakan flowchart preprocessing [4][5].



Gambar.3 Flowchat Preprocessing

3.2. Implementasi Perbandingan Algoritma

Berikut ini merupakan tahapan-tahapan dalam mengimplementasikan algoritma winnowing mendeteksi plagiarisme pada tugas mahasiswa yaitu :

Tabel 1. Data Uji

Percobaan 1. Tugas Mahasiswa dengan tingkat kemiripan 100 %	
Kriptografi adalah ilmu mengenai teknik enkripsi dimana “naskah asli” (plaintext) diacak menggunakan suatu kunci enkripsi menjadi “naskah acak yang sulit dibaca” (ciphertext) oleh seseorang yang tidak memiliki kunci dekripsi. Dekripsi menggunakan kunci dekripsi bisa mendapatkan kembali data asli. Probabilitas mendapat kembali naskah asli oleh seseorang yang tidak mempunyai kunci dekripsi dalam waktu yang tidak terlalu lama adalah sangat kecil.	Kriptografi adalah ilmu mengenai teknik enkripsi dimana “naskah asli” (plaintext) diacak menggunakan suatu kunci enkripsi menjadi “naskah acak yang sulit dibaca” (ciphertext) oleh seseorang yang tidak memiliki kunci dekripsi. Dekripsi menggunakan kunci dekripsi bisa mendapatkan kembali data asli. Probabilitas mendapat kembali naskah asli oleh seseorang yang tidak mempunyai kunci dekripsi dalam waktu yang tidak terlalu lama adalah sangat kecil.
Percobaan 2. Tugas Mahasiswa dengan tingkat kemiripan 50 %	
Kriptografi adalah ilmu mengenai teknik enkripsi dimana “naskah asli” (plaintext) diacak menggunakan suatu kunci enkripsi menjadi “naskah acak yang sulit dibaca” (ciphertext) oleh	Kriptografi adalah ilmu mengenai teknik enkripsi dimana “naskah asli” diacak menggunakan suatu kunci enkripsi menjadi “naskah acak yang sulit dibaca” oleh seseorang yang

seseorang yang tidak memiliki kunci dekripsi. Dekripsi menggunakan kunci dekripsi bisa mendapatkan kembali data asli. Probabilitas mendapat kembali naskah asli oleh seseorang yang tidak mempunyai kunci dekripsi dalam waktu yang tidak terlalu lama adalah sangat kecil.	tidak memiliki kunci dekripsi. Kriptanalisis adalah ilmu dan seni untuk memecahkan chiperteks menjadi plainteks tanpa mengetahui kunci yang diberikan. Pelakunya disebut kriptanalisis. Kriptologi (cryptology) adalah studi mengenai kriptografi dan kriptanalisis Kriptografi sudah lama digunakan oleh tentara Sparta di Yunani pada permulaan tahun 400 SM.
Percobaan 3. Tugas Mahasiswa dengan tingkat kemiripan 30 %	
Kriptografi adalah ilmu mengenai teknik enkripsi dimana “naskah asli” (plaintext) diacak menggunakan suatu kunci enkripsi menjadi “naskah acak yang sulit dibaca” (ciphertext) oleh seseorang yang tidak memiliki kunci dekripsi. Dekripsi menggunakan kunci dekripsi bisa mendapatkan kembali data asli. Probabilitas mendapat kembali naskah asli oleh seseorang yang tidak mempunyai kunci dekripsi dalam waktu yang tidak terlalu lama adalah sangat kecil.	Kriptografi adalah ilmu mengenai teknik enkripsi dimana “naskah asli” diacak menggunakan suatu kunci enkripsi menjadi “naskah acak yang sulit dibaca”. Kriptanalisis adalah ilmu dan seni untuk memecahkan chiperteks menjadi plainteks tanpa mengetahui kunci yang diberikan. Pelakunya disebut kriptanalisis. Kriptologi (cryptology) adalah studi mengenai kriptografi dan kriptanalisis Kriptografi sudah lama digunakan oleh tentara Sparta di Yunani pada

	permulaan tahun 400 SM. Mereka menggunakan alat yang namanya scytale.
--	---

1. Case Folding

Tahap pertama dalam preprocessing yaitu case folding yaitu proses dimana semua data uji akan dirubah menjadi huruf kecil semua karakter selain huruf dan angka dihilangkan dan dianggap delimiter. Hasil dari case folding dapat dilihat pada tabel 2

Tabel 2. Hasil Case Folding

Percobaan 1. Tugas Mahasiswa dengan tingkat kemiripan 100 %	
kriptografi adalah ilmu mengenai teknik enkripsi dimana “naskah asli” (plaintext) diacak menggunakan suatu kunci enkripsi menjadi “naskah acak yang sulit dibaca” (ciphertext) oleh seseorang yang tidak memiliki kunci dekripsi. dekripsi menggunakan kunci dekripsi bisa mendapatkan kembali data asli. probabilitas mendapat kembali naskah asli oleh seseorang yang tidak mempunyai kunci dekripsi dalam waktu yang tidak terlalu lama adalah sangat kecil.	kriptografi adalah ilmu mengenai teknik enkripsi dimana “naskah asli” (plaintext) diacak menggunakan suatu kunci enkripsi menjadi “naskah acak yang sulit dibaca” (ciphertext) oleh seseorang yang tidak memiliki kunci dekripsi. dekripsi menggunakan kunci dekripsi bisa mendapatkan kembali data asli. probabilitas mendapat kembali naskah asli oleh seseorang yang tidak mempunyai kunci dekripsi dalam waktu yang tidak terlalu lama adalah sangat kecil.
Percobaan 2. Tugas Mahasiswa dengan tingkat kemiripan 50 %	
kriptografi adalah ilmu mengenai teknik	kriptografi adalah ilmu mengenai teknik

enkripsi dimana “naskah asli” (plaintext) diacak menggunakan suatu kunci enkripsi menjadi “naskah acak yang sulit dibaca” (ciphertext) oleh seseorang yang tidak memiliki kunci dekripsi. dekripsi menggunakan kunci dekripsi bisa mendapatkan kembali data asli. probabilitas mendapat kembali naskah asli oleh seseorang yang tidak mempunyai kunci dekripsi dalam waktu yang tidak terlalu lama adalah sangat kecil.	enkripsi dimana “naskah asli” diacak menggunakan suatu kunci enkripsi menjadi “naskah acak yang sulit dibaca” oleh seseorang yang tidak memiliki kunci dekripsi. kriptanalisis adalah ilmu dan seni untuk memecahkan chiperteks menjadi plainteks tanpa mengetahui kunci yang diberikan. pelakunya disebut kriptanalisis.kriptologi (cryptology) adalah studi mengenai kriptografi dan kriptanalisis kriptografi sudah lama digunakan oleh tentara sparta di yunani pada permulaan tahun 400 sm.
Percobaan 3. Tugas Mahasiswa dengan tingkat kemiripan 30 %	
kriptografi adalah ilmu mengenai teknik enkripsi dimana “naskah asli” (plaintext) diacak menggunakan suatu kunci enkripsi menjadi “naskah acak yang sulit dibaca” (ciphertext) oleh seseorang yang tidak memiliki kunci dekripsi. dekripsi menggunakan kunci dekripsi bisa mendapatkan kembali data asli. probabilitas mendapat kembali naskah asli oleh seseorang yang tidak	kriptografi adalah ilmu mengenai teknik enkripsi dimana “naskah asli” diacak menggunakan suatu kunci enkripsi menjadi “naskah acak yang sulit dibaca”.kriptanalisis adalah ilmu dan seni untuk memecahkan chiperteks menjadi plainteks tanpa mengetahui kunci yang diberikan. pelakunya disebut kriptanalisis.kriptologi (cryptology) adalah studi mengenai kriptografi dan

mempunyai kunci dekripsi dalam waktu yang tidak terlalu lama adalah sangat kecil.	kriptanalisis kriptografi sudah lama digunakan oleh tentara sparta di Yunani pada permulaan tahun 400 sm. mereka menggunakan alat yang namanya scytale.
---	---

2. Filtering

Filtering merupakan proses menghapus kata-kata pada sebuah kalimat yang sering keluar dan dianggap tidak penting seperti penghubung, waktu, kata tanya dan lain sebagainya. Hasil dari Remove Stop Word dapat dilihat pada tabel 3

Tabel 3. Hasil Remove Stop Word

Percobaan 1. Tugas Mahasiswa dengan tingkat kemiripan 100 %	
kriptografi ilmu mengenai teknik enkripsi dimana naskah asli plaintext diacak menggunakan suatu kunci enkripsi menjadi naskah acak sulit dibaca ciphertext oleh seseorang tidak memiliki kunci dekripsi dekripsi menggunakan kunci dekripsi bisa mendapatkan kembali data asli probabilitas mendapat kembali naskah asli oleh seseorang tidak mempunyai kunci dekripsi dalam waktu tidak terlalu lama sangat kecil	kriptografi ilmu mengenai teknik enkripsi dimana naskah asli plaintext diacak menggunakan suatu kunci enkripsi menjadi naskah acak sulit dibaca ciphertext oleh seseorang tidak memiliki kunci dekripsi dekripsi menggunakan kunci dekripsi bisa mendapatkan kembali data asli probabilitas mendapat kembali naskah asli oleh seseorang tidak mempunyai kunci dekripsi dalam waktu tidak terlalu lama sangat kecil
Percobaan 2. Tugas Mahasiswa dengan tingkat kemiripan 50 %	

kriptografi ilmu mengenai teknik enkripsi dimana naskah asli plaintext diacak menggunakan suatu kunci enkripsi menjadi naskah acak sulit dibaca ciphertext oleh seseorang tidak memiliki kunci dekripsi dekripsi menggunakan kunci dekripsi bisa mendapatkan kembali data asli probabilitas mendapat kembali naskah asli oleh seseorang tidak mempunyai kunci dekripsi dalam waktu tidak terlalu lama sangat kecil	kriptografi ilmu mengenai teknik enkripsi dimana naskah asli diacak menggunakan suatu kunci enkripsi menjadi naskah acak sulit dibaca oleh seseorang tidak memiliki kunci dekripsi kriptanalisis ilmu seni memecahkan chiperteks menjadi plaintexts tanpa mengetahui kunci diberikan pelakunya kriptanalisis kriptologi cryptology studi mengenai kriptografi kriptanalisis kriptografi sudah lama digunakan oleh tentara sparta Yunani pada permulaan tahun sm
Percobaan 3. Tugas Mahasiswa dengan tingkat kemiripan 30 %	
kriptografi ilmu mengenai teknik enkripsi dimana naskah asli plaintext diacak menggunakan suatu kunci enkripsi menjadi naskah acak sulit dibaca ciphertext oleh seseorang tidak memiliki kunci dekripsi dekripsi menggunakan kunci dekripsi bisa mendapatkan kembali data asli probabilitas mendapat kembali naskah asli oleh seseorang tidak	kriptografi ilmu mengenai teknik enkripsi dimana naskah asli diacak menggunakan suatu kunci enkripsi menjadi naskah acak sulit dibaca kriptanalisis ilmu seni memecahkan chiperteks menjadi plaintexts tanpa mengetahui kunci diberikan pelakunya kriptanalisis kriptologi cryptology studi mengenai

mempunyai kunci dekripsi dalam waktu tidak terlalu lama sangat kecil	kriptografi kriptanalisis kriptografi sudah lama digunakan oleh tentara sparta yunani pada permulaan tahun sm mereka menggunakan alat namanya scytale
--	---

3. Tokenizing

Tokenizing merupakan proses pemotongan string menjadi potongan kata yang kemudian disusun menjadi baris. Hasil Tokenizing dapat dilihat pada tabel 4

Tabel 4. Hasil Tokenizing

Percobaan 1. Tugas Mahasiswa dengan tingkat kemiripan 100 %	
kriptografi ilmu	kriptografi ilmu
mengenai teknik	mengenai teknik
enkripsi dimana	enkripsi dimana
naskah asli	naskah asli
plaintext diacak	plaintext diacak
menggunakan suatu	menggunakan
kunci enkripsi	suatu kunci
menjadi naskah	enkripsi menjadi
acak sulit dibaca	naskah acak sulit
ciphertext oleh	dibaca ciphertext
seseorang tidak	oleh seseorang
memiliki kunci	tidak memiliki
dekripsi dekripsi	kunci dekripsi
menggunakan kunci	dekripsi
dekripsi bisa	menggunakan
mendapatkan	kunci dekripsi
kembali data asli	bisa mendapatkan
probabilitas	kembali data asli
mendapat kembali	probabilitas
naskah asli oleh	mendapat kembali
seseorang tidak	naskah asli oleh
mempunyai kunci	seseorang tidak
dekripsi dalam	mempunyai kunci
waktu tidak terlalu	dekripsi dalam
lama sangat kecil	waktu tidak

	terlalu lama
	sangat kecil
Percobaan 2. Tugas Mahasiswa dengan tingkat kemiripan 50 %	
kriptografi ilmu	kriptografi ilmu
mengenai teknik	mengenai teknik
enkripsi dimana	enkripsi dimana
naskah asli	naskah asli
plaintext diacak	diacak
menggunakan suatu	menggunakan
kunci enkripsi	suatu kunci
menjadi naskah	enkripsi menjadi
acak sulit dibaca	naskah acak sulit
ciphertext oleh	dibaca oleh
seseorang tidak	seseorang tidak
memiliki kunci	memiliki kunci
dekripsi dekripsi	dekripsi
menggunakan kunci	kriptanalisis ilmu
dekripsi bisa	seni memecahkan
mendapatkan	chiperteks
kembali data asli	menjadi plaintext
probabilitas	tanpa mengetahui
mendapat kembali	kunci diberikan
naskah asli oleh	pelakunya
seseorang tidak	kriptanalisis
mempunyai kunci	kriptologi
dekripsi dalam	cryptology studi
waktu tidak terlalu	mengenai
lama sangat kecil	kriptografi
	kriptanalisis
	kriptografi sudah
	lama digunakan
	oleh tentara
	sparta yunani
	pada permulaan
	tahun sm
Percobaan 3. Tugas Mahasiswa dengan tingkat kemiripan 30 %	
kriptografi ilmu	kriptografi ilmu
mengenai teknik	mengenai teknik
enkripsi dimana	enkripsi dimana
naskah asli	naskah asli
plaintext diacak	diacak
menggunakan suatu	menggunakan
kunci enkripsi	suatu kunci
menjadi naskah	enkripsi menjadi
acak sulit dibaca	naskah acak sulit
ciphertext oleh	dibaca

seseorang tidak memiliki kunci dekripsi dekripsi menggunakan kunci dekripsi bisa mendapatkan kembali data asli probabilitas mendapat kembali naskah asli oleh seseorang tidak mempunyai kunci dekripsi dalam waktu tidak terlalu lama sangat kecil	kriptanalisis ilmu seni memecahkan chiperteks menjadi plaintext tanpa mengetahui kunci diberikan pelakunya kriptanalisis kriptologi cryptology studi mengenai kriptografi kriptanalisis kriptografi sudah lama digunakan oleh tentara sparta yunani pada permulaan tahun sm mereka menggunakan alat namanya scytale
--	---

4. Stemming

Stemming merupakan proses menghilangkan imbuhan-imbuhan baik itu berupa prefix, sufiks maupun konfiks yang ada pada setiap kata. Hasil Stemming dapat dilihat pada tabel 5.

Tabel 5. Hasil Stemming

Percobaan 1. Tugas Mahasiswa dengan tingkat kemiripan 100 %	
kriptografi ilmu kena teknik enkripsi mana naskah asli plaintext acak guna suatu kunci enkripsi jadi naskah acak sulit baca ciphertext oleh orang tidak milik kunci dekripsi dekripsi guna kunci dekripsi bisa dapat kembali data asli probabilitas	kriptografi ilmu kena teknik enkripsi mana naskah asli plaintext acak guna suatu kunci enkripsi jadi naskah acak sulit baca ciphertext oleh orang tidak milik kunci dekripsi dekripsi guna kunci dekripsi bisa dapat kembali

dapat kembali naskah asli oleh orang tidak punya kunci dekripsi dalam waktu tidak lalu lama sangat kecil	data asli probabilitas dapat kembali naskah asli oleh orang tidak punya kunci dekripsi dalam waktu tidak lalu lama sangat kecil
Percobaan 2. Tugas Mahasiswa dengan tingkat kemiripan 50 %	
kriptografi ilmu kena teknik enkripsi mana naskah asli plaintext acak guna suatu kunci enkripsi jadi naskah acak sulit baca oleh orang tidak milik kunci oleh orang tidak milik kunci dekripsi dekripsi guna kunci dekripsi bisa dapat kembali data asli probabilitas dapat kembali naskah asli oleh orang tidak punya kunci dekripsi dalam waktu tidak lalu lama sangat kecil	kriptografi ilmu kena teknik enkripsi mana naskah asli acak guna suatu kunci enkripsi jadi naskah acak sulit baca oleh orang tidak milik kunci dekripsi kriptanalisis ilmu seni pecah chiperteks jadi plaintext tanpa tahu kunci beri pelaku kriptanalisis kriptologi cryptology studi kena kriptografi kriptanalisis kriptografi sudah lama guna oleh tentara sparta yunani pada mula tahun sm
Percobaan 3. Tugas Mahasiswa dengan tingkat kemiripan 30 %	
kriptografi ilmu kena teknik enkripsi mana naskah asli plaintext acak guna suatu kunci enkripsi jadi naskah acak sulit baca ciphertext oleh orang tidak	kriptografi ilmu kena teknik enkripsi mana naskah asli acak guna suatu kunci enkripsi jadi naskah acak sulit baca kriptanalisis ilmu seni pecah chiperteks jadi

milik kunci	plainteks tanpa
dekripsi dekripsi	tahu kunci beri
guna kunci	pelaku kriptanalisis
dekripsi bisa dapat	kriptologi
kembali data asli	cryptology studi
probabilitas dapat	kena kriptografi
kembali naskah	kriptanalisis
asli oleh orang	kriptografi sudah
tidak punya kunci	lama guna oleh
dekripsi dalam	tentara sparta
waktu tidak lalu	yunani pada
lama sangat kecil	mula tahun sm
	mereka guna alat
	nama scytale

Tabel.6 Percobaan 1. Tingkat kemiripan 100 % dengan winnowing

Uji oba	n-gram (n)	window (w)	kemiripan (%)
1	2	3	100
2	3	3	100
3	4	3	100
4	5	3	100
5	6	3	100
6	7	5	100
7	8	5	100
8	9	5	100
9	10	5	100

Tabel.7 Percobaan 2. Tingkat kemiripan 50 % dengan winnowing

uji coba	n-gram (n)	window (w)	kemiripan (%)
1	2	3	61,6
2	3	3	40,7
3	4	3	31,2
4	5	3	26,9
5	6	3	25,4
6	7	5	23,1
7	8	5	22,2
8	9	5	21,1
9	10	5	20,7

Tabel.8 Percobaan 3. Tingkat kemiripan 30 % dengan winnowing

uji coba	n-gram (n)	window (w)	kemiripan (%)
1	2	3	57,3
2	3	3	31,5
3	4	3	24,1
4	5	3	19,5
5	6	3	18,1
6	7	5	16,9
7	8	5	16,7
8	9	5	16,6
9	10	5	16,2

5. KESIMPULAN

Berdasarkan hasil percobaan yang telah dilakukan pada algoritma winnowing untuk mendeteksi plagiarisme tugas mahasiswa, maka dapat ditarik kesimpulan sebagai berikut;

1. Algoritma winnowing mendapatkan hasil terbaik pada uji coba dengan nilai n-gram (n) = 3 dan window (w) = 3.
2. Algoritma winnowing memiliki persentase akurasi yang baik.

DAFTAR PUSTAKA

- [1] Wali, M., & Safrizal, S. (2018). Similar text sebagai Pengkodean Aplikasi Plagiarisme. *Jurnal JTik (Jurnal Teknologi Informasi dan Komunikasi)*, 2(1), 11-19.
- [2] Astutik, Sariyanti, et al. "Sistem Penilaian Esai Otomatis Pada E-learning Dengan Algoritma Winnowing." *Jurnal Informatika University Petra Kristian*, vol. 12, no. 2, Nov. 2014, pp. 47-52

- [3] Parewe, Andi Maulidinnawati Ak; Aman, Andryanto; Dewang, Dody Putra M. Perbandingan Algoritma Winoing Dan Algoritma Manber Dalam Mendeteksi Berita Hoax Di Media Sosial. PROSIDING SEMANTIK, 2021, 3.1: 41-46.
- [4] Andhika, Fatardhi R. 2010. Penerapan String Suggestion dengan Algoritma Levenshtein Distance dan Alternatif Algoritma Lain dalam Aplikasi, Sekolah Tinggi Teknik Informatika, Institut Teknologi Bandung (ITB). Bandung.
- [5] Setiawan, A. (2017). Implementasi Algoritma Winoing Untuk Deteksi Kemiripan Judul Skripsi Studi Kasus Stmik Budidarma. Majalah Ilmiah INTI (Informasi dan Teknologi Ilmiah), 12(1).