

BAB II TINJAUAN PUSTAKA

2.1 String Metric

String metric atau *similarity metric* adalah kelas matriks berbasis tekstual yang dapat menghasilkan nilai kesamaan atau ketidaksamaan dari dua teks string untuk proses perbandingan dan penyamaan[1]. Beberapa algoritma yang berdasarkan kepada string metric diantaranya adalah Levenshtein distance, TF/IDF, Needleman-Wunsch distance, Jaro-Winkler distance, dan sebagainya. Dari algoritma yang telah disebutkan diatas Jaro-Winkler distance memiliki ketepatan yang baik didalam pencocokan string yang relatif pendek.

2.2 Algoritma Jaro-Winkler distance

Jaro-Winkler *distance* merupakan varian dari Jaro *distance metric* yaitu sebuah algoritma untuk mengukur kesamaan antara dua string, biasanya algoritma ini digunakan di dalam pendeteksian duplikat[2]. Semakin tinggi Jaro-Winkler *distance* untuk dua string, semakin mirip dengan string tersebut. Jaro-Winkler *distance* terbaik dan cocok untuk digunakan dalam perbandingan string singkat seperti nama orang. Skor normalnya seperti 0 menandakan tidak ada kesamaan, dan 1 adalah sama persis.

Algoritma jaro-winkler *distance* memiliki kompleksitas waktu *quadratic runtime complexity* yang sangat efektif pada *string* pendek dan dapat bekerja lebih cepat dari algoritma edit distance.

Algoritma Jaro-Winkler *distance* memiliki 3 bagian yaitu :

1. Menghitung panjang *string*
2. Menentukan jumlah karakter yang sama didalam dua *string*.
3. Menentukan jumlah *transposisi*.

Pada algoritma Jaro-Winkler distance digunakan rumus untuk menghitung jarak (d_j) antara dua *string* yaitu :

$$d_j = \frac{1}{3} \times \left(\frac{m}{|S_1|} + \frac{m}{|S_2|} + \frac{m-t}{m} \right)$$

Keterangan :

m = Jumlah karakter yang sama persis

$|S_1|$ = Panjang string 1

$|S_2|$ = Panjang string 2

t = Jumlah transposisi

Jarak teoritis dua buah karakter yang disamakan dapat dibenarkan jika tidak melebihi -1 :

$$\left(\frac{\max(|S_1|, |S_2|)}{S} \right) < -1$$

Akan tetapi bila mengacu kepada nilai yang akan dihasilkan oleh algoritma Jaro-Winkler maka nilai jarak maksimalnya adalah 1 yang menandakan kesamaan *string* yang dibandingkan mencapai 100% atau sama persis. Biasanya S_1 digunakan sebagai acuan untuk urutan di dalam mencari transposisi. Yang di maksud transposisi disini adalah karakter yang sama dari string yang dibandingkan akan tetapi tertukar urutannya. sebagai contoh, dalam membandingkan kata CRATE dengan TRACE, bisa dilihat seksama maka dapat dikatakan semua karakter yang ada di s_1 ada dan sama dengan karakter s_2 , tetapi dengan urutan yang berbeda. Dengan mengganti C dan T dapat dilihat perubahan data CRATE menjadi TRACE. Pertukaran dua elemen string ini adalah contoh nyata dari transportasi yang dijelaskan. Dalam pencocokan DWAYNE dan DUANE memiliki urutan yang sama D-A-N-E, jadi tidak ada transposisi.

Jaro-Winkler *Distance* menggunakan *prefix scale* (p) yang memberikan tingkat penilaian yang lebih, dan *prefix length* (l) yang menyatakan panjang awalan yaitu panjang karakter yang sama dari *string* yang dibandingkan sampai ditemukan ketidaksamaan. Bila *string* S_1 dan S_2 yang diperbandingkan, maka Jaro-Winkler *distance* (d_w) adalah :

$$d_w = d_j + (lp(1 - d_j))$$

keterangan :

d_j = Jaro distance untuk string s_1 dan s_2

l = Panjang prefix umum di awal string nilai maksimalnya 4 karakter

(panjang karakter yang sama sebelum ditemukan ketidaksamaan max 4)

p = konstanta scaling factor. Nilai standar untuk konstanta ini menurut Winkler adalah $p = 0,1$.

2.3 Algoritma Levenshtein

Levenshtein Distance atau Edit Distance merupakan perhitungan edit distance didapatkan dari matriks yang digunakan untuk menghitung jumlah perbedaan string antara dua string [15]. Pada algoritma Levenshtein Distance semakin kecil nilai skor yang dimiliki maka kemiripan yang dimiliki semakin tinggi. Perhitungan jarak antara dua string ini ditentukan dari jumlah minimum operasi perubahan untuk membuat string A menjadi string B.

Ada 3 macam operasi utama yang dapat dilakukan oleh algoritma ini :

1. Perubahan karakter
2. Penambahan karakter
3. Penghapusan karakter

Algoritma ini berjalan mulai dari pojok kiri atas sebuah array dua dimensi yang telah diisi sejumlah karakter string awal dan string target dan diberikan nilai cost. Nilai cost pada ujung kanan bawah menjadi nilai edit distance yang menggambarkan jumlah perbedaan dua string. Isian nilai pada matriks tersebut adalah jumlah operasi penghapusan, penyisipan dan penukaran yang dibutuhkan dalam mengubah string sumber ke string target. Rumus operasi penghapusan, penyisipan, dan penukaran karakter yang digunakan untuk mengisi nilai matriks adalah sebagai berikut:

$$(s,t) = \min(s - 1, t) + 1 \text{ (Penghapusan)}$$

$$(s,t) = \min(s, - 1) + 1 \text{ (Penyisipan)}$$

$$(s,t) = \min(s - 1, - 1) + 1, s_j \neq t_i \text{ (Penukaran)}$$

$$(s,t) = \min(s - 1, t - 1), s_j = t_i \text{ (Tidak ada perubahan)}$$

Keterangan :

- s = String Sumber
- (j) = Karakter String Sumber ke- j
- t = String Target
- (i) = Karakter String Target ke- i
- D = Jarak Levenshtein Distance

2.4 Prinsip Mengenal Nasabah (Know Your Customer Principles)

Peraturan Bank Indonesia yang mengatur tentang Penerapan Prinsip Mengenal Nasabah atau *Know Your Customer Principles*, dituangkan dalam Peraturan Bank Indonesia Nomor: 3/10/PBI/2001 pasal 1 ayat 2 yang berisi “Prinsip Mengenal Nasabah adalah prinsip yang diterapkan Bank untuk mengetahui identitas nasabah, memantau kegiatan transaksi nasabah termasuk pelaporan transaksi yang mencurigakan. Nasabah dalam hal ini adalah pihak yang menggunakan jasa Bank”[11].

Bank wajib menerapkan Prinsip Mengenal Nasabah . Dalam menerapkan Prinsip Mengenal Nasabah Bank wajib :

- a. Menetapkan kebijakan penerimaan Nasabah;
- b. Menetapkan kebijakan dan prosedur dalam mengidentifikasi Nasabah;
- c. Menetapkan kebijakan dan prosedur pemantauan terhadap rekening dan transaksi Nasabah;
- d. Menetapkan kebijakan dan prosedur manajemen risiko yang berkaitan dengan penerapan Prinsip Mengenal Nasabah.

Pada bab 2 PBI Nomor:3/10/PBI/2001 yang berisi tentang Kebijakan Penerimaan dan Identifikasi Nasabah. Sebelum melakukan hubungan usaha dengan Nasabah, Bank wajib meminta informasi mengenai:

- a. Identitas calon Nasabah;
- b. Maksud dan tujuan hubungan usaha yang akan dilakukan calon Nasabah dengan Bank;
- c. Informasi lain yang memungkinkan Bank untuk dapat mengetahui profil calon Nasabah; dan

- d. Identitas pihak lain, dalam hal calon Nasabah bertindak untuk dan atas nama pihak lain sebagaimana diatur dalam Pasal 6.

Dokumen pendukung sebagaimana dimaksud dalam pasal 4 ayat (2) bagi :

- 1) Identitas Nasabah yang memuat :
 - a. Nama;
 - b. Alamat tinggal tetap;
 - c. Tempat dan tanggal lahir;
 - d. Kewarganegaraan;
- 2) Keterangan mengenai pekerjaan;
- 3) Spesimen tanda tangan;
- 4) Keterangan mengenai sumber dana dan tujuan penggunaan dana.

2.5 Bank

Pada Undang-undang Nomor 10 Tahun 1998 tentang Perbankan, Bank disebutkan sebagai badan usaha yang menghimpun dana dari masyarakat dalam bentuk simpanan dan menyalurkannya kepada masyarakat dalam bentuk kredit dan atau bentuk-bentuk lainnya dalam rangka meningkatkan taraf hidup masyarakat[12].

Bank umum adalah bank yang melaksanakan kegiatan usaha secara konvensional dan atau berdasarkan prinsip syariah, yang dalam kegiatannya memberikan jasa dalam lalu lintas pembayaran.

Kegiatan Usaha Bank Umum

Kegiatan usaha yang dapat dilaksanakan oleh Bank Umum:

- Menghimpun dana dari masyarakat dalam bentuk simpanan berupa giro, deposito berjangka, sertifikat deposito, tabungan, dan atau bentuk lainnya yang dipersamakan dengan itu.
- Memberikan kredit.
- Menerbitkan surat pengakuan utang.
- Membeli, menjual, atau menjamin atas risiko sendiri maupun untuk kepentingan dan atas perintah nasabahnya: Surat-surat wesel termasuk wesel

yang diakseptasi oleh bank yang masa berlakunya tidak lebih lama daripada kebiasaan dalam perdagangan surat-surat dimaksud.

- Surat pengakuan utang dan kertas dagang lainnya yang masa berlakunya tidak lebih lama dari kebiasaan dalam perdagangan surat-surat dimaksud.
 - Kertas perbendaharaan negara dan surat jaminan pemerintah.
 - Sertifikat Bank Indonesia (SBI).
 - Obligasi.
 - Surat dagang berjangka waktu sampai dengan satu (1) tahun.
 - Instrumen surat berharga lain yang berjangka waktu sampai dengan satu (1) tahun
- Memindahkan uang baik untuk kepentingan sendiri maupun untuk kepentingan nasabah.
 - Menempatkan dana pada, meminjam dana dari, atau meminjamkan dana kepada bank lain, baik dengan menggunakan surat, sarana telekomunikasi maupun dengan wesel unjuk, cek atau sarana lainnya.
 - Menerima pembayaran dari tagihan atas surat berharga dan melakukan perhitungan dengan antar pihak ketiga.
 - Menyediakan tempat untuk menyimpan barang dan surat berharga.
 - Melakukan kegiatan penitipan untuk kepentingan pihak lain berdasarkan suatu kontrak.
 - Melakukan penempatan dana dari nasabah kepada nasabah lainnya dalam bentuk surat berharga yang tidak tercatat di bursa efek.
 - Melakukan kegiatan anjak piutang, usaha kartu kredit dan kegiatan wali amanat.
 - Menyediakan pembiayaan dan atau melakukan kegiatan lain berdasarkan Prinsip Syariah, sesuai dengan ketentuan yang ditetapkan oleh Bank Indonesia.

- Melakukan kegiatan lain yang lazim dilakukan oleh bank sepanjang tidak bertentangan dengan undang-undang ini dan peraturan perundang-undangan yang berlaku.

Selain itu Bank Umum dapat pula:

- Melakukan kegiatan dalam valuta asing dengan memenuhi ketentuan yang ditetapkan oleh Bank Indonesia.
- Melakukan kegiatan penyertaan modal pada bank atau perusahaan di bidang keuangan, seperti sewa guna usaha, modal ventura, perusahaan efek, asuransi serta lembaga kliring penyelesaian dan penyimpanan, dengan memenuhi ketentuan yang ditetapkan oleh Bank Indonesia.
- Melakukan kegiatan penyertaan modal sementara untuk mengatasi akibat kegagalan kredit atau kegagalan pembiayaan berdasarkan Prinsip Syariah, dengan syarat harus menarik kembali penyertaannya, dengan memenuhi ketentuan yang ditetapkan oleh Bank Indonesia, dan
- Bertindak sebagai pendiri dana pensiun dan pengurus pensiun sesuai dengan ketentuan dalam peraturan perundang-undangan dana pensiun yang berlaku.

2.6 Nasabah

Menurut undang-undang No. 10 Tahun 1998 tentang pokok perbankan pasal 1, mendefinisikan nasabah sebagai berikut: Nasabah adalah pihak yang menggunakan jasa bank. Nasabah penyimpan adalah nasabah yang menempatkan dananya di bank dalam bentuk simpanan berdasarkan perjanjian bank dengan nasabah yang bersangkutan“. Menurut Kasmir (2008:94) bahwa: “Nasabah merupakan konsumen yang membeli atau menggunakan produk yang dijual atau ditawarkan oleh bank”.

2.7 Penelitian Terkait

Tabel 2. 1 Tabel Literature Review

No	Sumber	Masalah dan Tujuan	Metode	Hasil
1	Leonardo B,Hansun S., <i>Text Documents Plagiarism Detection using Rabin-Karp and Jaro-Winkler Distance Algorithms.</i> Indonesia Journal of Electrical Engineering and Computer Science. 2017: 5	Perkembangan teknologi memudahkan mencari referensi, namun ada hal negatifnya yaitu plagiarisme. Dengan adanya aplikasi ini diharapkan bisa memudahkan dalam mendeteksi adanya plagiarisme dalam sebuah dokumen.	Menggunakan algoritma Rabin-Karp dan Jaro-Winkler Distance.	Kedua algoritma dapat melakukan deteksi plagiarisme dokumen, tetapi kinerja algoritma Rabin-Karp jauh lebih efektif daripada algoritma Jaro-Winkler Distance. Dalam hal kecepatan pemrosesan, algoritma Rabin-Karp lebih cepat daripada algoritma Jaro-Winkler Distance dengan selang waktu untuk semua skenario sekitar 0,389 menit. Menggunakan

				dokumen yang sama, algoritma Rabin-karp lebih cepata dibandingkan algoritma Jaro-Winkler distance.
2	Kurniawati A, Puspitodjati S, Rahman S., <i>Implementasi Algoritma Jaro-Winkler Distance untuk Membandingkan Kesamaan Dokumen Berbahasa Indonesia</i> . Repository Gunadarma.2014	Masalah: Dalam pembuatan karya tulis sering ditemukan kesamaan dengan karya tulis orang lain sehingga menimbulkan isu plagiarisme. Tujuan : Diharapkan dengan adanya sebuah aplikasi yang mampu mendeteksi kesamaan dokumen maka pada tulisan ilmiah tidak ditemukan lagi	Menggunakan Algoritma Jaro Winkler	Aplikasi berjalan baik untuk dokumen yang memiliki kemiripan dan urutan kata yang sama.

		kesamaan yang begitu tinggi dengan dokumen lainnya.		
3	Triani Srikandi Ratih. <i>Implementation Of Knowledge Management System at Nursing Section of Pelabuhan Hospital Cirebon</i> . Jurnal Ilmiah Komputer dan Informatika(KOMPUTA).2016: 1	Manajer keperawatan mengalami kesulitan dalam mengetahui apakah tindakan yang diambil oleh perawat sudah sesuai atau tidak sesuai dengan SOP, asisten manajer askep & kualitas mengalami kesulitan dalam penyebaran dan pembaruan SOP. Tujuan : Membuat aplikasi Knowledge	Menggunakan algoritma Nazief dan Adriani dan Algoritma Jaro-Winkler Distance.	Sistem manajemen pengetahuan telah dapat membantu manajer keperawatan dalam memantau tindakan perawat, membantu asisten manajer keperawatan askep dan kualitas dalam membuat dan memperbarui SOP, dan dapat membantu penyedia untuk berbagi pengalaman tindakan antara perawat.

		Management berdasarkan dokumen dari perawat yang sudah pengalaman, dokumen dibandingkan dengan menggunakan algoritma Jaro-Winkler Distance, hasil dari pengujian tersebut dijadikan SOP.		
4	Rochmawati Y, Kusumaningrum R. <i>Studi Perbandingan Algoritma Pencarian String dalam Metode Approximate String Matching untuk Identifikasi Kesalahan Pengetikan Teks.</i> Jurnal Buana Informatika.2016: 7,Nomor 2	Masalah : Kesalahan pengetikan mengakibatkan kata baku berubah menjadi kata tidak baku karena ejaan yang digunakan tidak sesuai. Tujuan : Untuk	Metode approximate string matching dengan menggunakan berbagai jenis algoritma pencarian string yaitu Lavenshtein Distance,	Hasil penelitian terhadap 50 kata salah menunjukkan bahwa algoritma Jaro Winkler Distance terbaik dalam melakukan pengecekan kata dengan nilai MAP sebesar 0,87.

		mengidentifikasi kesalahan pengetikan.	Hamming Distance, Damerau Levenshtein Distance dan Jaro Winkler Distance.	
5	Okta'mal F, Saptono R, Eko SM. <i>Jaro Winkler Distance dan Stemming untuk Deteksi Dini Hama dan Penyakit Padi</i> . Seminar Nasional Sistem Informasi Indonesia (SESINDO). 2015	Masalah: Terjadinya kesalahan deteksi hama dan penyakit pada tanaman padi. Tujuan : Membantu petani dalam mendeteksi hama padi .	Menggunakan algoritma Jaro Winkler, metode stemming dengan algoritma Nazief dan Andriani.	Pengujian identifikasi <i>input</i> percobaan pertama menghasilkan akurasi sebesar 100%. Dan identifikasi <i>input</i> percobaan kedua menghasilkan akurasi sebesar 98% dengan satu kesalahan identifikasi. Hasil identifikasi <i>output</i> percobaan pertama adalah 100 % dan percobaan kedua hanya 96% dikarenakan kesalahan hasil

				identifikasi <i>input</i> sebelumnya.
6	Kornain A, Yansen F & Tinaliah. <i>Penerapan Algoritma Jaro-Winkler Distance untuk Sistem Pendeteksi Plagiarisme pada Dokumen Teks Berbahasa Indonesia</i> . STMIK MDP. 2014	Masalah : Kesamaan dalam penulisan karya tulis. Tujuan : membandingkan kesamaan antar dokumen teks berbahasa Indonesia, sehingga dapat ditentukan sebuah dokumen tersebut plagiat atau tidak	Metode yang digunakan yaitu penerapan algoritma Jaro-Winkler	Pendeteksian plagiarisme terhadap dokumen korpus melalui tahapan stemming dan query melalui tahapan tanpa stemming memiliki nilai deteksi yang lebih baik sebesar 30.58% berdasarkan hasil pengujian pertama.
7	Rinusantoro S. <i>Aplikasi Deteksi Kemiripan Dokumen Teks</i> . Tesis Universitas Gajah Mada. Yogyakarta. 2014	Masalah : Mudah nya orang untuk melakukan plagiat. Tujuan : Meminimalisir tindak plagiat dan mengetahui	Menggunakan algoritma Jaro-Winkler Distance	Sistem mampu mendeteksi kemiripan dokumen teks dari artikel publikasi dan abstrak sivitas akademika.

		tingkat kesamaan dokumen dengan tema yang sama.		
8	Yulianingsih. <i>Implementasi Algoritma jaro-Winkler dan Levenstein Distance dalam Pencarian Data pada Database</i> . Jurnal String(Satuan tulisan riset dan inovasi teknologi). 2017 : 2	Masalah : data memiliki duplikasi. Tujuan : menentukan satu dari dua metode (algoritma jaro-winkler dan algoritma levenstein) yang memiliki tingkat validasi tertinggi dalam pencarian data yang memiliki kesamaan tinggi.	Menggunakan algoritma Jaro-Winkler dan algoritma Levenshtein distance.	1.Algoritma Jaro-Winkler memiliki kompleksitas waktu quadratic runtime complexity yang sangat efektif pada string pendek dan dapat memperoleh hasil lebih cepat dibandingkan dengan algoritma Levenshtein Distance. 2.Metode Jaro-Winkler mudah untuk diimplementasikan dan efektif dalam hasil yang dicapai.
9	Nufus K. <i>Implementasi Algoritma Fisher-Yates</i>	Masalah : Tidak sedikit	Menggunakan algoritma	Penerapan Algoritma Fisher-

<p><i>Shuffle dan Algoritma Jaro-Winkler Distance pada Aplikasi Uji Hafalan Tahfidz Quran berdasarkan Kemiripan Teks Menggunakan Metode Speech Recognition Berbasis Android.</i> Universitas Islam Negeri Sunan Gunung Djati.Bandung. 2017</p>	<p>mahasiswa Univesitas Islam Negeri Sunan Gunung Djati Bandung terhambat mengikuti ujian komprehensif dikarenakan belum lulus tahfidz. Tujuan: Diharapkan menjadi sarana pembelajaran untuk menghafal al-quran khususnya juz</p>	<p>Fisher-Yates Shuffle untuk pengacakan soal, algoritma jaro-winkler distance untuk pencocokan teks dan metode speech recognition untuk menjawab soal.</p>	<p>Yates Shuffle berhasil diterapkan dalam aplikasi tahfidz untuk menentukan solusi pengacakan objek yang variatif berdasarkan hasil pengujian sebanyak sepuluh kali percobaan. Metode speech recognition menggunakan Google speech API dapat diterapkan dengan baik untuk mengubah suara ke teks atau speech to text. Algoritma Jaro-Winkler Distance sebagai pencocokan teks dapat digunakan untuk</p>
--	---	---	--

				mencocokkan teks Al-quran dengan baik, dengan waktu proses pencocokan rata-rata 1.9 ms.
10	Prasetyo A, Maulana WB, Syam IH. <i>Algoritma Jaro-Winkler Distance : Fitur Autocorrect dan Spelling Suggestion pada Penulisan Naskah Bahasa Indonesia di BMS TV.</i> Jurnal Teknologi Informasi dan Ilmu Komputer(JTIK). 2018:5 No.4	Masalah :Tidak sedikit jumlah naskah berita yang mengalami kesalahan ejaan kata saat proses penulisan berita. Tujuan : Diharapkan dapat memeriksa dan memperbaiki kesalahan penulisan kata secara otomatis serta memberi saran penulisan ejaan kata yang benar dalam bahasa Indonesia	Menggunakan algoritma Jaro-Winkler Distance untuk pengembangan perangkat lunak menggunakan dalam menghitung nilai jarak kedekatan antara dua teks.	pengujian 60 kata yang terdiri dari berbagai skenario kesalahan penulisan kata fitur ini memperbaiki sepuluh kata secara otomatis dengan benar dan memunculkan saran ejaan kata pada 39 kata dengan tepat.

1 1	Joko Jatminto, I Kadek Dwi Nuryana. <i>Implementasi Spelling Checker dengan Algoritma Levensthein pada Ensiklopedia (Information Technology) Berbasis Website. Aplikasi Ensiklopedia Tentang Teknologi Informasi. 2016.</i>	Masalah : Kesalahan pengetikan dan kontrol pegecekan bahasa yang baik dan benar jika dilakukan secara manual pastinya akan menghabiskan banyak waktu. Tujuan : Diharapkan dapat mengatasi kesalahan penulisan pada umumnya disebabkan kedekatan letak keyboard, slip jari, ataupun karena dua karakter yang letaknya tertukar.	Membuat sistem pengecekan ejaan atau spelling checker di ensiklopedia tentang teknologi informasi yang dikembangkan dengan berbasis website menggunakan algoritma Levensthein.	Dapat membantu mengatasi permasalahan pada kesalahan pengetikan dengan mekanisme penambahan, penyisipan dan penghapusan karakter.
1 2	Chorin Maulina. <i>Pemeriksa Ejaan untuk</i>	Masalah : Kesalahan	Melakukan pengujian	Algoritma Levenshtein

	<p><i>Dokumen Berbahasa Indonesia dengan Algoritma Levensthein Distance.</i> Universitas Muhammadiyah Jember . 2015.</p>	<p>pengetikan dokumen memang sering sekali terjadi. Apalagi belakangan ini kesadaran masyarakat untuk menuangkan idenya ke dalam artikel, jurnal ilmiah, tugas kuliah ataupun dokumen lainnya mengalami peningkatan.</p> <p>Tujuan :</p> <p>Dibuat sebuah program yang dapat mendeteksi kesalahan penulisan Bahasa.</p>	<p>pada beberapa algoritma dapat diimplementasikan dalam memberikan kata saran yang paling mendekati dari kata yang salah penetikannya .</p>	<p>Distance yang dapat menghitung jarak keterbedaan antara 2 (dua) String. Dalam tugas akhir ini penulis akan mengimplementasikan algoritma Levenshtein Distance yang dapat membantu memeriksa kesalahan penulisan pada sebuah dokumen berbahasa Indonesia.</p>
--	--	---	--	---

1 3	<p>Nurul Fadhillah, Huzain Azis, Dirgahayu Lantara. <i>Validasi Pencarian Kata Kunci Menggunakan Algoritma Levenshtein Distance Berdasarkan Metode Approximate String Matching</i>. Prosiding Seminar Nasional Ilmu Komputer dan Teknologi Informasi. Vol. 3, No. 2, Desember 2018.</p>	<p>Masalah : Kesalahan dalam menuliskan kata kunci sehingga menghasilkan keluaran yang tidak sesuai dengan keinginan pengguna. Tujuan : Untuk mengatasi kesalahan dalam pencarian kata kunci perlu dilakukan optimasi proses pencarian pada aplikasi Kamus Besar Bahasa Indonesia (KBBI) digital.</p>	<p>Penelitian ini menggunakan metode Approximate String Matching pada algoritma Levenshtein Distance.</p>	<p>Metode ini akan diketahui jarak Levenshtein yang menjadi nilai kemiripan suatu objek bertipe string. Untuk mendapatkan nilai kemiripan dilakukan dengan menghitung jarak antar dua string dengan menghitung jumlah operasi yang terjadi seperti penambahan, penghapusan atau pengurangan karakter.</p>
1	B. P. Pratama, S. A.	Masalah :	Analisa	Hasil simulasi

4	<p>Pamungkas. <i>Analisa Kinerja Algoritma Levenshtein Distance Dalam Mendeteksi Kemiripan Dokumen Teks</i>. Jurnal “LOG!K@” , Jilid 6, No. 2, 2016, Hal. 131 – 143.</p>	<p>Plagiarisme pada dokumen teks merupakan salah satu permasalahan akademik yang semakin meningkat.</p> <p>Tujuan : dibangun sebuah sistem yang mampu mendeteksi tingkat kemiripan antar dokumen teks menggunakan algoritma Levenshtein distance dengan menambahkan proses case folding, tokenizing, stopword removal, stemming, dan</p>	<p>penggunaan stopword removal, stemming, dan sorting dilakukan untuk melihat pengaruhnya terhadap kinerja algoritma Levenshtein distance.</p>	<p>menunjukkan bahwa penggunaan sorting sangat berpengaruh bagi algoritma Levenshtein distance. Hasil terbaik pada data set 1 ditunjukkan pada proses yang menggunakan stopword removal, stemming, dan sorting sekaligus. Hasil terbaik pada data set 2 diperlihatkan pada proses yang B. P. Pratama dan S. A. Pamungkas 132 menggunakan stopword dan stemming yang digabungkan dengan sorting. Hasil terbaik</p>
---	--	--	--	---

		sorting.		pada data real diperlihatkan pada proses stemming-sorting.
1 5	Andry Hery Purba, Zakarias Situmorang. <i>Analisis Perbandingan Algoritma Rabin-Karp dan Levenshtein Distance dalam Menghitung Kemiripan Teks.</i> Jurnal Teknik Informatika Unika St. Thomas (JTIUST), Volume 02 Nomor 02, Desember 2017, ISSN : 2548-1916	Masalah : Terjadinya plagiarisme disebabkan karena kebiasaan peneliti yang malas dan ingin serba cepat dalam menyelesaikan tugas-tugasnya, sehingga kemungkinan peneliti akan mencari referensi yang terkait baik di internet maupun perpustakaan. Tujuan : Untuk	Proses yang harus dilakukan dalam sistem adalah proses preprocessing pada masing-masing algoritma dan kemudian menghitung tingkat kemiripan teks menggunakan algoritma Rabin-Karp dan Levenshtein Distance.	Hasil dari kedua algoritma yang ada pada sistem untuk dianalisis perbandingan hasilnya berupa grafik. Pada pengujian peneliti menggunakan dokumen teks dengan tipe .pdf dan berbahasa Indonesia.

		mendeteksi tingkat kemiripan teks berupa algoritma pencocokan string atau katadengan menggunakan algoritma Rabin-Karp dan Levenshtein Distance.		
--	--	---	--	--



UNIVERSITAS
MERCU BUANA