

Ensiklopedia Wanita Berbasis Mobile Apps dengan Penerapan Algoritma Jaro Winkler Distance

Ayu Aksari^{a,1,*}, Poetri Lestari Lokapitasari Belluano^{a,2}, Lilis Nur Hayati^{a,3}

^a *Uniyeritas Muslim Indonesia, Jl. Urip Sumoharjo KM.05, Makassar dan 90231, Indonesia*
¹ *ayuaksari178@gmail.com*; ² *poetrilestari@umi.ac.id*; ³ *lilis.nurhayati@umi.ac.id*
**corresponding author*

INFORMASI ARTIKEL	ABSTRAK
Diterima : 05 – 11 – 2020 Direvisi : 20 – 11 – 2020 Diterbitkan : 30 – 11 – 2020	Ensiklopedia Wanita Berbasis <i>Mobile Apps</i> merupakan aplikasi mobile yang menyajikan informasi mengenai wanita islami yang digolongkan kedalam 3 kategori yaitu sosial, psikologi, dan kesehatan. Aplikasi ini dibuat bukan hanya sebagai sumber informasi melainkan juga sebagai wadah berbagi informasi mengenai wanita islami yang secara otomatis akan melalui seleksi data menggunakan algoritma <i>Jaro winkler distance</i> dengan bantuan <i>text preprocessing</i> (<i>case folding</i> dan <i>stopword elimination</i>). Adapun pengelompokan kesamaan ada 2 yaitu kesamaan ringan (lebih kecil 1), dan kesamaan sama persis atau total (sama dengan 1). Berdasarkan hasil pengujian yang telah dilakukan dengan menggunakan teknik <i>blackbox</i> (<i>function analysis</i>) pada pengujian beta menghasilkan presentase tertinggi dari kuisisioner yaitu sebanyak 78% menyatakan setuju dengan adanya aplikasi Ensiklopedia Wanita berbasis <i>mobile app</i> dengan penerapan algoritma <i>Jaro winkler distance</i> . Oleh karena itu dapat ditarik kesimpulan bahwa aplikasi ini memang baik untuk diterapkan dan sangat mempermudah masyarakat dalam memperoleh informasi secara cepat dan menarik. Serta penggunaan seleksi data sangat membantu pada proses <i>insert</i> konten baru.
Kata Kunci: Ensiklopedia Wanita Algoritma <i>Jaro Winkler Distance</i> <i>Text Preprocessing</i>	

This is an open access article under the [CC-BY-SA](#) license.



I. Pendahuluan

Menurut data Badan Pusat Statistik (BPS) 2015, jumlah penduduk perempuan di Indonesia mencapai 131,9 juta jiwa [1], hal tersebut menunjukkan bahwa kebutuhan informasi tentang wanita secara cepat dan menarik semakin meningkat khususnya informasi wanita dalam agama islam. Berbagai media dapat digunakan sebagai sumber informasi seperti ensiklopedia. Ensiklopedia merupakan media pembelajaran yang berisikan berbagai hal seperti informasi dan pengetahuan suatu cabang ilmu tertentu, yang disusun ke dalam sebuah buku [2][3]. Namun, penggunaan buku teks dianggap kurang efektif karena buku tidak dapat di *update* sehingga mengharuskan pengguna untuk membeli edisi terbaru dengan harga yang tidak murah. Oleh karena itu diperlukan media sumber pengetahuan dan informasi khususnya tentang wanita islam yang menarik dan *up-to-date* sehingga dapat dengan mudah digunakan oleh semua kalangan dan mempermudah masyarakat untuk memperoleh informasi.

Dengan semakin berkembangnya teknologi, berbagai aplikasi yang menyajikan informasi tentang wanita hadir dengan berbagai aspek seperti sosial, psikologi, kecantikan, fashion, Kesehatan dan sebagainya. Untuk mendapatkan semua informasi tersebut, masyarakat perlu meng-*install* banyak aplikasi sehingga menyebabkan penggunaan memori telepon menjadi penuh dan lambat.

Penelitian terkait penggunaan algoritma Jaro-Winkler Distance telah dilakukan untuk melakukan autocorrect dan auto spelling pada penulisan naskah Bahasa Indonesia di BMS TV. Algoritma yang digunakan mampu menangani kesalahan penulisan ejaan kata, dengan pengujian 60 kata yang terdiri dari berbagai skenario kesalahan penulisan kata. Algoritma tersebut mampu memperbaiki sepuluh kata secara otomatis dengan benar dan memunculkan saran ejaan kata pada 39 kata dengan tepat [4]. Algoritma Jaro-Winkler Distance juga digunakan untuk mendeteksi plagiarisme dokumen skripsi dengan cara melakukan perbandingan antara dokumen asli dan dokumen uji sehingga dapat mendeteksi similarity pada dokumen. Algoritma tersebut dapat mendeteksi similarity pada dokumen dengan baik [5]. Penelitian lain mengusulkan pembuatan sistem analisis sentimen pada ulasan “Lazada” Berbahasa Indonesia menggunakan K-Nearest Neighbor (K-NN) Dengan Perbaikan Kata Menggunakan Jaro Winkler Distance. Pengujian berdasarkan nilai precision, recall, dan accuracy pada masing-masing analisis sentimen tanpa perbaikan kata, maupun dengan perbaikan kata.

Hasil pengujian dengan nilai accuracy yang baik terdapat pada analisis sentimen dengan perbaikan kata yakni 76 %, dengan nilai precision 0,76, dengan nilai recall yang optimal yaitu 1 [6].

Berdasarkan permasalahan yang telah dipaparkan, maka penulis merancang sebuah aplikasi ensiklopedia berbasis mobile apps yang akan memuat berbagai macam informasi tentang wanita dalam sisi islam yaitu sosial, psikologi, dan kesehatan sehingga masyarakat hanya perlu meng-*install* satu aplikasi. Pengguna juga dapat mengunggah pengetahuan yang dimiliki ke dalam aplikasi. Dalam penerapannya, penelitian ini menggunakan metode *Jaro Winkler Distance* untuk proses seleksi data pada proses pengunggahan pada aplikasi, jika nilai maksimalnya adalah 1 maka kesamaan string yang dibandingkan mencapai 100% atau sama persis sehingga dapat ditentukan range nilai atau tingkat kesamaan pada penelitian ini ada dua yaitu kesamaan sedang (dengan nilai <1) dan kesamaan total atau 100% sama (dengan nilai sama dengan 1) [7]. Pada proses pengolahan data menggunakan metode *case folding* dan *stopword elimination* untuk menghilangkan kata-kata yang tidak perlu atau yang tidak memiliki arti dari sebuah konten yang di upload [8].

II. Metode

Algoritma *Jaro Winkler Distance* digunakan untuk proses pengunggahan data karena memiliki akurasi *similarity* yang lebih tinggi dan waktu proses yang lebih cepat dibandingkan beberapa algoritma lainnya [9][10]. Pada algoritma Jaro-Winkler Distance digunakan rumus untuk menghitung jarak (d_j) antara dua string yaitu S1 dan S2 adalah:

$$d_j = \frac{1}{3} x \left(\frac{m}{s_1} + \frac{m}{s_2} + \frac{m-t}{m} \right) \quad (1)$$

Dimana :

m = Jumlah karakter yang sama persis
 $|S_1|$ = Panjang string 1
 $|S_2|$ = Panjang string 2
 t = Jumlah transposisi

Kemudian masuk ketahap selanjutnya dengan mengeksekusi rumus terakhir yaitu penghitungan nilai *Jaro winker distance* dengan rumus:

$$d_w = d_j + \left(lp(1 - d_j) \right) \quad (2)$$

Dimana :

d_j = untuk string S1 dan S2
 l = Panjang prefix nilai maksimum 4 karakter
 p = Konstanta *scaling factor* (Nilai standar untuk konstanta ini menurut Winkler adalah $p = 0.1$).

A. Tahapan Jaro distance Penelitian

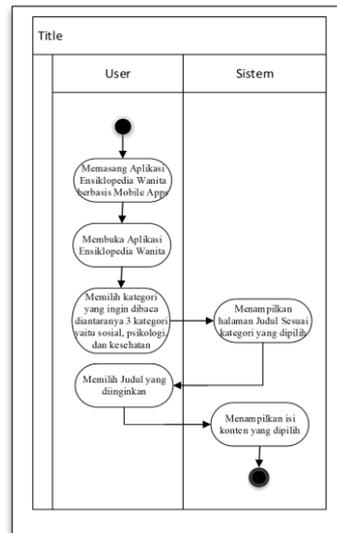
Pengumpulan data pada penelitian ini melalui 2 tahap yaitu wawancara (*interview*), pada tahap ini peneliti melakukan tanya jawab dengan mahasiswa Fakultas Ilmu Komputer UMI sehingga diharapkan dapat memberikan informasi dan solusi terbaik dalam membangun aplikasi ini. Tahap yang kedua dengan melakukan observasi dimana peneliti pada tahap ini melakukan pengamatan secara langsung kepada mahasiswi Fakultas Ilmu Komputer UMI terkait informasi mengenai wanita yang mereka butuhkan serta mengamati sistem yang berjalan sehingga memudahkan penulis untuk membuat aplikasi.

B. Teknik Analisis Data

Analisis data pada penelitian ini adalah:

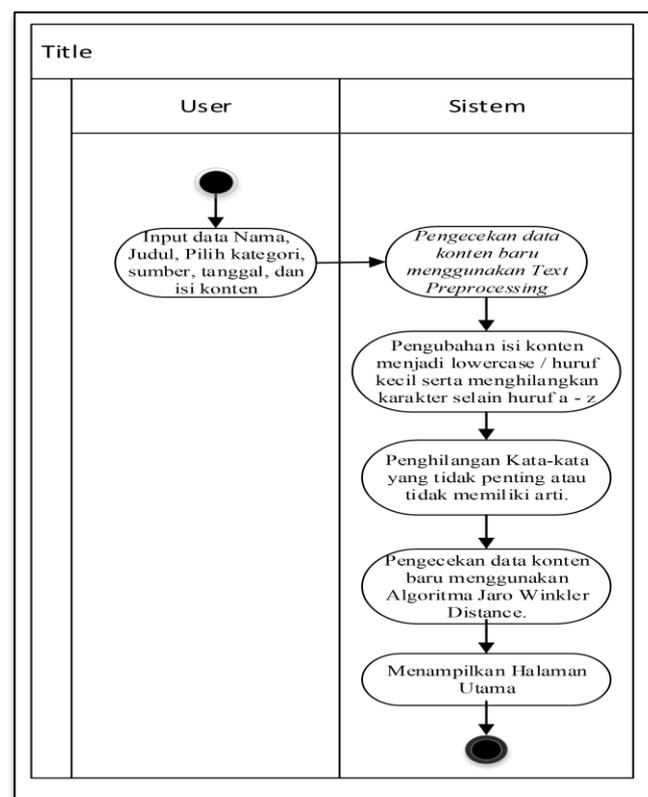
- Mengumpulkan data konten yang ingin diunggah kedalam aplikasi Ensiklopedia Wanita berbasis *mobile apps* berdasarkan kategori yang ada yaitu sosial, psikologi, dan kesehatan.
- Melakukan pengecekan data konten baru menggunakan text preprocessing (*case folding* dan *stopword elimination*) kemudian masuk ke tahap yang paling inti yaitu dengan algoritma jaro winkler distance.

C. Analisis Sistem Usulan



Gambar 1. Analisis Sistem Usulan

Gambar 1 menampilkan proses user memperoleh informasi dengan mudah, cepat dan menarik. Pengguna dapat memilih jenis kategori yang diinginkan yaitu sosial, psikologi, dan Kesehatan. Kemudian sistem akan menampilkan halaman judul sesuai dengan kategori yang telah dipilih. Lalu user memilih judul yang ingin dibaca maka sistem akan menampilkan halaman isi konten sesuai judul yang telah dipilih.

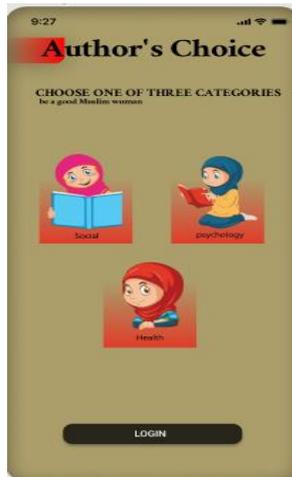


Gambar 2. Activity Diagram Pengolahan Data Konten Baru

Gambar 2 menampilkan alur proses pengolahan data konten baru dimulai dari user menginputkan nama, judul, pilih kategori, sumber, tanggal dan isi. Kemudian sistem akan melakukan pengecekan data konten baru menggunakan *text preprocessing* yang pertama adalah pengubahan isi konten menjadi *lowercase* / huruf kecil serta menghilangkan karakter selain huruf a – z. Kemudian tahap kedua melakukan penghilangan kata-kata yang tidak penting atau tidak memiliki arti selanjutnya pengecekan data menggunakan algoritma *Jaro winkler distance*.

III. Hasil dan Pembahasan

A. Interface Aplikasi Ensiklopedia Wanita Berbasis Mobile Apps



Gambar 1. Halaman Utama Aplikasi

Gambar 2 menunjukkan halaman utama dari aplikasi ensiklopedia Wanita berbasis mobile apps dengan menampilkan tiga kategori yaitu sosial, psikologi dan kesehatan.

B. Penerapan Algoritma Jaro Winkler Distance

Simulasi seleksi data yang akan dilakukan oleh sistem dengan menggunakan data dari artikel <https://muslim.or.id/9166-islam-menjaga-dan-memuliakan-wanita.html> memiliki hasil *pre-processing* (*case folding* dan *stopword elimination*) sebagai berikut:

Perhitungan algoritma jaro winkler distance dilakukan dengan persamaan (1) dan (2) sebagai berikut:

Diketahui:

$$\begin{aligned} m &= 561 \\ s_1 &= 991 \\ s_2 &= 799 \\ t &= 0 \end{aligned}$$

$$\begin{aligned} d_j &= \frac{1}{3} x \left(\frac{561}{991} + \frac{561}{799} + \frac{561 - 0}{561} \right) \\ d_j &= \frac{1}{3} x (0,56 + 0,70 + 1) \\ d_j &= 0,795 \end{aligned}$$

Berdasarkan hasil perhitungan di atas nilai *jaro distance* (d_j) sebesar 0,75.

$$\begin{aligned} d_w &= 0,75 + (3 x 0,1(1 - 0,75)) \\ d_w &= 0,75 + (0,3(0,25)) \\ d_w &= 0,75 + (0,07) \\ d_w &= 0,82 \end{aligned}$$

Berdasarkan hasil perhitungan diatas menghasilkan nilai *Jaro winkler distance* ialah 0,82. Maka dapat ditarik kesimpulan bahwa kedua teks diatas memiliki tingkat kesamaan rendah yaitu <1 maka akan berhasil melakukan insert data pada form pengunggahan konten baru.

C. Pengujian Sistem

Pengujian sistem pada penelitian ini menggunakan pengujian *blackbox* (*functional analysis*), yang dilakukan secara objektif dimana aplikasi diuji secara langsung ke lapangan dengan membuat kuisioner yang ditujukan pada pengguna aplikasi yang disebar sebanyak 17 responden dengan 6 pertanyaan dan 6 pilihan yang akan mewakili dari tujuan akhir penelitian. Berikut rekapitulasi perhitungan kuisioner pada penelitian ini sebagai berikut:

Tabel 1. Tabel Rekapitulasi Perhitungan Kuisisioner

Pertanyaan	Keterangan					
	SS	S	CS	KS	TS	STS
Apakah tampilan aplikasi menarik	-	11	4	1	-	1
Apakah konten dalam aplikasi menarik	2	8	7	-	-	-
Apakah aplikasi membantu pihak wanita dalam memperoleh informasi mengenai wanita islami	2	12	3	-	-	-
Apakah proses tidak membutuhkan waktu yang lama	1	12	4	-	-	-
Apakah aplikasi tidak membutuhkan ruang penyimpanan yang banyak	2	8	5	2	-	-
Apakah algoritma <i>Jaro winkler distance</i> cocok diimplementasikan pada aplikasi Ensiklopedia Wanita	1	10	6	-	-	-
TOTAL	8	61	29	3	-	1

Hasil perhitungan jawab responden sebagai berikut:

$$\begin{aligned} \text{Sangat Setuju} &= (8 \cdot 6) / 6 = 8 \\ \text{Setuju} &= (61 \cdot 5) / 6 = 50,83 \\ \text{Cukup Setuju} &= (29 \cdot 4) / 6 = 19,33 \\ \text{Kurang Setuju} &= (3 \cdot 3) / 6 = 1,5 \\ \text{Tidak Setuju} &= (0 \cdot 2) / 6 = 0 \\ \text{Sangat Tidak Setuju} &= (1 \cdot 1) / 6 = 0,16 \end{aligned}$$

Jadi total nilai yang didapat dari perhitungan diatas memperoleh nilai 79,82. Maka penilaian interpretasi responden terhadap media pembelajaran tersebut dapat diperoleh dengan cara:

$$\text{Hasil} = \text{total nilai} / Y \times 100 \dots (8)$$

Hasil = Hasil persentase keseluruhan

Y = jumlah nilai tertinggi pada item sangat setuju (nilai bobot dikali jumlah responden).

$$\begin{aligned} \text{Hasil} &= 79,82 / 102 \times 100 \\ &= 78\% . \end{aligned}$$

Tabel 2. Presentasi Nilai Kuisisioner

Jawaban	Keterangan
Sangat Setuju	86% – 100%
Setuju	69% - 85%
Cukup Setuju	52% – 68%
Kurang Setuju	35% - 51%
Tidak Setuju	18% – 34%
Sangat Tidak Setuju	1% - 17%

Pada Tabel 2 Presentase Nilai Kuisisioner, digunakan untuk mengukur nilai yang didapat pada perhitungan kuisisioner keseluruhan masuk kedalam kategori yang mana. Maka berdasarkan hasil diatas dengan nilai 78% menyatakan setuju, dapat disimpulkan bahwa aplikasi Ensiklopedia Wanita Berbasis Mobile Apps dengan Penerapan Algoritma *Jaro Winkler Distance* baik untuk diterapkan dan sangat mempermudah masyarakat dalam memperoleh informasi secara cepat dan menarik.

IV. Kesimpulan dan saran

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, maka penulis dapat menarik beberapa kesimpulan, antara lain:

- 1) Sistem ini telah berhasil menerapkan algoritma *Jaro winkler distance* dalam proses seleksi data, dengan bantuan *text preprocessing*. Ini dapat dibuktikan dengan adanya hasil presentase kuisisioner tertinggi yang menyatakan setuju pada kecocokan implementasi algoritma *Jaro winkler distance* pada aplikasi Ensiklopedia Wanita dengan memperoleh nilai sebesar 59%.
- 2) Sistem ini berhasil membandingkan text yang baru dengan text yang ada pada database sehingga apabila terdapat kemiripan, maka proses *insert* gagal dilakukan.

- 3) Memudahkan masyarakat khususnya wanita dalam mencari dan berbagi informasi mengenai wanita islami secara cepat dan menarik, baik dalam kategori sosial, psikologi maupun Kesehatan dengan hasil presentase paling tinggi dari kuisisioner yaitu sebesar 70% menyatakan setuju.
- 4) Berdasarkan hasil pengujian keseluruhan pada penelitian ini memperoleh nilai sebesar 78%, maka dapat di tarik kesimpulan bahwa aplikasi ini dapat diterapkan dan sangat mempermudah masyarakat dalam memperoleh informasi secara cepat dan menarik. Serta penggunaan seleksi data pada proses insert konten baru menggunakan *text preprocessing* dan algoritma *Jaro winkler distance* sangat membantu.

Daftar Pustaka

- [1] Badan Perencanaan Pembangunan Nasional, "Jumlah Penduduk Perempuan Indonesia pada 2018 Mencapai 131,9 Juta Jiwa," *Databoks.Co.Id*, p. 1, 2018.
- [2] A. Y. Ranius *et al.*, "Implementasi Metode Mobile-D Untuk Pembuatan Ensiklopedia Kebudayaan Kota Palembang Berbasis Android," *J. Ilm.*, vol. x, no. 12, pp. 1–12, 2014.
- [3] B. Permana, E. Wahyu Hidayat, and A. Rahmatulloh, "Aplikasi Ensiklopedia Pakaian Adat Dunia Berbasis Android," *J. Inform. J. Pengemb. IT*, vol. 3, no. 2, pp. 151–156, 2018.
- [4] Tinaliah, T. and Elizabeth, T., "Perbandingan Hasil Deteksi Plagiarisme Dokumen dengan Metode Jaro-Winkler Distance dan Metode Latent Semantic Analysis," *Jurnal Teknologi dan Sistem Komputer*, vol. 6, no.1, pp. 7-12, 2018.
- [5] Novantara, P., "Implementasi Algoritma Jaro-Winkler Distance Untuk Sistem Pendeteksi Plagiarisme Pada Dokumen Skripsi," *Buffer Informatika*, vol 3 no. 2, 2018
- [6] Febrianti, Y.M., "Analisis Sentimen Pada Ulasan "Lazada" Berbahasa Indonesia Menggunakan K-Nearest Neighbor (K-NN) Dengan Perbaikan Kata Menggunakan Jaro Winkler Distance ", *Universitas Brawijaya*, 2018.
- [7] P. Novantara and O. Pasruli, "Implementasi Algoritma Jaro-Winkler Distance Untuk Sistem Pendeteksi Plagiarisme Pada Dokumen Skripsi," *J. Buffer Inform.*, vol. 3, no. 2, 2017.
- [8] A. T. J. H., "Preprocessing Text untuk Meminimalisir Kata yang Tidak Berarti dalam Proses Text Mining," *Inform. UPGRIS*, vol. 1, pp. 1–9, 2015.
- [9] A. Kurniawati, "Implementasi Algoritma Jaro-Winkler Distance untuk Membandingkan Kesamaan Dokumen Berbahasa Indonesia," *Proceeding, Semin. Ilm. Nas. Komput. dan Sist. Intelijen KOMMIT 2008, Depok, Indones.*, 2010.
- [10] T. Tinaliah and T. Elizabeth, "Perbandingan Hasil Deteksi Plagiarisme Dokumen dengan Metode Jaro-Winkler Distance dan Metode Latent Semantic Analysis," *J. Teknol. dan Sist. Komput.*, vol. 6, no. 1, pp. 7–12, 2018.