

Koreksi Otomatis Ujian Esai Menerapkan Algoritma *Winnowing* Dan Metode *Cosine Similarity*

Diterima: 10 Juni 2024
Revisi: 10 Juli 2024
Terbit: 1 Agustus 2024

^{1*}Arwienda Kayan, ²Ardi Sanjaya, ³Umi Mahdiyah
¹⁻³Universitas Nusantara PGRI Kediri
^{1*}arwiendakayan@gmail.com, ²dersky@gmail.com,
³umimahdiyah@unpkediri.ac.id

Abstrak— Tujuan dari penelitian ini adalah untuk memudahkan dan mengurangi waktu koreksi dari ujian esai yang semula masih dilakukan secara manual dan membutuhkan waktu yang tidak sedikit, terlebih lagi jika terdapat jawaban yang tidak jelas pada saat siswa menjawab ujian esai. Proses yang digunakan dalam sistem berupa pra pemrosesan yaitu mengubah semua huruf kapital menjadi huruf kecil, dan menghapus karakter-karakter yang tidak diperlukan. Setelah melakukan pra pemrosesan maka akan dilakukan perhitungan menggunakan algoritma *Winnowing* dan metode *Cosine Similarity* yang akan menghasilkan presentase kemiripan antara jawaban siswa dengan kunci jawaban guru sehingga dapat digunakan untuk menilai apakah jawaban siswa benar atau salah.

Kata Kunci—Ujian Esai;Kemiripan Kata;*Winnowing*;*Cosine Similarity*

Abstract— *The purpose of this research is to simplify and reduce the time required for grading essay exams, which were previously done manually and consumed significant time, especially when students' answers were unclear. The system process includes preprocessing steps such as converting all uppercase letters to lowercase and removing unnecessary characters. After preprocessing, the system calculates the similarity using the Winnowing algorithm and Cosine Similarity method, which will result in a similarity percentage between the students' answers and the teachers' answer key.*

Keywords—*Essay Exam; Word Similarity; Winnowing; Cosine Similarity*

This is an open access article under the CC BY-SA License.



Penulis Korespondensi:

Arwienda Kayan,
Teknik Informatika,
Universitas Nusantara PGRI Kediri,
Email: arwiendakayan@gmail.com
ID Orcid: [<https://orcid.org/register>]
Handphone: 087880035531

I. PENDAHULUAN

Teknologi yang berkembang pesat memudahkan akses informasi oleh banyak orang, serta mendorong perkembangan sistem komunikasi dengan komputer. *Natural Language Processing* (NLP) atau Pemrosesan Bahasa Alami adalah salah satu sistem yang berkembang tersebut. Dengan bantuan NLP, manusia dapat berinteraksi dengan komputer dan menyelesaikan berbagai tugas terkait bahasa [1].

Metode ujian dalam pendidikan digunakan untuk menilai perkembangan siswa dalam pemahaman materi pelajaran yang telah di ajarkan. Ujian adalah alat ukur yang digunakan untuk menilai kinerja siswa dan menentukan tingkat pembelajaran yang telah di capai secara merata [2]. Ujian esai digunakan agar siswa dapat menjelaskan pemikiran dan pemahaman mereka secara langsung terhadap pertanyaan yang diajukan. Penggunaan ujian esai membantu mengidentifikasi pemahaman siswa dan mendorong pemikiran yang lebih mendalam [3].

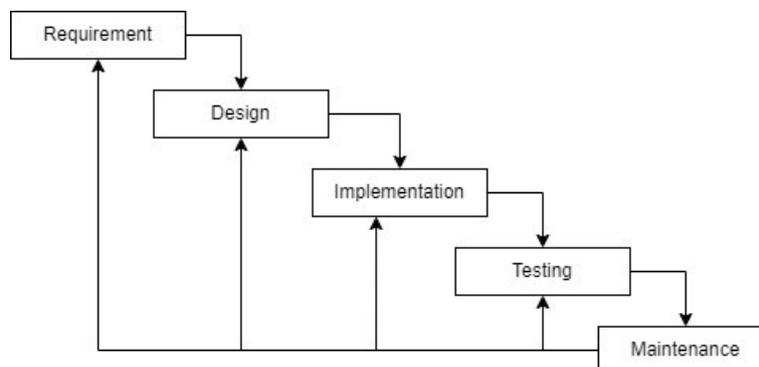
Meskipun ujian esai memiliki keunggulan, terdapat beberapa kekurangan terutama dalam proses koreksi. Koreksi masih dilakukan secara manual oleh pengajar, memakan waktu yang lama, terutama jika jawaban siswa tidak jelas dalam penulisannya. Untuk mengatasi hal tersebut, penelitian ini bertujuan menerapkan NLP. Sistem akan membandingkan jawaban siswa dengan kunci jawaban guru. Dengan demikian sistem dapat langsung menentukan apakah jawaban siswa tepat atau tidak.

Dalam penelitian [4] dapat disimpulkan bahwa algoritma *Winnowing* efektif digunakan untuk mengoreksi soal esai berupa isian pendek, dengan syarat ukuran window tidak melebihi nilai hash. Penelitian lain [5] [6] menunjukkan bahwa metode *Cosine Similarity* yang dilengkapi dengan pra-pemrosesan pada korelasi *Pearson* menghasilkan nilai korelasi tertinggi sebesar 0.62. Berdasarkan hasil penelitian sebelumnya, penelitian ini akan menggunakan algoritma *Winnowing* untuk menemukan nilai *fingerprnt* dari jawaban siswa dan kunci jawaban guru. Selanjutnya, metode *Cosine Similarity* akan digunakan untuk menghitung nilai kesamaan antar keduanya.

II. METODE

A. Metode Penelitian

Metode penelitian untuk mengembangkan sistem dalam penelitian ini menggunakan metode *Waterfall*. Metode ini dimulai dari fase *requirement* dan fase terakhir adalah *maintenance* [7]. Berikut adalah tahapan pengembangan sistem menggunakan metode *Waterfall*.



Gambar 1. Tahapan Metode *Waterfall*

1. *Requirement*

Pada tahap ini, langkah pertama yang dilakukan adalah pengumpulan kebutuhan untuk dianalisis. Tujuannya adalah untuk mengidentifikasi dan memahami kebutuhan apa yang diperlukan untuk mendukung penelitian ini.

- a. Data yang diperlukan yaitu berupa soal, jawaban siswa, dan kunci jawaban dari guru.
- b. Jenis data berupa teks yaitu soal, jawaban siswa, dan kunci jawaban dari guru. Lalu berupa numerik untuk nilai kemiripan.
- c. Tujuan pengambilan data soal adalah untuk penyusunan soal dan kunci jawaban, pengumpulan jawaban siswa untuk menilai kemiripan dengan kunci jawaban guru.

2. *Design*

Pada tahap desain, seluruh sistem dirancang mencakup perancangan *use case diagram*, proses *input* dan *output*, serta desain *database*, yang meliputi struktur data, alur proses, dan antarmuka pengguna yang akan digunakan dalam aplikasi.

3. *Implementation*

Pada tahap ini, dilakukan pengkodean menggunakan bahasa pemrograman PHP sesuai dengan rancangan desain sistem yang telah dibuat.

4. *Testing*

Pada tahap ini, dilakukan proses pengujian untuk memastikan sistem berjalan dengan fungsi dan desain yang telah ditetapkan.

5. *Maintenance*

Pada tahap ini, dilakukan pemeliharaan sistem, termasuk perbaikan dan evaluasi, agar sistem dapat berjalan sesuai dengan tujuannya.

B. Algoritma *Winnowing*

Algoritma *winnowing* adalah sebuah algoritma yang mengukur nilai kesamaan antara dua kata dan menghasilkan nilai *fingerprint* [8].

$$H(ck) = C_1 * b^{(k-1)} + C_2 * b^{(k-2)} + \dots + C_{(k-1)} * b^k + C_k \dots \dots \dots (1)$$

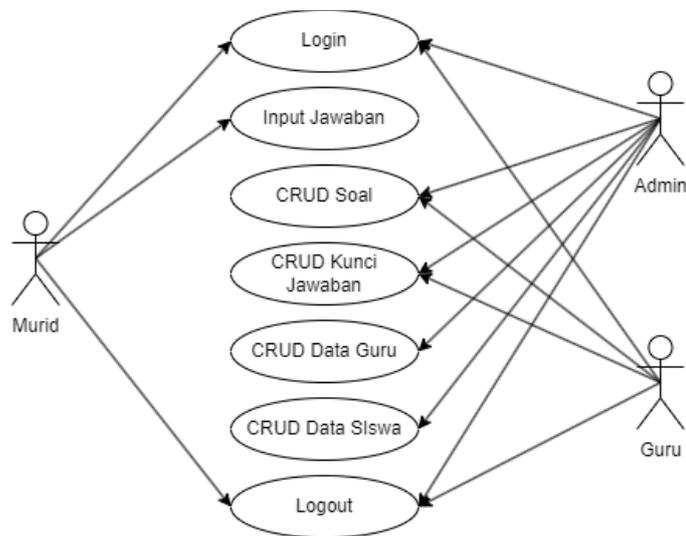
C. Metode *Cosine Similarity*

Metode *Cosine Similarity* adalah cara untuk mengukur kesamaan antara dua teks atau kata [9]. Penilaian dari perhitungan metode *Cosine Similarity* adalah antara 0-1, dimana nilai 1 berarti dua teks sangat mirip dan nilai 0 berarti dua teks tidak mirip.

$$\text{Similarity}(A, B) = \frac{\sum_{i=1}^n A \cdot B}{\sqrt{\sum_{i=1}^n A^2} \cdot \sqrt{\sum_{i=1}^n B^2}} \dots\dots\dots (2)$$

D. Use Case Diagram

Use case diagram dibawah merupakan gambaran aktifitas yang dapat dilakukan oleh siswa, guru dan admin pada saat menggunakan sistem. Dalam use case tersebut siswa dapat melakukan login, input jawaban, dan logout. Lalu guru dapat melakukan login, tambah soal, hapus soal, ubah soal, dan logout. Sedangkan admin dapat melakukan semua proses kecuali input jawaban.



Gambar 2. Use Case Diagram

III. HASIL DAN PEMBAHASAN

A. Pengujian dan Hasil

Pada tabel 1 adalah hasil penilaian guru pada setiap soal dari 18 orang siswa, jika bernilai 0 maka dianggap salah dan jika bernilai 10 maka dianggap benar.

Tabel 1. Hasil Penilaian Guru

| Jumlah Siswa | Penilaian Guru | | | | | | | | | |
|--------------|----------------|------|------|------|------|------|------|------|------|-------|
| | Ke-1 | Ke-2 | Ke-3 | Ke-4 | Ke-5 | Ke-6 | Ke-7 | Ke-8 | Ke-9 | Ke-10 |
| 1 | 10 | 0 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 10 | 10 |
| 3 | 10 | 0 | 10 | 0 | 10 | 0 | 10 | 0 | 10 | 10 |
| 4 | 10 | 0 | 10 | 0 | 0 | 0 | 10 | 10 | 10 | 10 |
| 5 | 10 | 0 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 6 | 10 | 0 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |

| | | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|----|----|
| 7 | 10 | 0 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 8 | 10 | 0 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 9 | 10 | 0 | 10 | 10 | 10 | 10 | 10 | 0 | 10 | 10 |
| 10 | 0 | 0 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 11 | 10 | 0 | 10 | 0 | 10 | 0 | 10 | 0 | 10 | 10 |
| 12 | 10 | 0 | 10 | 0 | 10 | 0 | 10 | 0 | 10 | 10 |
| 13 | 10 | 0 | 10 | 0 | 10 | 0 | 10 | 0 | 10 | 10 |
| 14 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 15 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 16 | 10 | 0 | 10 | 0 | 0 | 0 | 10 | 10 | 10 | 10 |
| 17 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 18 | 0 | 0 | 10 | 0 | 0 | 0 | 10 | 10 | 10 | 10 |

Lalu pada tabel 2 terdapat data hasil penilaian sistem. Sistem menilai dari rentang nilai 1 hingga 10.

Tabel 2. Hasil Penilaian Sistem

| Jumlah Siswa | Penilaian Sistem | | | | | | | | | |
|--------------|------------------|------|------|------|------|------|------|------|------|-------|
| | Ke-1 | Ke-2 | Ke-3 | Ke-4 | Ke-5 | Ke-6 | Ke-7 | Ke-8 | Ke-9 | Ke-10 |
| 1 | 10 | 1,9 | 10 | 5,4 | 10 | 7,4 | 10 | 7,5 | 10 | 10 |
| 2 | 9,3 | 9,1 | 8,4 | 7,3 | 6,1 | 5,2 | 7,7 | 2,9 | 10 | 9,9 |
| 3 | 7,9 | 7,1 | 9,9 | 7,7 | 10 | 7,1 | 10 | 7,2 | 10 | 10 |
| 4 | 10 | 1,9 | 10 | 3,1 | 7,1 | 9,4 | 10 | 7,5 | 10 | 9,9 |
| 5 | 10 | 1,9 | 10 | 6,5 | 10 | 10 | 10 | 7,5 | 10 | 10 |
| 6 | 10 | 1,9 | 8,7 | 5,4 | 10 | 7,4 | 10 | 7,5 | 10 | 10 |
| 7 | 10 | 1,9 | 10 | 5,4 | 10 | 7,4 | 10 | 7,5 | 10 | 10 |
| 8 | 10 | 1,9 | 10 | 5,4 | 9,9 | 7,4 | 10 | 7,5 | 10 | 9,9 |
| 9 | 7,9 | 1,9 | 10 | 6,9 | 8,7 | 7,4 | 9 | 4,2 | 10 | 10 |
| 10 | 4,9 | 1,9 | 10 | 6,5 | 9,9 | 7,4 | 10 | 7,5 | 10 | 9,9 |
| 11 | 10 | 7,1 | 10 | 7,3 | 10 | 7,1 | 10 | 6,9 | 10 | 10 |
| 12 | 10 | 4,6 | 10 | 4,4 | 10 | 4,2 | 10 | 4,2 | 10 | 10 |
| 13 | 7,9 | 7,1 | 9,9 | 7,7 | 10 | 7,4 | 9,5 | 7,2 | 10 | 10 |
| 14 | 10 | 10 | 10 | 9,7 | 10 | 10 | 10 | 10 | 10 | 10 |
| 15 | 10 | 9,2 | 10 | 9,7 | 10 | 7,4 | 10 | 10 | 10 | 10 |
| 16 | 10 | 1,9 | 10 | 3,1 | 7,1 | 8,5 | 9,5 | 7,5 | 10 | 10 |
| 17 | 7,9 | 7,9 | 9,9 | 9,9 | 8,7 | 7,4 | 9,5 | 6,6 | 10 | 10 |
| 18 | 9,3 | 1,9 | 10 | 3,1 | 8,7 | 6,4 | 10 | 7,5 | 10 | 10 |

Terdapat perbedaan penilaian guru dengan penilaian sistem, dimana guru menilai dengan nilai 10 dan 0 sementara sistem menilai dari rentang 1 hingga 10. Oleh sebab itu hasil penilaian dari sistem masih belum bisa menentukan apakah jawaban siswa benar atau salah. Maka, harus menggunakan *threshold* atau batas nilai. Misal menggunakan nilai

threshol 5 maka jika nilai pada tabel 2 di bawah 5 maka di nilai 0 dan jika di atas nilai 5 maka dinilai 10. Hasil penggunaan nilai *threshol* ada pada tabel 3.

$$f(x) = \begin{cases} 0, & \text{jika } x \leq \theta \\ 10, & \text{jika } x > \theta \end{cases} \dots\dots\dots(3)$$

Keterangan :

θ = *threshol*

x = hasil penilaian sistem

Tabel 3. Hasil Penilaian Sistem Setelah Penggunaan *Threshol*

| Jumlah Siswa | Nilai Setelah Penerapan <i>Threshol</i> | | | | | | | | | |
|--------------|---|------|------|------|------|------|------|------|------|-------|
| | Ke-1 | Ke-2 | Ke-3 | Ke-4 | Ke-5 | Ke-6 | Ke-7 | Ke-8 | Ke-9 | Ke-10 |
| 1 | 10 | 0 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 2 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 0 | 10 | 10 |
| 3 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 4 | 10 | 0 | 10 | 0 | 10 | 10 | 10 | 10 | 10 | 10 |
| 5 | 10 | 0 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 6 | 10 | 0 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 7 | 10 | 0 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 8 | 10 | 0 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 9 | 10 | 0 | 10 | 10 | 10 | 10 | 10 | 0 | 10 | 10 |
| 10 | 0 | 0 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 11 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 12 | 10 | 0 | 10 | 0 | 10 | 0 | 10 | 0 | 10 | 10 |
| 13 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 14 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 15 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 16 | 10 | 0 | 10 | 0 | 10 | 10 | 10 | 10 | 10 | 10 |
| 17 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 18 | 10 | 0 | 10 | 0 | 10 | 10 | 10 | 10 | 10 | 10 |

Pada tabel 3 adalah hasil penerapan nilai *threshol* 5. Digunakannya nilai *threshol* 5 karena pada saat pengujian dari nilai *threshol* dengan rentang 1 hingga 10 didapati bahwa nilai *threshol* 5 mendapat akurasi tertinggi yaitu 85.55%. Hasil pengujian nilai *threshol* menggunakan rumus berikut: [10]

$$Akurasi = \frac{\text{jumlah klasifikasi benar}}{\text{jumlah semua data}} \dots\dots\dots(4)$$

Tabel 4 merupakan hasil pengujian nilai *threshol* dengan menguji nilai *threshol* 1 hingga 10 mendapatkan hasil yaitu sebagai berikut:

Tabel 4. Hasil Pengujian Nilai *Threshol*

| <i>Threshold</i> | Sama | Tidak Sama |
|------------------|-------------|-------------------|
| 1 | 74,44 | 25,56 |
| 2 | 80,00 | 20,00 |
| 3 | 80,55 | 19,44 |
| 4 | 82,22 | 17,77 |
| 5 | 85,55 | 14,44 |
| 6 | 83,88 | 16,11 |
| 7 | 83,33 | 16,66 |
| 8 | 79,44 | 20,55 |
| 9 | 79,44 | 20,55 |
| 10 | 74,44 | 25,55 |

B. Analisis Hasil

Berdasarkan hasil pengujian dapat diketahui bahwa terdapat kesalahan dalam sistem dimana seharusnya jawaban siswa diberi nilai salah, namun oleh sistem jawaban tersebut diberi nilai benar. Hal tersebut terjadi karena algoritma *Winnowing* hanya menghitung atau mengolah pola karakter dan bukan mengolah makna dari kata yang di hitung kesamaannya. Sehingga kedua kata yang berbeda secara makna masih dianggap mirip oleh algoritma *Winnowing* jika punya kesamaan pola karakter.

IV. KESIMPULAN

Penggunaan algoritma *Winnowing* dan metode *Cosine Similarity* pada sistem koreksi otomatis ujian esai, dimana setelah melakukan pengujian di dapatkan nilai *threshold* 5 adalah nilai yang memiliki akurasi tertinggi yaitu sebesar 85.55%. Setelah penggunaan nilai *threshold* sistem lebih mendekati penilaian guru dalam menilai jawaban siswa. Namun, masih terdapat kesalahan sistem yaitu penilaian yang masih salah atau tidak sama dengan nilai guru disebabkan oleh terbatasnya algoritma *Winnowing* yang hanya mengolah pola karakter dan bukan mengolah makna dari kunci jawaban dan jawaban siswa, sehingga kata yang berbeda secara makna masih dapat dianggap mirip oleh sistem.

DAFTAR PUSTAKA

- [1] Goesderilidar, G., & Zainurrahman, D. (2021). Prospek Pengembangan Komputasional Linguistik Menggunakan Pemrosesan Bahasa Alamiah (Natural Language Processing/NLP). *IndraTech*, 2(2), 81-85..

- [2] Mufiid, I., & Lestanti, S. (2021). Aplikasi Penilaian Jawaban Esai Otomatis Menggunakan Metode Synonym Recognition dan Cosine Similarity Berbasis Web. *Jurnal Mnemonic*, 4(2), 31-37..
- [3] Ramadhani, R. A., Sanjaya, A., & Prayogo, Y. S. (2023, July). Pemanfaatan Metode Consine Similarity Untuk Koreksi Nilai UAS. In *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)* (Vol. 7, No. 1, pp. 122-130).
- [4] Kurniawati, E. F., & Pradnya, W. M. (2020). Implementasi Algoritma Wnnowing Pada Sistem Penilaian Otomatis Jawaban Esai Pada Ujian Online Berbasis Web. *Jurnal Khatulistiwa Informatika*, 6(2), 169-175..
- [5] Hasanah, U., & Mutiara, D. A. (2019, December). Perbandingan Metode Cosine Similarity dan Jaccard Similarity untuk Penilaian Otomatis Jawaban Pendek. In *SENSITIF: Seminar Nasional Sistem Informasi dan Teknologi Informasi* (pp. 1255-1263).
- [6] Sanjaya, A., Setiawan, A. B., Mahdiyah, U., Farida, I. N., & Prasetyo, A. R. (2024). Pengukuran Kemiripan Makna Menggunakan Cosine Similarity dan Basis Data Sinonim Kata. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 10(4), 747-752.
- [7] Wahid, A. A. (2020). Analisis metode waterfall untuk pengembangan sistem informasi. *J. Ilmu-ilmu Inform. dan Manaj. STMIK*, no. November, 1(1), 1-5.
- [8] Rihi, M. P., Pekuwali, A. A., & Sitaniapessy, D. A. (2022). Algoritma Wnnowing untuk Mendeteksi Kesamaan Judul Skripsi Teknik Informatika: Wnnowing Algorithm for Detecting the Similarity of Informatic Engineering Undergraduate Thesis Title. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 2(2), 42-52.
- [9] Arfandy, H., & Musdar, I. A. (2020). Rancang Bangun Sistem Cerdas Pemberian Nilai Otomatis Untuk Ujian Essai Menggunakan Algoritma Cosine Similarity. *Inspiration: Jurnal Teknologi Informasi dan Komunikasi*, 10(2), 123-136.
- [10] Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, 5(2), 697-711.