

Penerapan Algoritma *Winnowing* dan *Cosine Similarity* untuk Penilaian Jawaban Esai

Marwan Hadid^{1*}, Anggra Triawan²

^{1,2}Program Studi Teknik Informatika, Fakultas Informatika dan Komputer, Universitas Binaniaga Indonesia
e-mail: marwanhadid6@gmail.com

*Corresponding Author

ABSTRACT

Essay assessment is one of the evaluation methods frequently used to measure students' understanding of the material. However, manual essay assessment often takes a long time, is prone to errors, and sometimes the grading results are incorrect. Therefore, a computational approach is needed to facilitate lecturers in grading. The *Winnowing* and *Cosine Similarity* algorithms are proposed as solutions to improve the effectiveness and accuracy of essay assessments. The assessment process begins with text pre-processing, where important words are retained, and irrelevant words are discarded. Next, the essay text is broken into several substrings or turned into tokens using the *k-gram* method. Then, hashing is performed on these tokens using the *Rolling Hash* algorithm. The *Winnowing* algorithm is used to select the smallest hashing value from the set of generated hashing values, and the frequency of the occurrence of these values is calculated in both the student's essay and the answer key. Finally, the similarity between the essay and the answer key is calculated using *Cosine Similarity*. This assessment application has been tested by system experts and deemed "Highly Feasible" with a feasibility rate of 100%. A user test using the PPSUQ questionnaire also showed that this application is "Highly Feasible," with an overall result of 83%. Additionally, the MAPE (Mean Absolute Percentage Error) test result showed a grading error rate of 11%, which indicating that this application is effective and accurate in essay assessments.

Keywords: Essay Scoring, *Winnowing* Algorithm, *Cosine Similarity* Algorithm, Mean Absolute Percentage Error, Application Prototype

ABSTRAK

Penilaian esai merupakan salah satu bentuk evaluasi yang sering digunakan untuk mengukur kemampuan mahasiswa dalam memahami materi. Namun, penilaian esai secara manual sering kali memakan waktu yang lama, rentan terhadap kesalahan, dan terkadang hasil penilaiannya kurang tepat. Oleh karena itu, diperlukan pendekatan komputasi untuk mempermudah dosen dalam memberikan nilai. Algoritma *Winnowing* dan *Cosine Similarity* diusulkan sebagai solusi untuk meningkatkan efektivitas dan ketepatan penilaian esai. Proses penilaian dimulai dengan text pre-processing, di mana kata-kata penting dipertahankan dan kata-kata yang tidak relevan dibuang. Selanjutnya, teks jawaban dipecah menjadi beberapa potongan substring atau dijadikan sebagai token menggunakan metode *k-gram*. Kemudian, dilakukan hashing terhadap token-token menggunakan algoritma *Rolling Hash*. Algoritma *Winnowing* digunakan untuk memilih nilai hashing terkecil dari kumpulan nilai hashing yang dihasilkan, dan frekuensi kemunculan nilai tersebut dihitung pada jawaban mahasiswa dan kunci jawaban. Akhirnya, tingkat kemiripan antara jawaban dan kunci jawaban dihitung menggunakan *Cosine Similarity*. Aplikasi penilaian ini telah diuji oleh ahli sistem dan dinyatakan "Sangat Layak" dengan tingkat kelayakan 100%. Hasil pengujian kelayakan aplikasi kepada pengguna juga menunjukkan aplikasi ini "Sangat Layak" dengan hasil keseluruhan 83%. Selain itu, hasil uji Mean Absolute Percentage Error (MAPE) menunjukkan tingkat kesalahan nilai sebesar 11%, yang menunjukkan aplikasi ini efektif dan tepat dalam penilaian esai.

Kata Kunci: Penilaian Esai, Algoritma *Winnowing*, Algoritma *Cosine Similarity*, Mean Absolute Percentage Error, Prototype Aplikasi

A. PENDAHULUAN

1. Latar Belakang

Evaluasi merupakan proses yang bersifat komprehensif dan sistematis di mana ini meliputi pengukuran, penilaian, dan analitis untuk menentukan tingkat kemampuan peserta didik dalam mencapai tujuan pembelajaran serta menentukan keberhasilan suatu program pendidikan yang dilaksanakan (Widiyanto, 2018, hlm. 9). Ujian atau tes adalah salah satu bentuk evaluasi yang dilakukan secara sistematis, komprehensif, dan objektif di mana hasilnya akan dijadikan sebagai acuan dalam mengukur kemampuan siswa dan menentukan keputusan bagi dalam proses pembelajaran. Tes jenis uraian atau esai adalah jenis tes yang menuntut peserta didik untuk mengorganisasikan, mengekspresikan, serta menyatakan jawabannya menurut kata – kata sendiri. Proses penilaian tes uraian dilakukan dengan menyiapkan sejumlah kata kunci atau *keyword* sebagai kunci jawaban lalu akan dicocokkan dengan jawaban peserta ujian. Semakin banyak jumlah kata yang sama, semakin besar pula skor yang diberikan.

Proses penilaian jawaban ujian yang semula dilakukan secara penuh oleh manusia kini dapat dialihkan ke dalam bentuk sebuah aplikasi atau sistem. Sistem yang memiliki kecerdasan layaknya manusia. *Natural Language Processing* merupakan gabungan dari cabang ilmu komputer dan linguistik yang mempelajari serta mengembangkan kemampuan komputer untuk memahami bahasa alami manusia yang lebih kompleks.

Cosine Similarity adalah salah satu metode untuk mengukur kemiripan antara dua buah vektor. Secara spesifik, metode ini mengukur sudut cosinus di antara dua vektor dalam ruang multidimensi. Semakin kecil sudut antara dua vektor, semakin mirip pula kemiripan mereka. Penelitian sebelumnya menggunakan metode *Cosine Similarity* dengan pendekatan *string similarity* untuk melakukan penilaian otomatis jawaban esai ujian. dan didapat tingkat akurasi sebesar 85,4% (Kurniadi, Gernowo, Surarso, Wibowo, & Warsito, 2023).

Algoritma *Winnowing* adalah algoritma yang digunakan untuk melakukan *document fingerprinting*, yakni teknik mengubah setiap *string* menjadi susunan angka atau disebut juga dengan *hashing* (Purwaningrum dkk., 2023). Teknik ini dilakukan untuk mengukur tingkat kemiripan salinan antardokumen atau sebagian teks (Kurniawati & Pradnya, 2020). Berdasarkan beberapa pemaparan yang telah dijelaskan di atas, penyusun berkeinginan untuk melakukan penelitian yaitu menerapkan Algoritma *Winnowing* dan *Cosine Similarity* untuk penilaian jawaban esai ujian berbasis komputer.

2. Permasalahan

Proses penilaian jawaban esai ini cenderung menggunakan waktu yang tidak relatif singkat. Terdapat beberapa indikator yang memunculkan adanya suatu masalah yang dijabarkan pada tabel berisi tiga pertanyaan kuesioner sebagaimana dijelaskan pada Tabel 1.

Tabel 1. Pertanyaan Kuesioner Permasalahan

No	Pertanyaan	Jawaban	
		Iya	Tidak
1	Apakah selama melakukan penilaian esai terdapat kesulitan yang bapak / ibu alami?	58%	42%
2	Apakah bapak / ibu pernah melakukan kesalahan dalam melakukan penilaian esai?	54%	46%
3	Dalam melakukan penilaian esai, perlu adanya objektivitas selama melakukan penilaian seperti dengan cara tidak membedakan – bedakan antar mahasiswa, memberikan nilai sesuai dengan kompetensi yang dimiliki. Tetapi pelaksanaannya terdapat beberapa kasus penilaian dilakukan secara subjektif seperti memberikan nilai berdasarkan gender di mana mahasiswi akan lebih diutamakan dalam memberikan bobot nilai sebagaimana dijelaskan pada situs: https://persfe.com/menilik-objektivitas-dosen-dalam-memberikan-nilai/ . Dari kasus tersebut, apakah bapak atau ibu pernah melakukan penilaian esai secara subjektif?	21%	79%

Selain itu, dampak dari waktu dalam menilai esai yang relatif tidak singkat adalah adanya keterlambatan dosen dalam mengirimkan nilai. Hal ini dibuktikan dengan adanya tabel keterangan pengumpulan nilai mata kuliah dijelaskan pada Tabel 2.

Tabel 2. Keterangan Pengumpulan Nilai

Mata Kuliah	Keterangan Pengumpulan
Kelas A, B, C	
Statistik	Nilai untuk kelas C belum dikumpulkan
DBMS	Nilai sudah dikumpulkan
Pancasila	Nilai untuk kelas C belum dikumpulkan
Programming Language	Nilai untuk kelas C belum dikumpulkan
Business Process	Nilai sudah dikumpulkan
Platform Tech	Nilai sudah dikumpulkan
User Experience	Nilai sudah dikumpulkan
Kelas D	
Enterprise System	Nilai sudah dikumpulkan
Business Intelligence - Data Mining	Nilai sudah dikumpulkan
Project Proposal and Seminar	Nilai belum dikumpulkan
IT Service Management	Nilai sudah dikumpulkan
Management of Science	Nilai sudah dikumpulkan
Data Mining and Data Warehousing	Nilai untuk kelas C belum dikumpulkan
Information Visualization	Nilai sudah dikumpulkan
Enterprise System	Nilai sudah dikumpulkan
IS/IT Enterprise Architecture	Nilai sudah dikumpulkan
IS Strategy, Management, and Acquisition	Nilai sudah dikumpulkan
Web-based Programmang	Nilai untuk kelas C belum dikumpulkan
Bahasa Indonesia	Nilai untuk kelas C belum dikumpulkan
Information System Project (Minor)	Nilai untuk kelas A, B, dan C belum dikumpulkan
Data Integration	Nilai sudah dikumpulkan

Mata Kuliah	Keterangan Pengumpulan
Information System Application Testing	Nilai sudah dikumpulkan
Interpersonal Skill	Nilai sudah dikumpulkan
Technopreneurship	Nilai untuk kelas B belum dikumpulkan
Basic Culture	Nilai untuk kelas C belum dikumpulkan
Citizenship Studies	Nilai untuk kelas C belum dikumpulkan
Enterpreneurship	Nilai untuk kelas A dan B belum dikumpulkan
Religious Studies	Nilai sudah dikumpulkan
Social and Professionalism	Nilai sudah dikumpulkan
Bahasa Indonesia	Nilai sudah dikumpulkan
Interpersonal Skill	Nilai sudah dikumpulkan
Technopreneurship	Nilai belum dikumpulkan
Information System Project (Minor)	Nilai belum dikumpulkan
Distributed Computing	Nilai sudah dikumpulkan
Interactive System Development	Nilai belum dikumpulkan

Dari uraian permasalahan di atas, dapat diketahui bahwa proses penilaian esai cenderung sulit karena harus menyesuaikan jawaban mahasiswa yang kosakatanya beragam, adanya kesalahan selama penilaian esai seperti adanya poin penilaian yang tertukar dengan nilai soal poin lainnya, penilaian esai cenderung memakan waktu yang relatif tidak singkat dan adanya keterlambatan bagi dosen dalam mengirimkan nilai.

3. Tujuan

Menerapkan algoritma *Winnowing* dan *Cosine Similarity* untuk mendapatkan tingkat ketepatan, efektivitas, mengembangkan *prototype* aplikasi, dan mengukur seberapa tepat dan efektif algoritma yang diterapkan untuk penilaian esai.

B. METODE

Algoritma *Winnowing* adalah salah satu algoritma untuk melakukan *document fingerprinting* di mana *fingerprint* tersebut didapat dari dokumen teks yang diubah setiap karakter hurufnya menjadi susunan kode atau dengan istilah *hashing*. Pada setiap *window*, nilai *hash* yang dipilih adalah yang minimum, tetapi jika ada lebih dari satu nilai *hash* dengan nilai minimum, nilai yang dipilih berada posisinya paling kanan, kemudian simpan nilai *hash* yang dipilih untuk dijadikan sebagai *document fingerprinting* (Schleimer, Wilkerson, & Aiken, 2003, hlm. 78). Proses pembentukan *document fingerprinting* pada jawaban dan kunci jawaban terdiri dari beberapa langkah, yakni diawali dengan melakukan:

1. *Text pre-processing*, di mana teks akan diolah dengan cara mengubah huruf kapital menjadi huruf kecil, menghilangkan tanda baca, dan menghapus awalan dan akhiran teks;
2. Tokenisasi, dilakukan dengan cara memotong setiap kata pada teks menjadi beberapa potongan huruf; proses tokenisasi ini menggunakan metode k-gram;
3. *Hashing* atau mengonversi token jawaban esai menjadi susunan kode angka; pada dokumen token sejumlah karakter pertama pada token di-*hash* dengan cara karakter paling pertama dalam token dikonversi menjadi karakter ASCII; setelah nilai *hash* pertama didapatkan, *hashing* berlanjut dengan menggeser satu karakter ke kanan,
4. *Fingerprinting*, nilai *hashing* dikelompokkan ke dalam *window*; ukuran *window* menentukan berapa banyak nilai *hash* yang ditampung; dari kumpulan nilai – nilai *hash* dipilih nilai *hash* minimum,
5. Menghitung frekuensi kemunculan *fingerprinting* jawaban dan kunci jawaban. Nilai token pada *fingerprint* jawaban dan kunci jawaban digabungkan kemudian masing – masing token jawaban dan kunci jawaban dihitung frekuensi kemunculannya;

Cosine Similarity adalah algoritma untuk mengukur kesamaan antara dua dokumen yang mana dua dokumen tersebut masing – masing didefinisikan sebagai A_1 dan B_1 . Tingkat kesamaan antara dua dokumen diukur berdasarkan besar sudut yang menghubungkan dua dokumen tersebut, semakin kecil sudutnya, semakin besar tingkat kesamaan dua dokumen tersebut (Manning, Raghavan, & Schütze, 2009, hlm. 120–121). Rumus metode *cosine similarity* dapat dijelaskan sebagaimana yang ditunjukkan pada Gambar 1.

$$sim(d_1, d_2) = \frac{\vec{v}(d_1) \cdot \vec{v}(d_2)}{|\vec{v}(d_1)| |\vec{v}(d_2)|}$$

Gambar 1. Cosine Similarity

Keterangan:

θ adalah sudut antara dua vektor,

$A \cdot B$ adalah dot product pada vektor A dan B dan dikalkulasikan sebagai $A \cdot B = A^T B = \sum_1^n A_i B_i = A_1 B_1 + A_2 B_2 + A_3 B_3 + \dots + A_n B_n$,

$\|A\|$ dan $\|B\|$ adalah skala pada vektor yang dikalkulasikan sebagai $\|A\| = \sqrt{A_1^2 + A_2^2 + \dots + A_n^2}$ dan begitupun dengan $\|B\|$.

Setelah memperoleh *fingerprint* jawaban dan kunci jawaban, selanjutnya adalah menghitung tingkat similaritas antara dokumen jawaban esai dengan dokumen kunci jawaban dengan menggunakan algoritma *Cosine Similarity*.

Setelah mendapatkan tingkat kemiripan, pemberian nilai dilakukan dengan cara hasil tingkat kemiripan dikalikan dengan bobot soal sebagaimana dijelaskan pada rumus di bawah.

$$\text{nilai} = \text{tingkat kemiripan} \times \text{bobot soal}$$

C. HASIL DAN PEMBAHASAN

1. HASIL

Proses penilaian esai dilakukan dengan mencocokkan antara jawaban mahasiswa dengan kunci jawaban. Jawaban didapatkan dari soal esai yang telah dibuat sebelumnya beserta bobotnya.

Tabel 3. Soal Esai dan Bobot Soal

Soal Esai	Bobot Soal
Tuliskan 3 kerangka DSS dan jelaskan!	5

Setelah didapatkan soal beserta bobotnya, selanjutnya adalah mengumpulkan jawaban esai dan kunci jawabannya.

Tabel 4. Jawaban dan Kunci Jawaban Esai

Jawaban Esai	Kunci Jawaban
<ul style="list-style-type: none"> - Terstruktur, mengacu pada permasalahan rutin dan berulang untuk solusi standar yang ada - Tak terstruktur, adalah 'fuzzy', permasalahan kompleks dimana tak ada solusi serta merta - Semi terstruktur, terdapat beberapa keputusan terstruktur, tetapi tak semuanya dari fase-fase yg ada 	Terstruktur: mengacu kepada permasalahan yang rutin dan berulang untuk solusi standar yang ada. Tak terstruktur: mengacu kepada permasalahan kompleks di mana tidak ada solusi. Semi terstruktur: terdapat beberapa keputusan terstruktur, tetapi tak semuanya dari fase - fase yang ada.

Text pre-processing dilakukan baik pada jawaban esai maupun pada kunci jawaban esai.

Tabel 5. *Pre-processing* Jawaban dan Kunci Jawaban

Jawaban Esai	Kunci Jawaban
struktur acu masalah rutin ulang solusi standar struktur fuzzy masalah kompleks mana solusi merta semi struktur putus struktur fasefase yang	struktur acu masalah rutin ulang solusi standar struktur acu masalah kompleks solusi semi struktur putus struktur fase fase

Tokenisasi pada jawaban dan kunci jawaban dilakukan dengan cara memotong kalimat menjadi potongan substring sehingga potongan substring ini disebut sebagai k-gram. Proses tokenisasi ini dilakukan setelah melalui tahap *text preprocessing*. Untuk nilai k-gram yang digunakan adalah 2 sehingga pada setiap token terdapat dua buah huruf.

Tabel 6. Tokenisasi Jawaban dan Kunci Jawaban

Jawaban Esai	Kunci Jawaban
"st", "tr", "ru", "uk", "kt", "tu", "ur", "ra", "ac", "cu", "um", "ma", "as", "sa", "al", "la", "ah", "hr", "ru", "ut", "ti", "in", "nu", "ul", "la", "an", "ng", "gs", "so", "ol", "lu", "us", "si", "is", "st", "ta", "an", "nd", "da", "ar", "rs", "st", "tr", "ru", "uk", "kt", "tu", "ur", "rf", "fu", "uz", "zz", "zy", "ym", "ma", "as", "sa", "al",	"st", "tr", "ru", "uk", "kt", "tu", "ur", "ra", "ac", "cu", "um", "ma", "as", "sa", "al", "la", "ah", "hr", "ru", "ut", "ti", "in", "nu", "ul", "la", "an", "ng", "gs", "so", "ol", "lu", "us", "si", "is", "st", "ta", "an", "nd", "da", "ar", "rs", "st", "tr", "ru", "uk", "kt", "tu", "ur", "ra", "ac", "cu", "um",

$$= (14 * 14) + (7 * 14) + (3 * 2) + (14 * 14) + (2 * 1) + (1 * 1) + (23 * 14) + (2 * 2) + (4 * 6) + (2 * 2) + (1 * 1) + (6 * 0) + (19 * 6) + (4 * 7) + (3 * 0) + (2 * 7) + (4 * 4) + (2 * 2) + (1 * 1) + (0 * 1) = 1031$$

Menentukan nilai *magnitude* jawaban mahasiswa:

$$= 14^2 + 7^2 + 3^2 + 14^2 + 2^2 + 1^2 + 23^2 + 2^2 + 4^2 + 2^2 + 1^2 + 6^2 + 19^2 + 4^2 + 3^2 + 2^2 + 4^2 + 2^2 + 1^2 + 0^2 = 1456 = \sqrt{1456} = 38,1576.$$

Menentukan nilai *magnitude* kunci jawaban:

$$= 14^2 + 14^2 + 2^2 + 14^2 + 1^2 + 1^2 + 14^2 + 2^2 + 6^2 + 2^2 + 1^2 + 0^2 + 6^2 + 7^2 + 0^2 + 7^2 + 4^2 + 2^2 + 1^2 + 1^2 = 991 = \sqrt{991} = 31,4802$$

Menghitung nilai *similarity*:

$$= \frac{1031}{38,1576 * 31,4802};$$

$$= 0,858304 \text{ atau } 85,83\%.$$

Nilai *similarity* yang didapatkan dengan cara mengalikan nilai *similarity* dengan nilai bobot soal.

$$= 0,858304 \times 5$$

$$= 4,29152$$

nilai perolehan dibulatkan ke bilangan bulat terdekat sehingga nilai yang didapatkan adalah 4.

2. Pembahasan

Pengujian hasil dilakukan dengan cara membandingkan nilai yang diberikan dari dosen dengan nilai hasil prediksi dari algoritma menggunakan metode *Mean Absolute Percentage Error*. Pengujian ini dilakukan pada soal esai UTS mata kuliah *Desicion Support System* semester lima tahun akademik 2023/2024. Soal esai ini terdiri dari lima butir soal yang masing – masing butir soal memiliki jawaban sebanyak 17 jawaban dari mahasiswa. Untuk nilai bobot yang diberikan pada keseluruhan soal ini sebesar 25 di mana masing – masing butir soal diberi nilai bobot sebesar 5. Selanjutnya menghitung selisih antara nilai prediksi algoritma dengan nilai perolehan UTS yang diberikan oleh dosen.

Tabel 10. Perbandingan Nilai Prediksi Algoritma dengan Nilai Rekapitulasi UTS

No	NPM	Nama	Nilai Prediksi Algoritma	Nilai Rekapitulasi UTS
1	14200010	Rudy Sephiannudin	19	19
2	14200033	Diego Fauzi Nurcintya	18	15
3	14210001	Achmad Maulana	22	20
4	14210002	Sefira Salsabila	22	20
5	14210006	Dimas Permana Putra	19	20
6	14210007	Faris Safa Ramdani	22	19
7	14210009	Fitrinur Indriyana Salsabila	19	15
8	14210012	Meisa Rahma Siti Nurbani	24	20
9	14210014	Muhammad Reihan	21	20
10	14210015	Muhammad Rizky Pratama	18	16
11	14210016	Nadhira Faza Fatiha	21	19
12	14210017	Nisrina Salsabila	22	20
13	14210022	Yunisa Ahgnia Zahrah	19	20
14	14210024	Faisal Rifky Nugraha	20	20
15	14210025	Zacky Nur Ramadhan Bachtiar	18	19
16	14210028	Suparwan	22	19
17	14210030	Rindiyanti	15	18

Perhitungan nilai *Mean Absolute Percentage Error* didapatkan dengan menggunakan rumus:

$$= \frac{1}{N} \sum_{i=1}^n \left| \frac{(x_i - y_i)}{x_i} \right| \times 100\%;$$

Di mana nilai X_i adalah nilai sesungguhnya dan nilai Y_i adalah nilai prediksi, sehingga didapatkan:

$$\begin{aligned} &= \left| \frac{19 - 19}{19} \right| + \left| \frac{15 - 18}{15} \right| + \left| \frac{20 - 22}{20} \right| + \left| \frac{20 - 22}{20} \right| + \left| \frac{20 - 19}{20} \right| + \left| \frac{19 - 22}{19} \right| + \left| \frac{15 - 19}{15} \right| + \\ &\left| \frac{20 - 24}{20} \right| + \left| \frac{20 - 21}{20} \right| + \left| \frac{16 - 18}{20} \right| + \left| \frac{19 - 21}{20} \right| + \left| \frac{20 - 22}{20} \right| + \left| \frac{20 - 19}{20} \right| + \left| \frac{20 - 20}{20} \right| + \\ &\left| \frac{19 - 18}{19} \right| + \left| \frac{19 - 22}{19} \right| + \left| \frac{18 - 15}{18} \right| \\ &= 0,11 * 100\% = 11\% \end{aligned}$$

Berdasarkan penelitian yang dilakukan oleh (Sumari dkk., 2020) interpretasi dari nilai *Mean Absolute Percentage Error* tersebut menandakan bahwa kemampuan model dalam memberikan nilai masuk ke dalam kategori **Baik**.

D. KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan, dapat diuraikan beberapa kesimpulan, antara lain:

1. Mendapatkan hasil nilai perolehan jawaban esai secara tepat dengan menggunakan algoritma *Winnowing* dan *Cosine Similarity*.
2. Mendapatkan hasil nilai perolehan jawaban esai berdasarkan tingkat kemiripan *fingerprint* kata antara jawaban esai dengan kunci jawaban.
3. *Prototype* aplikasi penilaian esai berhasil dikembangkan untuk mempermudah dosen dalam proses pemberian nilai esai.
4. Hasil nilai *Mean Absolute Percentage Error* pada algoritma *Winnowing* dan algoritma *Cosine Similarity* dalam memberikan nilai saat dibandingkan selisihnya dengan nilai dosen sebesar 11% yakni ketepatan algoritma dalam memberikan nilai masuk ke dalam kategori **Baik**, kemudian hasil kuesioner kelayakan sistem dari pihak pengguna secara keseluruhan sebesar 83%.

E. DAFTAR PUSTAKA

- [1] Kurniadi, D., Gernowo, R., Surarso, B., Wibowo, A., & Warsito, B. (2023). Sistem Penilaian Jawaban Singkat Otomatis pada Ujian Online Berbasis Komputer Menggunakan Algoritma *Cosine Similarity*. *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, 9(2), 316.
- [2] Kurniawati, F. E., & Pradnya, W. M. (2020). Implementasi Algoritma *Winnowing* Pada Sistem Penilaian Otomatis Jawaban Esai Pada Ujian Online Berbasis Web. *Jurnal Teknik Komputer AMIK BSI*, VI(2). Diambil dari <https://ejournal.bsi.ac.id/ejournal/index.php/jtk/article/view/7838>
- [3] Lubis, M. R. S., Khairuna, K., & Rohani, R. (2023). Pengembangan Lembar Kerja Peserta Didik (LKPD) Berbasis Kemandirian Belajar bagi Siswa di SMA Cerdas Murni Tembung. *Pekobis : Jurnal Pendidikan, Ekonomi, dan Bisnis*, 8(1), 81. Diambil dari <http://openjournal.unpam.ac.id/index.php/Pekobis/article/view/29819>
- [4] Manning, C. D., Raghavan, P., & Schütze, H. (2009). *Introduction to Information Retrieval*. Cambridge: Cambridge University Press.
- [5] Purwaningrum, S., Susanto, A., & Kristiningsih, A. (2023). Pengaruh *Synonym Recognition* dalam Deteksi Kemiripan Teks Menggunakan *Winnowing* dan *Cosine Similarity*. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, 12(3), 219–226. Diambil dari <https://jurnal.ugm.ac.id/v3/JNTETI/article/view/6375>
- [6] Schleimer, S., Wilkerson, D. S., & Aiken, A. (2003). *Winnowing*. *Proceedings of the 2003 ACM SIGMOD international conference on Management of data* (hlm. 76–85). New York, NY, USA: ACM. Diambil dari <https://dl.acm.org/doi/10.1145/872757.872770>
- [7] Sumari, A. D. W., Musthafa, M. B., Ngatmari, & Putra, D. R. H. (2020). Comparative Performance of Prediction Methods for Digital Wallet Transactions in the Pandemic Period. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 4(4), 642–647. Diambil dari <https://jurnal.iaii.or.id/index.php/RESTI/article/view/2024>
- [8] Widiyanto, J. (2018). *EVALUASI PEMBELAJARAN (Sesuai dengan Kurikulum 2013)*. Madiun: UNIPMA Press. Diambil dari www.kwu.unipma.ac.id