

ALGORITMA *LATENT SEMANTIC ANALYSIS* DAN *COSINE SIMILARITY* DALAM PENILAIAN ESSAY OTOMATIS

Prawira Adhi Jatmiko; Husni Thamrin, M.T., Ph.D

**Program Studi Teknik Informatika, Fakultas Komunikasi dan Informatika,
Universitas Muhammadiyah Surakarta**

Abstrak

penilaian essay merupakan salah satu metode penilaian yang dilakukan untuk memenuhi nilai akademik siswa di lingkungan pembelajaran. Penilaian essay sering dipilih karena lebih mengukur sejauh mana siswa memahami dan menjelaskan permasalahan yang ada di dalam tes tersebut sesuai dengan materi yang telah diajarkan. penilaian essay dilakukan dengan siswa menjawab pertanyaan soal berupa teks uraian singkat menggunakan bahasa mereka sendiri. Oleh karena itu diperlukan algoritma untuk membantu menghemat waktu pengajar dalam melakukan koreksi. Tujuan dari penelitian ini adalah mengevaluasi dan menganalisis dua buah algoritma yang telah dipilih yaitu Latent Semantic Analysis dan Cosine Similarity dengan menggunakan perbandingan RMSE (Root Mean Square Error) untuk mengetahui algoritma mana yang cocok untuk diterapkan dalam penilaian esai otomatis. Penelitian ini menggunakan sebanyak 200 buah jawaban esai yang telah dikoreksi oleh pengajar. Penelitian ini menunjukkan bahwa LSA cenderung mengikuti pola pada penilaian manusia, dan Cosine Similarity memiliki nilai RMSE lebih besar dari algoritma LSA.

Kata Kunci : penilaian essay, *latent semantic analysis*, *Cosine Similarity*, *RMSE*

Abstracts

Essay assessment is one of the assessment methods carried out to meet students' academic values in the learning environment. Essay assessment is often chosen because it measures the extent to which students understand and explain the problems in the test according to the material that has been taught. Essay assessment is carried out by students answering questions in the form of short descriptive text using their own language. Therefore, an algorithm is needed to help save teachers' time in making corrections. The aim of this research is to evaluate and analyze the two selected algorithms, namely Latent Semantic Analysis and Cosine Similarity using RMSE (Root Mean Square Error) comparisons to find out which algorithm is suitable to be applied in automatic essay assessment. This research used 200 essay answers that had been corrected by the lecturers. This research shows that LSA tends to follow human scoring patterns, and Cosine Similarity has a greater RMSE score than the LSA algorithm

Keywords : essay assessment, latent semantic analysis, Cosine Similarity, RMSE

1. PENDAHULUAN

Penilaian (assesment) adalah salah satu aspek penting dalam sebuah proses pembelajaran (Mirda et al., 2022). Penilaian dilakukan untuk mencapai standar akademik di lingkungan pembelajaran. Terdapat beberapa metode untuk melakukan penilaian, salah satunya adalah tes

tertulis. Tes tertulis bisa berupa pilihan ganda, isian singkat, atau uraian (essay). Di tingkat Pendidikan Tinggi, sebagian besar penilaian dilakukan melalui tes tertulis berbentuk uraian (essay). Tes uraian sering dipilih oleh pengajar karena lebih mampu mengukur sejauh mana mahasiswa memahami dan menjawab permasalahan yang ada dalam tes tersebut sesuai dengan materi yang telah diajarkan (Rahutomo et al., 2020)

Pada tes uraian (*essay*), pengajar memberikan pertanyaan secara tertulis terkait dengan materi pembelajaran, lalu siswa menjawab secara tertulis berupa uraian singkat dengan bahasa mereka sendiri (Sihombing, 2022). Jawaban esai tersebut akan memberikan hasil bukan hanya benar dan salah, akan tetapi bisa menghasilkan nilai yang hampir benar, atau setengah benar. Setelah mahasiswa memberikan jawaban, pengajar akan melakukan koreksi soal secara manual yang membutuhkan ketelitian yang maksimal sehingga memakan waktu yang lama (Mufiid et al., 2021).

Waktu yang diperlukan untuk mengoreksi jawaban esai merupakan salah satu permasalahan karena setiap jawaban harus dibaca sebelum diberi nilai. Selain itu terdapat permasalahan lain seperti bias dan subjektivitas ketika melakukan koreksi (Bahri et al., 2020). Oleh karena itu, telah dilakukan penelitian penilaian esai secara otomatis. Dengan penilaian otomatis, waktu yang diperlukan untuk mengoreksi menjadi cepat. Selain itu, penilaian dengan mesin tidak mengalami bias dan subjektivitas.

Beberapa penelitian telah metode Latent Semantic Analysis (LSA) dan Cosine Similarity. Penelitian dari Amalia dkk. menganalisis penggunaan algoritma Cosine Similarity untuk diterapkan pada aplikasi essay otomatis yang dibuat dengan beberapa tahap pengujian. Kesimpulan yang mereka dapat dalah dengan menggunakan metode Cosine Similarity didapat kemiripan hasil penilaian dengan penilaian manusia. Penelitian menghasilkan nilai rata-rata *precision* 93%, *recall* 86%, dan *f-measure* 89% (Amalia et al., 2021)

Penelitian Arifuddin dkk. juga membuat sistem penilaian esai otomatis menggunakan metode cosine similarity. Penelitian tersebut bertujuan untuk meminimalisir terjadinya human error dikarenakan penilaian esai yang dilakukan secara manual. Penelitian tersebut berhasil membuat sistem yang dapat secara langsung menilai jawaban yang telah dimasukkan kedalam sistem tersebut secara cepat dan penggunaan algoritma cosine similarity terbukti akurat (Arifuddin et al., 2023)

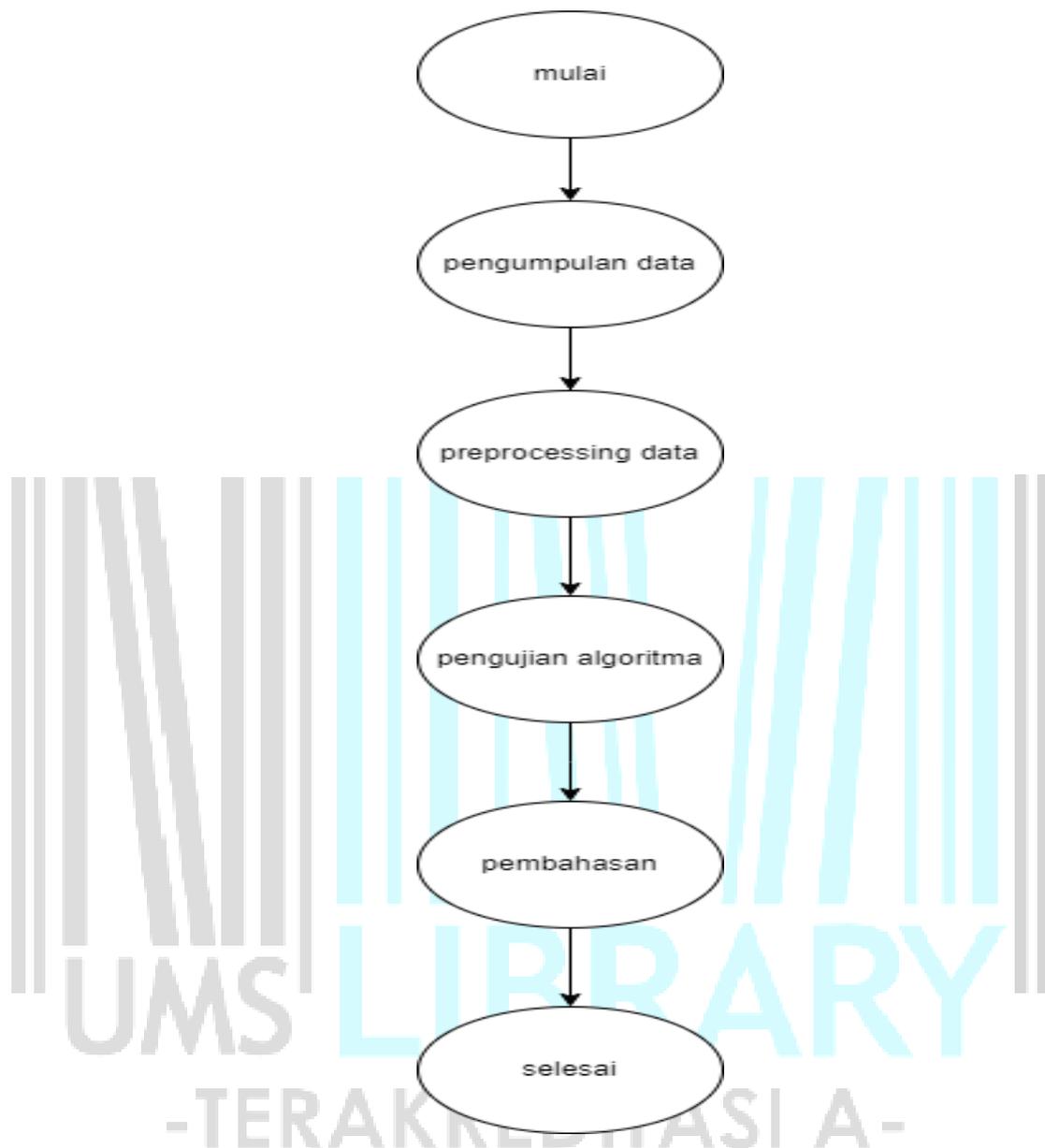
Penelitian lain tentang penilaian esai otomatis dilakukan oleh Bharata dkk. menggunakan metode yang berbeda yaitu Support Vector Machine (SVM) dan Latent Semantic Analysis (LSA). Tujuan penelitian yaitu untuk memudahkan dalam melakukan koreksi esai secara cepat

dan akurat. Dengan menggunakan kedua metode tersebut didapatkan nilai akurasi yang cukup tinggi dibandingkan dengan nilai manual yaitu 96,23 persen (Bharata, 2020)

Penelitian yang dilakukan oleh Susanto dkk, menggunakan 5 algoritma sekaligus untuk mengetahui kedekatan antara jawaban siswa dengan jawaban kunci. Dengan membandingkan 5 algoritma tersebut yang salah satunya adalah LSA. Pada penelitian tersebut, LSA memiliki pola yang lebih dekat dengan penilaian yang dilakukan oleh manusia, dan LSA memiliki tingkat RMSE yang cukup tinggi (Susanto et al., 2023)

Tulisan ini membahas perbandingan kinerja metode LSA dan Cosine Similarity ketika digunakan dalam penilaian esai otomatis. LSA dan Cosine Similarity dipilih karena mudah dan cepat untuk diterapkan didalam penilaian esai otomatis. Jika dibandingkan dengan algoritma lain seperti Support Vector Machine (SVM), meskipun algoritma tersebut lebih kuat dalam klasifikasi, algoritma tersebut tidak terlalu praktis dalam hal penilaian esai otomatis karena algoritma SVM membutuhkan data yang berlabel, sedangkan LSA dan Cosine Similarity adalah algoritma yang dapat bekerja secara *unsupervised* atau tidak memerlukan data berlabel. Kebaruan penelitian ini adalah pada penggunaan dataset yang diambil dari ujian semester pada sebuah matakuliah di Program Studi Ilmu Komunikasi. Hasil penelitian ini mendukung temuan sebelumnya bahwa metode LSA lebih akurat dibanding metode Cosine Similarity.

2. METODE



Gambar 1. metode penelitian

Gambar 1 merupakan representasi dari tahapan penelitian mengenai perbandingan metode Cosine Similarity dan Latent Semantic Analysis (LSA) untuk mengetahui RMSE dan keakuratan dalam penilaian esai otomatis. Berikut adalah uraian dari tahapan penelitian.

2.1 Pengumpulan data

Data yang digunakan adalah jawaban dan kunci jawaban UTS mahasiswa Ilmu Komunikasi Universitas Muhammadiyah Surakarta. Proses pengumpulan data dimulai dengan mengumpulkan kunci jawaban dan jawaban ujian mahasiswa. Setiap jawaban ujian mahasiswa diberikan nilai atau rating oleh manusia atau konteksnya disini adalah penguji, dalam rentang 0 sampai 25 yang berarti setiap jawaban yang tidak benar diberikan nilai 0, dan jawaban yang

benar diberikan nilai 25. Sedangkan jawaban yang mendekati benar diberikan nilai diantara kedua nilai tersebut. Sejumlah 200 data dikumpulkan dan dilakukan perubahan dari data mentah berbentuk kertas jawaban menjadi file *csv* menggunakan fitur eja aplikasi *Google Document*.

2.2 Data preprocessing

Setelah mendapatkan data yang cukup, maka dilakukan data *preprocessing*. Data *preprocessing* adalah serangkaian langkah yang diterapkan pada dokumen teks sebelum dilakukan perhitungan kemiripan. Tahapan ini bertujuan untuk membersihkan teks sehingga hanya kata-kata yang penting yang digunakan, sementara kata-kata yang kurang relevan tidak dimasukkan ke dalam proses berikutnya (Arfandy & Alwiah Musdar, 2020). Penulis melakukan *reprocessing*, yaitu Tokenisasi. Tokenisasi atau *Tokenization* adalah sebuah metode *preprocessing* data untuk memisahkan teks yang berasal dari kalimat atau paragraf untuk menjadi bagian-bagian tertentu yang dapat disebut juga dengan token. Pada konteks ini, Tokenisasi adalah memisahkan kalimat atau paragraf tertentu menjadi kata – kata yang lebih kecil sehingga dapat dianalisis dengan lebih baik (Rahma Firmansyah et al., 2024). Setelah *preprocessing* data dilakukan, data tersebut sudah siap untuk diolah menggunakan dua metode algoritma, yaitu Cosine Similarity dan LSA.

2.3 Pengujian algoritma

2.3.1 Cosine Similarity

Cosine Similarity adalah ukuran kesamaan antara dua vektor dalam ruang dimensi yang dihitung dari nilai kosinus sudut hasil perkalian kedua vektor tersebut. Karena kosinus dari sudut 0° bernilai 1 dan kurang dari 1 untuk sudut lainnya, maka dua vektor dianggap mirip jika nilai cosine similarity-nya adalah 1 (Amalia et al., 2021)

Cosine similarity digunakan dalam ruang positif, dengan hasil yang dibatasi antara 0 dan 1. Jika nilainya 0, maka dokumen tersebut dianggap mirip, sedangkan jika hasilnya 1, dokumen tersebut dianggap tidak mirip. Batasan ini berlaku untuk berbagai dimensi, dan cosine similarity paling sering diterapkan dalam ruang positif berdimensi tinggi. Misalnya, dalam klasifikasi teks, setiap kata atau istilah dianggap sebagai dimensi yang berbeda, dan dokumen direpresentasikan sebagai vektor di mana nilai setiap dimensi sesuai dengan frekuensi kemunculan istilah dalam dokumen (Park et al., 2020). Berikut adalah rumus dari Cosine Similarity :

$$Similarity = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i x B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \quad (1)$$

Keterangan : A = vektor

B = vektor

A_i = bobot term i dalam blok A_i

B_i = bobot term i dalam blok B_i

I = jumlah term dalam kalimat

n = jumlah vektor

Dimana A merupakan bobot setiap ciri pada vektor A, dan B merupakan bobot setiap ciri B, maka A merupakan bobot setiap istilah pada dokumen A, dan B adalah bobot setiap istilah pada dokumen B. Pengukuran kemiripan dilakukan dengan membandingkan dua buah dokumen yaitu dokumen 1 dengan dokumen 2 kemudian sistem akan menghitung kemiripannya (Dewa Ayu Rai Ariantini, 2016)

2.3.2 Latent Semantic Analysis (LSA)

Latent Semantic Analysis merupakan sebuah metode yang memanfaatkan model statistik matematis untuk menganalisis struktur semantik suatu teks. LSA dapat digunakan untuk menilai esai dengan mengonversi esai menjadi matriks-matriks yang diberi bobot pada setiap term untuk dicari kesamaan dengan term referensi. Dalam perhitungannya, LSA menggunakan Singular Value Decomposition (SVD). SVD merepresentasikan ruang semantik ke dalam bentuk matriks dengan ordo yang lebih kecil dibandingkan matriks aslinya, namun perhitungan matriks tetap menghasilkan nilai yang hampir sama (Syahfrialief & Muh Aliyazid Mude, 2023). Sedangkan Singular Value Decomposition (SVD) adalah salah satu teknik reduksi dimensi yang bermanfaat untuk memperkecil nilai kompleksitas dalam pemrosesan term-document matrix. (Setyaningsih et al., 2020)

Pada SVD, sebuah matriks persegi panjang diuraikan menjadi perkalian tiga matriks berbeda. Matriks pertama mewakili entitas baris dari matriks asli sebagai vektor dari nilai faktor ortogonal. Matriks kedua juga mewakili entitas kolom dari matriks asli dengan cara yang serupa. Matriks ketiga adalah matriks diagonal yang mengandung nilai skala, sehingga ketika ketiga matriks ini dikalikan, matriks asli akan terbentuk.

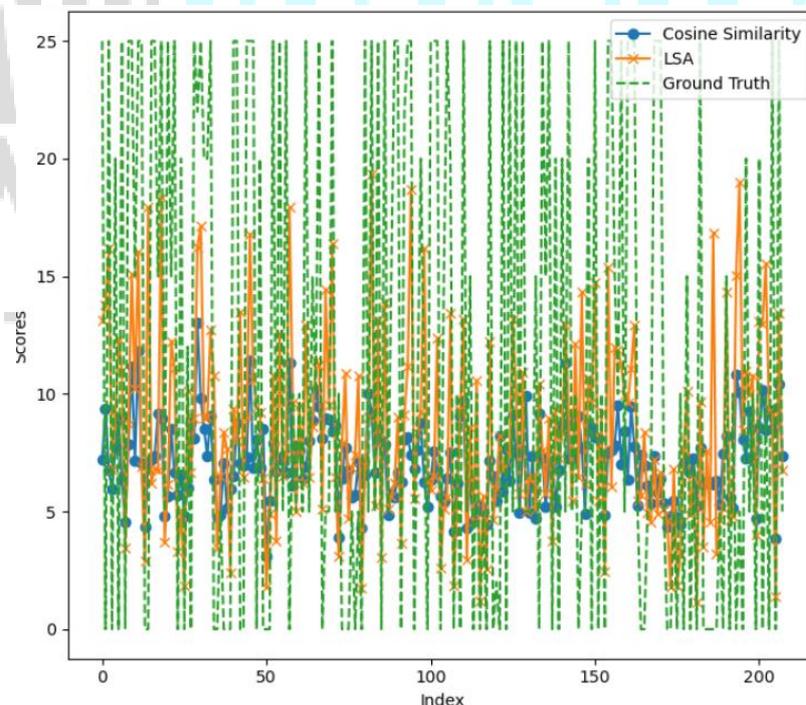
$$A_{m \times n} = U_{m \times m} \times S_{m \times n} \times V_{n \times n}^T \quad (2)$$

Pada persamaan diatas, A adalah matriks asli dengan m baris dan n kolom. Matriks U dan V adalah matriks singular kanan dan kiri, sementara S adalah matriks diagonal. Dalam proses

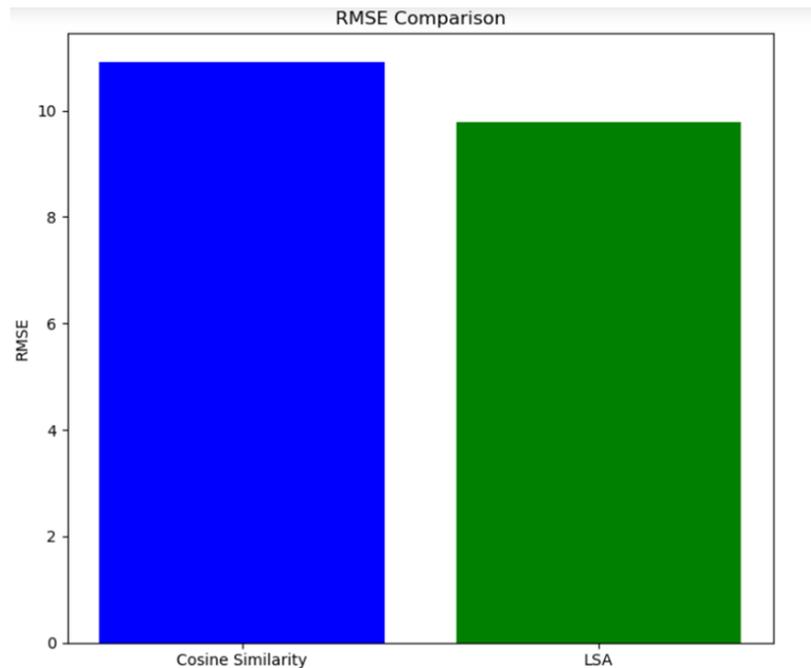
LSA, vektor kata dihasilkan, di mana setiap teks jawaban dan kunci dapat direpresentasikan menggunakan vektor tersebut (Susanto et al., 2023).

3. HASIL DAN PEMBAHASAN

Penelitian ini berfokus pada RMSE (Root Mean Square Error) dari kedua metode tersebut. RMSE adalah nilai akar kuadrat dari mean square error (MSE) dimana MSE adalah nilai rata rata dari jumlah error kuadrat (Pradani et al., 2023). Perhitungan RMSE dilakukan untuk mengetahui ukuran kuantitatif dari kesalahan prediksi yang dapat membantu dalam memahami seberapa baik model algoritma memprediksi hasil sebenarnya. RMSE dipilih karena lebih sensitive terhadap outlier jika dibandingkan dengan MAE (Mean Absolute Error) yang hanya menghitung outlier secara linier atau tidak dikuadratkan sehingga outlier tidak berpengaruh secara signifikan. RMSE dilihat dari perbedaan nilai atau skor antara perhitungan algoritma dengan perhitungan manusia, dalam hal ini setiap jawaban siswa dan jawaban kunci dihitung dengan menggunakan algoritma, setelah didapatkan hasil prediksinya, kemudian hasil prediksi tersebut dan nilai sebenarnya dihitung dengan dikuadratkan, dijumlah lalu diambil akarnya sehingga ditemukan RMSE-nya. Lebih lanjut hasil dari penelitian ini ditunjukkan dengan grafik dibawah ini :



Gambar 2. Grafik prediksi skor antara Cosine Similarity dengan LSA



Gambar 3. Grafik perbandingan RMSE Cosine Similarity dengan LSA

Gambar 2 merupakan grafik hasil prediksi skor antara Cosine Similarity dengan LSA terhadap *Ground Truth* atau nilai sebenarnya. Setiap kalimat pada jawaban mahasiswa dihitung oleh algoritma dan dibandingkan dengan *ground truth* atau nilai sebenarnya. Konteks nilai sebenarnya disini adalah nilai yang diberikan oleh manusia. Terlihat pada garis orange menunjukkan algoritma LSA, garis biru menunjukkan Cosine Similarity, dan garis hijau menunjukkan *ground truth*. Algoritma LSA memiliki rentang nilai yang fluktuatif dimulai dari 13,19 bahkan turun hingga ke 1, yang mana grafik LSA cenderung lebih mengikuti pola *ground truth* walaupun tidak terlalu mendekati *ground truth* yang memiliki rentang nilai antara 0 sampai 25. Sedangkan algoritma Cosine Similarity, menunjukkan grafik yang secara konstan di antara 5 sampai 10. Dapat diartikan bahwa algoritma cosine kurang dapat mengikuti pola pada *ground truth*. Sejalan dengan penelitian yang dilakukan oleh (Susanto et al., 2023) yang menyatakan bahwa LSA lebih dapat mengikuti pola penilaian manusia daripada algoritma Cosine

Gambar 3 adalah grafik perbandingan RMSE antara cosine similarity dengan LSA. grafik tersebut ditunjukkan dengan algoritma Cosine berwarna biru dan LSA berwarna hijau. Grafik tersebut menunjukkan algoritma Cosine berada pada angka diatas 10, atau lebih tepatnya 10.9022. sedangkan pada grafik LSA berada pada angka 9, atau lebih tepatnya 9.7869. Maka dapat disimpulkan bahwa nilai RMSE algoritma LSA lebih kecil dari pada Cosine Similarity.

4. PENUTUP

Hasil dari penelitian ini menunjukkan bahwa algoritma LSA cenderung mengikuti pola penilaian pada manusia dan memiliki nilai RMSE yaitu 9.7865, lebih rendah daripada Cosine Similarity yang memiliki nilai RMSE yaitu 10.9022. Dapat disimpulkan bahwa algoritma LSA cocok digunakan dalam penilaian esai otomatis dikarenakan nilai RMSE yang rendah sehingga cukup akurat. Saran untuk penelitian selanjutnya adalah mencoba untuk menggunakan algoritma yang lain seperti algoritma Jaccard, TF-IDF, ataupun algoritma SVM untuk mendapatkan hasil yang berbeda dan mungkin akan lebih maksimal

DAFTAR PUSTAKA

- Amalia¹, E. L., Jumadi, A. J., Mashudi³, I. A., Wibowo⁴, W., Malang, P. N., & Korespondensi, P. (2021). Analisis Metode Cosine Similarity Pada Aplikasi Ujian Online Esai Otomatis (Studi Kasus Jti Polinema) Cosine Similarity Method Analysis On Automatic Esai Online Test Application. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIK)*, 8, 343–348. <https://doi.org/10.25126/jtiik.202184356>
- Arfandy, H., & Alwiah Musdar, I. (2020). Rancang Bangun Sistem Cerdas Pemberian Nilai Otomatis Untuk Ujian Essai Menggunakan Algoritma Cosine Similarity. <https://doi.org/http://dx.doi.org/10.35585/inspir.v10i2.2580>
- Arifuddin, M. R., Rafiq, I. A., Mubarak, R., & Susilo, P. H. (2023). Sistem Cerdas Penilaian Ujian Essay Menggunakan Metode Cosine Similarity. In *Generation Journal* (Vol. 7, Issue 1). <https://doi.org/https://doi.org/10.29407/gj.v7i1.18318>
- Bahri, S., Wajhillah, R., Nusamandiri, S., Damai No, J., Jati Barat, W., & Minggu Jakarta Selatan, P. (2020). Optimalisasi Algoritma Rabin Karp Menggunakan Tf-Idf Dalam Pencocokan Text Pada Penilaian Ujian Essay Otomatis. 4(2). <https://doi.org/10.30743/infotekjar.v4i2.2294>
- Dewa Ayu Rai Ariantini, A. S. M. L. A. Jacobus. (2016). Pengukuran Kemiripan Dokumen Teks Bahasa Indonesia Menggunakan Metode Cosine Similarity. *Jurnal Teknik Informatika*, 2016 *Ejournal.Unsrat.Ac.Id*, vol.9 No.1. <https://doi.org/https://doi.org/10.35793/jti.9.1.2016.13752>
- Hudi Kusuma Bharata, H. S. (2020). Optimasi Sistem Penilaian Ujian Essay Online Menggunakan Support Vector Machine (Svm) Dan Latent Semantic Analysis (Lsa) Dengan Bahasa R. *Scientific Journal of Information Systems Technology and Applied Computer Engineering*, 10, 63–70. jurnal.stmik.banisaleh.ac.id
- Mirda, M., Sarjan, M., & Khairat, U. (2022). Aplikasi Ujian Essay Koreksi Otomatis Menggunakan Metode Cosine Similarity. *Journal Pegguruang: Conference Series*, 4(1), 418. <https://doi.org/10.35329/jp.v4i1.2344>
- Mufiid, I., Lestanti, S., & Kholila, ma. (2021). Aplikasi Penilaian Jawaban Esai Otomatis Menggunakan Metode Synonym Recognition Dan Cosine Similarity Berbasis Web. In *Jurnal MNEMONIC* (Vol. 4, Issue 2).
- Park, K., Hong, J. S., & Kim, W. (2020). A Methodology Combining Cosine Similarity With Classifier For Text Classification. *Applied Artificial Intelligence*, 34(5), 396–411. <https://doi.org/10.1080/08839514.2020.1723868>
- Pradani, K. L. S. (2023). Automated Essay Scoring Menggunakan Semantic Textual Similarity Berbasis Transformer Untuk Penilaian Ujian Esai. *Jtiik.Ub.Ac.Id*, 10, 1177–1185. <https://doi.org/10.25126/jtiik.2023107338>

- Rahma Firmansyah, D., Lestariningsih, E., & Homepage, J. (2024). Analisis Sentimen Ulasan Aplikasi Smart Campus Unisbank Di Google Playstore Menggunakan Algoritma Naive Bayes Jurnal Jtik (Jurnal Teknologi Informasi Dan Komunikasi). *Jurnal Teknologi Informasi Dan Komunikasi*, 8(2), 2024. <https://doi.org/10.35870/jti>
- Rahutomo, F., Novita Sari, D., Studi Teknik Informatika, P., Teknologi Informasi, J., & Negeri Malang, P. (2020). Implementasi Library Deep Learning Keras Pada Sistem Ujian Essay Online. *Jurnal.Polinema.Ac.IdF Rahutomo, DN SariJurnal Informatika Polinema*, 2020•*jurnal.Polinema.Ac.Id.* <http://jurnal.polinema.ac.id/index.php/jip/article/view/2559>
- Setyaningsih, L., Setyati, E., Tjandra, S., Informasi, T., Sains, I., Teknologi, D., Surabaya, T., Ngagel, J., Tengah, J., & 73-77, N. (2020). Klasifikasi Soal Pilihan Ganda Berbahasa Indonesia Berdasarkan Level Kognitif Puspendik Dengan Support Vector Machine. *Jurnal Insand Comtech*, 6(2). <https://dreamlikechild.weebly.com/uploads/3/9/0>
- Sihombing, D. O. (2022). Implementasi Natural Language Processing (Nlp) Dan Algoritma Cosine Similarity Dalam Penilaian Ujian Esai Otomatis. *Jurnal Sistem Komputer Dan Informatika (JSON)*, 4(2), 396. <https://doi.org/10.30865/json.v4i2.5374>
- Susanto, M. R. R., Husni Thamrin, & Naufal Azmi Verdikha. (2023). Performance Of Text Similarity Algorithms For Essay Answer Scoring In Online Examinations. *Jurnal Teknik Informatika (Jutif)*, 4(6), 1515–1521. <https://doi.org/10.52436/1.jutif.2023.4.6.1025>
- Syahfrialief, I., & Muh Aliyazid Mude, dan. (2023). Impelementasi Metode Latent Semantic Analysis Pada Sistem Informasi Artikel Abstrak. 4(2), 106–111.

UMS

LIBRARY

-TERAKREDITASI A-