

Penerapan Metode Rubrik Berbasis Algoritma Natural Language Processing dalam Penilaian Pembelajaran Siswa

Muhamad Mahda Vickia^{1*}, Hidola Syamsito²

Program Studi Teknik Informatika, Fakultas Informatika dan Komputer, Universitas Binaniaga Indonesia
email: Muhammadmahda02@gmail.Com

*Corresponding Author

ABSTRACT

The study titled "Implementation of Rubric-Based NLP Algorithm in Student Learning Assessment" aims to develop an automatic assessment system based on Natural Language Processing (NLP) algorithms, utilizing a rubric method to assist teachers in evaluating students' essay answers objectively, consistently, and efficiently. One of the main challenges teachers face is providing fair assessments that align with rubric criteria, especially when dealing with a large number of responses with high variability. This research adopts a prototype-based web application development approach, where the NLP algorithm is designed to automatically match students' answers with rubric criteria. Data for this study was collected through field observations and system trials, involving teachers and students from SMK Kosgoro as research subjects. The results indicate that this system improves assessment efficiency with a high level of accuracy in matching answers to predefined rubrics. The system also helps teachers provide faster and more relevant feedback, enabling students to better understand their level of mastery of the subject matter. Based on the conducted evaluations, the system achieved an expert validation score of 100%, a user evaluation score of 87.27%, and an overall performance score of 86.75%, demonstrating that the application of rubric-based NLP algorithms has a high level of effectiveness. This study proves that the implementation of rubric-based NLP algorithms can significantly enhance the overall quality of learning.

Keywords: natural language processing, rubric method, automatic assessment.

ABSTRAK

Penelitian ini berjudul "Penerapan Metode Rubrik Berbasis Algoritma NLP dalam Penilaian Pembelajaran Siswa", bertujuan untuk mengembangkan sistem penilaian otomatis berbasis algoritma Natural Language Processing (NLP) yang menggunakan metode rubrik untuk membantu guru menilai jawaban esai siswa secara objektif, konsisten, dan efisien. Tantangan utama yang dihadapi guru adalah memberikan penilaian yang adil dan sesuai dengan bobot kriteria rubrik, terutama saat menghadapi banyak jawaban dengan keragaman tinggi. Penelitian ini mengadopsi metode pengembangan prototipe aplikasi berbasis web, di mana algoritma NLP dirancang untuk mencocokkan jawaban siswa dengan kriteria rubrik secara otomatis. Data penelitian dikumpulkan melalui observasi lapangan dan uji coba sistem, melibatkan guru dan siswa dari SMK Kosgoro sebagai subjek penelitian. Hasilnya menunjukkan bahwa sistem ini mampu meningkatkan efisiensi penilaian dengan tingkat akurasi tinggi dalam mencocokkan jawaban dengan rubrik yang telah ditentukan. Sistem ini juga membantu guru memberikan umpan balik yang lebih cepat dan relevan, sehingga siswa dapat memahami tingkat penguasaan mereka terhadap materi pelajaran dengan lebih baik. Berdasarkan evaluasi yang dilakukan, sistem ini memperoleh hasil uji ahli sebesar 100%, uji hasil pengguna sebesar 87,27%, dan uji hasil keseluruhan sebesar 86,75%, yang menunjukkan bahwa penerapan metode rubrik berbasis algoritma NLP memiliki tingkat efektivitas yang tinggi. penelitian ini membuktikan bahwa penerapan metode rubrik berbasis algoritma NLP dapat meningkatkan kualitas pembelajaran secara keseluruhan.

Kata Kunci: natural language processing, metode rubrik, penilaian otomatis

1. PENDAHULUAN

Penilaian merupakan bagian penting dalam proses pembelajaran, karena membantu guru dalam mengukur tingkat pemahaman siswa terhadap materi yang diajarkan. Namun, penilaian terhadap jawaban esai sering kali menjadi tantangan bagi guru, terutama dalam memastikan bahwa penilaian dilakukan secara objektif dan konsisten. Metode penilaian tradisional yang mengandalkan koreksi manual sering kali memakan waktu lama dan rentan terhadap subjektivitas, terutama ketika jumlah siswa yang dinilai cukup banyak. Hambatan ini dapat mengurangi efektivitas proses pembelajaran dan memberikan beban kerja tambahan bagi guru.

Dalam beberapa tahun terakhir, perkembangan teknologi Natural Language Processing (NLP) telah memungkinkan penerapan sistem penilaian otomatis yang dapat membantu guru dalam menilai jawaban esai siswa secara lebih efisien. Salah satu pendekatan yang dapat digunakan adalah metode rubrik, yang memungkinkan penilaian dilakukan berdasarkan kriteria yang telah ditentukan sebelumnya. Dengan menerapkan algoritma NLP, sistem dapat mencocokkan jawaban siswa dengan kriteria rubrik secara otomatis, sehingga menghasilkan penilaian yang lebih objektif, konsisten, dan efisien.

Salah satu tantangan dalam dunia pendidikan adalah memastikan bahwa setiap siswa mendapatkan umpan balik yang cepat dan relevan terhadap jawaban mereka. Penilaian manual yang memakan waktu lama dapat memperlambat proses pembelajaran dan menghambat siswa dalam memahami tingkat penguasaan mereka terhadap materi. Oleh karena itu, pengembangan sistem penilaian otomatis berbasis metode rubrik dan NLP menjadi solusi yang potensial untuk meningkatkan efektivitas pembelajaran. Dengan sistem ini, guru dapat menghemat waktu dalam proses penilaian, sementara siswa dapat menerima umpan balik yang lebih cepat dan akurat.

Selain itu, penelitian ini juga bertujuan untuk mengukur efektivitas sistem penilaian otomatis ini dalam meningkatkan akurasi dan efisiensi penilaian jawaban esai siswa. Dengan melibatkan guru dan siswa sebagai subjek penelitian, diharapkan sistem ini dapat menjadi inovasi yang bermanfaat dalam mendukung pembelajaran di sekolah, terutama dalam konteks evaluasi hasil belajar siswa.

2. BAHAN DAN METODE

a. Smart system

“Smart system adalah sistem berbasis teknologi yang mampu mengambil keputusan secara otomatis berdasarkan analisis data yang dikumpulkan, serta dapat beradaptasi terhadap perubahan lingkungan tanpa campur tangan manusia yang signifikan.” McKinsey & Company (2018)

b. Natural Language Processing

Menurut Lee (2023, p.10), NLP adalah cabang AI yang memungkinkan komputer menganalisis dan menafsirkan bahasa manusia untuk berbagai tugas seperti ekstraksi informasi, analisis sentimen, serta terjemahan mesin. Salah satu penerapannya dalam penilaian adalah sistem otomatis untuk jawaban esai siswa, di mana algoritma NLP mencocokkan teks jawaban dengan kriteria rubrik. Teknik seperti *word embedding*, *TF-IDF*, dan *semantic similarity* digunakan untuk menilai relevansi serta kesesuaian makna jawaban. Pendekatan ini mengurangi subjektivitas penilaian, mempercepat umpan balik, dan meningkatkan efektivitas pembelajaran melalui analisis tata bahasa serta struktur argumen siswa.

- 1) tokenization adalah langkah awal dalam pemrosesan teks yang bertujuan untuk mengubah teks mentah menjadi struktur yang lebih terorganisir agar dapat dianalisis lebih lanjut, proses dalam Natural Language Processing (NLP) yang membagi teks menjadi unit-unit kecil yang disebut token, token dapat berupa kata, frasa, atau bahkan karakter, tergantung pada tingkat tokenisasi yang digunakan.
- 2) Text Length Analysis adalah metode untuk mengukur panjang teks berdasarkan jumlah kata atau karakter dalam suatu dokumen, yang dapat menjadi indikator kualitas jawaban siswa dalam penilaian berbasis rubrik. Metrik yang digunakan meliputi jumlah token (word count) untuk menghitung kata, jumlah karakter (character count) termasuk spasi, panjang kalimat (sentence length) untuk menilai kompleksitas, serta rasio panjang jawaban terhadap standar rubrik guna memastikan kesesuaian dengan ekspektasi penilaian.
- 3) Field Observation: yaitu dengan cara mengobservasi pengguna secara langsung bagaimana mereka menggunakan aplikasi tersebut, tingkat efektivitas diukur menggunakan tingkat keberhasilan pengguna (Neilsen, 2001). Pada penelitian sebelumnya, metode F1 Score juga digunakan untuk menilai efektivitas sistem, seperti dalam penelitian berjudul "Implementasi Seleksi Fitur Pada Algoritma Klasifikasi Machine Learning Untuk Prediksi Penghasilan Pada Adult Income Dataset" (Serafim Clara, Dhea Laksmi Prianto, Rizal Al Habsi, Ester Friscila Lumbantobing, Nurul Chamidah, S. Kom, M. Kom, 2021). Nilai efektivitas sistem dihitung dengan rumus sebagai berikut: Efektivitas=($\sum_{i=1}^n x_i$)/(n x Skala Maksimum)

Keterangan:

x_i : Nilai kepuasan yang diberikan oleh responden ke-i (dengan skala 1–5)

N : Jumlah responden

Skala Maksimum: Nilai tertinggi pada skala Likert (dalam hal ini, 5)

Selanjutnya, untuk menghitung nilai kepuasan pengguna, digunakan persentase perbandingan antara total nilai yang diberikan oleh setiap responden (x_i) dengan nilai maksimum yang dapat dicapai berdasarkan skala Likert yang digunakan, dikalikan dengan jumlah responden (n). Rumusnya adalah sebagai berikut:

Tabel 1. Skor Responden

Pernyataan	Skor
Sangat Setuju (SS)	5
Setuju (S)	4
Ragu-Ragu (RG)	3
Tidak Setuju (TS)	2
Sangat Tidak Setuju (STS)	1

Interpretasi hasil nilai kepuasan berdasarkan presentase pencapaian menggunakan kategori berikut:

Tabel 2. Kategori Kelayakan

No	Presentase Kelayakan	Keterangan Interpretasi
1	Dibawah 40%	Sangat Tidak Efektif
2	40 – 59,99%	Tidak Efektif
3	60 – 79,99%	Cukup Efektif
4	Diatas 80%	Sangat Efektif

Hasil analisis menggunakan acuan standar sesuai dengan (Litbang Depdagri, 1991). Analisis ini memberikan gambaran tingkat kepuasan pengguna terhadap efektivitas kerja sistem yang diuji, sehingga dapat dijadikan panduan untuk perbaikan dan pengembangan sistem di masa mendatang.

3. PEMBAHASAN

Hasil uji coba sistem memuat data pengujian keseluruhan, mulai dari nilai responden hingga tingkat kepuasan, melalui langkah-langkah uji coba yang telah dijelaskan pada sub bagian sebelumnya. Nilai responden dalam sistem ini diperoleh berdasarkan penilaian dari 10 orang guru yang berperan sebagai responden. Pengukuran dilakukan menggunakan metode field observation, di mana guru mengevaluasi sistem berdasarkan pengalaman langsung dalam penggunaannya. Hasil kepuasan responden dianalisis dengan membandingkan kesesuaian sistem terhadap kriteria rubrik yang telah ditetapkan, sehingga diperoleh gambaran mengenai efektivitas sistem dalam membantu proses penilaian.

Tabel 3 Tabel Hasil Responden PSSUQ

No	Responden	Pertanyaan															
		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16
1	R1	4	4	4	3	5	5	4	5	5	3	5	3	5	4	5	4
2	R2	4	4	5	5	5	3	5	5	4	5	5	5	3	5	5	5
3	R3	5	5	5	5	3	5	4	5	4	5	5	5	3	5	5	4
4	R4	5	4	3	5	4	5	4	5	3	5	4	5	5	3	4	5
5	R5	5	3	4	5	3	5	4	4	5	5	4	4	3	4	5	5
6	R6	4	5	5	4	3	5	4	5	5	4	5	4	5	4	5	5
7	R7	5	4	4	5	4	3	5	5	3	5	4	5	4	5	5	4
8	R8	4	3	5	5	5	5	4	4	5	3	5	3	4	5	4	5
9	R9	5	5	4	4	3	5	5	4	5	5	5	4	3	4	5	5
10	R10	4	4	3	5	5	4	5	5	4	5	4	5	5	3	4	4
Jumlah		694															

Langkah selanjutnya adalah Perhitungan skor pada kuesioner kepada pengguna dibagi menjadi empat kategori, yaitu keseluruhan (Overall), kegunaan sistem (System Usability), kualitas informasi (Information Quality), dan kualitas antarmuka (Interface Quality). Maka diketahui perolehan rata-rata skor kelayakan setiap kategori sebagai berikut:

a. Overall

Dalam kategori overall skor yang diobservasi sebesar 694, dari skor yang di harapkan sebesar 800 dari 16 pertanyaan kuesioner PSSUQ sehingga rata-rata presentasi kelayakan yaitu:

$$\text{Presentase Kelayakan (\%)} = \frac{\text{Skor yang di observasi}}{\text{Skor yang diharapkan}} \times 100\%$$

$$\text{Presentase Kelayakan (\%)} = \frac{694}{800} \times 100\%$$

$$\text{Presentase Kelayakan (\%)} = 86.75\%$$

b. System Usability

Dalam kategori overall skor yang diobservasi sebesar 257, dari skor yang di harapkan sebesar 300 dari 6 pertanyaan kuesioner (P1 – P6) PSSUQ sehingga rata-rata presentasi kelayakan yaitu:

$$\text{Presentase Kelayakan (\%)} = \frac{\text{Skor yang di observasi}}{\text{Skor yang diharapkan}} \times 100\%$$

$$\text{Presentase Kelayakan (\%)} = \frac{257}{300} \times 100\%$$

$$\text{Presentase Kelayakan (\%)} = 85,67\%$$

c. Information Quality

Dalam kategori overall skor yang diobservasi sebesar 270, dari skor yang di harapkan sebesar 300 dari 6 pertanyaan kuesioner (P7 – P12) PSSUQ sehingga rata-rata presentasi kelayakan yaitu:

$$\text{Presentase Kelayakan (\%)} = \frac{\text{Skor yang di observasi}}{\text{Skor yang diharapkan}} \times 100\%$$

$$\text{Presentase Kelayakan (\%)} = \frac{270}{300} \times 100\%$$

$$\text{Presentase Kelayakan (\%)} = 90\%$$

d. Interface Quality

Dalam kategori overall skor yang diobservasi sebesar 130, dari skor yang di harapkan sebesar 150 dari 6 pertanyaan kuesioner (P13 – P15) PSSUQ sehingga rata-rata presentasi kelayakan yaitu:

$$\text{Presentase Kelayakan (\%)} = \frac{\text{Skor yang di observasi}}{\text{Skor yang diharapkan}} \times 100\%$$

$$\text{Presentase Kelayakan (\%)} = \frac{130}{150} \times 100\%$$

$$\text{Presentase Kelayakan (\%)} = 86,67\%$$

Hasil perhitungan rata-rata persentase dari seluruh kategori yaitu “87,27%” dan dinyatakan “Sangat Layak” sebagai sebuah smart system untuk Penerapan Metode Rubrik Berbasis Algoritma Natural Language Processing Dalam Penilaian Pembelajaran Siswa. Setelah data dari observasi dikumpulkan, skor total dari semua responden dihitung, kemudian dibandingkan dengan skor maksimal untuk menentukan persentase efektivitas. Dalam penelitian ini, skor total yang diperoleh adalah 694 sehingga perhitungan tingkat efektivitasnya adalah sebagai berikut:

$$\text{Efektivitas} = \frac{\sum i = 1nXi}{n \times \text{Skala Maksimum}}$$

$$\text{Efektivitas (\%)} = \frac{694}{800} \times 100\%$$

$$\text{Efektivitas (\%)} = 86,75\%$$

Dengan demikian, hasil perhitungan menunjukkan bahwa efektivitas prototipe aplikasi yang dikembangkan mencapai 86%. Nilai ini berada dalam kategori **sangat efektif**, yang menunjukkan bahwa aplikasi berhasil memenuhi sebagian besar kebutuhan pengguna sesuai dengan kriteria yang ditetapkan dalam evaluasi. Hal ini mengindikasikan bahwa sistem penilaian berbasis rubrik yang menggunakan algoritma NLP pada aplikasi ini tidak hanya memudahkan guru dalam memberikan penilaian, tetapi juga dapat diadopsi dengan baik oleh pengguna dalam lingkungan pendidikan yang sesungguhnya

4. KESIMPULAN

- a. Berdasarkan hasil pengujian, sistem penilaian otomatis berbasis algoritma NLP dengan metode rubrik menunjukkan tingkat keberhasilan yang tinggi dalam mencocokkan jawaban siswa dengan kriteria rubrik yang telah ditetapkan. Hasil perhitungan rata-rata persentase dari seluruh kategori menunjukkan nilai sebesar 87,27%, sedangkan skor total dari seluruh responden mencapai 86,75%. Selain itu, sistem mendapatkan respons positif dari guru sebagai pengguna, terutama dalam aspek akurasi, kemudahan penggunaan, dan efektivitas dalam membantu proses penilaian. Hasil ini mengindikasikan bahwa penerapan metode rubrik berbasis algoritma NLP telah memenuhi kebutuhan pengguna dalam memberikan penilaian yang lebih objektif, konsisten, dan efisien.
- b. Sistem penilaian otomatis berbasis algoritma NLP dengan metode rubrik dapat diandalkan untuk menilai jawaban siswa secara objektif dan efisien. Namun, pengembangan lebih lanjut diperlukan untuk meningkatkan akurasi pencocokan jawaban dengan rubrik serta mengoptimalkan sistem dalam menangani variasi jawaban siswa.

5. DAFTAR PUSTAKA

- [1] Alamanda, R., Suhery, C., Brianorman, Y., Sistem Komputer, J., & MIPA Universitas Tanjungpura Jl Hadari Nawawi, F. H. (2020). Aplikasi Pendekripsi Plagiat Terhadap Karya Tulis Berbasis Web Menggunakan Natural Language Processing Dan Algoritma Knuth-Morris-Pratt. *Jurnal Coding, Sistem Komputer Untan*, 04(1), 33–44.
- [2] Ambrose Cobb. (2022). *PHP MySQL*.
- [3] Bernd Klein. (2021). *Python Tutorial*.
- [4] Chamidah, N., & Mega Santoni, M. (2021). *Pencocokan Berbasis Kata Kunci pada Penilaian Esai Pendek Otomatis Berbahasa Indonesia Keyword Matching-Based on Short Essay Autograding in Indonesian*. 20(1), 19–27.
- [5] Creations, N., & Ltd, P. (2004). *MACHINE LEARNING Intelligence Derived from Data Prabhu TL*. www.nestfamecreations.com
- [6] Houde Hill. (1997). *What Do Prototypes Prototype*.
- [7] Icshan, M., & Teknologi Informasi dan Komputer Politeknik Negeri Lhokseumawe Jln, J. B. (2024). *JAISE: Journal of Artificial Intelligence and Software Engineering Sistem Penilaian Mahasiswa Terhadap Fasilitas Kampus Politeknik Negeri Lhokseumawe Menggunakan Natural Language Processing*.

- [8] Ilmiah Oleh, A., Satya Adi Kristiantoro, C., Dewi, C., & Kristen Satya Wacana Salatiga, U. (2023). *Peningkatan Kualitas Layanan Melalui Klasifikasi Pesan Chatbot Berbasis Natural Language Processing (NLP): Model Bidirectional Long Short-Term Memory Network (LSTM)*.
- [9] Lee, R. S. T. (2023). *Natural Language Processing A Textbook with Python Implementation*.
- [10] Merry Anggraeni. (2021). Interactive question-answering menggunakan nlp dan algoritma naïve bayes untuk literasi tentang gangguan dengar (I-CHAT BOT). *Universitas Budi Luhur Jakarta*.
- [11] Mochamad, E., Ardian, R., & Suryadi, G. A. (2024). *Systematic Literature Review: Pemanfaatan Metode Nlp Dalam Sistem Penilaian Esai Otomatis*. 3(1), 8–16. <https://doi.org/10.69808>
- [12] Muktafin, E. H., Kusrini, K., & Luthfi, E. T. (2020). Analisis Sentimen pada Ulasan Pembelian Produk di Marketplace Shopee Menggunakan Pendekatan Natural Language Processing. *Jurnal Eksplora Informatika*, 10(1), 32–42. <https://doi.org/10.30864/eksplora.v10i1.390>
- [13] Muzayyid, M., Hakim, A., Kadek, I., & Nuryana, D. (2023). *Rancang Bangun Sistem Informasi Penerimaan Peserta Magang Menggunakan Natural Language Processing & Algoritma K-Nearest Neighbors Pada Pt. Ide Jualan Creative*.
- [14] Prasetya, M. A., Wulandari, M., & Nikmah, S. A. (2024). Implementasi NLP (Natural Language Processing) Dasar pada Analisis Sentiment Review Spotify. *Universitas Nusantara PGRI Kediri*, 3, 2828–299.
- [15] Sheba Nedianna, U., Putera Permana, E., Nurlaila Zunaidah, F., Pendidikan Guru Sekolah Dasar, P., Nusantara PGRI Kediri, U., Ahmad Dahlan No, J., Mojoroto, K., Kediri, K., & Timur, J. (2023). Pengembangan Media Kadobudi (Kartu Domino Budaya Indonesia) pada Materi Kebudayaan Indonesia Siswa Kelas IV Sekolah Dasar. *Journal on Education*, 05(02), 2290–2296.
- [16] 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>