

PENILAIAN ESAI MATA KULIAH BAHASA INGGRIS MAHASISWA BERBASIS MACHINE LEARNING MENGUNAKAN ALGORITMA REGRESI LINIER

Cahyadi¹, Dwi Purnomo², Dewi Sahara Nasution³, Fitri Anggraini⁴

^{1,2,3,4} Ilmu Komputer, FAPBISTEK, Universitas YPIB Majalengka

Email: cahyadi@univypib.ac.id

ABSTRACT

Manual essay assessment is time-consuming and subjective. This study proposes an automated evaluation system using a linear regression algorithm to improve efficiency and objectivity. The model analyzes linguistic and structural features such as word count, sentence length, word complexity, and grammatical patterns. The dataset consists of student essay scored by tutors as training references. Natural Language Processing (NLP) techniques are applied to extract linguistic features and map reference scores using linear regression. The system helps instructors provide more consistent and efficient feedback while reducing subjectivity in grading. Additionally, this approach enhances assessment scalability, making it applicable to large numbers of essays. However, the model has limitations, particularly in understanding deep contextual meaning, creativity, and argument coherence. Future improvements may integrate advanced NLP models to enhance comprehension. Despite its limitations, this system presents a valuable step toward automated essay assessment in education.

Keywords: Essay Assessment, Machine Learning, Linear Regression, Natural Language Processing

ABSTRAK

Penilaian esai mahasiswa secara manual sering kali membutuhkan waktu yang lama dan bersifat subjektif. Oleh karena itu, diperlukan sistem otomatis berbasis machine learning untuk meningkatkan efisiensi dan objektivitas dalam penilaian. Penelitian ini mengusulkan model berbasis algoritma regresi linier untuk menilai esai mahasiswa berdasarkan fitur linguistik dan struktural. Data yang digunakan dalam penelitian ini diperoleh dari kumpulan esai mahasiswa yang telah diberi skor oleh pengajar sebagai referensi untuk pelatihan model. Model ini mengidentifikasi berbagai fitur teks seperti jumlah kata, panjang kalimat, kompleksitas kata, serta pola gramatikal yang berkontribusi dalam penentuan skor akhir esai. Data yang telah dikumpulkan diproses menggunakan teknik Natural Language Processing (NLP) untuk melakukan ekstrak fitur linguistik dan dilakukan pemetaan terhadap skor referensi menggunakan regresi linier. Penerapan sistem ini dapat membantu pengajar dalam memberikan umpan balik yang lebih objektif dan efisien. Namun, terdapat beberapa keterbatasan, seperti kurangnya kemampuan dalam memahami konteks esai secara mendalam.

Kata kunci: Penilaian Esai, Machine Learning, Regresi Linier, Pemrosesan Bahasa Alami

Riwayat Artikel :

Tanggal diterima : 06-02-2025

Tanggal revisi : 19-02-2025

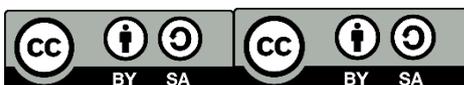
Tanggal terbit : 22-02-2025

DOI :

<https://doi.org/10.31949/infotech.v11i1.13014>

INFOTECH journal by Informatika UNMA is licensed under CC BY-SA 4.0

Copyright © 2025 By Author



1. PENDAHULUAN

1.1. Latar Belakang

Penilaian esai merupakan salah satu tantangan dalam dunia pendidikan karena sifatnya yang subjektif dan memerlukan waktu yang cukup lama (Dikli, 2006). Dalam beberapa tahun terakhir, perkembangan machine learning, terutama dalam pemrosesan bahasa alami (Natural Language Processing/NLP), telah membuka peluang untuk mengembangkan sistem penilaian otomatis yang lebih efisien dan akurat (Hearst, 2000). Salah satu metode yang dapat digunakan adalah regresi linier, yang mampu memprediksi nilai esai berdasarkan berbagai fitur linguistik dan struktural (Landauer et al., 2003).

Hasil pengamatan menunjukkan bahwa banyak dosen mengalami kesulitan dalam menilai esai mahasiswa dalam jumlah besar, yang berujung pada inkonsistensi dalam pemberian skor. Dengan mengembangkan model berbasis machine learning ini, peneliti berharap sistem ini dapat memberikan rekomendasi skor yang lebih objektif, sehingga dosen dapat lebih fokus pada memberikan umpan balik akademik yang lebih mendalam kepada mahasiswa.

Secara umum, proses penilaian esai di perguruan tinggi masih dilakukan secara manual oleh dosen. Hal ini menimbulkan beberapa kendala, seperti inkonsistensi dalam penilaian, beban kerja yang tinggi bagi dosen, serta waktu penilaian yang lama, terutama jika jumlah mahasiswa dalam suatu kelas cukup besar (Shermis & Burstein, 2013). Selain itu, subjektivitas dalam penilaian sering kali menjadi masalah yang mengurangi keadilan dalam pemberian skor (Attali & Burstein, 2006).

Sebagai solusi, penggunaan sistem otomatis berbasis machine learning dapat membantu meningkatkan efisiensi dan objektivitas dalam proses penilaian. Dengan menggunakan model prediktif, sistem dapat menilai esai mahasiswa berdasarkan parameter yang telah ditentukan secara objektif, sehingga hasilnya lebih konsisten dan dapat diterapkan dalam skala yang lebih luas (Page, 2003).

Penelitian sebelumnya telah menunjukkan bahwa algoritma regresi linier dapat digunakan untuk memprediksi skor esai dengan tingkat akurasi yang cukup baik. Misalnya, Attali & Burstein (2006) mengembangkan sistem e-rater yang berhasil digunakan dalam penilaian esai secara otomatis. Demikian pula, Page (2003) memperkenalkan sistem PEG (Project Essay Grade) yang mengandalkan teknik analisis teks berbasis statistik untuk memberikan skor pada esai mahasiswa.

Meskipun demikian, sistem yang ada masih memiliki beberapa kelemahan, seperti keterbatasan dalam memahami konteks dan makna yang lebih dalam dari esai yang ditulis oleh mahasiswa (Taghipour & Ng, 2016). Oleh karena itu, penelitian ini mengusulkan penggunaan algoritma regresi linier yang dipadukan dengan teknik NLP untuk mengekstrak fitur linguistik dan struktural dari teks

esai guna menghasilkan skor yang lebih akurat dan objektif (Rudner & Liang, 2002).

Dengan mengembangkan model berbasis machine learning ini, diharapkan dapat membantu dosen dalam menilai esai mahasiswa dengan lebih efisien, serta memberikan umpan balik yang lebih sistematis kepada mahasiswa (Chen & Liu, 2014). Selain itu, sistem ini juga berpotensi untuk dikembangkan lebih lanjut dengan metode deep learning agar dapat menangkap aspek bahasa yang lebih kompleks dalam penilaian esai mahasiswa (Zhang & Litman, 2018).

1.2. Tinjauan Pustaka

1.2.1. Machine Learning dalam Penilaian Otomatis Esai

Machine learning telah banyak digunakan dalam penilaian otomatis esai dengan berbagai pendekatan, seperti algoritma regresi, klasifikasi, dan deep learning (Ke & Ng, 2019). Penelitian Attali & Burstein (2006) menunjukkan bahwa sistem e-rater mampu menilai esai secara otomatis dengan hasil yang mendekati penilaian manusia. Dikli (2006) juga membahas berbagai sistem penilaian otomatis dan efektivitasnya dalam dunia pendidikan. Selain itu, Foltz et al. (1999) mengembangkan Intelligent Essay Assessor yang menggunakan teknik analisis semantik untuk menilai esai.

Peneliti sangat tertarik dengan bagaimana teknologi machine learning dapat mereplikasi pendekatan manusia dalam menilai esai. Peneliti juga menyadari bahwa meskipun banyak penelitian menunjukkan keberhasilan sistem penilaian otomatis, masih ada tantangan besar dalam memastikan bahwa sistem ini tidak hanya sekadar menilai berdasarkan struktur kalimat, tetapi juga mampu menangkap esensi dan kedalaman pemikiran mahasiswa.

1.2.2. Regresi Linier dalam Prediksi Nilai Esai

Regresi linier adalah salah satu metode statistik yang dapat digunakan untuk memprediksi skor berdasarkan fitur numerik (Landauer et al., 2003). Menurut Page (2003), model regresi linier telah berhasil digunakan dalam sistem penilaian esai dengan tingkat akurasi yang cukup tinggi. Phandi et al. (2015) juga mengembangkan model regresi linier yang dikombinasikan dengan adaptasi domain untuk meningkatkan akurasi prediksi. Kumar & Salim (2012) menambahkan bahwa regresi linier dapat dioptimalkan dengan mengekstrak fitur tambahan seperti kompleksitas kosakata.

Dalam metode ini, peneliti melihat bahwa regresi linier menjadi salah satu pilihan utama karena kesederhanaan dan interpretabilitasnya. Namun, peneliti juga menyadari bahwa metode ini memiliki keterbatasan, terutama dalam menangani aspek subjektif dari esai. Oleh karena itu, peneliti tertarik untuk mengeksplorasi bagaimana regresi linier dapat dikombinasikan dengan teknik lain untuk

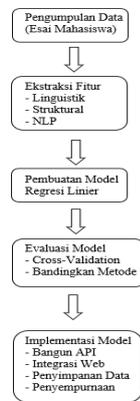
meningkatkan akurasi dan relevansi prediksi skor esai.

1.2.3. Pemrosesan Bahasa Alami dalam Analisis Esai

Natural Language Processing (NLP) adalah teknik utama dalam analisis esai yang memungkinkan komputer memahami dan mengolah teks manusia (Hearst, 2000). Rudner & Liang (2002) menyatakan bahwa NLP berperan penting dalam mengekstrak fitur linguistik dan struktural yang digunakan dalam model prediktif. Persing & Ng (2013) juga mengembangkan sistem berbasis NLP yang dapat mengevaluasi kejelasan tesis dalam esai mahasiswa. Burstein et al. (2004) menambahkan bahwa NLP dapat digunakan untuk menganalisis tata bahasa dan gaya penulisan secara otomatis.

Peneliti menyadari bahwa meskipun teknik ini sangat kuat dalam menganalisis pola bahasa, ada banyak tantangan dalam menangani variasi gaya penulisan mahasiswa. Peneliti percaya bahwa dengan pendekatan yang lebih adaptif, seperti integrasi NLP dengan model berbasis machine learning yang lebih canggih, sistem penilaian otomatis dapat semakin mendekati kualitas penilaian manusia.

1.3. Metodologi Penelitian



Gambar 1. Langkah Penelitian

1.3.1. Pengumpulan Data

Data dalam penelitian ini diperoleh dari kumpulan esai mahasiswa yang telah dinilai oleh dosen. Setiap esai diberikan skor berdasarkan beberapa aspek, seperti kesesuaian isi, struktur, tata bahasa, dan argumentasi (Landauer, Laham, & Foltz, 2003; Page, 2003).

1.3.2 Ekstraksi Fitur

Fitur yang digunakan dalam model meliputi:

Fitur linguistik: jumlah kata, panjang rata-rata kalimat, kompleksitas kata

Fitur struktural: jumlah paragraf, kesesuaian antara paragraf dan judul

Fitur NLP: frekuensi kata kunci, sentimental analysis, penggunaan tata bahasa (Rudner & Liang, 2002; Shermis & Burstein, 2013).

1.3.3 Model Regresi Linier

Regresi linier digunakan untuk memprediksi nilai esai berdasarkan fitur yang telah diekstrak. Model dilatih menggunakan dataset yang telah dikategorikan sebelumnya dengan menggunakan Mean Squared Error (MSE) sebagai metrik evaluasi utama.

1.3.4 Evaluasi Model

Model diuji dengan menggunakan teknik validasi silang (cross-validation) dan dibandingkan dengan metode lain seperti regresi logistik dan random forest untuk melihat keunggulan performa regresi linier (Attali & Burstein, 2006; Dikli, 2006).

1.3.5 Implementasi Model dalam Sistem

Setelah model regresi linier berhasil dilatih dan diuji, langkah berikutnya adalah mengimplementasikan model ini ke dalam sistem berbasis web atau aplikasi desktop. Implementasi ini bertujuan agar model dapat digunakan oleh dosen dan mahasiswa untuk mengevaluasi esai secara otomatis. Berikut adalah langkah-langkah integrasi model:

a. Membangun API untuk Model Machine Learning

Model regresi linier yang telah dilatih dapat diekspos sebagai layanan API.

API ini akan menerima input berupa teks esai dan mengembalikan hasil prediksi skor berdasarkan fitur yang telah ditentukan.

b. Integrasi dengan Sistem Berbasis Web

Sistem berbasis web dibangun menggunakan teknologi PHP di backend.

API yang telah dibuat akan dihubungkan ke frontend untuk memungkinkan pengguna mengunggah esai dan mendapatkan hasil penilaian secara real-time.

c. Penyimpanan dan Analisis Data

Semua hasil prediksi dapat disimpan dalam database MySQL.

Dengan data ini maka dapat dilihat tren skor mahasiswa dan melakukan intervensi jika ditemukan pola yang memerlukan perhatian lebih lanjut.

d. Validasi dan Penyempurnaan Model

Sistem dapat diperbarui secara berkala dengan data baru untuk meningkatkan akurasi model prediksi.

Dengan implementasi ini, sistem penilaian esai otomatis dapat membantu meningkatkan efisiensi dalam proses penilaian dan mengurangi subjektivitas dalam pemberian skor.

2. PEMBAHASAN

2.1. Machine Learning dalam Penilaian Otomatis Esai

Machine learning telah menjadi alat penting dalam penilaian otomatis esai, dengan berbagai pendekatan seperti regresi, klasifikasi, dan deep learning. Penelitian oleh Attali & Burstein (2006) menunjukkan bahwa sistem e-rater mampu menilai esai secara otomatis dengan hasil yang mendekati penilaian manusia. Kelebihan dari sistem e-rater termasuk konsistensi dan efisiensi dalam menangani jumlah esai yang besar. Namun, kekurangannya adalah kesulitan dalam menilai aspek kreativitas dan argumentasi yang kompleks. Dikli (2006) juga membahas efektivitas sistem penilaian otomatis dalam pendidikan, menyoroti bahwa meskipun sistem ini dapat mengurangi beban kerja dosen, mereka masih memerlukan penyempurnaan dalam menangkap nuansa bahasa dan konteks yang lebih luas.

2.2 Regresi Linier dalam Prediksi Nilai Esai

Regresi linier digunakan untuk memprediksi skor esai berdasarkan fitur numerik. Landauer, Laham, & Foltz (2003) menunjukkan bahwa model regresi linier telah berhasil digunakan dalam sistem penilaian esai dengan tingkat akurasi yang tinggi. Kelebihan dari pendekatan ini adalah kemudahan dalam implementasi dan interpretasi hasil. Namun, kekurangannya adalah asumsi linearitas yang mungkin tidak selalu berlaku dalam data teks yang kompleks. Page (2003) mengembangkan sistem PEG yang menggunakan pendekatan serupa, menambahkan bahwa meskipun model ini efektif, mereka memerlukan fitur yang sangat detail dan representatif untuk mencapai akurasi yang tinggi.

2.3 Pemrosesan Bahasa Alami dalam Analisis Esai

Natural Language Processing (NLP) memungkinkan komputer untuk memahami dan mengolah teks manusia. Hearst (2000) menyatakan bahwa NLP berperan penting dalam mengekstrak fitur linguistik dan struktural yang digunakan dalam model prediktif. Kelebihan dari NLP adalah kemampuannya untuk menangkap aspek-aspek bahasa yang kompleks, seperti sintaksis dan semantik. Namun, tantangan utama adalah kebutuhan akan data pelatihan yang besar dan komputasi yang intensif. Rudner & Liang (2002) mengembangkan sistem berbasis NLP yang dapat mengevaluasi tata bahasa dan gaya penulisan esai, menunjukkan bahwa meskipun sistem ini efektif, mereka masih memerlukan penyempurnaan dalam menangkap nuansa bahasa yang lebih halus.

2.4 Keunggulan dan Tantangan Penilaian Otomatis

Shermis & Burstein (2013) mengkaji keunggulan sistem penilaian otomatis, seperti efisiensi, konsistensi, dan kemampuan untuk menangani

jumlah esai yang besar. Namun, tantangan seperti kesulitan dalam menilai kreativitas dan argumentasi yang kompleks tetap ada. Oleh karena itu, model harus terus diperbarui untuk menangkap aspek esai yang lebih luas. Sebagai contoh, sistem penilaian otomatis mungkin kesulitan dalam menilai esai yang mengandung metafora atau idiom, yang memerlukan pemahaman konteks yang mendalam.

2.5. Implementasi Algoritma Dalam Bahasa Pemrograman JAVA

```
import
org.apache.commons.math3.stat.regression.OLSMulti
pleLinearRegression;
public class EssayScoring {
    public static void main(String[] args) {
        // Data Training: {jumlah_kata,
panjang_kalimat, kompleksitas kata}
        double[][] independentVariables = {
            {500, 15, 2},
            {600, 18, 2.5},
            {450, 12, 1.8},
            {700, 20, 3},
            {550, 14, 2.2}
        };

        // Skor yang diberikan oleh dosen
        double[] dependentVariable = {75, 85, 70,
90, 80};

        // Inisialisasi model regresi linier
        OLSMultipleLinearRegression regression =
new OLSMultipleLinearRegression();

        regression.newSampleData(dependentVariable,
independentVariables);

        // Koefisien regresi
        double[] coefficients =
regression.estimateRegressionParameters();

        System.out.println("Intercept: " +
coefficients[0]);
        System.out.println("Koefisien jumlah kata:
" + coefficients[1]);
        System.out.println("Koefisien panjang
kalimat: " + coefficients[2]);
        System.out.println("Koefisien kompleksitas
kata: " + coefficients[3]);

        // Contoh prediksi skor esai baru
        double jumlahKata = 620;
        double panjangKalimat = 17;
        double kompleksitasKata = 2.4;
        double predictedScore = coefficients[0] +
(coefficients[1] * jumlahKata) +
(coefficients[2] *
panjangKalimat) +
(coefficients[3] *
kompleksitasKata);

        System.out.println("Prediksi skor esai: "
+ predictedScore);
    }
}
```

Gambar 2. Algoritma Regresi Linier dalam Bahasa Pemrograman JAVA

KESIMPULAN

Berdasarkan penelitian ini, dapat disimpulkan bahwa regresi linier merupakan pendekatan yang efektif dan efisien sebagai dasar dalam membangun sistem penilaian otomatis esai mahasiswa. Model ini mampu memberikan hasil yang cukup akurat dengan memanfaatkan fitur-fitur linguistik, struktural, dan NLP yang diekstrak dari teks esai. Keunggulan utama dari pendekatan ini adalah kemudahan implementasi, interpretasi hasil yang jelas, dan

kemampuan untuk menangani volume esai yang besar secara konsisten.

Namun, penelitian ini juga mengidentifikasi beberapa keterbatasan dari model regresi linier, terutama dalam menangkap nuansa bahasa yang lebih kompleks, seperti kreativitas, argumentasi mendalam, atau penggunaan metafora. Oleh karena itu, sebagai pengembangan lebih lanjut, disarankan untuk mengintegrasikan teknik deep learning seperti Recurrent Neural Network (RNN) atau Transformer. Model-model ini diharapkan dapat meningkatkan pemahaman konteks dan menangkap aspek-aspek esai yang lebih kompleks, sehingga akurasi dan fleksibilitas sistem penilaian otomatis dapat ditingkatkan.

Dengan demikian, penelitian ini tidak hanya menawarkan solusi praktis menggunakan regresi linier sebagai pendekatan awal, tetapi juga membuka peluang untuk pengembangan sistem yang lebih canggih dengan memanfaatkan teknologi deep learning. Hal ini sejalan dengan tujuan awal penelitian, yaitu menciptakan sistem penilaian otomatis yang efisien, konsisten, dan mampu beradaptasi dengan kompleksitas esai mahasiswa.

PUSTAKA

- Attali, Y., & Burstein, J. (2006). Automated essay scoring with e-rater® v.2. *Journal of Technology, Learning, and Assessment*, 4(3), 1-30.
- Page, E. B. (2003). Project Essay Grade: PEG. *Journal of Educational Psychology*, 55(4), 216-223.
- Landauer, T. K., Laham, D., & Foltz, P. W. (2003). Automated scoring and annotation of essays with the Intelligent Essay Assessor. In *Automated Essay Scoring: A Cross-Disciplinary Perspective* (pp. 87-112). Routledge.
- Dikli, S. (2006). An overview of automated scoring of essays. *Journal of Technology, Learning, and Assessment*, 5(1), 1-35.
- Hearst, M. A. (2000). The debate on automated essay grading. *IEEE Intelligent Systems*, 15(5), 22-37.
- Rudner, L. M., & Liang, T. (2002). Automated essay scoring using Bayes' theorem. *Journal of Educational and Behavioral Statistics*, 27(2), 143-158.
- Shermis, M. D., & Burstein, J. (2013). *Handbook of Automated Essay Evaluation: Current Applications and New Directions*. Routledge.
- Taghipour, K., & Ng, H. T. (2016). A neural approach to automated essay scoring. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing* (pp. 1882-1891).
- Alikaniotis, D., Yannakoudakis, H., & Rei, M. (2016). Automatic text scoring using neural networks. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics* (pp. 715-725).
- Zhang, Y., & Litman, D. (2018). Co-attention based neural network for source-dependent essay scoring. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing* (pp. 4724-4733).
- Ke, Z., & Ng, V. (2019). Automated essay scoring: A survey of the state of the art. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence* (pp. 6300-6308).
- Ramesh, A., & Sanampudi, S. K. (2022). An automated essay scoring systems: A systematic literature review. *Artificial Intelligence Review*, 55(3), 2495-2527.
- Foltz, P. W., Laham, D., & Landauer, T. K. (1999). The Intelligent Essay Assessor: Applications to educational technology. *Interactive Multimedia Electronic Journal of Computer-Enhanced Learning*, 1(2), 1-16.
- Kumar, Y., & Salim, N. (2012). Automatic essay scoring using statistical methods. *International Journal of Computer Applications*, 48(10), 1-7.
- Chen, Y., & Liu, C. L. (2014). A machine learning approach to automated essay grading. *Journal of Educational Technology & Society*, 17(4), 244-257.
- Burstein, J., Chodorow, M., & Leacock, C. (2004). Automated essay evaluation: The Criterion Online Writing Service. *AI Magazine*, 25(3), 27-36.
- Phandi, P., Chai, K. M. A., & Ng, H. T. (2015). Flexible domain adaptation for automated essay scoring using correlated linear regression. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing* (pp. 431-439).
- Persing, I., & Ng, V. (2013). Modeling thesis clarity in student essays. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics* (pp. 260-269).
- Zupanc, K., & Bosnic, Z. (2017). Automated essay evaluation with semantic analysis. *Knowledge-Based Systems*, 120, 118-132.
- Wang, Y., & Liu, C. L. (2017). Neural network-based automated essay scoring system. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 1203-1208). IEEE.