

ABSTRAK

Automatic Essay Scoring adalah sistem yang dirancang untuk menilai esai secara otomatis dengan membandingkan jawaban siswa terhadap kunci jawaban guru. Penilaian manual sering kali menimbulkan masalah subjektivitas dan inkonsistensi, sehingga diperlukan sistem penilaian otomatis yang lebih efisien dan objektif. Penelitian ini bertujuan untuk mengembangkan sistem penilaian esai otomatis berbasis *Cosine Similarity* dan *N-Gram* guna meningkatkan akurasi dan efisiensi evaluasi jawaban siswa. *Cosine Similarity* digunakan untuk mengukur tingkat kesamaan antara jawaban siswa dan kunci jawaban, sedangkan *N-Gram* berperan dalam menganalisis pola urutan kata dalam teks. Sistem ini juga menerapkan teknik *Natural Language Processing* (NLP) seperti *cleanning*, *case folding*, *tokenization*, *stopword removal*, dan *stemming* untuk mempersiapkan teks sebelum dianalisis. Dataset yang digunakan berasal dari jawaban esai siswa pelajaran IPA kelas IX-A SMPN 1 Kalianget. Berdasarkan hasil evaluasi, representasi Unigram menghasilkan nilai Mean Absolute Error (MAE) paling rendah sebesar 8.84, dibandingkan dengan Bigram (21.85) dan Trigram (31.57), yang menunjukkan bahwa Unigram memberikan hasil prediksi paling mendekati penilaian guru. Dengan adanya sistem ini, diharapkan institusi pendidikan dapat menghemat waktu dalam proses evaluasi, meningkatkan kualitas umpan balik kepada siswa, serta mengurangi kemungkinan bias dalam penilaian.

Kata kunci: *Automatic Essay Scoring, Cosine Similarity, N-Gram, Natural Language Processing, Penilaian Esai*

ABSTRACT

Automatic Essay Scoring is a system designed to evaluate essays automatically by comparing students' answers with the teacher's answer key. Manual assessment often leads to issues of subjectivity and inconsistency, making it necessary to develop a more efficient and objective automatic scoring system. This study aims to develop an automatic essay scoring system based on Cosine Similarity and N-Gram methods to improve the accuracy and efficiency of student answer evaluations. Cosine Similarity is used to measure the degree of similarity between the student's answer and the answer key, while N-Gram plays a role in analyzing word sequence patterns within the text. The system also applies Natural Language Processing (NLP) techniques such as cleaning, case folding, tokenization, stopword removal, and stemming to prepare the text before analysis. The dataset used in this study consists of essay answers from Grade IX-A science students at SMPN 1 Kalianget. Based on the evaluation results, the Unigram representation produced the lowest Mean Absolute Error (MAE) of 8.84, compared to Bigram (21.85) and Trigram (31.57), indicating that Unigram provided predictions closest to the teacher's scoring. This system is expected to help educational institutions save time in the evaluation process, improve the quality of feedback provided to students, and reduce the potential for bias in assessments.

Keywords: Automatic Essay Scoring, Cosine Similarity, N-Gram, Natural Language Processing, Essay Assessment