

ABSTRAK

Penilaian esai merupakan salah satu metode untuk mengevaluasi kemampuan siswa dalam memahami materi pelajaran, namun proses penilaian ini sering kali memakan waktu lama karena harus dilakukan secara manual. Untuk mengatasi masalah ini, sistem penilaian esai otomatis (*Automatic Essay Scoring*) menjadi solusi yang efektif. Penelitian ini membandingkan dua teknik representasi teks, yaitu TF-IDF dan Word2Vec, dengan menerapkan metode *Generalized Latent Semantic Analysis* (GLSA). Dalam penelitian ini, jarak kesamaan dokumen diukur menggunakan *Cosine Similarity*, *Dice Similarity*, dan *Jaccard Similarity*. Metode GLSA digunakan untuk mengatasi masalah urutan kata dengan menggunakan teknik N-gram, serta menghitung nilai vektor dan eigen dari dokumen-dokumen tersebut dengan algoritma *Singular Value Decomposition* (SVD). Penelitian ini bertujuan untuk menentukan nilai *Root Mean Square Error* (RMSE) terbaik dari metode GLSA dengan TF-IDF dan Word2Vec pada penilaian esai otomatis jawaban pendek Ilmu Pengetahuan Alam (IPA) yang dilakukan di MTS Negeri 1 Sumenep. Hasil pengujian menunjukkan bahwa pada metode TF-IDF, trigram menghasilkan RMSE terbaik untuk *Cosine Similarity* sebesar 0.0861%, sementara bigram unggul untuk *Dice Similarity* dengan RMSE 0.0697%, dan unigram memberikan performa terbaik untuk *Jaccard Similarity* dengan RMSE 0.0613%. Sebaliknya, pada metode Word2Vec, bigram menunjukkan performa terbaik untuk *Cosine* dan *Jaccard Similarity* dengan nilai RMSE masing-masing sebesar 0.0684%, sedangkan trigram unggul untuk *Dice Similarity* dengan RMSE 0.0811%. Hasil ini menunjukkan bahwa Word2Vec lebih efektif dalam menangkap hubungan semantik antar kata pada konteks yang kompleks, sedangkan TF-IDF lebih unggul pada representasi sederhana. Penelitian ini memberikan kontribusi penting untuk mendukung otomatisasi penilaian esai pendek dengan hasil yang akurat dan menjadi acuan untuk penelitian lanjutan.

Kata Kunci: *Automatic Essay Scoring, Generalized Latent Semantic Analysis, N-gram, Similarity, TF-IDF, Word2Vec*

ABSTRACT

Essay scoring is one of the methods to evaluate students' ability to understand the subject matter, but this scoring process often takes a long time because it must be done manually. To overcome this problem, an automatic essay scoring system is an effective solution. This research compares two text representation techniques, namely TF-IDF and Word2Vec, by applying the Generalized Latent Semantic Analysis (GLSA) method. In this research, document similarity distance is measured using Cosine Similarity, Dice Similarity, and Jaccard Similarity. The GLSA method is used to overcome the word order problem by using the N-gram technique and calculating the vector and eigenvalues of the documents with the Singular Value Decomposition (SVD) algorithm. This study aims to determine the best Root Mean Square Error (RMSE) value of the GLSA method with TF-IDF and Word2Vec on automatic short answer essay assessment of Natural Sciences (IPA) conducted at MTS Negeri 1 Sumenep. The test results show that in the TF-IDF method, trigram produces the best RMSE for Cosine Similarity of 0.0861%. In comparison, bigram excels for Dice Similarity with RMSE 0.0697%, and Unigram gives the best performance for Jaccard Similarity with RMSE 0.0613%. In contrast, in the Word2Vec method, bigrams performed best for Cosine and Jaccard Similarity with an RMSE value of 0.0684% each, while trigrams excelled for Dice Similarity with an RMSE of 0.0811%. These results show that Word2Vec captures semantic relationships between words in complex contexts more effectively, while TF-IDF excels in simple representations. This research makes an important contribution to supporting the automation of short essay assessments with accurate results and serves as a reference for further study.

Kata Kunci: *Automatic Essay Scoring, Generalized Latent Semantic Analysis, N-gram, Similarity, TF-IDF, Word2Vec*