

ABSTRAK

Dalam proses pembelajaran, sebuah metode pengukuran hasil proses belajar diperlukan. Salah satu metode pengukuran dalam proses belajar adalah dengan menggunakan sistem ujian. Beberapa metode ujian digunakan untuk mengukur pemahaman siswa, yaitu metode pilihan ganda, metode *essay*, metode jawaban singkat, dan metode berbasis proyek. Jenis ujian *essay* memiliki kekurangan dalam hal waktu untuk pemeriksaan ujian secara manual. Meskipun memiliki kekurangan, metode *essay* memiliki keuntungan dalam melatih daya ingatan serta meningkatkan kompetensi siswa dalam kemampuan tata bahasa. Metode ujian *essay* juga memiliki jawaban yang merupakan data tidak terstruktur. Untuk mengatasi teks yang tidak terstruktur, maka diperlukan metode *text mining* untuk pengolahan teks tersebut. *Text mining* merupakan metode penemuan informasi dengan mengekstrak informasi berbagai sumber. Dengan *text mining*, proses untuk prediksi penilaian otomatis terhadap ujian berbasis *essay* adalah dengan menggunakan sistem AES. Sistem AES (*Automatic Essay Scoring*) merupakan sebuah sistem yang bekerja dalam menilai ujian berbasis *essay* secara otomatis. Dalam penelitian ini, penerapan *text mining* tersebut dilakukan dengan simulasi aplikasi penilaian *essay* di Universitas Universal, Batam, Kepulauan Riau. Rancang bangun aplikasi dipilih terkait permasalahan waktu pemeriksaan jawaban dan ketidakakuratan sistem ujian *essay* yang digunakan saat ini. Berdasarkan permasalahan tersebut, diperlukan aplikasi ujian *essay* dan memanfaatkan metode *text mining* untuk melakukan penilaian terhadap jawaban *essay*. Metode yang digunakan dalam penelitian ini adalah *text mining* dengan bantuan pustaka NLTK untuk data *stopwords* dan proses *tokenizing*, Sastrawi untuk proses *stemming*, dan *query expansion* dengan kamus data Tesaurus, *cosine similarity* dan pembobotan TF-IDF menggunakan scikit-learn, serta normalisasi untuk konversi *cosine similarity* menjadi hasil penilaian sistem. Rata-rata akurasi dalam penelitian ini adalah 89,93% dengan kondisi jawaban mahasiswa yang mengandung *typo*. Sistem juga dapat mempermudah dan mempercepat pemeriksaan jawaban dengan menghasilkan nilai berdasarkan similaritas jawaban mahasiswa dan kunci jawaban.

Kata Kunci: *Text Mining, Cosine Similarity, TF-IDF, Query Expansion*

ABSTRACT

In the learning process, a method of measuring the results of learning process is needed. One of the methods is by using an exam system. Several methods are used to measure students' understanding which are multiple choice method, essay method, short answer method, and project-based method. Essay based method has a drawback in time needed for manual examination. Although of its drawback, the essay method has the advantage in training memorization skills and improving student's competence in grammatical skills. Essay method also has answers which are unstructured data. To overcome this problem, a text mining method is needed for processing the text. Text mining is an information retrieval method which extracts information from several sources. With text mining, the examination process for essay questions is by utilizing the AES system. AES (Automatic Essay Grading) is a system which works in automatic-based essay scoring. In this study, the implementation of text mining is done by simulating an essay assessment at Universal University, Batam, Riau Islands. The design of the application is chosen based on the drawback in assessment duration and inaccuracy of the current system. Based on these problems, it is necessary for an essay assessment application and implementation of text mining method for essay assessment. The method used in this research is text mining with the help of NLTK library for stopwords data and tokenizing, Sastrawi for stemming process, query expansion with Tesauro dictionary, cosine similarity and TF-IDF weighting using scikit-learn, and normalization for converting cosine similarity results into system score. The average of accuracy in this study is 89,93% with the condition of students' answers containing typos. The system can also simplify and accelerate the examination by generating scores based on the similarity of student answers and answer keys.

Keywords: *Text Mining, Cosine Similarity, TF-IDF, Query Expansion*