

ABSTRACT

In the recovery period after the world experienced a pandemic like today, e-learning has become widely used by various educational institutions ranging from high schools to universities. However, in practice, the implementation of e-learning still encounters several challenges, one of those challenges is how to measure the student's achievement of learning objectives. This achievement is usually measured by an assessment or scoring of the work on an assignment or exam done by students.

Therefore, the assessment process becomes one of the most important factors in the use of e-learning. The results of a good quality assessment can be used to measure student abilities accurately and as evaluation material for further learning. To support this assessment process, an assessment method that is more effective in terms of time and energy is needed. The purpose of this research is to assist the process of scoring or correcting answers in an automatic way.

This study used the “Kuis Mata Kuliah Pemrograman Dasar” dataset taken from the Department of Electrical Engineering and Information Technology. The method used is to modify the transformer-based model, namely the Bidirectional Encoder Representation from Transformers (BERT) model with fine-tuning and hyperparameter optimization techniques used as encoders. Furthermore, the results of the encoding of each essay will be measured by semantic textual similarity to get the final score. The results of this study are able to improve the performance of the pre-trained model for essay assessment, so that it can provide an Automated Essay Scoring model with good performance and accuracy on a small dataset. The measurement results are based on evaluation metrics, obtaining a QWK score of 93% and an accuracy score of 92%.

Keywords: *Automated Essay Scoring, Natural Language Processing, Transformer, BERT, Semantic Textual Similarity*

INTISARI

Di masa pemulihan setelah dunia mengalami pandemi seperti saat ini, penggunaan *e-learning* menjadi lebih banyak digunakan oleh berbagai institusi pendidikan mulai dari sekolah menengah hingga perguruan tinggi. Namun pada pelaksanaannya, penerapan *e-learning* masih menemui beberapa tantangan, salah satunya yaitu bagaimana cara untuk mengukur ketercapaian tujuan pembelajaran. Ketercapaian ini biasa diukur dengan asesmen atau penilaian atas pengerjaan suatu assignment atau ujian yang dikerjakan oleh siswa.

Sehingga proses penilaian menjadi salah satu faktor yang sangat penting dalam penggunaan *e-learning*. Hasil dari kualitas penilaian yang baik bisa digunakan untuk mengukur kemampuan siswa secara akurat dan sebagai bahan evaluasi untuk pembelajaran selanjutnya. Untuk mendukung proses penilaian ini, maka dibutuhkan metode asesmen yang lebih efektif dari segi waktu ataupun tenaga. Tujuan dari penelitian ini adalah untuk membantu proses penilaian atau koreksi jawaban dengan cara otomatis.

Penelitian ini menggunakan dataset Kuis Mata Kuliah Pemrograman Dasar Departemen Teknik Elektro dan Teknologi Informasi. Metode yang digunakan adalah memodifikasi model dengan basis transformer yaitu model *Bidirectional Encoder Representation from Transformers* (BERT) yang digunakan untuk task *Natural Language Processing* (NLP) dengan teknik *fine-tuning* dan *hyperparameter optimization* yang digunakan sebagai *encoder*. Selanjutnya hasil dari *encoding* dari setiap esai akan diukur dengan *semantic textual similarity* untuk mendapatkan skor akhir. Hasil dari penelitian ini mampu memperbaiki kinerja model *pre-trained* untuk penilaian esai, sehingga bisa memberikan model *Automated Essay Scoring* dengan kinerja dan akurasi yang baik pada dataset yang kecil. Hasil pengukuran berdasarkan metrik evaluasi, mendapatkan skor QWK 93% dan skor Akurasi 92%.

Kata kunci – *Automated Essay Scoring, Natural Language Processing, Transformer, BERT, fine-tuning, Semantic Textual Similarity.*