

Perkembangan teknologi informasi telah mempengaruhi berbagai aspek kehidupan termasuk pada sektor pendidikan. Perkembangan ini memungkinkan munculnya *e-Learning* yang perkembangannya semakin didorong akibat pandemi Covid-19. Pada *e-Learning*, evaluasi performa siswa biasanya dilakukan melalui kuis dengan tipe pilihan ganda atau esai. Namun, penilaian esai, khususnya esai jawaban pendek, pada *platform e-Learning* saat ini masih dilakukan secara manual yang memiliki kerentanan terhadap bias dan inkonsistensi yang diakibatkan oleh kelelahan. Salah satu cara yang dapat digunakan untuk mengatasi permasalahan penilaian manual pada soal esai jawaban pendek adalah dengan menggunakan sistem penilai jawaban pendek otomatis.

Penelitian ini mengeksplorasi penggunaan salah satu model *deep learning*, yaitu *large language model* yang memiliki pemahaman umum mengenai bahasa yang kemudian dilakukan proses *finetuning*. Secara spesifik, penelitian ini menggunakan model *transformers*, khususnya BERT, dengan metode *contrastive learning* untuk membentuk sistem penilai jawaban pendek otomatis untuk soal esai dan membandingkan hasilnya dengan sistem sejenis. Model disusun dalam dua bagian, model *body* berupa variasi dari BERT, yaitu 'bert-base-multilingual-cased' dan model *head* berupa model regresi logistik. Model *body* disusun dalam arsitektur *siamese*.

Hasil penelitian ini menunjukkan peningkatan performa model dengan memanfaatkan *grid search* untuk optimisasi *hyperparameter* dan batas skor *contrastive* dengan *10-fold cross validation*. Penurunan kesalahan prediksi pada metrik MAE sebesar 21.72% dan 9.90%, sedangkan pada metrik RMSE sebesar 17.79% dan 13.80%, jika dibandingkan pada model BERT *pretrained* (model *baseline*) dan BERT dengan *finetuning cosine similarity* (model *finetuning*). Model berbasis *transformers* dengan *contrastive learning* mampu mencapai metrik sebesar 0.191 untuk MAE dan 0.231 RMSE. Hal ini menunjukkan potensi penggunaan metode *contrastive learning* pada model *transformers* untuk mengembangkan sistem penilai jawaban pendek otomatis

Kata kunci: *contrastive learning*, *transformers*, sistem penilai jawaban pendek otomatis, *e-Learning*, *deep learning*

ABSTRACT

The rapid development of technology has impacted various sectors, including education. These developments have enabled e-Learning to thrive, especially during the Covid-19 pandemic. Evaluating student performance and understanding in e-Learning is typically done through quizzes in multiple-choice or essay formats. However, these evaluations, especially in essay grading, still require manual effort. This can lead to exhaustion and introduce bias and inconsistency into the scoring process. To address this issue, one possible solution is to develop an automated short-answer grading system.

This research explores one of the deep learning models, which is a large language model that has a general understanding of language. This model is then subjected to a finetuning process. Specifically, this study employs a transformer model, especially BERT, with contrastive learning method to develop an automated short-answer scoring system and compare its performance with similar systems. The model is composed of two components, namely the model body which utilizes BERT variation, which is 'bert-base-multilingual-cased', and the model head which employs logistic regression. The body model is structured in a siamese architecture.

The results of this research demonstrate an improvement in model performance by utilizing grid search for hyperparameter and contrastive score limit optimization with 10-fold cross-validation. When compared to the pretrained BERT (baseline model) and BERT with cosine similarity finetuning (finetuning model), the reduction in prediction errors, measured by the MAE metric, is 21.72% and 9.90%, while for the RMSE metric, it is 17.79% and 13.80%. The transformers-based model with contrastive learning achieves metrics of 0.191 for MAE and 0.231 for RMSE. These findings indicate the potential of using the contrastive learning method in transformers models to develop an automated short-answer scoring system.

Keywords : contrastive learning, transformers, automated short answer grading, e-Learning, deep learning