

ABSTRAK

Proses penilaian manual untuk soal jawaban singkat sering kali memakan waktu yang lama dan rentan terhadap subjektivitas. *Automatic Short Answer Grading* (ASAG) telah banyak diteliti sebagai solusi otomatis dengan sebagian besar pendekatan memanfaatkan dua masukan, yaitu referensi dan jawaban mahasiswa. Namun, pendekatan ini belum sepenuhnya optimal dalam menangkap hubungan semantik yang melibatkan soal sebagai konteks. Dalam penelitian ini, digunakan tiga masukan berupa soal, referensi, dan jawaban mahasiswa yang diimplementasikan melalui *fine-tuning triplet loss*. Sebagian besar penelitian ASAG yang menggunakan *triplet loss* bergantung pada data berlabel dalam proses *fine-tuning*. Namun, tantangan muncul ketika bekerja dengan data jawaban mahasiswa yang tidak berlabel (*unsupervised*), karena sulitnya menentukan pasangan sampel positif dan negatif secara langsung. Oleh karena itu, penelitian ini mengusulkan metode *clustering* menggunakan algoritma *K-Means* untuk mengelompokkan jawaban mahasiswa berdasarkan kedekatan semantik dengan referensi. Soal kemudian digunakan sebagai *anchor*, *cluster* jawaban yang dekat dengan referensi digunakan sebagai sampel positif, sedangkan *cluster* jawaban yang jauh dari referensi digunakan sebagai sampel negatif. Model berbasis *Sentence BERT* (SBERT), yaitu IndoSBERT dan *paraphrase-multilingual-mpnet-v2* yang telah di *fine-tuning* digunakan untuk representasi teks, dengan perhitungan relevansi semantik antara soal, referensi, dan jawaban mahasiswa menggunakan *cosine similarity*. Penyesuaian parameter α dan β dilakukan untuk mengoptimalkan performa empat *dataset* yang dievaluasi menggunakan SMAPE. Hasil penelitian menunjukkan bahwa strategi *clustering* menggunakan algoritma *K-Means* mampu mengatasi tantangan *unsupervised* dalam *fine-tuning*, serta secara signifikan meningkatkan performa model. Model IndoSBERT yang di *fine-tuning* dengan $\alpha = 0,7$ dan $\beta = 0,3$ menghasilkan rata-rata nilai SMAPE sebesar 8,9418%, mengungguli model *paraphrase-multilingual-mpnet-v2* yang di *fine-tuning* dengan $\alpha = 0,6$ dan $\beta = 0,4$, dengan nilai SMAPE sebesar 9,8390%. Penelitian ini tidak hanya membuktikan efektivitas model berbasis korpus bahasa Indonesia untuk tugas ASAG (IndoSBERT), tetapi juga menawarkan solusi *fine-tuning* yang inovatif untuk data tidak berlabel.

Kata kunci : *Automatic Short Answer Grading*, SBERT, *Clustering*, *Triplet Loss*, SMAPE