

ABSTRAK

Uraian jawaban singkat merupakan bentuk soal ujian yang dinilai tepat untuk mengevaluasi kognitif, pemahaman, dan pengetahuan. Sayangnya, jawaban yang dihasilkan sangat variatif sehingga penilaian manual akan memakan waktu dan memunculkan bias atau inkonsistensi. *Automated Short Answer Grading* (ASAG) adalah penilaian jawaban singkat menggunakan teknik komputasi dengan keluaran nilai maupun kategori kognitif. Input ASAG berupa soal dan jawaban membantu mengenali jawaban benar yang bervariasi berdasarkan relevansi konteks atau atribut semantik dua representasi teks. Metode representasi teks pada ASAG yang populer saat ini, *sentence embedding Siamese-BERT*, mampu menghasilkan vektor semantik berukuran tetap. Representasi soal dan jawaban dari metode tersebut dapat dinilai relevansinya dengan metode *matching matrix* dan dipahami konteksnya dengan BiLSTM. Oleh karena itu, dilakukan penelitian ASAG menggunakan *matching matrix* dan BiLSTM pada representasi soal dan jawaban hasil *sentence embedding Siamese-BERT*. IndoSBERT, model *Siamese-BERT* Indonesia, di-*fine-tuning* dan digunakan untuk merepresentasikan soal dan jawaban dari empat set data uraian jawaban singkat berbahasa Indonesia. Empat arsitektur model ASAG dibangun untuk mengetahui pengaruh *matching matrix* dan BiLSTM. Model tanpa *matching matrix* dengan BiLSTM menjadi model terbaik untuk soal uraian terbatas dengan jawaban analitis (set data Siscer) dengan SMAPE 13,2262%. Model dengan *matching matrix* dan BiLSTM menjadi model terbaik untuk soal uraian terbuka dengan jawaban analitis (set data MPI dan Ukara B) dengan SMAPE 14,2271% dan 30,7918%. Model dengan *matching matrix* tanpa BiLSTM menjadi model terbaik untuk soal uraian terbuka dengan jawaban nonanalitis (set data Ukara A) dengan SMAPE 19,7149%. *Matching matrix* terpantau berpengaruh baik pada soal uraian terbuka, sedangkan BiLSTM terpantau berpengaruh baik pada jawaban analitis.

Kata kunci : *Automated Short Answer Grading*, BiLSTM, *Matching Matrix*, SMAPE