ABSTRACT

The manual grading process for short-answer questions is often time-consuming and prone to subjectivity. Automatic Short Answer Grading (ASAG) has been widely studied as an automated solution, with most approaches utilizing two inputs: the reference answer and the student's response. However, this approach is not yet fully optimal in capturing semantic relationships that involve the question as context. In this study, three inputs—question, reference answer, and student response—are used and implemented through fine-tuning with triplet loss. Most ASAG studies employing triplet loss rely on labeled data during fine-tuning. However, challenges arise when dealing with unlabeled student responses (unsupervised learning), as it is challenging to directly determine positive and negative sample pairs. Therefore, this study proposes a clustering method using the K-Means algorithm to group student responses based on their semantic proximity to the reference answer. The question is then used as an anchor, clusters of answers close to the reference are used as positive samples, while clusters far from the reference are used as negative samples. A Sentence BERT (SBERT)-based model, specifically IndoSBERT and paraphrase-multilingual-mpnet-v2, fine-tuned for text representation, is used. Semantic relevance between the question, reference answer, and student response is calculated using cosine similarity. Parameter adjustments for α and β are performed to optimize performance on four datasets evaluated using SMAPE. The results indicate that the clustering strategy using the K-Means algorithm effectively addresses the challenges of unsupervised fine-tuning and significantly improves model performance. The fine-tuned IndoSBERT model with $\alpha = 0.7$ and $\beta = 0.3$ achieved an average SMAPE score of 8.9418%, outperforming the fine-tuned paraphrase-multilingual-mpnet-v2 model, which achieved 9.8390% with $\alpha = 0.6$ and $\beta = 0.4$. This study demonstrates the effectiveness of an Indonesian-language corpus-based model (IndoSBERT) for ASAG tasks and offers an innovative fine-tuning solution for unlabeled data.

Keywords : Automatic Short Answer Grading, SBERT, Clustering, Triplet Loss, SMAPE