# ABSTRACT

One of the challenges in implementing Automatic Short Answer Grading (ASAG) is developing a system that assesses answers across various domains without relying on domain-specific datasets. However, the biggest obstacle is the limited availability of training data, which hinders optimal model performance. A common solution is few-shot learning (FSL), but its application in ASAG remains rare. Most studies have yet to fully replicate real-world scenarios where only a small amount of training data is available. A previous study applied FSL through transfer learning by pre-training on Wikipedia data but still requiring a large amount of training data. This study aims to develop an FSL-based ASAG model using transfer learning with a more limited amount of training data. Three models—Pseudo-Siamese Neural Network (PSNN), Siamese BiLSTM, and BiLSTM-Parallel CNN—were compared. The models were evaluated in two scenarios: 20% training+validation data and 10% training+validation data, using four different datasets. SMAPE values were averaged for comparison. Experimental results show that the optimal scenario is 20% training+validation data. PSNN consistently outperformed the other models, with only a 1.6133% SMAPE increase when reducing training data. Meanwhile, Siamese BiLSTM's performance declined significantly with a 4.6597% SMAPE increase, while BiLSTM-Parallel CNN performed the worst, with an 11.3213% increase. This study demonstrates that PSNN with FSL is an effective ASAG approach, overcoming training data limitations while offering a more practical real-world solution.

**Keywords** : Automatic Short Answer Grading, BiLSTM, CNN, few shot learning, Pseudo-Siamese Neural Network