

Machine Reading Comprehension untuk Teks Bahasa Indonesia

Zendy Bramantia Alfareza¹, Ade Romadhony²

^{1,2}Fakultas Informatika, Universitas Telkom, Bandung

¹zendybramantia@student.telkomuniversity.ac.id, ²aderomadhony@telkomuniversity.ac.id

Abstract

Reading is a crucial skill that is essential for everyone to master. A commonly used method to enhance reading comprehension abilities is through reading comprehension lessons. To facilitate the creation of reading comprehension assessments, there exists a task in the field of Natural Language Processing (NLP) known as machine reading comprehension (MRC). machine reading comprehension serves as a fundamental task in question answering (QA), where each question is provided with relevant context to predict its answer. The goal of machine reading comprehension is to predict the correct answers from the given context, and in some cases, even generate more complex answers based on the provided context. The model employed in this study is the BERT model, which has been pre-trained to understand Indonesian language texts, also known as IndoBERT. This research employs a dataset consisting of 99 reading comprehension questions. The testing results demonstrate that using a learning rate of $1e-5$, the answer predictions in this reading comprehension task achieve the best performance with an accuracy of 75% and an F1 score of 92%.

Keywords: BERT, IndoBERT, Reading Comprehension, QA, NLP

