

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

Abstrak

Membaca merupakan suatu keterampilan yang sangat penting dan harus dikuasai oleh semua orang. Metode yang sering digunakan dalam meningkatkan kemampuan pemahaman membaca adalah pelajaran pemahaman membaca (reading comprehension). Dalam mempermudah pembuatan *assessment reading comprehension* terdapat *task* pada bidang *Natural Language Processing (NLP)* yakni *Machine Reading Comprehension (MRC)*. MRC adalah *task* dasar dari *question answering (QA)*, di mana pada setiap pertanyaan diberikan konteks terkait untuk memprediksi jawabannya. Tujuan MRC adalah untuk memprediksi jawaban yang benar dari konteks yang diberikan atau bahkan menghasilkan jawaban yang lebih kompleks berdasarkan konteks yang diberikan. Model yang digunakan merupakan model BERT yang sudah dilatih untuk memahami teks bahasa Indonesia atau bisa disebut dengan IndoBERT. Penelitian ini menggunakan dataset berupa 99 soal reading comprehension. Hasil pengujian menunjukkan bahwa pada learning rate $1e-5$, prediksi jawaban pada tugas pemahaman baca ini mencapai kinerja terbaik dengan akurasi sebesar 75% dan skor F1 sebesar 92%.

Kata kunci : BERT, IndoBERT, *Reading Comprehension*, QA, NLP

