

Abstrak

Proses pencarian informasi pada dokumen dengan cara membaca seluruh isi dokumen membutuhkan waktu yang relatif lama. Untuk membantu pencarian informasi yang relevan pada dokumen atau teks, penelitian ini mengembangkan Question Answering System (QAS) pada dokumen bahasa Indonesia dengan arsitektur *retriever-reader*. *Retriever* bertujuan untuk mengambil dokumen yang relevan dengan pertanyaan yang diberikan, sementara *reader* bertujuan untuk mengekstrak jawaban dari dokumen yang diberikan *retriever*. Pada penelitian ini, komponen *retriever* dibangun dengan TF-IDF yang akan mengembalikan dokumen paling relevan dengan pertanyaan berdasarkan nilai *cosine similarity*, sedangkan komponen *reader* dibangun dengan *fine-tuned* IndoBERT untuk *question answering*. Evaluasi dilakukan pada dataset SQuAD yang diterjemahkan ke Bahasa Indonesia, dengan metrik evaluasi EM dan F1-score.

Kata Kunci: *Question Answering System, IndoBERT, TF-IDF, cosine similarity*