

ABSTRAK

Pertanyaan memainkan peran penting dalam proses belajar-mengajar. Namun, menyusun pertanyaan secara manual memerlukan banyak waktu. Salah satu jenis pertanyaan, yaitu pilihan ganda (Multiple Choice Questions/MCQs), telah menjadi bidang penelitian yang aktif dalam beberapa tahun terakhir. Meskipun studi di bidang ini bukanlah hal baru, sebagian besar penelitian lebih berfokus pada bahasa Inggris. Berdasarkan studi literatur penulis, penelitian yang secara khusus membahas tentang pembuatan MCQ dalam bahasa Indonesia masih terbatas. Dalam penelitian ini, kami mengusulkan sistem MCQ bahasa Indonesia secara menyeluruh dengan mengeksplorasi pendekatan pembelajaran single-task dan multi-task. Sistem ini terdiri dari dua komponen yang saling terhubung: Question Generation (QG) dan Distractor Generation (DG). Performa dari kedua komponen ini dievaluasi menggunakan metrik BLEU, ROUGE-L, SBERT, dan analisis kesalahan. Selain itu, hasil MCQ yang dihasilkan juga dievaluasi oleh 22 responden manusia. Hasil penelitian menunjukkan bahwa skor ROUGE-L dan persentase relevansi untuk Single-Task Question Generation dan Multi-Task Question Generation masing-masing adalah 37 dengan 46% dari 100 pertanyaan dianggap relevan, serta 19 dengan 65% dari 100 pertanyaan dianggap relevan. Sementara itu, Single-Task Distractor Generator dan Multi-Task Distractor Generator masing-masing memperoleh rata-rata skor ROUGE-L sebesar 14.38 dengan 32% dari 300 distraktor dianggap relevan, serta 14.36 dengan 27.3% dari 300 distraktor dianggap relevan. Selain itu, hasil evaluasi manusia terhadap konfigurasi MCQ terbaik menunjukkan skor 1.38/5 untuk tingkat kesulitan, 3.99/5 untuk kejelasan, dan 2.07/5 untuk kekakuan (semakin rendah, semakin baik).

Kata kunci: Soal Pilihan Ganda dalam Bahasa Indonesia, Pembuatan Pertanyaan, Pembuatan Distraksi, Pemrosesan Bahasa Alami, Pembuatan Teks, T5, Pengujian oleh Manusia