

Pembangkit Pertanyaan Otomatis Berbahasa Indonesia dengan Metode IndoBERT dan Transformasi T5

Oleh: andaru Jati, Nurkhamid, Ratna Wardani, Yuniar Indrihapsari, Pradana Setialana, Dhista Dwi Nur Ardiansyah, Satya Adhiyaksa Ardy, Widya Ardiyanto, Maria Charlotta Bernadetha Wonda Tiala, Andi Hakim Al-khawarizmi, Damar Albaribin Syamsu, Zainal Ma'ruf Abidin

ABSTRAK

Pembuatan pertanyaan secara manual adalah proses yang membosankan dan rumit. Metode Pembuatan Pertanyaan Otomatis (AQG) berupaya mengurangi biaya-biaya ini dan memenuhi kebutuhan akan pasokan pertanyaan baru yang terus-menerus. Teknik AQG saat ini menggunakan arsitektur yang rumit, yang memerlukan sumber daya komputasi intensif serta pemahaman subjek yang lebih mendalam. Dalam makalah ini kami mengusulkan sistem AQG end-to-end yang memanfaatkan kekuatan transformator yang baru diperkenalkan, Text-to-Text Transfer Transformer (T5). Kami menggunakan model T5 yang telah dilatih sebelumnya dan menyempurnakannya untuk tugas hilir pembuatan pertanyaan. Model kami berkinerja sangat baik pada data yang tidak terlihat dan menghasilkan pertanyaan yang disusun dengan baik dan benar secara tata bahasa. Pertanyaan-pertanyaan ini dapat digunakan secara langsung oleh siswa, untuk menguji tingkat pemahaman mereka sendiri, dan guru untuk dengan cepat memperkuat konsep-konsep kunci kapan pun diperlukan. Model tersebut juga telah diterapkan dalam bentuk aplikasi web untuk akses publik. Aplikasi ini berfungsi sebagai alat pendidikan yang dapat digunakan oleh setiap individu untuk menilai pengetahuan mereka dan mengidentifikasi bidang-bidang perbaikan. Model ini mencapai skor BLEU-1, BLEU-2, BLEU-3, BLEU-4, dan ROUGE-L masing-masing sebesar 36,54, 28,24, 22,61, 18,44, dan 39,57. Model kami berkinerja baik dan menghasilkan pertanyaan dalam bahasa Indonesia yang mudah dipahami dengan pilihan kata dan tata bahasa yang baik berdasarkan validasi manual.

Kata Kunci: AQG, NLP, pembangkit pertanyaan, IndoBERT, transformasi, T5