

Pembangkitan Parafrase Untuk Reading Comprehension

Faishal Januarahman¹, Ade Romadhony²

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

¹faishaljanuarahman@students.telkomuniversity.ac.id, ²aderomadhony@telkomuniversity.ac.id,

Abstrak

Reading comprehension merupakan assessment yang menguji pembaca untuk memahami suatu konsep dari suatu bacaan. Proses pengujian dilakukan dengan memberikan pertanyaan yang berhubungan dengan isi pada konteks bacaan. Tujuan dari penelitian ini adalah untuk menciptakan variasi pertanyaan baru dari pertanyaan yang sudah ada, yang dimana salah satu caranya adalah melakukan proses parafrase pertanyaan dengan task *paraphrase generation*. Hal tersebut dapat memastikan bahwa pembaca telah sepenuhnya memenuhi suatu konsep bacaan yang diberikan. Penelitian ini menerapkan pendekatan tradisional yakni *thesaurus-based* yang dimana proses nya adalah melakukan *synonym substitution* yang diambil dari kamus tesaurus Bahasa Indonesia. Data yang digunakan merupakan daftar pertanyaan *assessment reading comprehension* berbahasa Indonesia dari level SD sampai SMA. Untuk mengukur hasil pertanyaan parafrase yang dibangkitkan dilakukan dua proses evaluasi, yakni *automatic evaluation* dengan skala 0-1, dan *human evaluation* dengan skala 1-4. *Automatic evaluation* terdiri dari metric BLEU-4 yang menghasilkan nilai 0.044 dan ROUGE metric yang menghasilkan nilai F1-score 0.421. Pada proses *human evaluation*, nilai *relevancy* yang didapat adalah 2.533 dan *fluency* bernilai 3.186. Hasil dari kedua evaluation metric tersebut merepresentasikan bahwa pertanyaan parafrase yang dibangkitkan memiliki keberagaman kata-kata baru, namun memiliki makna yang cenderung sedikit berbeda dari pertanyaan referensi.

Kata kunci : BLEU, Human Evaluation, Paraphrase Generation, ROUGE, Reading Comprehension, Tesaurus
