

## ABSTRACT

Automatic Item Generation (AIG) is a system to generate questions automatically with a model using computer technology. AIG can help teachers to provide questions in education assessments more efficiently. It has been implemented in various subjects and specifically researched in vocabulary tests. While most AIG studies focus on semantic similarity between correct answers and distractors, vocabulary tests need to prioritize paradigmatic relations due to their strong connection to cognitive ability.

The generation of distractors with paradigmatic relations in English vocabulary tests involves three phases: candidate generation, candidate filtering using Part of Speech Tagging (POS Tagging), and candidate ranking based on cosine similarity with the correct answer. WordNet is used to retrieve candidates, which are then converted into vectors using fastText for calculating cosine similarity.

The study has effectively developed a method to generate distractors for English vocabulary tests that exhibit paradigmatic relations with the correct answer. Expert evaluation confirmed that 102 out of 105 generated distractors demonstrated paradigmatic relations. The overall accuracy of the questions was 100%. Accuracy refers to the condition where none of the distractors in a particular question can be deemed as correct answers. Additionally, the evaluation of the fastText model demonstrated impressive performances in word similarity (70,06%), word analogy with 3CosAdd technique (20,78%), and word analogy with 3CosMul technique (55,62%). The word analogy results indicate that the fastText model is capable of predicting word equivalents, ultimately leading to the generation of accurate distractors.

**Keywords:** assessment, automatic item generation, English vocabulary test, fastText, WordNet

## INTISARI

*Automatic Item Generation* (AIG) merupakan sebuah sistem yang dapat membangkitkan soal secara otomatis menggunakan model dengan bantuan teknologi komputer. AIG dapat membantu pendidik menyiapkan soal yang akan digunakan dalam asesmen menjadi lebih efisien. Saat ini, AIG telah diimplementasikan pada berbagai bidang dan secara khusus diteliti dalam *vocabulary test*. Mayoritas penelitian AIG dalam *vocabulary test* berfokus pada *semantic similarity* antara *correct answer* dengan *distractor*. Di sisi lain, *vocabulary test* perlu memprioritaskan relasi *paradigmatic* antara *correct answer* dengan *distractor* karena relasi tersebut memiliki keterkaitan yang kuat dengan kemampuan kognitif.

Pembangkitan *distractor* yang memiliki relasi *paradigmatic* dengan *correct answer* dilakukan dalam tiga tahap: *candidate generation*, *candidate filtering* menggunakan *Part of Speech Tagging* (POS Tagging), dan *candidate ranking* berdasarkan nilai *cosine similarity* dengan *correct answer*. *WordNet* digunakan untuk mencari kandidat *distractor* yang kemudian kandidat dikonversi ke dalam vektor menggunakan fastText untuk menghitung nilai *cosine similarity*.

Penelitian ini menghasilkan metode yang dapat secara efektif membangkitkan *distractor* pada *English vocabulary test* yang memiliki relasi *paradigmatic* dengan *correct answer*. Penilaian ahli menunjukkan 102 dari 105 *distractor* yang dibangkitkan memiliki relasi *paradigmatic*. Secara keseluruhan soal memiliki tingkat akurasi 100%. Akurasi mengacu pada kondisi di mana tidak ada *distractor* yang dapat dipertimbangkan sebagai *correct answer*. Selain itu, evaluasi model fastText menunjukkan performa yang baik dalam *word similarity* (70,06%), *word analogy* dengan teknik 3CosAdd (20,78%), dan *word analogy* dengan teknik 3CosMul (55,62%). Hasil evaluasi *word analogy* tersebut menunjukkan bahwa model fastText mampu memprediksi padanan kata, yang pada akhirnya menghasilkan *distractor* yang akurat.

**Kata kunci** – asesmen, *automatic item generation*, *English vocabulary test*, *fastText*, *WordNet*