

ABSTRAK

ALFINA DAMAYANTI, Penerapan *Large Language Model* (LLM) untuk Pencarian Hubungan Makna Antar Kata Pada Kamus Bahasa Bugis.

(Dibimbing oleh Chyquitha Danuputri, S.Kom., M.Kom. dan Muhyiddin A M Hayat, S.Kom., M.T.).

Pelestarian bahasa daerah merupakan bagian penting dalam menjaga identitas budaya, termasuk Bahasa Bugis yang memiliki kekayaan kosakata dan nilai budaya yang tinggi. Namun, kamus Bahasa Bugis yang ada saat ini umumnya masih bersifat leksikal dan belum mampu merepresentasikan hubungan makna antar kata secara semantik. Penelitian ini bertujuan untuk menerapkan *Large Language Model* (LLM) dalam pencarian hubungan makna antar kata pada Kamus Bahasa Bugis guna menghasilkan pencarian yang lebih kontekstual dan bermakna. Metode penelitian yang digunakan meliputi pengumpulan dataset kamus Bahasa Bugis dari Balai Bahasa Provinsi Sulawesi Selatan, pra-pemrosesan data teks, serta penerapan model LLM berbasis sentence embedding untuk merepresentasikan kata dalam bentuk vektor semantik. Hubungan makna antar kata dianalisis menggunakan perhitungan kesamaan semantik berbasis cosine similarity yang menghasilkan relevance score sebagai indikator tingkat kedekatan makna. Evaluasi sistem dilakukan dengan menganalisis distribusi relevance score dan penerapan threshold evaluasi untuk menyaring hasil pencarian yang paling relevan. Hasil penelitian menunjukkan bahwa penerapan LLM mampu menghasilkan representasi semantik yang efektif dalam menangkap hubungan makna antar kata Bahasa Bugis, seperti relasi keluarga, arah mata angin, dan waktu. Sistem mampu menampilkan kata-kata yang memiliki keterkaitan makna secara kontekstual, tidak hanya berdasarkan kesamaan ejaan atau terjemahan literal. Dengan demikian, penelitian ini membuktikan bahwa LLM berpotensi menjadi pendekatan yang efektif dalam pengembangan kamus digital Bahasa Bugis yang lebih semantik-aware dan mendukung upaya pelestarian bahasa daerah di era kecerdasan buatan.

Kata kunci: Large Language Model, Bahasa Bugis, Kamus Digital, Hubungan Makna, Semantic Similarity

ABSTRACT

ALFINA DAMAYANTI, *Implementation of Large Language Models (LLM) for Semantic Relationship Search Between Words in the Bugis Language Dictionary.*
(Supervised by Chyquitha Danuputri, S.Kom., M.Kom. and Muhyiddin A M Hayat, S.Kom., M.T.).

The preservation of regional languages is an essential aspect of maintaining cultural identity, including the Bugis language, which possesses rich vocabulary and strong cultural values. However, existing Bugis dictionaries are generally lexical in nature and lack the ability to represent semantic relationships between words. This study aims to apply Large Language Models (LLMs) to semantic relationship retrieval in a Bugis language dictionary in order to provide more contextual and meaningful search results. The research methodology includes collecting Bugis dictionary data from the Provincial Language Center of South Sulawesi, performing text preprocessing, and applying an LLM-based sentence embedding model to represent words as semantic vectors. Semantic relationships between words are analyzed using cosine similarity to generate relevance scores that indicate the degree of semantic closeness. System evaluation is conducted by analyzing the distribution of relevance scores and applying an evaluation threshold to filter the most relevant results. The results demonstrate that the application of LLMs effectively captures semantic relationships between Bugis words, including family relations, cardinal directions, and temporal concepts. The system is able to retrieve contextually related words rather than relying solely on lexical similarity or literal translations. Therefore, this research confirms that LLMs are a promising approach for developing semantic-aware Bugis digital dictionaries and contribute to the preservation of regional languages in the era of artificial intelligence.

Keywords: *Large Language Model, Bugis Language, Digital Dictionary, Semantic Relationship, Semantic Similarity*