

## ABSTRAK

**Nama** : Mohammad Adzriel Ramdhani  
**NIM** : 1197010044  
**Judul** : *Semantic Similarity Word Embeddings* dengan *Clustering CLARANS* menggunakan Data Terjemahan Al-Qur'an Bahasa Inggris

Al-Qur'an adalah kitab suci agama Islam, yang diwahyukan dalam bahasa Arab dan dianggap sebagai panduan hidup yang sempurna. Namun, bagi mereka yang tidak mahir dalam bahasa Arab, memahami makna asli dari Al-Qur'an bisa menjadi tantangan, sehingga penerjemahan ke dalam bahasa Inggris menjadi sangat penting. Terjemahan ini juga menghadirkan tantangan dalam memahami makna kata-kata, yang dapat diatasi melalui sistem pengelompokan kata. *Word embeddings* digunakan untuk mengubah kata menjadi *vector* dalam ruang *vector* berdimensi tinggi, dengan melihat kemiripan kata dalam Al-Qur'an *vector* akan memiliki nilai dan menempatkan titik *vector* berdekatan dengan *vector* yang memiliki konteks kata yang mirip. Algoritma CLARANS (*Clustering Large Applications based on Randomized Search*) dipilih karena kemampuannya untuk menangani dataset yang besar dan kompleks dengan melakukan pencarian acak secara efisien untuk menemukan pusat kluster yang optimal. Penelitian ini bertujuan untuk menerapkan algoritma *word2vec* dan CLARANS pada pengelompokan kata dalam terjemahan bahasa Inggris Al-Qur'an dan mengevaluasi efektivitas pengelompokan CLARANS dengan melihat kesamaan semantik dengan nilai *wordnet*. Dengan demikian, penerapan algoritma ini diharapkan dapat memberikan kontribusi yang signifikan terhadap studi Al-Qur'an, terutama dalam memfasilitasi pemahaman makna dalam terjemahan bahasa Inggris Al-Qur'an.

**Kata Kunci** : *Word embeddings, Word2vec, Algoritma CLARANS, Semantic Similarity, Wordnet*

## ABSTRACT

**Name** : Mohammad Adzriel Ramdhani

**NIM** : 1197010044

**Thesis Title** : *Semantic Similarity Word Embeddings dengan Clustering CLARANS menggunakan Data Terjemahan Al-Qur'an Bahasa Inggris*

The Qur'an is the holy book of Islam, revealed in Arabic and regarded as a perfect guide for life. However, for those who are not proficient in Arabic, understanding the original meaning of the Qur'an can be challenging, making translation into English very important. This translation also presents challenges in understanding the meaning of words, which can be addressed through word grouping systems. *Word embeddings* are used to convert words into *vectors* in a high-dimensional *vector* space, where similar words in the Qur'an will have *vector* values that place their points close to other *vectors* with similar word contexts. The CLARANS algorithm (*Clustering Large Applications based on Randomized Search*) was chosen for its ability to efficiently handle large and complex datasets by conducting randomized searches to find optimal cluster centers. This study aims to apply the *word2vec* and CLARANS algorithms to word clustering in the English translation of the Qur'an and to evaluate the effectiveness of CLARANS clustering by examining semantic similarity with *wordnet* values. Thus, the application of these algorithms is expected to make a significant contribution to Qur'anic studies, particularly in facilitating the understanding of meanings in the English translation of the Qur'an.

**Keywords** : *Word Embeddings, Word2vec, CLARANS algorithm, Semantic Similarity, Wordnet*