

ABSTRACT

Nama : Muhammad Saqti Romadhon

NIM : 1187010057

Title : Mengukur Kemiripan Kata (*Semantic Similarity*) Terjemahan Al-Qur'an Bahasa Indonesia menggunakan Penyajian *Word2Vec*

This study aims to analyze the Semantic Similarity between words in the Indonesian translation of the Qur'an using the Word2Vec model with the Continuous Bag of Words (CBOW) architecture. The research is motivated by the need to understand semantic relationships in religious texts with unique structure and contexts. The dataset consists of the official Indonesian Qur'an translation released by the Ministry of Religious Affairs, sourced from tanzil.net. The model is trained using the Hierarchical Softmax approach to reduce computational complexity, and experiments are conducted with various training parameters such as vector size and context window size. Preprocessing steps include tokenization, case folding, symbol removal, and stemming using the Sastrawi library. The evaluation includes calculating cosine similarity scores, average similarity, and standard deviation for selected Islamic keywords. Results indicate that increasing the vector size from 50 to 500 improves the average similarity from 0.7006 to 0.7100, with the optimal window size found at WS=11. The t-SNE visualization reveals clustering of semantically related words, confirming the model's ability to capture meaningful associations. In conclusion, the CBOW-based Word2Vec model with optimal parameters effectively represents semantic meaning in the context of the Qur'an translation corpus semantically and consistently.

Keywords: Word2Vec, CBOW, Semantic Similarity, Qur'an, Hierarchical Softmax, Cosine Similarity.

KATA PENGANTAR

Bismillaahirrahmaanirrahiim

Puji serta syukur penulis sampaikan kehadiran Allah SWT yang telah memberikan rahmat, taufik, hidayah dan inayah-Nya. Shalawat serta salam semoga selalu tercurah limpahkan kepada Nabi Muhammad SAW beserta keluarga, sahabat dan kita semua sebagai umatnya hingga akhir zaman.

Alhamdulillah atas izin dan karunia Allah SWT penulis dapat menyelesaikan skripsi yang berjudul **“Kemiripan Kata (*Semantic Similarity*) pada Terjemahan Al-Qur’an Berbahasa Indonesia dengan Kata yang Disajikan dalam Bentuk Vektor dari *Word2Vec*”**. Dalam penulisan skripsi ini, penulis sangat menyadari bahwa tidak akan tercapainya hasil yang maksimal tanpa bimbingan, motivasi, saran dan do’a dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis mengucapkan banyak terima kasih kepada :

1. Allah *subhannahu wa ta’ala*, yang telah melimpahkan rahmat, nikmat serta karunianya yang tak terhingga.
2. Kedua orang tua yang selalu memberikan do’a, semangat, dukungan, dan kasih sayang tiada henti.
3. Keluarga tercinta yang tidak pernah lelah memberikan doa, motivasi dan semangat.
4. Bapak Asep Solih Awalluddin, M.Si, selaku dosen pembimbing akademik, dosen pembimbing II, dan Ketua Jurusan Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Gunung Djati Bandung yang telah memberikan bimbingan, motivasi, serta nasihatnya kepada penulis.
5. Bapak Dr. Arief Fatchul Huda, S.Si., M.Kom, selaku dosen pembimbing I yang selalu memberikan bimbingan, arahan, motivasi, serta nasihatnya kepada penulis.
6. Seluruh dosen dan staf jurusan Matematika Fakultas Sains dan Teknologi UIN Sunan Gunung Djati Bandung yang telah menyampaikan ilmunya kepada penulis.