

ABSTRAK

Zaki Maulana Hidayat
NPM 140110190049

Pada zaman ini, penggunaan teknologi dalam membantu aktivitas manusia semakin meningkat. Hal tersebut memicu penelitian-penelitian yang berguna untuk meningkatkan kualitas dari teknologi tersebut. Salah satu topik penelitian pada bidang tersebut yang sering dibahas adalah *word embedding*. *Word embedding* merupakan salah satu metode dalam memetakan sebuah kata ke dalam vektor representasinya dengan mempertimbangkan jarak dengan vektor representasi kata yang lain. Dalam *word embedding*, semakin dekat jarak antara dua vektor representasinya, maka makna antara dua kata yang direpresentasikan akan semakin dekat. Penelitian pada skripsi ini berfokus pada penerapan *word embedding* yang terjadi antar bahasa. Hal tersebut dapat dilakukan dengan menerapkan konsep transformasi linear ke dalam pemetaan vektor-vektor representasi kata bahasa pertama ke vektor-vektor terjemahannya dalam bahasa kedua. Transformasi linear dalam penelitian ini diperoleh dengan menggunakan metode *singular value decomposition*. Proses dimulai dengan memetakan kata-kata ke vektor representasinya yang kemudian disusun ke dalam sebuah matriks untuk setiap bahasa. Transformasi linear yang diperoleh kemudian diterapkan pada contoh kasus sederhana dan penerapannya dengan bantuan bahasa pemrograman Python. Hasil dari kajian penelitian ini adalah transformasi linear yang memetakan vektor representasi kata antara bahasa pertama dengan bahasa kedua dapat diperoleh dengan menggunakan metode *singular value decomposition* dengan terlebih dahulu memperoleh matriks yang berisikan vektor-vektor baris dari masing-masing bahasa.

Kata kunci : transformasi linear; *word embedding*; *singular value decomposition*.

ABSTRACT

Zaki Maulana Hidayat
NPM 140110190049

In this era, the use of technology in helping human activities is increasing. This triggers many useful research to improve the quality of that technology. One of the research topics in this field that is often discussed is word embedding. Word embedding is one of the methods that can map a word into its representation vector by considering the distance with representation vector of the other words. In word embedding, the closer the distance between the two representation vectors, the closer the meaning between the two words represented will be. The research in this thesis focuses on the application of word embedding that occurs between languages. This can be done by applying the concept of linear transformation to the mapping of first-language word representation vectors to their translation word vectors in the second language. The linear transformation in this research is obtained by using the singular value decomposition method. The process starts by mapping words to their representation vectors which are then organized into a matrix for each language. The linear transformation that has been obtained then applied to a simple case example and its application with the help of Python. The result of this study research is a linear transformation that maps the word representation vectors between the first language and the second language can be obtained using the singular value decomposition method by first obtain a matrix that contain the representation vectors of each language.

Keywords : *linear transformation; word embedding; singular value decomposition*