WORD EMBEDDING PADA SENTIMEN ANALASIS UNTUK MEREPRESENTASIKAN KATA MENJADI VEKTOR

SKRIPSI

disusun oleh

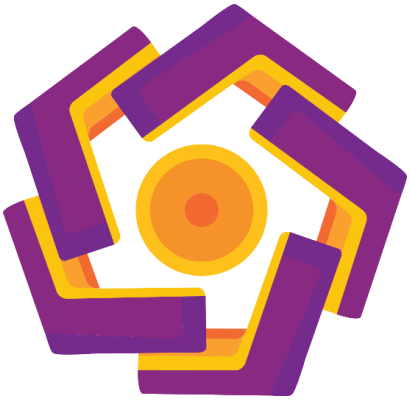
Dimas Midyan Alam
17.11.0982

PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021

WORD EMBEDDING PADA SENTIMEN ANALASIS UNTUK MEREPRESENTASIKAN KATA MENJADI VEKTOR

SKRIPSI

Untuk memenuhi sebagaian persyaratan
mencapai gelar Sarjana
pada Program Studi Informatika



disusun oleh

Dimas Midyan Alam
17.11.0982

PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021

PERSETUJUAN

SKRIPSI

WORD EMBEDDING PADA SENTIMEN ANALISIS UNTUK MEREPRESENTASIKAN KATA MENJADI VEKTOR

yang dipersiapkan dan disusun oleh

Dimas Midyan Alam

17.11.0982

telah disetujui oleh Dosen Pembimbing Skripsi

pada tanggal 29 Januari 2021

Dosen Pembimbing,

Mardhiya Hayaty, S.T., M.Kom.,

NIK : 190302108

PENGESAHAN

SKRIPSI

WORD EMBEDDING PADA SENTIMEN ANALASIS UNTUK MEREPRESENTASIKAN KATA MENJADI VEKTOR

yang dipersiapkan dan disusun oleh

Dimas Midyan Alam

17.11.0982

telah dipertahankan di depan Dewan Penguji

pada tanggal 17 Februari 2021

Susunan Dewan Penguji

Nama Penguji	**Tanda Tangan**
Dina Maulina, M.Kom. **NIK : 190302250**	
Nuraini, M.Kom. **NIK : 190302066**	
Mardhiya Hayaty, S.T., M.Kom. **NIK. 190302108**	

Skripsi ini telah diterima sebagai salah satu persyaratan

untuk memperoleh gelar Sarjana Komputer

Tanggal 17 Februari 2021

DEKAN FAKULTAS ILMU KOMPUTER

Krisnawati, S.Si, M.T.
NIK. 190302038

PERNYATAAN

Saya yang bertanda tangan di bawah ini menyatakan bahwa skripsi ini merupakan karya saya sendiri (ASLI) dan isi pada skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar Pustaka.

Segala sesuatu yang terkait dengan masalah dan karya yang telah dibuat adalah menjadi tanggung jawab saya pribadi.

Yogyakarta, 5 Maret 2021

Dimas Midyan Alam

NIM 17.11.0982

MOTTO

“Setiap manusia memiliki jalan ceritanya masing-masing, jangan pernah samakan jalan ceritamu dengan jalan cerita orang lain”

- Dimas Midyan Alam -

PERSEMBAHAN

Puji syukur saya panjatkan kepada ALLAH SWT atas berkah dan karunia yang telah diberikan, serta junjungan Nabi Besar kita Nabi Muhammad SAW, sehingga skripsi ini terselesaikan dengan baik dan lancar. Dengan ini saya persembahkan skripsi ini kepada semua pihak yang terlibat secara langsung maupun tidak langsung, yaitu kepada :

1. Kedua orang tua saya (Jeng Yanti dan Abdul Hamid), kedua kakak saya (Dewi dan Aby), bude saya (Bude Nuk), dan orang-orang disekitar saya yang tidak dapat disebutkan satu persatu. Selalu mendoakan saya, selalu mensuport saya dalam mengerjakan skripsi ini, dan selalu memberikan motivasi untuk maju terus pantang mundur.
2. Dosen pembimbing saya Ibu Mardhiya Hayaty, S.T., M.Kom., yang telah membimbing saya dari awal sampai akhir pembuatan skripsi.
3. Dosen - dosen Universitas Amikom Yogyakarta yang telah memberikan banyak ilmu pengetahuan dari semester awal hingga akhir.
4. Teman - teman khususnya Kelas 17-IF-01 yang telah menemani dan selalu memberikan banyak cerita setiap harinya dalam perkuliahan selama 6 semester.
5. Teman teman Uwuhnisty (Tedy, Seto, Anif, Sileh, Alwi, Bijis, Daniel, Manul, Agus, Farhan, Dito, Uwuh, Deri, dan Andri) yang sxesalu memberikan support kepada saya,

KATA PENGANTAR

Penulis panjatkan puji dan syukur kepada ALLAH SWT atas berkat dan rahmat-Nya, sehingga penulis dapat menyelesaikan skripsi dengan baik yang berjudul **"WORD EMBEDDING PADA SENTIMEN ANALISIS UNTUK MEREPRESENTASIKAN KATA MENJADI VEKTOR"** disusun sebagai salah satu syarat utama untuk menyelesaikan program sarjana pada Universitas AMIKOM Yogyakarta. Penyelesaian skripsi ini juga tidak lepas dari bantuan berbagai pihak, karena itu pada kesempatan ini penulis ingin menyampaikan rasa hormat dan terima kasih kepada :

1. Prof. Dr. M. Suyanto, MM. selaku Rektor Universitas AMIKOM Yogyakarta.
2. Ibu Krisnawati, S.Si, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Bapak Sudarmawan, M.T. selaku Ketua Program Studi Informatika Universitas AMIKOM Yogyakarta.
4. Ibu Mardhiya Hayaty, S.T., M.Kom. selaku dosen pembimbing yang selalu meluangkan tenaga, waktu, dan pikirannya dalam membimbing saya.
5. Ibu Dina Maulina, M.Kom. dan Ibu Nuraini, M.Kom. selaku dosen penguji. Terimakasih atas segala saran yang diberikan selama pengujian untuk memperbaiki penelitian menjadi lebih baik lagi.

Penulis menyadari bahwa skripsi ini masih banyak kekurangan. Maka, penulis menerima segala kritik dan saran yang membangun dari semua pihak. Semoga skripsi ini bisa bermanfaat baik bagi penulis serta pembaca. Atas saran dan kritik, penulis ucapkan terima kasih.

Yogyakarta, 17 Februari 2021

Dimas Midyan Alam

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

INTISARI

Analisis sentimen adalah sebuah proses yang memahami, mengekstraksi, dan mengolah data teks secara otomatis untuk menemukan jenis sentimen pada teks tersebut, merupakan salah satu solusi mengatasi masalah untuk mengelompokan opini atau review menjadi opini positif atau negatif secara otomatis Analisis sentimen berguna untuk memudahkan pengguna pada proses memahami sentimen sehingga dapat melakukan penentuan keputusan pada suatu objek.

Pada penelitian ini menerapkan Word Embedding dengan Metode Word2Vec dan menggunakan algoritma klasifikasi Long Short Term Memory (LSTM). Metode Word2Vec digunakan untuk merepresentasikan kata kata dalam bentuk matematis. Word2Vec merupakan sebuah algoritma untuk mempelajari posisi kedekatan semantic antar kata dari sebuah teks masukan. Penelitian ini bertujuan untuk mengetahui berapa hasil akurasi *Word Embedding* dengan metode *Word2Vec* dan tanpa menggunakan *Word Embedding*.

Berdasarkan penelitian yang dilakukan dengan pembagian data *training* 80% dan data *testing* 20% dari total jumlah dataset 10022, diperoleh nilai akurasi analisis sentimen yang menggunakan word embedding adalah 0,92 atau 92%, sedangkan nilai akurasi tanpa menggunakan word embedding adalah 0,88 atau 88%.

Kata-kunci: Sentimen Analisis, Word2Vec, Word Embedding, LSTM, Cosine Similarity

ABSTRACT

Sentiment analysis is a process that understands, extracts, and processes text data automatically to find the type of sentiment in the text, which is a solution to the problem of grouping opinions or reviews into positive or negative opinions automatically Sentiment analysis is useful for making it easier for users to process Understand sentiment so that you can make decisions on an object.

This research applies Word Embedding with the Word2Vec method and uses the Long Short Term Memory (LSTM) classification algorithm. The Word2Vec method is used to represent words in mathematical form. Word2Vec is an algorithm for learning semantic closeness between words from an input text. This study aims to see the accuracy results of Word Embedding using the Word2Vec method and without using Word Embedding.

Based on research conducted by sharing 80% training data and testing data 20% of the total number of 10022 datasets, the value of sentiment analysis using word embedding was 0.92 or 92%, while the value without using word embedding was 0.88 or 88%.

Keywords: Sentimen Analisis, Word2Vec, Word Embedding, LSTM, Cosine Similarity