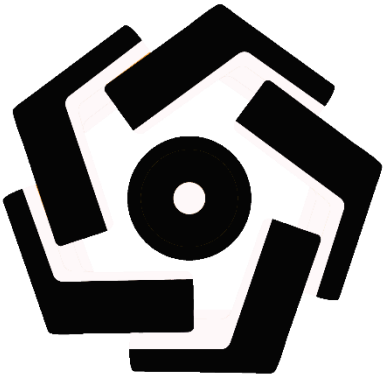
REPRESENTASI KATA MENJADI VECTOR PADA ULASAN HOTEL DI INDONESIA MENGGUNAKAN METODE WORD2VEC

SKRIPSI



Disusun oleh

Andrian Tri Muryanto

17.11.0977

PROGRAM SARJANA

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

2020

REPRESENTASI KATA MENJADI VECTOR PADA ULASAN HOTEL DI INDONESIA MENGGUNAKAN METODE WORD2VEC

SKRIPSI

Untuk memenuhi sebagaian persyaratan

Mencapai gelar Sarjana

pada Program Studi Informatika

Disusun oleh

Andrian Tri Muryanto

17.11.0977

PROGRAM SARJANA

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

2020

PERSETUJUAN

SKRIPSI

REPRESENTASI KATA MENJADI VECTOR PADA ULASAN HOTEL DI INDONESIA MENGGUNAKAN METODE WORD2VEC

yang dipersiapkan dan disusun oleh

Andrian Tri Muryanto

17.11.0977

telah disetujui oleh Dosen Pembimbing Skripsi

pada tanggal 26 Juni 2020

Dosen Pembimbing,

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

NIK : 190302108

PENGESAHAN

SKRIPSI

REPRESENTASI KATA MENJADI VECTOR PADA ULASAN HOTEL DI INDONESIA MENGGUNAKAN METODE WORD2VEC

yang dipersiapkan dan disusun oleh

Andrian Tri Muryanto

17.11.0977

telah dipertahankan di depan Dewan Penguji

pada tanggal 18 November 2020

Susunan Dewan Penguji

Nama Penguji	Tanda Tangan
Ainul Yaqin, M. Kom. **NIK : 190302098**	
M. Rudyanto Arief, S.T, M.T. **NIK : 190302098**	
Mardhiya Hayaty, S.T., M.Kom. **NIK. 190302108**	

Skripsi ini telah diterima sebagai salah satu persyaratan

untuk memperoleh gelar Sarjana Komputer

Tanggal .....

DEKAN FAKULTAS ILMU KOMPUTER

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

PERNYATAAN

Saya yang bertandatangan 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 tanggungjawab saya pribadi.

Yogyakarta, 21 November 2020

Andrian Tri Muryanto

NIM 17.11.0977

MOTTO

“What will people say, and I don’t care about that”

- deddy corbuzier –

“Biarkan orang lain menganggapmu rendah, karena mereka hanya memandang dirimu dari covernya saja”

- Andrian Tri Muryanto -

PERSEMBAHAN

Puji syukur saya panjatkan kepada Allah SWT atas berkah dan karunia-NYA 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 dan kakak kakak saya yang selalu mendoakan dan selalu mensupport saya dalam mengerjakan skripsi ini, dan selalu memberikan motivasi untuk maju terus.
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 Yogyalarta yang telah memberikan ilmu dari semester awal sampai akhir selama kuliah
4. Teman teman khususnya Kelas 17-IF-01 yang telah menemani dan selalu memberikan semangat juang dalam mengerjakan skripsi ini
5. Teman teman SRAWUNG CLUB (Damar Komir, Aldo bill bae, Dhimskut, Lutfi edan, Fariz paryono, Rizky Roy, Bappocx POY, Dalipon, Padhe Dedidot, Alif Part-min, Anjasopo anjay, Vina mbak pina pathok, IFA ipak ipol, RIA Riaw, Puspa jaelani, Ema Lia) yang sesalu memberikan support kepada saya.
6. Teman teman Uwuhnisty (Tedi beer, Setoin, Bangdim, Anip, Hendra Sileh, Alwi, Bijis, Daniel, A.K.A Manul, Farhan Koton, Deri OC) yang sesalu memberikan support kepada saya,

KATA PENGANTAR

Puji syukur kehadirat Allah SWT atas segala rahmat dan hidayah-NYA yang telah diberikan kepada peneliti sehingga dapat menyelesaikan penelitian ini. Dan tidak lupa kita panjatkan shalawat seta salam kepada junjungan nabi besar kita Nabi Muhammad SAW, yang telah menjadi suri tauladan yang baik bagi umatnya dan untuk berbuat kebijakan.

Skripsi yang berjudul **"REPRESENTASI KATA MENJADI VECTOR PADA ULASAN HOTEL DI INDONESISA MENGUNAKAN METODE WORD2VEC"** ini disusun sebagai salah satu syarat kelulusan bagi setiap mahasiswa Universitas AMIKOM Yogyakarta. Selain itu juga merupakan suatu bukti bahwa mahasiswa telah menyelesaikan kulah jenjang program Strata-1 dan untk memperoleh gelar Sarjana Komputer.

Penyelesaian skripsi ini juga tidak lepas dari bantukan berbagai pihak, karena itu pada kesempatan ini penulis menyampaikan rasa hormat dan terimakasih 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 bijaksana memberikan bimbingan, nasehat serta waktunya selama penulisan skripsi ini.
5. Bapak Ainul Yaqin, M. Kom dan Bapak M. Rudyanto Arief, S.T, M.T selaku dosen penguji, terimakasih atas saran yang diberikan selama pengujian untuk memperbaiki penelitian ini menjadi lebih baik lagi.

Peneliti menyadari bahwa pembuatan skripsi ini masih banyak kekurangan dan kelemahannya. Oleh karena itu peneliti berharap kepada semua pihak agar dapat menyampaikan kritik dan saran yang membangun untuk menambah kesempurnaan skripsi ini. Namun peneliti tetap berharap skripsi ini akan bermanfaat bafi semua pihak yang membacanya. Apa bila terdapat kesalahan semoga Allah SWT melimpahkan magfirah-NYA. *Aamiin yaa Kholi.*

Yogyakarta, 21 November 2020

Andrian Tri Muryanto

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

INTISARI

Hotel merupakan salah satu produk pariwisata yang sangat penting untuk dipertimbangkan baik dari segi fasilitas dan pelayanan. Saat ini sudah banyak website wisata yang menyediakan fasilitas internet untuk menuliskan opini dan pengalaman pribadinya secara online. Analisa sentimen atau opinion mining merupakan salah satu solusi mengatasi masalah untuk mengelompokan opini atau review menjadi opini positif atau negatif secara otomatis.

Analisis sentimen adalah sebuah proses yang memahami, mengekstraksi, dan mengolah data teks secara otomatis untuk menemukan jenis sentimen pada teks tersebut. Analisis sentimen berguna untuk memudahkan pengguna pada proses memahami sentiment sehingga dapat melakukan penentuan keputusan pada suatu objek.

Penelitian ini menerapkan Metode Word2Vec untuk dilakukannya kinerja sentiment analisis. Metode Word2Vec dapat digunakan untuk merepresentasikan kata kata dalam bentuk matematis. Word2Vec merupakan sebuah algoritma untuk mempelajari posisi kedekatan semantic antar kata dari sebuah teks masukan. Word2vec memiliki dua arsitektur yaitu Skip-gram dan CBOW .

Berdasarkan hasil penelitian diperoleh nilai akurasi analisis sentimen untuk ulasan hotel di Indonesia yang menggunakan word embedding adalah 0,90 atau 90%, sedangkan nilai akurasi dari tanpa word embedding adalah 0,88 atau 88%. Namun saat menggunakan word embedding, kecepatan epoch membutuhkan waktu proses yang lebih lama. Jika tidak menggunakan word embedding, kecepatan pemrosesan akan lebih cepat.

Kata Kunci : Analisis Sentimen, *Word Embedding*, *Word2Vec*, *LSTM*

ABSTRACT

Hotels are one of the tourism products that are very important to consider both in terms of facilities and services. Currently, there are many tourist websites that provide internet facilities to write opinions and personal experiences online. Sentiment analysis or opinion mining is one solution to overcome problems to group opinions or reviews into positive or negative opinions automatically.

Sentiment analysis is a process that understands, extracts, and processes text data automatically to find the type of sentiment in the text. Sentiment analysis is useful for making it easier for users in the process of understanding sentiment so that they can make decisions on an object.

This study applies the Word2Vec method to perform sentiment analysis performance. The Word2Vec method can be used to represent words in mathematical form. Word2Vec is an algorithm for studying semantic proximity positions between words from an input text. Word2vec has two architectures namely Skip-gram and CBOW.

Based on the results of the study, the accuracy value of sentiment analysis for hotel reviews in Indonesia that uses word embedding is 0.90 or 90%, while the accuracy value for no word embedding is 0.88 or 88%.

However, when using word embedding, epoch speed requires a longer processing time. If you don't use word embedding, the processing speed will be faster.

Keywords: *Sentiment Analysis, Word Embedding, Word2Vec, LSTM*