

**PERBANDINGAN METODE *WORD EMBEDDING* UNTUK ANALISIS  
SENTIMEN PADA DATA ULASAN MARKETPLACE**

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**Arfian Yogi Ferianto**

**20.21.1465**

Kepada

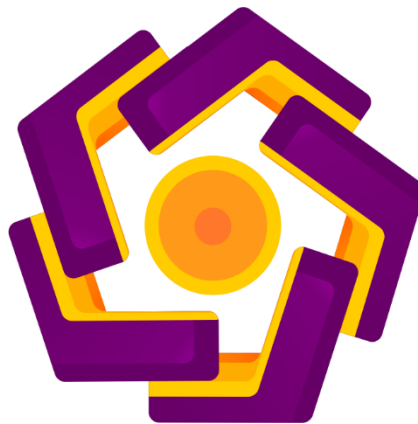
**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2022**

**PERBANDINGAN METODE *WORD EMBEDDING* UNTUK ANALISIS  
SENTIMEN PADA DATA ULASAN MARKETPLACE**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**Arfian Yogi Ferianto**

**20.21.1465**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2022**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**PERBANDINGAN METODE WORD EMBEDDING UNTUK ANALISIS  
SENTIMEN PADA DATA ULASAN MARKETPLACE**

yang disusun dan diajukan oleh

**Arfian Yogi Ferianto**

**20.21.1465**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 17 November 2022

**Dosen Pembimbing,**



**Mardhiya Hayaty, S.T., M.Kom.**  
**NIK. 190302108**

**HALAMAN PENGESAHAN**

**SKRIPSI**

**PERBANDINGAN METODE WORD EMBEDDING UNTUK ANALISIS  
SENTIMEN PADA DATA ULASAN MARKETPLACE**

yang disusun dan diajukan oleh

**Arfian Yogi Ferianto**

**20.21.1465**

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 17 November 2022

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Irma Rofni Wulandari, S.Pd., M.Eng.**  
**NIK. 190302329**

**Erni Seniwati, S.Kom., M.Cs.**  
**NIK. 190302231**

**Mardhiya Hayaty, S.T., M.Kom.**  
**NIK. 190302108**

Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 17 November 2022

**DEKAN FAKULTAS ILMU KOMPUTER**

**Hanif Al Fatta, S.Kom., M.Kom.**  
**NIK. 190302096**

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : **Arfian Yogi Ferianto**  
NIM : **20.21.1465**

Menyatakan bahwa Skripsi dengan judul berikut:

**Perbandingan Metode Word Embedding untuk Analisis Sentimen pada Data Ulasan Marketplace**

Dosen Pembimbing : **Mardhiya Hayaty, S.T., M.Kom.**

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 17 November 2022

Yang Menyatakan,



Arfian Yogi Ferianto

## HALAMAN PERSEMBAHAN

Puji syukur penulis panjatkan kehadirat Allah SWT yang telah melimpahkan rahmat dan hidayah serta karunia-Nya sehingga penulis dapat menyelesaikan Tugas Akhir ini dengan sebaik-baiknya dan mendapatkan hasil yang maksimal.

Tidak lepas dari bantuan berbagai pihak, oleh karena itu penulis ingin menyampaikan terima kasih kepada:

1. Allah SWT yang telah memberikan kelancaran kepada penulis dalam menyelesaikan Tugas Akhir ini.
2. Kedua orang tua, kakak, dan saudara-saudara yang telah memberikan do'a terbaiknya.
3. Bapak Ibu Dosen Universitas Amikom Yogyakarta yang telah ikhlas dan selalu sabar memberikan ilmu mulai dari semester awal hingga akhir.
4. Ibu Mardhiya Hayaty, S.T., M.Kom. sebagai dosen pembimbing yang telah mendampingi dalam penyusunan Tugas Akhir ini.
5. Keluarga besar kelas 17-D3TI-01 dan Informatika Transfer '20  
Terimakasih atas canda tawanya selama perkuliahan serta memberi do'a dan dukungannya.
6. Teman-teman TKJ016 Skagata yang telah memberikan semangat dan doanya selama ini.
7. Terimakasih untuk kekasih saya Ifa Dwi Puspitasari yang telah menemani saya selama kuliah di AMIKOM hingga selesai.

**-Arfian Yogi Ferianto**

## KATA PENGANTAR

Puji dan syukur penulis persembahkan untuk Allah SWT yang telah memberikan rahmat, hidayah dan kekuatan sehingga penulis dapat menyelesaikan Tugas Akhir ini sesuai dengan waktu yang diinginkan penulis. Tidak lupa sholawat serta salam penulis haturkan pada junjungan Nabi besar kita yakni Muhammad SAW.

Skripsi ini disusun sebagai salah satu syarat kelulusan bagi setiap mahasiswa Universitas AMIKOM Yogyakarta. Selain itu juga merupakan suatu bukti bahwa mahasiswa telah menyelesaikan kuliah jenjang Strata I dan untuk memperoleh gelar Sarjana Komputer.

Dengan selesainya Tugas Akhir ini, maka penulis tidak lupa mengucapkan terima kasih kepada :

1. Universitas AMIKOM Yogyakarta selaku pihak yang telah mengizinkan penulis melakukan penelitian.
2. Bapak Prof.Dr.M. Suyanto, MM selaku Ketua Universitas AMIKOM Yogyakarta.
3. Bapak Hanif Al Fatta,S.Kom., M.Kom. selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
4. Ibu Windha Mega Pradnya Duhita, M.Kom. selaku Ketua Program Studi S1 Informatika Universitas AMIKOM Yogyakarta.
5. Ibu Mardhiya Hayaty, S.T., M.Kom. selaku dosen pembimbing yang telah memberikan petunjuk, bimbingan dan nasihatnya dalam proses penulisan Skripsi ini.
6. Kedua orang tua, kerabat, dan rekan-rekan yang telah memberikan motivasi, doa, dukungan, maupun semangat.

Yogyakarta, November 2022

Penulis

## DAFTAR ISI

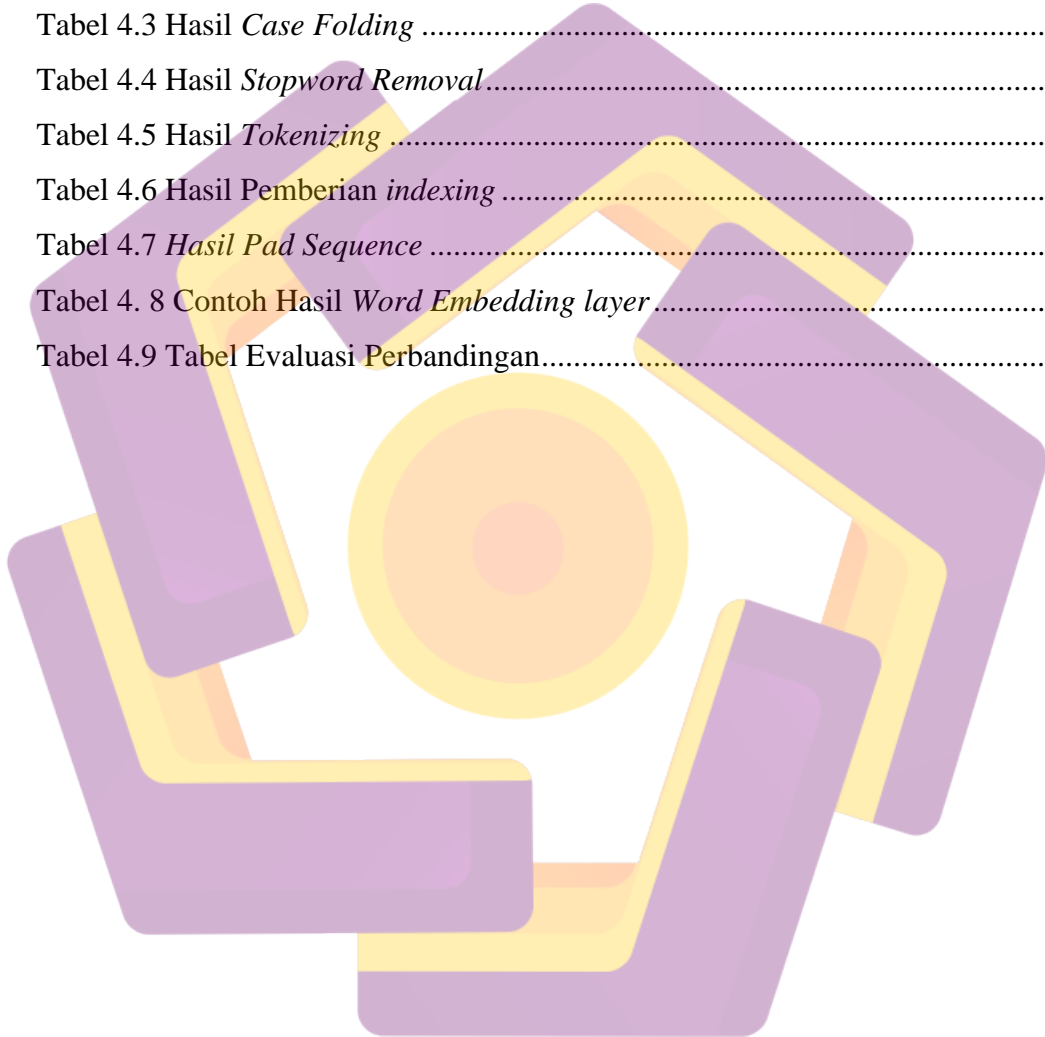
HALAMAN JUDUL .....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN .....	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI .....	iv
HALAMAN PERSEMBAHAN .....	v
KATA PENGANTAR .....	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR.....	x
INTISARI .....	xi
ABSTRACT.....	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah .....	2
1.4 Tujuan Penelitian .....	2
1.5 Manfaat Penelitian .....	3
1.6 Sistematika Penulisan .....	3
BAB II TINJAUAN PUSTAKA .....	4
2.1 Studi Literatur .....	4
2.2 Dasar Teori .....	8
2.2.1 Natural Language Processing .....	8
2.2.2 Deep Learning .....	8
2.2.3 Analisis Sentimen .....	11
2.2.4 <i>Review</i> .....	11
2.2.5 <i>Word Embedding</i> .....	12
2.2.6 <i>Word2Vec CBOW (continous bag-of-word)</i> .....	13
2.2.7 <i>Global Vector (GloVe)</i> .....	14
2.2.8 Scraping .....	15
BAB III METODE PENELITIAN .....	17



3.1. Alur Penelitian .....	17
3.2. Alat Penelitian.....	17
3.2.1. Perangkat Keras .....	17
3.2.2. Perangkat Lunak .....	17
3.3. Pengumpulan Data.....	18
3.4. Annotation.....	18
3.5. Preprocessing Data .....	18
3.6. Data Training dan Data Testing.....	19
3.7. Ekstraksi Fitur.....	19
3.8. Implementasi Algoritma Klasifikasi.....	20
3.9. Evaluasi.....	20
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>21</b>
4.1. Pengumpulan .....	21
4.2. Dataset.....	22
4.3. Preprocessing Data .....	23
4.3.1. Data Cleaning .....	23
4.3.2. Case Folding .....	24
4.3.3. Stop Word Removal.....	24
4.3.4. Tokenizing .....	25
4.4. Ekstraksi Fitur.....	27
4.5. Implementasi Algoritma LSTM.....	31
4.6. Hasil Training Data Word2Vec .....	33
4.7. Hasil Training Data Glove .....	34
4.8. Pengukuran Algoritma.....	35
4.9. Pembahasan .....	36
<b>BAB V PENUTUP .....</b>	<b>38</b>
5.1 Kesimpulan .....	38
5.2 Saran .....	38
<b>REFERENSI .....</b>	<b>xiii</b>

## DAFTAR TABEL

Tabel 2.1 Perbandingan Penelitian .....	6
Tabel 4.1 Contoh <i>dataset</i> ulasan <i>marketplace shopee</i> .....	22
Tabel 4.2 Hasil <i>Data Cleaning</i> .....	23
Tabel 4.3 Hasil <i>Case Folding</i> .....	24
Tabel 4.4 Hasil <i>Stopword Removal</i> .....	25
Tabel 4.5 Hasil <i>Tokenizing</i> .....	26
Tabel 4.6 Hasil Pemberian <i>indexing</i> .....	28
Tabel 4.7 Hasil <i>Pad Sequence</i> .....	28
Tabel 4. 8 Contoh Hasil <i>Word Embedding layer</i> .....	30
Tabel 4.9 Tabel Evaluasi Perbandingan.....	36



## DAFTAR GAMBAR

Gambar 2.1 Arsitektur <i>RNN</i> .....	9
Gambar 2.2 <i>Memory RNN</i> .....	10
Gambar 2.3 <i>Memory Cell</i> .....	11
Gambar 2.4 Arsitektur <i>Word2Vec CBOW</i> .....	14
Gambar 3.1 Diagram Alur Penelitian .....	17
Gambar 4.1 <i>Scrapping Data</i> .....	21
Gambar 4.2 <i>Dataset</i> .....	22
Gambar 4.3 <i>Script Case Folding</i> .....	24
Gambar 4.4 <i>Script Stopword Removal</i> .....	25
Gambar 4.5 <i>Script Tokenizing</i> .....	26
Gambar 4.6 <i>Script Pembagian Dataset</i> .....	27
Gambar 4.7 <i>Script Generate Word2Vec</i> .....	29
Gambar 4.8 <i>Script</i> menampilkan pendekatan antar kata .....	30
Gambar 4.9 <i>Script Train LSTM</i> .....	31
Gambar 4.10 <i>Total Parameter</i> .....	32
Gambar 4.11 <i>Architecture Model LSTM</i> .....	33
Gambar 4.12 <i>Script fit network LSTM</i> .....	33
Gambar 4.13 Hasil <i>training</i> dan <i>validasi Word2Vec</i> .....	34
Gambar 4.14 Hasil <i>training</i> dan <i>validasi Glove</i> .....	34
Gambar 4.15 Hasil <i>Confusion Matrix Word2Vec</i> .....	35
Gambar 4.16 Hasil <i>Confusion Matrix Glove</i> .....	35

## INTISARI

*Marketplace* merupakan *platform* untuk jual beli barang secara *online*, salah satunya yaitu *shopee*. *Platform* tersebut menyediakan banyak data teks singkat mengenai ulasan berbagai macam produk yang dijual. Karenanya dilakukan sentimen analisis untuk klasifikasi ulasan dengan memperhatikan faktor pada objek sentimennya. Dalam analisis sentimen, ada sebuah metode yang lebih maju yaitu menggunakan *word embedding*, representasi kata dalam vektor, banyak peneliti yang telah menggunakan metode ini dalam penelitiannya. Oleh sebab itu penelitian ini menggunakan data ulasan yang diperoleh dari marketplace *shopee* untuk dilakukan sentimen analisis.

Pada penelitian ini data diklasifikasikan menggunakan *Long Short Term Memory* (LSTM). Ulasan yang diklasifikasi akan memiliki 2 label yaitu positif dan negatif. Penelitian ini bertujuan untuk mengetahui akurasi akhir dan vocabulary yang dihasilkan oleh *word embedding* yang diklasifikasikan menggunakan LSTM dalam menganalisis sentimen ulasan *shopee* bahasa Indonesia.

Metode *word embedding* yang digunakan adalah *Word2Vec* dan *Global Vector (Glove)*. Penelitian ini menggunakan jumlah dataset 10.000 menghasilkan vocabulary sebanyak 18004 kata. Dari dataset dilakukan pembagian *data training* 80% dan *data test* 20%. Akurasi yang dihasilkan metode *word embedding word2vec* yaitu 83% dan metode *word embedding Glove* mendapatkan hasil akurasi 86%.

**Kata kunci:** Sentimen Analisis, *Word Embedding*, *Word2Vec*, *Glove*, LSTM.

## **ABSTRACT**

*Marketplace is a platform for buying and selling goods online, one of which is shopee. The platform provides a lot of short text data about reviews of various products being sold. Therefore, sentiment analysis is carried out for the classification of reviews by taking into account the factors in the sentiment object. In sentiment analysis, there is a more advanced method, namely using word embedding, word representation in vectors, many researchers have used this method in their research. Therefore, this study uses review data obtained from the shopee marketplace for sentiment analysis.*

*In this study, data is classified using Long Short Term Memory (LSTM). Reviews that are classified will have 2 labels namely positive and negative. This study aims to determine the final accuracy and vocabulary generated by word embedding which is classified using LSTM in analyzing sentiment in Indonesian shopee reviews.*

*Word embedding methods used are Word2Vec and Global Vector (Glove). This study uses a dataset of 10,000 to produce a vocabulary of 18004 words. From the dataset, 80% training data and 20% test data were distributed. The accuracy of the word embedding word2vec method is 83% and the word embedding Glove method gets 86% accuracy.*

**Keyword:** Sentiment Analysis, Word Embedding, Word2Vec, Glove, LSTM.