



**SKRIPSI**

**ANALISIS SENTIMEN KEPUASAN LAYANAN  
TRANSPORTASI *ONLINE* PADA TWITTER  
MENGUNAKAN EKSTRAKSI FITUR  
*WORD2VEC TEXT EMBEDDING* DAN  
ALGORITMA *EXTREME LEARNING MACHINE***

**AMELIYAH RISKIYAH**  
NPM 20083010022

**DOSEN PEMBIMBING**

Kartika Maulida Hindrayani, S.Kom., M.Kom  
Aviolia Terza Damaliana, S.Si., M.Stat

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI  
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR  
FAKULTAS ILMU KOMPUTER  
PROGRAM STUDI SAINS DATA  
SURABAYA  
2025**



**SKRIPSI**

**ANALISIS SENTIMEN KEPUASAN LAYANAN  
TRANSPORTASI *ONLINE* PADA TWITTER  
MENGUNAKAN EKSTRAKSI FITUR  
*WORD2VEC TEXT EMBEDDING* DAN  
ALGORITMA *EXTREME LEARNING MACHINE***

**AMELIYAH RISKIYAH**  
NPM 20083010022

**DOSEN PEMBIMBING**

Kartika Maulida Hindrayani., S.Kom., M.Kom  
Aviolla Terza Damaliana, S.Si., M.Stat

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI  
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR  
FAKULTAS ILMU KOMPUTER  
PROGRAM STUDI SAINS DATA  
SURABAYA  
2025**



**SKRIPSI**

**ANALISIS SENTIMEN KEPUASAN PELAYANAN  
TRANSPORTASI *ONLINE* PADA TWITTER  
MENGUNAKAN EKSTRAKSI FITUR  
*WORD2VEC TEXT EMBEDDING* DAN  
ALGORITMA *EXTREME LEARNING MACHINE***

**AMELIYAH RISKIYAH**  
NPM 20083010022

**DOSEN PEMBIMBING**

Kartika Maulida Hindrayani., S.Kom., M.Kom  
Aviolia Terza Damaliana, S.Si., M.Stat

**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI  
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR  
FAKULTAS ILMU KOMPUTER  
PROGRAM STUDI SAINS DATA  
SURABAYA  
2025**

*Halaman ini sengaja dikosongkan*

**LEMBAR PENGESAHAN**

**ANALISIS SENTIMEN KEPUASAN LAYANAN TRANSPORTASI ONLINE  
PADA TWITTER MENGGUNAKAN EKSTRAKSI FITUR *WORD2VEC* TEXT  
EMBEDDING DAN ALGORITMA *EXTREME LEARNING MACHINE***

Oleh :

**AMELIYAH RISKIYAH**

**NPM. 21083010022**

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi Sains Data Fakultas Ilmu  
Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal 21 Januari 2025.

**Kartika Maulida Hindrayani, S.Kom., M.Kom**  
**NIP. 199209092022032009**



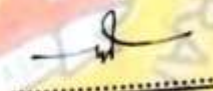
**(Pembimbing I)**

**Aviolla Terza Damalliana, S.Si., M.Stat**  
**NIP. 199408022022032015**



**(Pembimbing II)**

**Dr. Ir. Mohammad Idhom, S.P., S.Kom., M.T.**  
**NIP. 198303102021211006**



**(Ketua Penguji)**

**Trimono, S.Si., M.Si.**  
**NIP. 199509082022031003**



**(Penguji I)**

**Mengetahui,**

**Dekan Fakultas Ilmu Komputer**



**Prof. Dr. Jf. Novirina Hendrasarie, MT**  
**NIP. 19681126 199403 2 001**

*Halaman ini sengaja dikosongkan*

**LEMBAR PERSETUJUAN**

**ANALISIS SENTIMEN KEPUASAN LAYANAN TRANSPORTASI ONLINE  
PADA TWITTER MENGGUNAKAN EKSTRAKSI FITUR WORD2VEC TEXT  
EMBEDDING DAN ALGORITMA EXTREME LEARNING MACHINE**

Oleh :

**AMELIYAH RISKIYAH**

**NPM. 20083010022**

Telah disetujui untuk mengikuti Ujian Skripsi



**Menyetujui,**

**Koordinator Program Studi Sains Data**

**Fakultas Ilmu Komputer**

**Dr. Eng. Ir. Dwi Arman Prasetya, ST., MT., IPU., Asean., Eng.**

**NIP. 19801205 200501 1 002**

*Halaman ini sengaja dikosongkan*



## SURAT PERNYATAAN ORISINALITAS

Yang bertandatangan di bawah ini:

Nama Mahasiswa : Ameliah Riskiyah

Program Studi : Sains Data

Dosen Pembimbing : 1. Kartika Maulida Hindrayani., S.Kom., M.Kom

2. Aviolla Terza Damaliana, S.Si., M.Stat

dengan ini menyatakan bahwa isi sebagian maupun keseluruhan buku skripsi yang berjudul:

**ANALISIS SENTIMEN KEPUASAN LAYANAN TRANSPORTASI *ONLINE*  
PADA TWITTER MENGGUNAKAN EKSTRASI FITUR *WORD2VEC TEXT*  
*EMBEDDING* DAN ALGORITMA *EXTREME LEARNING MACHINE***

adalah benar-benar hasil karya intelektual mandiri, diselesaikan tanpa menggunakan bahan-bahan yang tidak diizinkan dan bukan merupakan karya pihak lain yang saya akui sebagai karya sendiri. Semua referensi yang dikutip maupun dirujuk telah ditulis secara lengkap pada daftar pustaka. Apabila ternyata pernyataan ini tidak benar, saya bersedia menerima sanksi sesuai peraturan yang berlaku.

Surabaya, 21 Januari 2025  
Yang Membuat Pernyataan,

  
**AMELIYAH RISKIYAH**  
20083010022

*Halaman ini sengaja dikosongkan*

## ABSTRAK

Nama Mahasiswa / NPM : Ameliyah Riskiyah / 20083010022  
Judul Skripsi : Analisis Sentimen Kepuasan Layanan Transportasi  
*Online* Menggunakan Ekstraksi Fitur *Word2vec Text Embedding* dan Algoritma *Extreme Learning Machine*  
Dosen Pembimbing : 1. Kartika Maulida Hindrayani., S.Kom., M.Kom  
2. Aviolla Terza Damaliana, S.Si., M.Stat

Perkembangan teknologi di Indonesia jumlahnya semakin banyak, salah satunya transportasi *online*. Meskipun banyak perkembangan pada transportasi *online* sering kali ditemukan permasalahan yang mempengaruhi kualitas layanan misalkan pada perilaku pengemudi, perilaku pengguna, dan kondisi kendaraan. Oleh karena itu, penelitian ini dilakukan dengan tujuan untuk menganalisis opini publik atau sentimen mengenai pelayanan yang diberikan oleh transportasi *online* tersebut. Penelitian ini menggunakan algoritma *Extreme Learning Machine* dengan ekstraksi fitur *Word2Vec*. *Extreme Learning Machine* adalah algoritma *machine learning* yang hanya mempunyai satu *hidden layer*. Penggunaan metode ini memberikan keuntungan yaitu kecepatan pada proses pembelajaran dibandingkan dengan metode tradisional berbasis gradien. Evaluasi model yang digunakan yaitu menggunakan *Confusion Matrix*. Transportasi *online* yang mempunyai sentimen positif terbanyak adalah GRAB sebanyak 844 tweet lalu diikuti dengan GOJEK 799 tweet dan terakhir adalah MAXIM 724 tweet. Sentimen negatif terbanyak adalah MAXIM yaitu 724 tweet, GOJEK 701 tweet, dan GRAB 656 tweet. Model ELM yang terbaik yaitu akurasi pada model ELM GOJEK dengan akurasi sebesar 0,8533, MAXIM mempunyai akurasi yang sama dengan GOJEK tetapi model ELM GOJEK terbaik secara keseluruhan karena mempunyai nilai *F1-Score* 0,8544 dan *Recall* 0,8529 tertinggi. Model MAXIM, *F1-Score* 0,8523 dan *recall* 0,8514. Sedangkan terakhir adalah GRAB mempunyai *accuracy* 0,84, *F1 Score* 0,8385 dan *recall* 0,8378.

**Kata kunci :** Transportasi *Online*, Twitter, Analisis Sentimen, *Extreme Learning Machine* (ELM), *Word2Vec*

*Halaman ini sengaja dikosongkan*

## ABSTRACT

Student Name / NPM : Ameliyah Riskiyah / 20083010022  
Thesis Title : Sentiment Analysis of Online Transportation Service Satisfaction using Word2Vec Text Embedding Feature Extraction and Extreme Learning Machine Algorithm  
Advisor : 1. Kartika Maulida Hindrayani., S.Kom., M.Kom  
2. Aviolla Terza Damaliana, S.Si., M.Stat

There are more and more technological developments in Indonesia, one of which is online transportation. Although there are many developments in online transportation, there are often problems that affect service quality, for example in driver behavior, user behavior, and vehicle conditions. Therefore, this research was conducted with the aim of analyzing public opinion or sentiment about the services provided by online transportation. This research uses Extreme Learning Machine algorithm with Word2Vec feature extraction. Extreme Learning Machine is a machine learning algorithm that has only one hidden layer. The use of this method provides the advantage of speed in the learning process compared to traditional gradient-based methods. Evaluation of the model used is using Confusion Matrix. The online transportation that has the most positive sentiment is GRAB with 844 tweets, followed by GOJEK with 799 tweets and MAXIM with 724 tweets. The most negative sentiment is MAXIM with 724 tweets, GOJEK with 701 tweets, and GRAB with 656 tweets. The best ELM model is the accuracy of the GOJEK ELM model with an accuracy of 0.8533, MAXIM has the same accuracy as GOJEK but the GOJEK ELM model is the best overall because it has the highest F1-Score value of 0.8544 and Recall 0.8529. MAXIM model, F1-Score 0.8523 and recall 0.8514. While the last is GRAB has accuracy 0.84, F1 Score 0.8385 and recall 0.8378.

**Keywords:** *Online Transportation, Twitter, Sentiment Analysis, Extreme Learning Machine (ELM), Word2Vec*

*Halaman ini sengaja dikosongkan*

## KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas segala rahmat, hidayah dan karunia-Nya kepada penulis sehingga skripsi dengan judul **“Analisis Sentimen Kepuasan Layanan Transportasi Online Pada Twitter menggunakan Ekstraksi Fitur *Word2Vec Text Embedding* dan Algoritma *Extreme Learning Machine*”** dapat terselesaikan dengan baik. Dalam penyusunan Skripsi ini tidak terlepas dari bantuan berbagai pihak, dan dalam kesempatan ini penulis ingin mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Ir. Akhmad Fauzi, MMT., IPU., selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Dr. Ir. Novirina Hendrasarie, MT. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur
3. Bapak Dr.Eng. Ir. Dwi Arman Prasetya, S.T., M.T., IPU selaku Koordinator Program Studi Sains Data Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. Tresna Maulana Fahrudin, S.ST., M.T Selaku Dosen Wali.
5. Ibu Kartika Maulida Hindrayani, S.Kom., M.Kom selaku Dosen Pembimbing 1 yang memberikan arahan dan masukan selama pengerjaan skripsi.
6. Ibu Aviolla Terza Damaliana, S.Si., M.Stat selaku Dosen Pembimbing 2 yang memberikan masukan dan saran selama pengerjaan skripsi
7. Kedua orang tua saya dan keluarga yang selalu memberikan doa dan dukungan selama pengerjaan skripsi
8. Teman-teman kost yang memberikan dukungan, doa, dan bantuan saat mengalami kesulitan

Penulis menyadari bahwa penulisan skripsi ini masih belum sempurna dan terdapat beberapa kekurangan. Untuk itu, penulis mohon kritik dan saran yang bersifat membangun demi kesempurnaan penulisan di masa mendatang. Akhir kata, semoga skripsi ini dapat berguna bagi penulis dan bagi semua pihak.

Surabaya, Januari 2025

Ameliyah Riskiyah

## DAFTAR ISI

HALAMAN SAMPUL .....	i
LEMBAR PENGESAHAN .....	iii
LEMBAR PERSETUJUAN.....	v
ABSTRAK.....	ix
ABSTRACT .....	xi
KATA PENGANTAR .....	xiii
DAFTAR ISI .....	xiv
DAFTAR GAMBAR .....	xvii
DAFTAR TABEL .....	xviii
DAFTAR LAMPIRAN.....	xxi
DAFTAR NOTASI .....	xxiii
<b>BAB I PENDAHULUAN .....</b>	<b>1</b>
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	3
1.3. Batasan Masalah .....	4
1.4. Tujuan Penelitian .....	4
1.5. Manfaat Penelitian .....	4
<b>BAB II TINJAUAN PUSTAKA .....</b>	<b>7</b>
2.1. Dasar Teori .....	7
2.1.1. Transportasi <i>Online</i> .....	7
2.1.2. X .....	7
2.1.3. Analisis Sentimen .....	8
2.1.4. <i>Crawling</i> Data.....	8
2.1.5. <i>Preprocessing</i> Data.....	8
2.1.6. Ekstraksi Fitur.....	10
2.1.7. <i>SMOTE</i> .....	10
2.1.8. <i>WorldCloud</i> .....	11
2.1.9. <i>Extreme Learning Machine</i> .....	11
2.1.10. Evaluasi Model .....	17
2.2. Penelitian Terdahulu.....	18
<b>BAB III METODOLOGI PENELITIAN.....</b>	<b>26</b>



3.1. Variabel Penelitian dan Sumber Data.....	26
3.2. Langkah Analisis .....	26
3.2.1. Studi Literatur.....	27
3.2.2. <i>Crawling</i> Data Pada Twitter .....	28
3.2.3. <i>Preprocessing</i> Data.....	29
3.2.4. Labeling .....	33
3.2.5. Ekstraksi Fitur.....	34
3.2.6. Extreme Learning Machine .....	35
3.2.7. Evaluasi Model .....	35
3.3 Desain Sistem .....	36
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>39</b>
4.1. Hasil Pengujian.....	39
4.2. Hasil Pembahasan.....	63
4.3. Analisis Sentimen .....	70
<b>BAB V PENUTUP .....</b>	<b>76</b>
5.1. Kesimpulan.....	76
5.2. Saran Pengembangan.....	77
<b>DAFTAR PUSTAKA .....</b>	<b>78</b>
<b>LAMPIRAN .....</b>	<b>81</b>
<b>Lampiran 1.</b> Hasil Uji Plagiasi.....	81
<b>Lampiran 2.</b> Data Penelitian .....	88
<b>Lampiran 3.</b> Source Code yang digunakan untuk analisis.....	89

*Halaman ini sengaja dikosongkan*

## DAFTAR GAMBAR

Gambar 2. 1 Arsitektur <i>Multi Layer Perceptron</i> (MLP) .....	12
Gambar 2. 2 Arsitektur <i>Extreme Learning Machine</i> (ELM) .....	13
Gambar 2. 3 Arsitektur <i>Multi Hidden Layer</i> ELM.....	16
Gambar 3. 1 Langkah analisis penelitian.....	27
Gambar 3. 2 Tahapan <i>Preprocessing</i> .....	29
Gambar 3. 3 Desain Sistem .....	36
Gambar 4. 1 <i>WordCloud</i> .....	57
Gambar 4. 2 <i>Confussion Matrix</i> Gojek.....	64
Gambar 4. 3 <i>Confussion Matrix</i> Grab .....	66
Gambar 4. 4 <i>Confussion Matrix</i> Maxim.....	68
Gambar 4. 5 <i>WordCloud</i> positif dan negatif GOJEK.....	71
Gambar 4. 6 Jumlah Analisis Sentimen GOJEK.....	71
Gambar 4. 7 <i>Wordcloud</i> positif dan negatif GRAB .....	72
Gambar 4. 8 Jumlah Analisis sentimen GRAB .....	72
Gambar 4. 9 <i>Wordcloud</i> positif dan negatif MAXIM .....	73
Gambar 4. 10 Jumlah analisis sentimen MAXIM .....	73

## DAFTAR TABEL

Tabel 2.1. Penelitian Terdahulu.....	18
Tabel 3. 1 Hasil <i>Crawling</i> Data (1) .....	28
Tabel 3. 2 Hasil <i>Crawling</i> Data (2) .....	29
Tabel 3. 3 <i>Case Folding</i> .....	30
Tabel 3. 4 <i>Cleaning</i> .....	30
Tabel 3. 5 <i>Tokenizing</i> .....	31
Tabel 3. 6 Normalisasi.....	31
Tabel 3. 7 <i>Convert Negasi</i> .....	32
Tabel 3. 8 <i>Stopword</i> .....	33
Tabel 3. 9 <i>Stemming</i> .....	33
Tabel 3. 10 <i>Labeling</i> .....	34
Tabel 3. 11 <i>Confussion Matrix</i> .....	36
Tabel 4. 1 Hasil <i>Crawling</i> Data.....	40
Tabel 4. 2 <i>Load</i> Data .....	41
Tabel 4. 3 Hasil <i>Case Folding</i> .....	42
Tabel 4. 4 Hasil <i>Cleaning</i> .....	44
Tabel 4. 5 Hasil <i>Tokenizing</i> .....	46
Tabel 4. 6 Hasil Normalisasi .....	49
Tabel 4. 7 Hasil <i>Convert Negasi</i> .....	52
Tabel 4. 8 Hasil <i>Stopword</i> .....	54
Tabel 4. 9 Hasil <i>Stemming</i> .....	55
Tabel 4. 10 Hasil <i>Labelling</i> .....	58
Tabel 4. 11 Hasil Ekstraksi Fitur .....	60
Tabel 4. 12 Hasil Konfigurasi <i>hidden neuron GOJEK</i> .....	63
Tabel 4. 13 Akurasi 150 neuron GOJEK.....	64
Tabel 4. 14 Hasil Konfigurasi Hidden Neuron Grab.....	65
Tabel 4. 15 Akurasi 150 neuron Grab .....	66
Tabel 4. 16 Hasil Konfigurasi <i>Hidden Neuron Maxim</i> .....	67
Tabel 4. 17 Akurasi 150 <i>Neuron Maxim</i> .....	68
Tabel 4. 18 Perbandingan Hasil Performa Model MELM pada GOJEK, GRAB, dan MAXIM.....	69

Tabel 4. 19 Hasil Prediksi.....	70
---------------------------------	----

*Halaman ini sengaja dikosongkan*

## DAFTAR LAMPIRAN

<b>Lampiran 1.</b> Hasil Uji Plagiasi.....	81
<b>Lampiran 2.</b> Data Penelitian .....	88
<b>Lampiran 3.</b> Source Code yang digunakan untuk analisis.....	89

*Halaman ini sengaja dikosongkan*



## DAFTAR NOTASI

$n$	:	<i>Neuron input layer</i>
$l$	:	<i>Neuron hidden layer</i>
$m$	:	<i>Neuron output layer</i>
$\beta_{jk}$	:	<i>Bobot antara input layer dengan neuron hidden layer</i>
$g(x)$	:	<i>Fungsi aktivasi</i>
$T$	:	<i>Output matriks</i>
$T'$	:	<i>Transpose T</i>
$H$	:	<i>Output dari hidden layer</i>
$\beta_{new}^+$	:	<i>Invers <math>\beta_{new}</math></i>
$H_2$	:	<i>Output actual dari hidden layer 2</i>
$H_2$	:	<i>Bobot antara hidden layer kedua dengan hidden layer ketiga</i>
$\beta_2$	:	<i>Bias dari hidden layer ketiga</i>
$H_{new}^+$	:	<i>Generational invers dari <math>H_{e1}</math></i>
$g^{-1}(x)$	:	<i>Invers fungsi aktivasi</i>

*Halaman ini sengaja dikosongkan*