

**IMPLEMENTASI SUPPORT VECTOR MACHINE DAN
NAÏVE BAYES UNTUK ANALISIS SENTIMEN PADA SITUS
GOOGLE REVIEW WARJO SUMMARECON BEKASI**

SKRIPSI

Oleh:

Helen Emiliana Marpaung

202010225278



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA**

2025

LEMBAR PERSETUJUAN PEMBIMBING
(TUGAS AKHIR)

Judul Tugas Akhir : Implementasi Support Vector Machine Dan Naïve Bayes Untuk Analisis Sentimen Pada Situs Google Review Warjo Summarecon Bekasi.

Nama Mahasiswa : Helen Emiliana Marpaung

Nomor Pokok Mahasiswa : 202010225278

Program Studi/Fakultas : Informatika / Ilmu Komputer

Jakarta, 15 Juli 2025

MENYETUJUI,

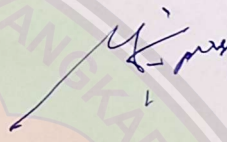
Pembimbing I



(Wowon Priatna, S.T., M.TI)

NIDN : 0429118007

Pembimbing II



(Mugiarso, S.Kom., M.Kom)

NIDN : 0420117403

Ketua Program Studi



(Ahmad Fathurrozi, S.E., M.M.S.I)

NIP : 2012786

Program Studi Informatika
Fakultas Ilmu Komputer
Universitas Bhayangkara Jakarta Raya

2025

LEMBAR PERSETUJUAN PENGESAHAN

Judul Proposal Tugas Akhir : Implementasi Support Vector Machine Dan Naïve Bayes Untuk Analisis Sentimen Pada Situs Google Review Warjo Summarecon Bekasi.

Nama Mahasiswa : Helen Emiliana Marpaung
Nomor Pokok Mahasiswa : 202010225278
Program Studi/Fakultas : Informatika / Ilmu Komputer
Tanggal Lulus Tugas Akhir : 28 Juli 2025

Jakarta, 04 Agustus 2025
MENGESAHKAN

Ketua Tim Penguji : Prima Dina Atika, S.Kom., M.Kom
NIDN: 0311037107

Penguji I : Agus Hidayat, S.Kom., M.Si
NIDN: 0304088303

Penguji II : Wowon Priatna, S.T., M.TI
NIDN: 0429118007

MENGETAHUI

Ketua
Program Studi Informatika

Dekan
Fakultas Ilmu Komputer

(Ahmad Fathurrozi, S.E., M.M.S.I)

NIP : 2012486

(Dr. Robertus Suraji, S.S., M.A.)

NIP : 1904406



UNIVERSITAS BHAYANGKARA JAKARTA RAYA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA

LEMBAR PERNYATAAN BUKAN PLAGIASI

Nama : Helen Emiliana Marpaung
NPM : 202010225278
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Implementasi Support Vector Machine Dan Naïve Bayes Untuk Analisis Sentimen Pada Situs Google Review Warjo Summarecon Bekasi

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun.

Bekasi, 04 Agustus 2025
Penulis



Helen Emiliana Marpaung

ABSTRAK

Helen Emiliana Marpaung 202010225278. Implementasi Support Vector Machine Dan Naïve Untuk Analisis Sentimen Pada Situs Google Review Warjo Summarecon Bekasi. Bekasi : Fakultas Ilmu Komputer. Universitas Bhayangkara Jakarta Raya. 2025.

Pada era digital saat ini, ulasan konsumen secara daring memainkan peranan penting dalam membentuk citra dan keputusan pembelian masyarakat. Salah satu contohnya adalah ulasan terhadap restoran Warjo Summarecon Bekasi yang dibagikan melalui platform Google Review. Banyaknya jumlah ulasan yang masuk membuat proses analisis secara manual menjadi tidak efisien. Oleh karena itu, diperlukan metode yang dapat secara otomatis mengklasifikasikan opini pelanggan ke dalam kategori positif dan negatif melalui proses analisis sentimen. Penelitian ini dilakukan dengan tahapan preprocessing data (case folding, tokenizing, stopword removal, dan stemming), pembobotan kata menggunakan metode TF-IDF, penerapan algoritma klasifikasi, serta evaluasi hasil klasifikasi. Dua algoritma yang digunakan dalam penelitian ini adalah Support Vector Machine (SVM) dan Naïve Bayes. Data dikumpulkan melalui proses scraping terhadap 1.102 ulasan pelanggan Warjo, kemudian dilakukan pelabelan manual untuk keperluan pelatihan model. Berdasarkan pengujian yang dilakukan, diperoleh bahwa algoritma Support Vector Machine menghasilkan akurasi tertinggi sebesar 88%, sementara Naïve Bayes memperoleh akurasi sebesar 80%. Dengan demikian, algoritma SVM dinilai lebih efektif dalam melakukan klasifikasi sentimen terhadap ulasan pelanggan dan dapat menjadi solusi bagi pelaku usaha untuk memahami persepsi pelanggan secara efisien dan objektif.

Kata Kunci: Analisis Sentimen, Google Review, TF-IDF, Naïve Bayes, Support Vector Machine.

ABSTRACT

Helen Emiliana Marpaung 202010225278. Implementation of Support Vector Machine and Naive for Sentiment Analysis on Google Review Site Warjo Summarecon Bekasi. Bekasi: Faculty of Computer Science. Bhayangkara University of Jakarta Raya. 2025.

In today's digital era, online customer reviews play a crucial role in shaping public perception and influencing purchasing decisions. One such example is customer reviews of Warjo Summarecon Bekasi, which are widely shared through the Google Review platform. The large volume of incoming reviews makes manual sentiment analysis inefficient and prone to bias. Therefore, a method is needed to automatically classify customer opinions into positive and negative categories through sentiment analysis. This study involves several stages: data preprocessing (case folding, tokenizing, stopword removal, and stemming), word weighting using the TF-IDF method, application of classification algorithms, and evaluation of classification results. The two algorithms used in this research are Support Vector Machine (SVM) and Naïve Bayes. Data was collected by scraping 1,102 customer reviews of Warjo, followed by manual labeling to train the models. Based on testing results, the Support Vector Machine algorithm achieved the highest classification accuracy of 88%, while Naïve Bayes reached an accuracy of 80%. These findings indicate that SVM is more effective in classifying customer sentiment and can be a valuable tool for business owners to gain insights into customer perceptions efficiently and objectively.

Keywords: Sentiment Analysis, Google Review, TF-IDF, Naïve Bayes, Support Vector Machine.

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK**

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini :

Nama : Helen Emiliana Marpaung
NPM : 202010225278
Program Studi : Ilmu Komputer
Fakultas : Informatika
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya **Hak Bebas Royalti Non-Eksklusif (*Non-Exclusive Royalty-Free Right*)**, atas karya ilmiah saya yang berjudul :

Implementasi Support Vector Machine Dan Naïve Bayes Untuk Analisis Sentimen Pada Situs Google Review Warjo Summarecon Bekasi

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Bekasi
Pada tanggal : 04 Agustus 2025
Yang Menyatakan



Helen Emiliana Marpaung

KATA PENGANTAR

Puji syukur penulis panjatkan sebagai penulis ke hadirat Tuhan Yang Maha Esa atas berkat Rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi sesuai dengan waktu yang telah ditentukan. Skripsi ini disusun sebagai salah satu syarat untuk mencapai gelar sarjana Strata Satu (S1) pada program studi Informatika Universitas Bhayangkara Jakarta Raya. Adapun judul skripsi yang penulis gunakan adalah “Implementasi Support Vector Machine dan Naive Bayes Untuk Analisis Sentimen Pada Situs Google Review Warjo Summarecon Bekasi”.

Dalam penyusunan skripsi ini penulis banyak sekali mendapatkan bantuan dari berbagai pihak, Terutama kedua orang tua saya yang selalu memberikan doa, semangat serta dukungan dalam proses penulisan skripsi. Oleh sebab itu atas bantuan dan dukungannya, penulis mengucapkan terimakasih kepada :

1. Bapak Irjen Pol. (Purn) Prof. Dr. Drs. H. Bambang Karsono, S.H., M.M., Ph.D., D.Crim(HC) Selaku Rektor Universitas Bhayangkara Jakarta Raya.
2. Bapak Dr. Robertus Suraji, S.S., M.A selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
3. Bapak Ahmad Fathurrozi, SE., M.M.S.I. selaku Ketua Program Studi Informatika Universitas Bhayangkara Jakarta Raya Bekasi.
4. Bapak Wowon Priatna, S.T., M.TI. selaku Dosen Pembimbing pertama dalam penyusunan tugas akhir yang senantiasa memberikan arahan dan bimbingan yang luar biasa selama penyusunan tugas akhir.
5. Bapak Mugiarto, S.Kom., M.Kom selaku Dosen Pembimbing kedua tugas akhir ini yang telah memberikan bimbingan dan arahan selama

proses penyusunan tugas akhir ini sehingga tugas akhir ini dapat terselesaikan dengan baik.

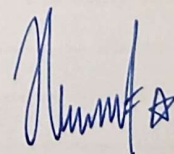
6. S.B selaku support sistem penulis yang selalu mengasih semangat selama penulisan tugas akhir berlangsung.
7. Nadia hafiza dan Fadhila Alfiyatul Amri teman seperjuangan saya dari SMA-sekarang mendukung saya dalam penyusunan tugas akhir.
8. Teman-teman seperjuangan pada Program Studi Informatika yang selalu menghibur dalam kondisi apapun. Selalu membantu dan mendukung dalam penulisan skripsi ini.

Penulis menyadari bahwa penyusunan Skripsi ini, masih terdapat banyak kekurangan dan masih jauh dari kata sempurna, oleh karena itu penulis dengan hati yang tulus menerima masukan baik berupa kritik, maupun saran-saran yang dapat membangun untuk memperbaiki kekurangan-kekurangan yang ada. Semoga Skripsi ini bermanfaat, baik bagi penulis maupun bagi orang lain yang membacanya.

Akhir kata penulis mengucapkan banyak terimakasih, semoga skripsi ini dapat berguna dan bermanfaat bagi semua pembaca dan semoga tuhan yang maha esa selalu melimpahkan Rahmat dan karunia-Nya kepada kita semua.

Jakarta, ...

Hormat Saya



Helen Emiliana Marpaung

202010225278

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PERSETUJUAN PENGESAHAN	ii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iii
ABSTRAK	iv
ABSTRACT	v
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI	vi
KATA PENGANTAR.....	vi
DAFTAR ISI.....	viii
DAFTAR GAMBAR.....	x
DAFTAR TABEL	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	4
1.3 Rumusan Masalah	4
1.4 Batasan Masalah.....	5
1.5 Tujuan Penelitian	5
1.6 Manfaat Penelitian	6
1.7 Sistematika Penulisan	7
BAB II TINJAUAN PUSTAKA.....	7
2.1 Penelitian Terdahulu	7
2.2 Analisis Sentimen	11
2.3 Google Review	11
2.4 Knowledge Discovery In Database	12
2.5 Pembobotan TF-IDF	14
2.6 Data Mining	15
2.6.1 Text Mining.....	17
2.7 Machine Learning	19
2.8 Algoritma Support Vector Machine (SVM)	21
2.9 Algoritma Naïve Bayes	22
2.10 Split Data.....	23
2.11 Confusion Matrix	23

2.12 Diagram Alir (Flowchart)	24
2.13 Python	26
2.14 Google Colab	27
BAB III METODOLOGI PENELITIAN	28
3.1 Kerangka Penelitian	28
3.2 Metode Pengumpulan Data	30
3.3 Metode Analisis Data	30
3.4 Tabel Instrumen Analisis Data	49
BAB IV HASIL DAN PEMBAHASAN	56
4.1 Seleksi Data (<i>Selection Data</i>)	56
4.1.1 Sumber Data	56
4.1.2 Pengumpulan Data	59
4.2 Pembersihan Data (<i>Preprocessing Data</i>)	61
4.2.1 <i>Case folding</i>	61
4.2.2 <i>Tokenizing</i>	65
4.2.3 <i>Filtering / Stopword Removing</i>	68
4.2.4 <i>Stemming</i>	71
4.3 Transformasi Data (<i>Transformation Data</i>)	74
4.3.1 Klasifikasi Data Sentimen Positif Dan Negatif	74
4.3.2 Pembobotan (TF-IDF)	78
4.4 Data Mining	94
4.4.1 Implementasi Algoritma SVM	94
4.4.2 Implementasi Algoritma Naïve Bayes	100
4.5 Evaluasi Dan Interpretasi (<i>Evaluation And Interpretation</i>)	122
4.5.1 Evaluasi Dan Interpretasi Algoritma SVM	122
4.5.2 Evaluasi Dan Interpretasi Algoritma Naïve Bayes	137
4.6 Kesimpulan Evaluasi Algoritma SVM dan Algoritma Naïve Bayes ..	146
4.6.1 Kesimpulan Evaluasi Algoritma SVM	147
4.6.2 Kesimpulan Evaluasi Algoritma Naïve Bayes	150
4.7 Knowledge	153
4.8 Perbandingan SVM dan Naïve Bayes	155
BAB V PENUTUP	159
5.1 Kesimpulan	159
5.2 Saran	160
DAFTAR PUSTAKA	161



DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	7
Tabel 3. 1 Hasil Labelling Manual.....	31
Tabel 3. 2 Aturan Tahapan Case Folding.....	33
Tabel 3. 3 Ilustrasi Tahapan Case Folding.....	34
Tabel 3. 4 Aturan Tahap Tokenizing	37
Tabel 3. 5 Ilustrasi Tahapan Tokenizing.....	38
Tabel 3. 6 Ilustrasi Tahapan Filtering / Stopword Removing.....	40
<i>Tabel 3. 7 Ilustrasi Tahapan Stemming</i>	<i>45</i>
Tabel 3. 8 Pengelompokan Data	50
Tabel 3. 9 Tabel Validasi Opini	54



DAFTAR GAMBAR

Gambar 2. 1 Tahapan Proses pada Data mining	12
Gambar 3. 1 Kerangka Penelitian	29
Gambar 3. 2 Tahapan Case Folding	33
Gambar 3. 3 Tahapan Tokenizing	36
Gambar 3. 4 Tahapan Filterring / Stopword	40
Gambar 3. 5 Tahapan Stemming.....	43
Gambar 3. 6 Tahapan Naïve Bayes.....	47
Gambar 3. 7 Tahapan SVM	48
Gambar 4. 1 Contoh Data Ulasan Positif.....	57
Gambar 4. 2 Contoh Data Dengan Ulasan Negatif.....	58
Gambar 4. 3 Extension Scraper.....	59
Gambar 4. 4 Scraping Data Google Review	60
Gambar 4. 5 Syntax Impelementasi Case Folding	63
Gambar 4. 6 Review Text Sebelum Case Folding	63
Gambar 4. 7 Review Text Setelah Case Folding	64
Gambar 4. 8 Syntax Tahapan Tokenizing.....	67
Gambar 4. 9 Hasil Tahapan Tokenizing	67
Gambar 4. 10 Syntax Tahapan Filtering / Stopword Removing.....	70
Gambar 4. 11 Hasil Syntax Filtering / Stopword Removing	70
Gambar 4. 12 Syntax Tahapan Stemming.....	73
Gambar 4. 13 Hasil Syntax Stemming.....	73
Gambar 4. 14 Syntax Klasifikasi Sentimen	77
Gambar 4. 15 Hasil Count Sentimen	77
Gambar 4. 16 Syntax Implementasi Algoritma SVM.....	99
Gambar 4. 17 Syntax Implementasi Algoritma Naive Bayes	122
Gambar 4. 18 Syntax menampilkan Confusion Matrix model SVM.....	123
Gambar 4. 19 Confusion Matrix Model SVM	124
Gambar 4. 20 Syntax Evaluasi Akurasi Model SVM	129
Gambar 4. 21 Akurasi Model SVM	130
Gambar 4. 22 Syntax Confusion Matrix Naive Bayes	137

Gambar 4. 23 Confusion Matriks Model Naive Bayes	138
Gambar 4.24 Syntax Evaluasi Akurasi Model Naive Bayes.....	143
Gambar 4.25 Akurasi Model Naive Bayes	143
Gambar 4. 26 ROC Curve Model SVM dan Naive Bayes.....	157
Gambar 4.27 Diagram Batang Perbandingan Akurasi SVM dan Naive Bayes..	158



DAFTAR LAMPIRAN

Lampiran 1 Kartu Bimbingan Pembimbing 1	167
Lampiran 2 Kartu Bimbingan Pembimbing 2	168
Lampiran 3 Cek Plagiasi	169
Lampiran 4 Biodata Mahasiswa	170

