

**PERBANDINGAN ALGORITMA *NAIVE BAYES* DAN
SUPPORT VECTOR MACHINE PADA ANALISIS
SENTIMENULASAN MALL PAKUWON
BEKASI DI *REVIEW* GOOGLE MAPS**

SKRIPSI

**Oleh:
Natasya Marito T
202110715197**



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2025**

LEMBAR PERSETUJUAN PEMBIMBING

Judul Skripsi : Perbandingan Algoritma *Naïve Bayes* dan *SVM*
pada Analisis Sentimen Ulasan Mall Pakuwon
Bekasi di *Review Google Maps*

Nama Mahasiswa : Natasya Marito T

Nomor Pokok Mahasiswa : 202110715197

Program Studi/Fakultas : Informatika/Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 24 Juli 2025

Jakarta, 31 Juli 2025

MENYETUJUI,

Pembimbing I

Pembimbing II

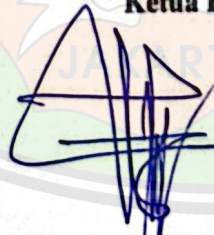


Prima Dina Atika, S.Kom., M.Kom.
NIDN: 0311037107



Khairunnisa Fadhillah Ramdhan, S.Si., M.Si.
NIDN: 0328039201

Ketua Program Studi



Ahmad Fathurrozi, S.E., M.M.S.I
NIP: 2012486

Program Studi Informatika
Fakultas Ilmu Komputer
Universitas Bhayangkara Jakarta Raya
2025

LEMBAR PENGESAHAN

Judul Skripsi : Perbandingan Algoritma *Naïve Bayes* dan *SVM* pada Analisis Sentimen Ulasan Mall Pakuwon Bekasi di Review Google Maps
Nama Mahasiswa : Natasya Marito T
Nomor Pokok Mahasiswa : 202110715197
Program Studi/Fakultas : Informatika / Ilmu Komputer
Tanggal Lulus Ujian Skripsi : 24 Juli 2025

Jakarta, 31 Juli 2025

MENGESAHKAN,

Ketua Tim Penguji : Dr. Herlawati, S.Si., M.M., M.Kom.
NIDN: 0311097302

Penguji I : Muhammad Yasir, S.Si., M.Kom.
NIDN: 0317129002

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

MENGETAHUI,

Ketua

Dekan

Program Studi Informatika

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

Yang bertanda tangan dibawah ini :

Nama : Natasya Marito T
NPM : 202110715197
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Perbandingan Algoritma *Naive Bayes* dan *SVM*
Pada Analisis Sentimen Ulasan Mall Pakuwon Bekasi
Di Review Google Maps

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, 31 Juli 2025

Penulis



Natasya Marito T

ABSTRAK

Natasya Marito T. 202110715197. Perbandingan Algoritma *Naïve Bayes* dan *SVM*

Pada Analisis Sentimen Ulasan Mall Pakuwon Bekasi di *Review* Google Maps.

Bekasi: Fakultas Ilmu Komputer. Universitas Bhayangkara Jakarta Raya. 2025.

Penelitian ini bertujuan untuk membandingkan performa dua algoritma klasifikasi teks, yaitu Naive Bayes dan Support Vector Machine (SVM), dalam menganalisis sentimen ulasan pengguna terhadap Mall Pakuwon Bekasi yang diambil dari platform Google Maps. Data diperoleh melalui teknik web scraping, dengan total sebanyak 1.037 ulasan yang berhasil dikumpulkan. Selanjutnya, data dibagi menggunakan skema 80% data latih dan 20% data uji, kemudian diproses melalui tahapan pra-pemrosesan teks dan vektorisasi menggunakan TF-IDF. Untuk mengatasi ketidakseimbangan kelas pada data latih, digunakan teknik Synthetic Minority Over-sampling Technique (SMOTE) sebelum model Naive Bayes dan SVM dilatih, guna memastikan distribusi data yang lebih proporsional antar kelas sentimen. Pengujian dilakukan untuk mengetahui algoritma mana yang menghasilkan kinerja klasifikasi terbaik. Berdasarkan hasil evaluasi, algoritma SVM memperoleh nilai akurasi sebesar 75%, precision 0.86, recall 0.79, dan f1-score 0.79. Sementara itu, algoritma Naive Bayes mencapai akurasi sebesar 65%, dengan precision 0.89, recall 0.71, dan f1-score 0.79. Meskipun Naive Bayes mencatat precision yang lebih tinggi, hasil keseluruhan menunjukkan bahwa SVM memiliki performa klasifikasi yang lebih konsisten dan seimbang, terutama dalam konteks data ulasan yang bersifat tidak terstruktur.

Kata Kunci: Analisis Sentimen, *Naïve Bayes*, *Support Vector Machine*, Google Maps, Mall Pakuwon Bekasi

ABSTRACT

Natasya Marito T. 202110715197. *Comparison of Naïve Bayes and SVM Algorithms on Sentiment Analysis of Pakuwon Bekasi Mall. Bekasi: Faculty of Computer Science. University of Bhayangkara Jakarta Raya. 2025.*

This study aims to compare the performance of two text classification algorithms, namely Naive Bayes and Support Vector Machine (SVM), in conducting sentiment analysis on user reviews of Mall Pakuwon Bekasi, obtained from Google Maps. A total of 1,037 reviews were collected using web scraping techniques. The dataset was then split into 80% training data and 20% testing data, followed by preprocessing and text vectorization using the TF-IDF method. To address the class imbalance in the training data, the Synthetic Minority Over-sampling Technique (SMOTE) was applied prior to training both the Naive Bayes and SVM models, ensuring a more proportional distribution of sentiment classes. The goal of this evaluation is to determine which algorithm delivers the most accurate classification results. Based on the findings, the SVM algorithm achieved an accuracy of 75%, with precision of 0.86, recall of 0.79, and an f1-score of 0.79. Meanwhile, Naive Bayes attained an accuracy of 65%, with precision of 0.89, recall of 0.71, and the same f1-score of 0.79. Although Naive Bayes recorded a higher precision, the overall results indicate that SVM offers more balanced and consistent classification performance, especially when dealing with unstructured user-generated content.

Keywords: *Sentiment Analysis, Naïve Bayes, Support Vector Machine, Google Maps, Mall Pakuwon Bekasi*

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

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

Nama : Natasya Marito T
NPM : 202110715197
Program Studi : Informatika
Fakultas : Ilmu Komputer
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 :

“Perbandingan Algoritma *Naïve Bayes* dan *SVM* pada Analisis Sentimen Ulasan Mall Pakuwon Bekasi di *Review Google Maps*”

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 izin 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 : 31 Juli 2025

Yang Menyatakan



Natasya Marito T

KATA PENGANTAR

Segala puji dan syukur ke hadirat Allah SWT atas limpahan rahmat, karunia, dan hidayah-Nya sehingga tugas akhir dengan judul: “Perbandingan Algoritma Naïve Bayes dan SVM pada Analisis Sentimen Ulasan Mall Pakuwon Bekasi di *Review Google Maps*” dapat diselesaikan sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya. Ucapan terima kasih disampaikan kepada semua pihak yang telah memberikan bimbingan, dukungan, dan doa selama proses penyusunan skripsi ini. Secara khusus, ucapan terimakasih yang tulus diberikan kepada:

1. Bapak Irjen Polisi (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, S.E., M.M.S.I., selaku Ketua Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
4. Ibu Prima Dina Atika, S.Kom., M.Kom., dan Mrs. Khairunnisa Fadhilla Ramdhanian, S.Si., M.Si., selaku Dosen Pembimbing I dan Dosen Pembimbing II atas bimbingan, arahan, dan dukungan yang konsisten selama proses penyusunan skripsi ini.
5. Bapak Agus Hidayat, S.Kom., M.Si., selaku Dosen Pembimbing Akademik, atas pendampingan dan dukungannya selama masa perkuliahan.

6. Kedua orang tua tercinta, atas kasih sayang, doa, dukungan moral, dan semangat yang tiada henti dalam setiap langkah, serta menjadi sumber kekuatan selama proses studi dan penyusunan skripsi ini.
7. Teman-teman seperjuangan yang selalu memberikan semangat dan dukungan yaitu “Princess A5” atas kebersamaan, tawa, serta dukungan moril yang berarti selama proses perkuliahan hingga penyusunan skripsi ini.

Jakarta, 31 Juli 2025

Hormat Saya



Natasya Marito T
202110715197



DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN BUKAN PLAGIASI.....	iv
ABSTRAK	v
ABSTRACT	vi
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI	vii
KATA PENGANTAR	viii
DAFTAR ISI.....	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	5
1.3 Rumusan Masalah	6
1.4 Batasan Masalah	6
1.5 Tujuan Penelitian	7
1.6 Manfaat Penelitian	7
1.7 Sistematika Penulisan.....	8
BAB II TINJAUAN PUSTAKA.....	10
2.1 Penelitian Terkait	10
2.1.1 Tabel <i>Related Research /State Of The Art</i>	10
2.2 Mall Pakuwon Bekasi.....	13
2.3 Google Maps.....	13
2.4 Data Mining	14
2.5 Machine Learning	16
2.6 Analisis Sentimen	17
2.7 Web Scraping	18
2.8 Lexicon Based	19

2.9	<i>SMOTE</i>	20
2.10	<i>Text Mining</i>	21
2.11	<i>Confusion Matrix</i>	23
2.12	<i>Naïve Bayes</i>	25
2.13	<i>Support Vector Machine</i>	27
2.14	<i>Knowledge Discovery in Database (KDD)</i>	28
2.15	Pembobotan Kata	31
BAB III METODOLOGI PENELITIAN		33
3.1	Objek Penelitian.....	33
3.2	Kerangka Penelitian	34
3.3	Metode Pengumpulan Data	41
3.4	<i>Data Scraping</i>	42
3.5	<i>Preprocessing</i>	44
3.5.1	Case Folding	44
3.5.2	Cleaning Data.....	45
3.5.3	Tokenizing	46
3.5.4	Normalization.....	47
3.5.5	Filtering/Stopword Removal.....	48
3.5.6	Stemming	50
3.5.7	Labelling.....	52
3.5.8	Transformation.....	54
3.6	<i>Splitting Data</i>	62
3.7	Analisis Kebutuhan Sistem.....	63
BAB IV HASIL DAN PEMBAHASAN		65
4.1	Data Mining	65
4.1.1	Naïve Bayes	65
4.1.2	Support Vector Machine	72
4.1.3	Pengujian Data Teks.....	74
4.2	Evaluation.....	76
4.2.1	<i>Evaluation Naïve Bayes</i>	77
4.2.1	<i>Evaluation SVM</i>	79

4.3	<i>WordCloud</i>	82
4.3.1	<i>WordCloud Naïve Bayes</i>	83
4.3.2	<i>WordCloud SVM</i>	85
BAB V PENUTUP		88
5.1	Kesimpulan	88
5.2	Saran	89
DAFTAR PUSTAKA.....		91
LAMPIRAN.....		94



DAFTAR TABEL

Tabel 2. 1 Tabel Related Research/ State Of The Art	10
Tabel 2. 2 Tabel Confusion Matrix	24
Tabel 3. 1 Hasil Tahapan Case Folding	44
Tabel 3. 2 Hasil Proses Tokenisasi	47
Tabel 3. 3 Hasil Proses Normalisasi	48
Tabel 3. 4 Hasil Proses Filtering/Stopword Removal	49
Tabel 3. 5 Tabel perhitungan manual vector	57
Tabel 3. 6 Tabel perhitungan tf-idf	58
Tabel 3. 7 Tabel Perhitungan	59
Tabel 3. 8 Tabel perhitungan bobot	61
Tabel 4. 1 Tabel Sampel Data Train	66
Tabel 4. 2 Tabel Sampel Data Test	66
Tabel 4. 3 Tabel Penentuan Prediksi dan Aktual	67
Tabel 4. 4 Perhitungan Prior Probability	67
Tabel 4. 5 Hasil Uji Kategori	71
Tabel 4. 6 Ulasan Perhitungan Manual	73
Tabel 4. 7 Multiclass Confusion Matrix	77
Tabel 4. 8 Perbandingan Confusion Matrix Naïve Bayes dan SVM	82

DAFTAR GAMBAR

Gambar 1. 1 Grafik Data Urgensi Analisis Mall.....	2
Gambar 2. 1 Mall Pakuwon Bekasi.....	33
Gambar 2. 2 Google Maps Mall Pakuwon Bekasi	33
Gambar 2. 3 Hyperplane SVM.....	27
Gambar 2. 4 Tahapan KDD	29
Gambar 3. 1 Kerangka Penelitian.....	35
Gambar 3. 2 Tahapan Naïve Bayes Sumber: Hasil Penelitian (2025)	39
Gambar 3. 3 Tahapan SVM	41
Gambar 3. 4 Parameter Crawling	42
Gambar 3. 5 Script Crawling Ulasan.....	43
Gambar 3. 6 Hasil Crawling Data	43
Gambar 3. 7 Script Case Folding	44
Gambar 3. 8 Script Cleaning.....	45
Gambar 3. 9 Hasil Proses Cleaning.....	46
Gambar 3. 10 Script Proses Tokenisasi	46
Gambar 3. 11 Script Proses Normalisasi	47
Gambar 3. 12 Script Proses Filtering/Stopword Removal.....	49
Gambar 3. 13 Script Mencari Kata Paling Sering Muncul	50
Gambar 3. 14 Hasil 15 Token Paling Umum.....	50
Gambar 3. 15 Script Proses <i>Stemming</i>	51
Gambar 3. 16 Hasil Proses <i>Stemming</i>	52
Gambar 3. 17 Script Labelling	53
Gambar 3. 18 Jumlah Hasil Labelling	53
Gambar 3. 19 Grafik Kelas Sentimen.....	54
Gambar 3. 20 Hasil Labelling	54
Gambar 3. 21 Penerapan SMOTE	55
Gambar 3. 22 Hasil Distribusi Setelah SMOTE	55
Gambar 3. 23 Script Proses TF-IDF	56
Gambar 3. 24 Hasil TF-IDF	56

Gambar 3. 25 Kode dan Hasil Splitting Data Rasio 90:10	63
Gambar 3. 26 Kode dan Hasil Splitting Data Rasio 80:20	63
Gambar 3. 27 Kode dan Hasil Splitting Data Ratio 90:10.....	63
Gambar 4. 1 Import Library Sklearn Naïve Bayes.....	65
Gambar 4. 2 Penerapan SMOTE.....	66
Gambar 4. 3 Import Library Sklearn Support Vector Machine	72
Gambar 4. 4 Penerapan SMOTE.....	72
Gambar 4. 5 Hasil Evaluasi Naïve Bayes	78
Gambar 4. 6 Confusion Matrix Naïve Bayes	79
Gambar 4. 7 Hasil Evaluasi SVM	80
Gambar 4. 8 Confusion Matrix SVM	81
Gambar 4. 9 WordCloud Positive Naïve Bayes	83
Gambar 4. 10 WordCloud Neutral Naïve Bayes.....	84
Gambar 4. 11 WordCloud Negative Naïve Bayes	85
Gambar 4. 12 WordCloud Positive SVM	85
Gambar 4. 13 WordCloud Neutral SVM	86
Gambar 4. 14 WordCloud Negative SVM.....	87

DAFTAR LAMPIRAN

Uji Plagiarisme

Biodata Mahasiswa

Kartu Konsultasi Bimbingan

Kartu Bimbingan Skripsi Dosen Pembimbing I

Kartu Bimbingan Skripsi Dosen Pembimbing II

Surat Rekomendasi Dari Pembimbing I

Surat Rekomendasi Dari Pembimbing II

Dataset

