

**ANALISIS PERBANDINGAN ALGORITMA SUPPORT VECTOR
MACHINE (SVM) DAN NAIVE BAYES CLASSIFIER (NBC) DALAM
MELAKUKAN ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI
SHOPEE PADA GOOGLE PLAY STORE**

SKRIPSI

**Untuk Memenuhi Persyaratan Memperoleh Gelar Sarjana (S1) Pada
Program Studi Sistem Informasi Fakultas Sains Dan Teknologi
Universitas Labuhanbatu**



OLEH :

SANNY KHAIRANI LUBIS, S.Kom

2309105195

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS LABUHANBATU

RANTAUPRAPAT

2025

HALAMAN PERSETUJUAN

JUDUL TUGAS AKHIR : ANALISIS PERBANDINGAN
ALGORITMA SUPPORT VECTOR
MACHINE (SVM) DAN NAIVE BAYES
CLASSIFIER (NBC) DALAM
MELAKUKAN ANALISIS SENTIMEN
ULASAN PENGGUNA APLIKASI
SHOPEE PADA GOOGLE PLAY STORE

NAMA : SANNY KHAIRANI LUBIS

NPM : 2309105195

PRODI : SISTEM INFORMASI

Disetujui Pada Tanggal 07 Maret 2025

Pembimbing I



Budianto Bangun, S.Sos., M.Kom
NIDN : 0113028702

Pembimbing II



Irmayanti, S.Si., M.Pd
NIDN : 0124088404

LEMBAR PENGESAHAN NASKAH SKRIPSI

JUDUL TUGAS AKHIR : ANALISIS PERBANDINGAN
ALGORITMA SUPPORT VECTOR
MACHINE (SVM) DAN NAIVE BAYES
CLASSIFIER (NBC) DALAM
MELAKUKAN ANALISIS SENTIMEN
ULASAN PENGGUNA APLIKASI
SHOPEE PADA GOOGLE PLAY STORE

NAMA : SANNY KHAIRANI LUBIS
NPM : 2309105195
PRODI : SISTEM INFORMASI
KONSENTRASI : UMUM

Telah Diuji Dan Dinyatakan Lulus Ujian Sarjana Pada Tanggal 07 Maret 2025

TIM PENGUJI

Penguji 1(Ketua)

Nama : Masrizal, S.Kom., M.Kom
NIDN : 0105039401

Tanda Tangan

Penguji II(Anggota)

Nama : Budianto Bangun, S.Sos., M.Kom
NIDN : 0113028702

Penguji III(Anggota)

Nama : Irmayanti, S.Si., M.Pd
NIDN : 0124088404

Rantauprapat, 07 Maret 2025

Dekan
Fakultas Sains Dan Teknologi



Dr. Iwan Purnama, S.Kom., M.Kom
NIDN.0112029202

Kaprodi Sistem Informasi
Fakultas Sains Dan Teknologi



Budianto Bangun, S.Sos., M.Kom
NIDN.0124047003

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

NAMA : SANNY KHAIRANI LUBIS
NPM : 2309105195
PROGRAM STUDI : SISTEM INFORMASI
JUDUL SKRIPSI : Analisis Perbandingan Algoritma Support Vector
Machine (SVM) Dan Naive Bayes Classifier (NBC)
Dalam Melakukan Analisis Sentimen Ulasan
Pengguna Aplikasi Shopee Pada Google Play Store

Dengan ini penulis menyatakan bahwa skripsi ini disusun sebagai syarat untuk memperoleh gelar sarjana pada program studi sistem informasi fakultas sains dan teknologi universitas labuhanbatu adalah hasil karya penulis sendiri. Semua kutipan maupun rujukan dalam penulisan skripsi ini telah penulis cantumkan sumber dengan benar sesuai dengan ketentuan yang berlaku.

Jika dikemudian hari ternyata ditemukan seluruh atau sebagian skripsi ini bukan hasil karya penulis atau plagiat, penulis bersedia menerima sanksi pencabutan gelar akademik yang disandang dengan sanksi-sanksi lainnya sesuai dengan peraturan perundang-undangan yang berlaku.

Rantauprapat, 07 Maret 2025

Yang membuat pernyataan



SEPUKUH RIBU RUPIAH
10000
TSL
METERAI
TEMPEL
5E5BCAMX162693984

Sanny Khairani Lubis

KATA PENGANTAR

Puji syukur penulis haturkan kehadiran Allah SWT atas berkat rahmatnya dan karunia-Nya penulis dapat menyelesaikan penulisan proposal ini dengan judul “Analisis Perbandingan Algoritma Support Vector Machine (Svm) Dan Naive Bayes Classifier (Nbc) Dalam Melakukan Analisis Sentimen Ulasan Pengguna Aplikasi Shopee Pada Google Play Store” yang merupakan salah satu persyaratan kelulusan untuk memperoleh gelar Sarjana Strata S1 di Program Studi Sistem Informasi Fakultas Sains Dan Teknologi Universitas Labuhanbatu. Penulisan proposal ini dapat terselesaikan dengan baik berkat dukungan dari berbagai pihak. Pada kesempatan ini penulis mengucapkan banyak terimakasih kepada :

1. Bapak Assoc. Prof. Ade Parlaungun Nasution, S.B., M.Si., Ph.d. selaku rektor Universitas Labuhanbatu.
2. Bapak Dr. Iwan Purnama, S.Kom., M.Kom. selaku dekan Fakultas Sains dan Teknologi.
3. Bapak Budianto Bangun, S.Sos., M.Kom. selaku Ketua Program Prodi Sistem Informasi serta Dosen Pembimbing Pendamping 1 yang telah memberikan dukungan selama masa perkuliahan di jurusan Sistem Informasi.
4. Ibu Irmayanti, S.Si., M.Pd. selaku Dosen Pembimbing Pendamping II yang telah meluangkan waktu memberikan dukungan, bimbingan dan motivasi dalam menyelesaikan proposal penelitian saya ini.
5. Bapak Masrizal, S.Kom., M.Kom. selaku Dosen Penguji yang telah meluangkan waktu memberikan dukungan, bimbingan dan motivasi dalam menyelesaikan proposal penelitian saya ini.
6. Saya juga mengucapkan terimakasih kepada semua dosen-dosen fakultas sains dan teknologi khususnya dosen-dosen sistem informasi, seluruh staf/pegawai fakultas sains dan labuhanbatu yang telah memberikan amanat, bantuan, dan saran selama saya melakukan penelitian ini.
7. Teristimewa kepada Kedua Orang Tua Saya, Bapak Bangun Lubis dan Ibu Purnama Br.Ginting, S.E. yang telah bekerja keras untuk pendidikan saya dan senantiasa mendidik, membimbing, memberi dukungan baik moral maupun material serta doa-doa yang tak terhingga kepada saya selama berkuliah di Universitas Labuhanbatu, saya sangat berharap agar kelak saya menjadi anak berguna dan bermanfaat kepada kedua orang tua, saudara dan semua orang.
8. Teristimewa kepada saudara saya Febby Erna Sari Lubis, S.Pd, Bisma Salsa Fitri Lubis, Jihan Assyifa Lubis, beserta seluruh Keluarga Besar yang selalu memberikan doa, semangat dan motivasi kepada saya.
9. Teman seperjuangan yang telah memberikan motivasi dan dukungan bagi penulis, tetap solid dan semoga kita semua wisuda tepat waktu.
10. Terimakasih kepada Suwegiono Putra yang selalu menemani, membantu dan memberikan support selama saya menyelesaikan proposal penelitian

ini. Terimakasih karena telah ikhlas selalu saya repotkan dalam menjalani proses skripsian ini.

Saya juga sangat berterimakasih kepada semua pihak yang telah membantu penulis dalam penyusunan proposal penelitian ini dan tidak dapat penulis sebutkan satu persatu. Penulis menyadari bahwa penulisan ini masih jauh dari kesempurnaan. Oleh karena itu, kritik dan saran yang memotivasi sangat penulis harapkan dari pembaca guna perbaikan penulis dikemudian hari. Semoga proposal ini bermanfaat bagi pembaca dan bagi perkembangan di sistem informasi. Sebagai ungkapan rasa terimakasih yang mendalam penulis hanya dapat berdoa semoga segala perhatian, bantuan, dukungan serta semangat yang telah diberikan mendapat balasan yang lebih baik dari Allah SWT. Kekurangan akan selalu menjadi bagian manusia, kesempurnaan hanya miliknya, untuk itu saran dan kritik akan lebih menyempurnakan pikiran penulis untuk melangkah lebih lanjut. Semoga skripsi ini dapat memberikan manfaat bagi kita semua.

Rantauprapat, 07 Maret 2025



Sanhy Khairani Lubis

ABSTRAK

Perkembangan *e-commerce* di Indonesia semakin pesat, dengan Shopee sebagai salah satu platform terbesar yang memiliki jutaan ulasan dari pengguna. Namun, jumlah ulasan yang besar membuat analisis manual menjadi tidak efisien. Oleh karena itu, penelitian ini bertujuan untuk menerapkan dan membandingkan kinerja algoritma Support Vector Machine dan Naïve Bayes Classifier dalam melakukan analisis sentimen terhadap ulasan pengguna aplikasi Shopee yang diperoleh dari Google Play Store. Penelitian ini dilakukan dengan beberapa tahapan utama, yaitu pengumpulan data, pelabelan sentimen, *preprocessing* teks, pembobotan fitur menggunakan TF-IDF, klasifikasi menggunakan Support Vector Machine dan Naïve Bayes Classifier, serta evaluasi kinerja model. Dataset yang digunakan terdiri dari 5000 ulasan yang dikategorikan ke dalam tiga kelas sentimen: Positif, Netral, dan Negatif. Hasil penelitian menunjukkan bahwa Support Vector Machine memiliki akurasi lebih tinggi 72% dibandingkan Naïve Bayes Classifier, dengan nilai 71,45%. Dari hasil ini, dapat disimpulkan bahwa Support Vector Machine lebih cocok untuk analisis sentimen ulasan pengguna Shopee karena kemampuannya dalam mengenali pola kompleks dalam data.

Kata Kunci: Analisis Sentimen, Google Play Store, Naïve Bayes Classifier, Shopee, Support Vector Machine.

ABSTRACT

The development of e-commerce in Indonesia is increasingly rapid, with Shopee as one of the largest platforms that has millions of reviews from users. However, the large number of reviews makes manual analysis inefficient. Therefore, this study aims to apply and compare the performance of the Support Vector Machine and Naïve Bayes Classifier algorithms in conducting sentiment analysis of Shopee application user reviews obtained from the Google Play Store. This study was conducted with several main stages, namely data collection, sentiment labeling, text preprocessing, feature weighting using TF-IDF, classification using Support Vector Machine and Naïve Bayes Classifier, and model performance evaluation. The dataset used consists of 5000 reviews categorized into three sentiment classes: Positive, Neutral, and Negative. The results of the study showed that the Support Vector Machine has a higher accuracy of 72% compared to the Naïve Bayes Classifier, with a value of 71.45%. From these results, it can be concluded that the Support Vector Machine is more suitable for sentiment analysis of Shopee user reviews because of its ability to recognize complex patterns in data.

Keywords: Sentiment Analysis, Google Play Store, Naïve Bayes Classifier, Shopee, Support Vector Machine.

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN.....	iv
KATA PENGANTAR.....	v
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI.....	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah	1
1.2 Perumusan Masalah.....	3
1.3 Ruang Lingkup Masalah	4
1.4 Tujuan dan Manfaat Penelitian.....	4
1.4.1 Tujuan Penelitian	4
1.4.2 Manfaat Penelitian	5
1.5 Sistematika Penulisan.....	5
BAB II LANDASAN TEORI	7
2.1 Analisis Sentimen.....	7
2.2 Text Preprocessing	8
2.3 Support Vector Machine (SVM)	11
2.4 Naïve Bayes Classifier (NBC)	14
2.5 Confusion Matrix	18

2.6	Alat Bantu Pemrograman	20
2.6.1	Python	20
2.6.2	Google Colaboratory	22
2.6.3	Jupyter Notebook	24
2.7	Penelitian Terdahulu.....	25
BAB III ANALISA DAN PERANCANGAN.....		30
3.1	Alat Dan Perangkat Lunak Yang Digunakan	31
3.2	Alur Penelitian.....	31
3.2.1	Pengumpulan Data	31
3.2.2	Pelabelan Data.....	31
3.2.3	Text Preprocessing	32
3.3	Visualisasi Data	34
3.4	Pembobotan TF-IDF.....	34
3.5	Klasifikasi.....	34
3.6	Evaluasi Kinerja	34
BAB IV IMPLEMENTASI DAN PEMBAHASAN		36
4.1	Implementasi Sistem	36
4.1.1	Alat dan Perangkat Lunak yang Digunakan.....	36
4.1.2	Pengumpulan Data	38
4.1.3	Pelabelan Data.....	41
4.1.4	Text Preprocessing	42
4.1.5	Visualisasi Data.....	48
4.1.6	Pembobotan TF-IDF	49
4.1.7	Klasifikasi	50
4.1.8	Evaluasi	52
4.2	Pembahasan	53

BAB V KESIMPULAN DAN SARAN	69
5.1 Kesimpulan.....	69
5.2 Saran	70
DAFTAR RIWAYAT HIDUP	78
A. Identitas Pribadi	78
B. Riwayat Pendidikan	78
DAFTAR PUSTAKA.....	72

DAFTAR TABEL

Tabel 2.2 Penelitian Terdahulu	25
Tabel 4.1 Frekuensi Kemunculan Kata Hasil TF-IDF	64
Tabel 4.2 Hasil Evaluasi Kinerja SVM vs NBC	67

DAFTAR GAMBAR

Gambar 2.1 Tahapan Text Preprocessing	9
Gambar 2.2 Model Algoritma SVM	12
Gambar 2.3 Penemuan Hyperlane pada SVM	12
Gambar 2.4 Fungsi Pemisahan optimal SVM.....	14
Gambar 2.5 Alur Kerja Algoritma Naïve Bayes.....	17
Gambar 2.6 Logo Python	20
Gambar 2.7 Google Colab	23
Gambar 2.8 Jupyter Notebook	24
Gambar 3.1 Alur Penelitian	30
Gambar 4.1 Instalasi Pustaka <i>google-play-scraper</i>	38
Gambar 4.2 Import Pustaka	38
Gambar 4.3 Pengambilan Data Ulasan	39
Gambar 4.4 Konversi Data ke DataFrame	40
Gambar 4.5 Proses Pelabelan Dataset.....	41
Gambar 4.6 Proses Case Folding	43
Gambar 4.7 Proses <i>Tokenizing</i>	44
Gambar 4.8 Proses Formalisasi.....	45
Gambar 4.9 Proses Stopword Removal	47
Gambar 4.10 Proses Stemming.....	48
Gambar 4.11 Proses Visualisasi Data	49
Gambar 4.12 Proses Pembobotan Kata dengan TF-IDF	50
Gambar 4.13 Proses Klasifikasi dengan SVM.....	51
Gambar 4.14 Proses Klasifikasi dengan NBC	52
Gambar 4.15 Proses Evaluasi.....	53
Gambar 4.16 Cuplikan Dataset	53
Gambar 4.17 Hasil Pelabelan Dataset.....	55

Gambar 4.18 Hasil <i>Text Preprocessing</i>	56
Gambar 4.19 <i>WordCloud</i> Sentimen Negatif	58
Gambar 4.20 <i>WordCloud</i> Sentimen Netral	60
Gambar 4.21 <i>WordCloud</i> Sentimen Positif.....	62
Gambar 4.22 Grafik Jumlah Sentimen.....	63
Gambar 4.23 Perbandingan Akurasi SVM dan NBC	66

DAFTAR LAMPIRAN

A. BIODATA PENULIS.....	72
-------------------------	----