

**KOMPARASI PERFORMA ALGORITMA NAÏVE BAYES
DAN SUPPORT VECTOR MACHINE PADA ANALISIS
SENTIMEN PROGRAM KAMPUS MENGAJAR 7**

SKRIPSI SARJANA KOMPUTER

Oleh :

Shereen Putri Djunaidi

217006516086



PROGRAM STUDI SISTEM INFORMASI

FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA

UNIVERSITAS NASIONAL

2025

**PERFORMANCE COMPARISON OF NAÏVE BAYES AND
SUPPORT VECTOR MACHINE ALGORITHMS ON
SENTIMENT ANALYSIS OF KAMPUS
MENGAJAR 7 PROGRAM**

SKRIPSI SARJANA KOMPUTER

Oleh :

Shereen Putri Djunaidi

217006516086



PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA
UNIVERSITAS NASIONAL

2025

**KOMPARASI PERFORMA ALGORITMA NAÏVE BAYES
DAN SUPPORT VECTOR MACHINE PADA ANALISIS
SENTIMEN PROGRAM KAMPUS MENGAJAR 7**

SKRIPSI SARJANA

Karya ilmiah sebagai salah satu syarat untuk memperoleh gelar
Sarjana Komputer dari Fakultas Teknologi Komunikasi dan Informatika

Oleh :

Shereen Putri Djunaidi

217006516086

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA

UNIVERSITAS NASIONAL

2025

HALAMAN PENGESAHAN

TUGAS AKHIR

KOMPARASI PERFORMA ALGORITMA NAÏVE BAYES DAN SUPPORT
VECTOR MACHINE PADA ANALISIS SENTIMEN PROGRAM KAMPUS
MENGAJAR 7



Dosen Pembimbing 1

Moch. Firmansyah, S. Kom., M. Kom

PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

KOMPARASI PERFORMA ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MACHINE PADA ANALISIS SENTIMEN PROGRAM KAMPUS MENGAJAR 7

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasional maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber informasi atau acuan yang dicantumkan sebagaimana mestinya.

Jakarta, 04 Maret 2025

The block contains an official stamp of Universitas Nasional, featuring a red and white emblem with the text 'UNIVERSITAS NASIONAL' and '1000'. Below the stamp is a handwritten signature in black ink.

Shereen Putri Djunaidi

217006516086

HALAMAN PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya saya sendiri, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Bilamana di kemudian hari ditemukan bahwa karya tulis ini menyalahi peraturan yang ada berkaitan etika dan kaidah penulisan karya ilmiah yang berlaku, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku

Yang menyatakan,

Nama : Shereen Putri Djunaidi

NIM : 217006516086

Tanda Tangan : 

Tanggal : 04 Maret 2025

Mengetahui

Pembimbing I : Moch. Firmansyah, S.Kom., M.Kom (

)



HALAMAN PENGESAHAN

TUGAS SARJANA

**KOMPARASI PERFORMA ALGORITMA NAÏVE BAYES
DAN SUPPORT VECTOR MACHINE PADA ANALISIS
SENTIMEN PROGRAM KAMPUS MENGAJAR 7**

Oleh :

Shereen Putri Djunaidi

217006516086

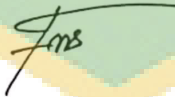
Program Studi Sistem Informasi

Fakultas Teknologi Komunikasi dan Informatika

Universitas Nasional

Disetujui pada Tanggal :

Pembimbing I



(Moch. Firmansyah, S.Kom., M.Kom)

NIP. 0317128602

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa yang telah memberikan rahmat dan karunia sehingga penulis dapat menyelesaikan skripsi dengan judul **“Komparasi Performa Algoritma Naïve Bayes dan Support Vector Machine Pada Analisis Sentimen Program Kampus Mengajar 7”** sebagai salah satu syarat kelulusan Program Studi Sarjana Sistem Informasi Fakultas Teknologi Komunikasi dan Informatika.

Penelitian dan penulisan skripsi ini tidak terlepas dari bantuan berbagai pihak, oleh karena itu penulis menyampaikan banyak terima kasih terutama kepada dosen pembimbing Tugas Akhir, Bapak Moch. Firmansyah, S.Kom., M.Kom yang telah meluangkan banyak waktu, tenaga, pikiran, bimbingan, arahan, motivasi serta memaklumi segala kekurangan penulis selama penelitian tugas akhir dan penyusunan skripsi. Penulis juga mengucapkan banyak terima kasih kepada :

1. Para pengguna media sosial X yang telah membagikan pendapatnya mengenai program Kampus Mengajar 7 sehingga mendukung penelitian dalam bentuk data.
2. Ayah dan Ibu selaku orang tua penulis yang telah banyak memberi dukungan salam segala bentuk yang tak terhitung.
3. Seluruh dosen pengajar di Program Studi Sistem Informasi FTKI maupun dosen di Program Studi lain yang memberikan banyak ilmu.
4. Teman-teman seangkatan yang telah membantu serta mendukung dengan saling berbagi informasi dan diskusi.
5. Rekan tim program Kampus Mengajar 7 sekolah penugasan SD Negeri Mekarjaya 12 yang telah memberikan banyak dukungan dan semangat.

Akhir kata, semoga Tuhan Yang Maha Esa membalas kebaikan dan bantuan yang telah diberikan dengan hal yang lebih baik. Penulis mengharapkan kritik dan saran yang bersifat membangun dan semoga skripsi ini dapat memberikan manfaat di bidang pertanian.

Jakarta, 20 Februari 2025

Penulis

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika, saya yang bertanda tangan di bawah ini:

Nama : Shereen Putri Djunaidi

NIM : 217006516086

demikian demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Fakultas Teknologi Komunikasi dan Informatika, Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalti Free Right*) atas karya ilmiah saya yang berjudul:

KOMPARASI PERFORMA ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MACHINE PADA ANALISIS SENTIMEN PROGRAM KAMPUS MENGAJAR 7

Berserta perangkat yang ada (jika diperlukan). Dengan Hak ini Fakultas Teknologi Komunikasi dan Informatika berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal : 04 Maret 2025

Yang menyatakan



(Shereen Putri Djunaidi)

ABSTRAK

Algoritma Naïve Bayes dan Support Vector Machine (SVM) merupakan jenis algoritma supervised learning yang sering digunakan pada analisis sentimen, kedua algoritma tersebut memiliki karakteristik dan cara kerja yang berbeda, oleh karena itu dilakukan penelitian untuk membandingkan performanya pada suatu proses analisis sentimen dengan data yang sama. Program Kampus Mengajar 7 memiliki beberapa perubahan kebijakan yang memicu beragam opini publik, untuk mengetahui sentimen publik terhadap perubahan kebijakan tersebut maka dilakukan analisis sentimen dengan menggunakan algoritma naïve bayes dan SVM. Data yang digunakan berasal dari opini publik yang dibagikan melalui media sosial X (twitter). Penelitian ini mendapatkan hasil bahwa algoritma naïve bayes memiliki nilai akurasi 56,9% dan SVM 51,7%. Dalam waktu pengelolaan data naïve bayes cenderung lebih cepat dibandingkan dengan SVM, sedangkan pada AUC score naïve bayes mendapat 0.6695 dan SVM adalah 0.6719. Distribusi sentimen yang dihasilkan oleh model klasifikasi naïve bayes maupun SVM mendapatkan hasil bahwa kelas sentimen mayoritas adalah positif, yang berarti meskipun memiliki perubahan kebijakan, mayoritas publik memiliki respons positif terhadap program Kampus Mengajar 7. Berdasarkan hasil penelitian ini naïve bayes cukup unggul dibandingkan dengan SVM, sehingga dalam pengelolaan data dengan karakteristik serupa algoritma naïve bayes lebih direkomendasikan.

Kata Kunci : Naïve Bayes, Support Vector Machine, *Supervised Learning*, Analisis Sentimen, Kampus Mengajar.

ABSTRACT

Naïve Bayes and Support Vector Machine (SVM) are types of supervised learning algorithms that are often used in sentiment analysis, both algorithms have different characteristics and ways of working, therefore research was conducted to compare their performance in a sentiment analysis process with the same data. The Kampus Mengajar 7 program has several policy changes that spark various public opinions, to find out public sentiment towards these policy changes, sentiment analysis is carried out using the naïve bayes and SVM algorithms. The data used comes from public opinions shared through social media X (twitter). This study found that the naïve bayes algorithm has an accuracy value of 56.9% and SVM 51.7%. In data management time naïve bayes tends to be faster than SVM, while in the AUC score naïve bayes gets 0.6695 and SVM is 0.6719. The sentiment distribution generated by both the naïve bayes and SVM classification models found that the majority sentiment class is positive, which means that despite the policy changes, the majority of the public has a positive response to the Teaching Campus 7 program. Based on the results of this study, naïve bayes is quite superior compared to SVM, so in managing data with similar characteristics, the naïve bayes algorithm is more recommended.

Keyword : Naïve Bayes, Support Vector Machine, Supervised Learning, Sentiment Analysis, Kampus Mengajar.

DAFTAR ISI

HALAMAN PERNYATAAN ORISINALITAS	i
HALAMAN PENGESAHAN	ii
KATA PENGANTAR	iii
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI	iv
ABSTRAK	v
ABSTRACT	vi
DAFTAR ISI	vii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
BAB I	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Kontribusi	4
BAB II	5
TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Landasan Teori	11
2.2.1 Kampus Mengajar	11
2.2.2 Analisis Sentimen	12
2.2.3 Naïve Bayes	12
2.2.4 Support Vector Machine	13
BAB III	14
METODOLOGI PENELITIAN	14
3.1 Lokasi dan Waktu Penelitian	14
3.2 Tahapan Penelitian	14
3.3.1 Pengumpulan Data	15

3.3.2	<i>Preprocessing Data</i>	16
3.3.3	Pelabelan Data	17
3.3.4	Pembagian Data.....	18
3.3.5	<i>Vectorization</i>	18
3.3.6	Pembobotan TF-IDF	18
3.3.7	SMOTE	19
3.3.8	Pembuatan Model Analisis Sentimen	19
3.3.9	Pengujian Model.....	20
3.3.10	Evaluasi Performa Model	21
3.3.11	Analisis Komparasi Performa Model	22
BAB IV		23
HASIL DAN PEMBAHASAN		23
4.1	Pengumpulan Data.....	23
4.1.1	Masa Pendaftaran.....	23
4.1.2	Masa Pra-pelaksanaan	23
4.1.3	Masa Pelaksanaan	23
4.1.4	Masa Pasca-pelaksanaan	24
4.1.5	Penggunaan Hastag	24
4.2	<i>Preprocessing Data</i>	26
4.2.1	<i>Cleansing</i>	26
4.2.2	<i>Tokenization</i>	27
4.2.3	<i>Stopword Removal</i>	27
4.2.4	<i>Stemming</i>	27
4.3	Pelabelan Data.....	27
4.4	Pembagian Data	28
4.5	<i>Vectorization</i>	28
4.6	Pembobotan TF-IDF.....	28
4.7	SMOTE	29
4.8	Pembuatan Model	29
4.5.1	Naïve Bayes	29
4.5.2	Support Vector Machine (SVM)	33

4.9	Evaluasi Performa Model.....	35
4.7.1	Naïve Bayes	35
4.7.2	SVM	38
4.10	Analisis Komparasi Performa Model.....	40
BAB V		44
KESIMPULAN DAN SARAN.....		44
5.1.	Kesimpulan.....	44
5.2.	Saran.....	45
DAFTAR PUSTAKA		46



DAFTAR GAMBAR

Gambar 2.1 Perkembangan Program Kampus Mengajar	11
Gambar 3.1 Flowchart Tahap Penelitian	14
Gambar 4.1 Hasil Pengumpulan Data	24
Gambar 4.2 Flowchart Tahapan Preprocessing Data	26
Gambar 4. 3 Distribusi Sentimen Labeling Kamus Lexicon	27
Gambar 4. 4 Distribusi Sentimen Naive Bayes	29
Gambar 4. 5 Distribusi Sentimen SVM.....	34
Gambar 4. 6 Komparasi Nilai Evaluasi Naive Bayes & SVM.....	40
Gambar 4. 7 Komparasi Waktu Kerja Naive Bayes & SVM	41
Gambar 4. 8 Komparasi AUC Score Naive Bayes & SVM	42
Gambar 4. 9 Komparasi Distribusi Sentimen.....	42



DAFTAR TABEL

Tabel 3.1 Dataset Awal	15
Tabel 4.1 Sample Dataset Hasil Pengumpulan Data	25
Tabel 4.2 Distribusi Sentimen Hasil Pelabelan Data	28
Tabel 4.3 Sample Frekuensi Kata Dataset	31



DAFTAR RUMUS

Rumus 3.1 Akurasi	21
Rumus 3.2 Precision	21
Rumus 3.3 Recall.....	22
Rumus 3.4 F-1 Score	22
Rumus 4.1 Klasifikasi Naïve Bayes	32
Rumus 4.2 Klasifikasi SVM.....	34

