



**Evaluasi Performa Naive Bayes dan SVM dalam Analisis Sentimen
Kendaraan Listrik di Media Sosial Twitter**



GIGIH NUR HENDRAWAN 41518120005

**UNIVERSITAS
MERCU BUANA**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA 2023**



**Evaluasi Performa Naive Bayes dan SVM dalam Analisis Sentimen
Kendaraan Listrik di Media Sosial Twitter**

TUGAS AKHIR

GIGIH NUR HENDRAWAN 41518120005

UNIVERSITAS

MERCU BUANA
Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA 2023**

LEMBAR PERNYATAAN ORISINILITAS

Saya yang bertanda tangan di bawah ini:

Nama : GIGIH NUR HENDRAWAN
NIM : 41518120005
Program Studi : Teknik Informatika

Judul TUGAS AKHIR : Evaluasi Performa Naive Bayes dan SVM
dalam Analisis Sentimen Kendaraan Listrik di
Media Sosial Twitter

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan didalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

UNIVERSITAS
MERCU BUANA

Jakarta, 5 Desember 2023



GIGIH NUR HENDRAWAN

HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Gigih Nur Hendrawan
NIM : 41518120005
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Evaluasi Performa Naive Bayes dan SVM dalam Analisis Sentimen Kendaraan Listrik di Media Sosial Twitter

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana

Disahkan oleh:

Pembimbing : Harni Kusniyati, M.Kom
NIDN : 0324068101
Ketua Penguji : Dr. Ir. Eliyani
NIDN : 0321026901
Penguji 1 : Umniy Salamah, S.Kom., MMSI
NIDN : 0306098104
Penguji 2 : Dr. Harwikarya, MT
NIDN : 0014076808

UNIVERSITAS
MERCU BUANA

Jakarta, 29 Januari 2023

Mengetahui,

Dekan



Dr. Bambang Jokonowo, S.Si., M.T.I

Ketua Program Studi



Dr. Hadi Santoso, S.Kom., M.Kom

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan Tugas Akhir ini dengan judul “Evaluasi Performa Naive Bayes dan SVM dalam Analisis Sentimen Kendaraan Listrik di Media Sosial Twitter “.

Penulisan TUGAS AKHIR ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer pada Fakultas Ilmu Komputer Universitas Mercu Buana. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan ini, sangatlah sulit bagi saya untuk menyelesaikan Penelitian ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Prof. Dr. Ir. Andi Adriansyah, M.Eng selaku Rektor Universitas Mercu Buana
2. Ibu Harni Kusniyati, M.kom sebagai Pembimbing TA yang telah membimbing dalam penyusunan laporan ini.
3. Kedua Orang Tua atas doa serta dukungan moral yang tiada henti kepada penulis dalam menyelesaikan tugas akhir ini.
4. Ibu Dr. Ir. Eliyani, selaku dosen penguji atas koreksi dan arahan serta masukannya
5. Bapak Dr. Hadi Santoso, S.kom., M.Kom selaku kepala program studi Teknik informatika Universitas Mercubuana.
6. Semua dosen mata kuliah yang tidak dapat disebutkan satu persatu yang telah memberikan ilmu yang bermanfaat bagi penulis.
7. Teman-teman seperjuangan serta semua pihak yang tidak bisa disebutkan satu persatu, terima kasih atas dukungan serta kerjasama yang telah diberikan kepada penulis.

Akhir kata , Terima kasih atas dukungan dan doa dari semua pihak. Penulis menyadari masih banyak kekurangan dalam penyusunan tugas akhir ini. Saya berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Laporan Skripsi ini membawa manfaat bagi pengembangan ilmu sekaligus menambah pengetahuan bagi berbagai pihak.

Jakarta, 5 Desember 2023

Gigih Nur Hendrawan

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS
AKHIR UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Gigih Nur Hendrawan

NIM : 4158120005

Program Studi : Teknik Informatika

Judul Laporan Skripsi : Evaluasi Performa Naive Bayes dan SVM dalam Analisis Sentimen Kendaraan Listrik di Media Sosial Twitter

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

UNIVERSITAS
MERCU BUANA

Jakarta, 28 Januari 2024

enyatakan,



Gigih Nur Hendrawan

ABSTRAK

Nama : Gigih Nur Hendrawan
NIM : 41518120005
Program Studi : Teknik Informatika
Evaluasi Performa Naive Bayes dan SVM
Judul TUGAS AKHIR : dalam Analisis Sentimen Kendaraan Listrik di
Media Sosial Twitter
Pembimbing : Harni Kusniyati, M.Kom

Penelitian ini bertujuan untuk mengetahui hasil perbandingan performa tingkat akurasi, recall dan precision dari Algoritma *Naïve Bayes* dan *Support Vector Machine (SVM)* dalam mengklasifikasikan sentimen masyarakat pada media sosial *Twitter* terhadap kendaraan listrik. Jenis penelitian yang dipergunakan adalah penelitian kuantitatif yang mengacu pada pendekatan penelitian dengan cara mengumpulkan data yang dapat diukur secara numerik atau menggunakan metode statistik untuk menganalisis data tersebut. Metode yang digunakan yaitu pengumpulan data, labelling, preprocessing, pembobotan TF-IDF, Pemodelan dengan Algoritma *Naïve Bayes* dan *Support Vector Machine (SVM)* hingga proses visualisasi dan analisis data. Alat yang digunakan untuk pengumpulan data dari *Twitter* yaitu *Web-Crawling*. Hasil penelitian menunjukkan bahwa model SVM mengungguli *Naive Bayes* dengan akurasi keseluruhan yang signifikan yaitu 95.79%, berbanding dengan akurasi *Naive Bayes* yang tercatat sebesar 87.39%.

Kata Kunci : Analisis Sentimen, *Support vector Machine*, *Naïve Bayes*, kendaraan listrik

ABSTRACT

Nama : Gigih Nur Hendrawan
NIM : 41518120005
Program Studi : Teknik Informatika
Evaluasi Performa Naive Bayes dan SVM
Judul TUGAS AKHIR : dalam Analisis Sentimen Kendaraan Listrik di
Media Sosial Twitter
Pembimbing : Harni Kusniyati, M.Kom

The research aims to find out the comparison of the performance levels of accuracy, recall and precision of the Naïve Bayes algorithm and the Support Vector Machine (SVM) in classifying public sentiment on social media Twitter towards electric vehicles. The type of research used is quantitative research that refers to research approaches by collecting data that can be measured numerically or using statistical methods to analyze the data. Methods used include data collection, labelling, preprocessing, TF-IDF grinding, modeling with Naïve Bayes algorithm and Support Vector Machine (SVM) to the process of data visualization and analysis. The tool used to collect data from Twitter is Web Crawling. The research results show that the SVM model outperforms Naive Bayes with a significant overall accuracy of 95.79%, compared to Naive Bayes' recorded accuracy of 87.39%.

Keywords : Sentiment Analysis, Suppor vector Machine, Naïve Bayes, electric vehicle

DAFTAR ISI

LEMBAR PERNYATAAN ORISINILITAS	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS	
AKHIR UNTUK KEPENTINGAN AKADEMIS	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN.....	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian.....	3
1.4 Batasan Masalah.....	3
1.5 Sistematika Penulisan.....	3
BAB II LANDASAN TEORI.....	5
2.1 Teori Utama.....	5
2.1.1 Analisis Sentimen	5
2.1.2 <i>Text Mining</i>	5
2.1.3 Kendaraan Listrik.....	6
2.1.4 <i>Twitter</i>	6
2.2 Teori Pendukung	6
2.2.1 TF – IDF.....	6
2.2.2 <i>Naïve Bayes</i>	7
2.2.3 SVM.....	8
2.2.4 Python	9

2.3	<i>Literatur Review</i>	10
2.4	<i>Gap Penelitian</i>	17
BAB III METODOLOGI PENELITIAN		25
3.1	Pendekatan Penelitian	25
3.2	Desain Penelitian	25
3.2.1	Pengumpulan Data	25
3.2.2	Labelling	25
3.2.3	Preprocessing	26
1.	Penghapusan Data Duplikat	26
2.	Case Folding	26
3.	Penghapusan Tanda Baca	26
4.	Penghapusan Spasi Berlebih	26
5.	<i>Stopword Removal</i>	26
6.	Tokenisasi	27
7.	Stemming	27
8.	Pembobotan Kata dengan TF-IDF	27
9.	Oversampling dengan SMOTE	27
3.2.4.	Pembagian Data	28
1.	Data Latih	28
2.	Data Uji	28
3.2.5.	Pemodelan Dengan <i>Naïve Bayes</i> dan SVM	28
3.3.	Subjek Penelitian	29
3.4.	Instrumen Penelitian	29
3.5.	Teknik Pengumpulan Data	29
3.6.	Analisis Data	30
3.7.	Prosedur Penelitian	30
3.8.	Evaluasi	31
3.9.	Timeline Penelitian	32
BAB IV		33
HASIL DAN PEMBAHASAN		33
4.1.	Pengumpulan data	33
4.2.	Labelling	35

4.3.	<i>Preprocessing</i>	36
4.3.1	<i>Remove Duplicate Data</i>	36
4.3.2	<i>Casefolding</i>	37
4.3.3	<i>Remove Punctuation</i>	38
4.3.4	<i>Remove Double Whitespace</i>	38
4.3.5	<i>Tokenization</i>	39
4.3.6	<i>Stopword Removal</i>	40
4.3.7	<i>Stemming</i>	40
4.3.8	<i>TF iDF</i>	41
4.3.9	<i>Oversampling dengan SMOTE</i>	41
4.3.10	<i>Dataset Splitting</i>	42
4.4.	<i>Processing</i>	43
4.4.1	<i>Pemodelan dengan algoritma SVM</i>	43
4.4.2	<i>Pemodelan dengan Naïve Baiyes</i>	44
4.5	<i>Evaluasi Model SVM</i>	45
4.6	<i>Evaluasi Model Naïve Baiyes</i>	46
BAB V	47
KESIMPULAN DAN SARAN	47
5.1.	<i>Kesimpulan</i>	47
5.2.	<i>Saran</i>	47
DAFTAR PUSTAKA	49
LAMPIRAN	55

DAFTAR TABEL

Tabel 2.1 Literatur Riview	10
Tabel 2.2 Gap Penelitian	17
Tabel 3.1 Confusion Matrix	31
Tabel 3.2 Timeline Penelitian	32
Tabel 4. 1 Sampel Hasil Crawling	34
Tabel 4. 2 Sample Hasil Labelling	35



DAFTAR GAMBAR

Gambar 2. 1 Margin Hyperplane	9
Gambar 3. 1 Diagram Alir Penelitian	31
Gambar 4.1 Perbandingan Jumlah Data Sebelum dan Sesudah Penghapusan Duplikat Data	37
Gambar 4. 2 Hasil Casefolding	37
Gambar 4. 3 Hasil Remove Punctuation	38
Gambar 4. 4 Hasil Remove Double Whitespace.....	39
Gambar 4. 5 Hasil Tokenisasi	39
Gambar 4. 6 Hasil Stopword Removal	40
Gambar 4. 7 Hasil Stemming	40
Gambar 4. 8 Hasil TF iDF	41
Gambar 4. 9 Distribusi Kelas Sebelum Oversampling	41
Gambar 4. 10 Distribusi Kelas Setelah Oversampling	42
Gambar 4. 11 Skor Latih dan Skor Uji SVM.....	43
Gambar 4. 12 Parameter yang diuji.....	44
Gambar 4. 13 Parameter terbaik yang dihasilkan	44
Gambar 4. 14 Skor Uji dan Skor Latih Naïve Baiyes	44
Gambar 4. 15 Confusion matrix (SVM)	45
Gambar 4. 16 Confusion Matrix (Naïve Baiyes)	46



DAFTAR LAMPIRAN

LAMPIRAN BIMBINGAN.....	55
LAMPIRAN BUKTI SUBMIT JURNAL	57
CURICULUM VITAE.....	58
BUKTI UJI SERTIFIKASI BNSP	59
LOA JURNAL	60

