

**PENGAJIAN PERBANDINGAN PERFORMANSI ALGORITMA *NAÏVE*
BAYES DAN *SUPPORT VECTOR MACHINES* (SVM) PADA ANALISIS
SENTIMEN ULASAN APLIKASI GOBIS SUROBOYO BUS**

SKRIPSI



Oleh:

AVIA ARISTA

19081010077

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
2023**

LEMBAR PENGESAHAN SKRIPSI

**Judul : PENGKAJIAN PERBANDINGAN PERFORMANSI ALGORITMA
NAÏVE BAYES DAN SUPPORT VECTOR MACHINES (SVM)
PADA ANALISIS SENTIMEN ULASAN APLIKASI GOBIS
SUROBOYO BUS**

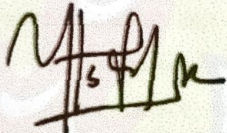
**Oleh : Avia Arista
NPM : 19081010077**

**Telah Diseminarkan Dalam Ujian Skripsi Pada :
Hari Senin, Tanggal 22 Mei 2023**

Mengetahui

Dosen Pembimbing

1.



**Yisti Vita Via, S.ST., M.Kom.
NIP. 19860425 2021212 001**

2.



**Made Hanindia Prami S., S.Kom, M.Cs
NIP: 19890205 2018032 001**

Dosen Penguji

1.



**Achmad Junaidi, S.Kom, M.Kom.
NPT. 3 7811 04 0199 1**

2.



**Fawwaz Ah Akbar, S.Kom, M.Kom
NIP: 19920317 2018031 002**

Menyetujui

Dekan

Fakultas Ilmu Komputer



**Dr. Ir. Nuvirina Hendrasari, MT.
NIP: 19681126 199403 2 001**

**Koordinator Program Studi
Informasi**



**Fetty Tri Anggraeny, S.Kom, M.Kom.
NIP: 19820211 2021212 005**



SURAT PERNYATAAN ANTI PLAGIAT

Saya mahasiswa program studi Informatika UPN "Veteran" Jawa Timur, yang bertandatangan dibawah ini:

Nama : Avia Arista

NPM : 19081010077

Menyatakan bahwa Judul Skripsi/Tugas Akhir yang Saya ajukan dan akan dikerjakan, yang berjudul:

"PENGKAJIAN PERBANDINGAN PERFORMANSI ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MACHINES (SVM) PADA ANALISIS SENTIMEN ULASAN APLIKASI GOBIS SUROBOYO BUS"

Bukan merupakan plagiat dari Skripsi/Tugas Akhir/Penelitian orang lain juga bukan merupakan produk dan atau software yang beli dari pihak lain. Saya juga menyatakan bahwa Skripsi/Tugas Akhir ini adalah pekerjaan Saya sendiri, kecuali yang dinyatakan dalam Daftar Pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di UPN "Veteran" Jawa Timur maupun di institusi Pendidikan lain.

Jika ternyata di kemudian hari pernyataan ini terbukti tidak benar, maka Saya siap menerima segala konsekuensinya.

Surabaya, 25 Mei 2023

Hormat Saya,



Avia Arista

NPM. 19081010077

PENGAJIAN PERBANDINGAN PERFORMANSI ALGORITMA *NAÏVE BAYES* DAN *SUPPORT VECTOR MACHINES* (SVM) PADA ANALISIS SENTIMEN ULASAN APLIKASI GOBIS SUROBOYO BUS

Nama : Avia Arista

NPM : 19081010077

Jurusan : Informatika

Dosen Pembimbing : Yisti Vita Via, S.ST., M.Kom.

Made Hanindia Prami Swari, S.Kom., M.Cs.

Abstrak

Berdasarkan Data INRIX salah satu perusahaan analisis data dan penyedia data lalu lintas, pada tahun 2021 mencatat kota Surabaya berada di peringkat 41 kota paling macet se asia tenggara. Atas permasalahan tersebut Dinas Perhubungan (Dishub) kota surabaya berinovivasi mengembangkan aplikasi GOBIS suroboyo Bus untuk menunjang penumpang dalam menggunakan Suroboyo Bus.

Pada penelitian ini akan dilakukan analisis sentimen terkait aplikasi GOBIS Suroboyo Bus dengan menggunakan algoritma *Naïve Bayes* dan *Support Vector Machines (SVM)* dan juga akan dilakukan analisis tingkat kepuasan pengguna. Data yang berhasil didapatkan adalah sebanyak 600 data. Dimana pada penelitian ini bertujuan untuk mengetahui kepuasan pengguna aplikasi terkait fitur dan juga tingkat perbandingan algoritma *Naïve Bayes* dan *Support Vector Machines (SVM)*. Setelah dilakukan uji coba, algoritma *Naives Bayes* mendapatkan nilai akurasi 81% dan algoritma *Support Vector Machines (SVM)* mendapatkan nilai *accuracy* 73%. Selain itu algoritma *Naïve Bayes* lebih akurat dalam mengenali data teks dibandingkan SVM yang menggunakan tipe data *float*. Sehingga algoritma *Naïve Bayes* akan digunakan sebagai algoritma utama dalam proses analisis sentimen dari tahun 2018 hingga April 2023, dan hasilnya fitur yang ada di aplikasi GOBIS Suroboyo Bus masih kurang baik dalam menunjang penggunaan Suroboyo Bus dimana sesuai dengan rating aplikasi yang hanya 3.9 dari 5.

Kata Kunci: Analisis Sentimen, Kemacetan, Suroboyos Bus, *Naive Bayes*, *Support Vector Machines*.

KATA PENGANTAR

Alhamdulillahirobbil'alamin, segala puji bagi Allah SWT atas segala limpahan rahmat, taufik, hidayah serta inayah-NYA. Sholawat serta salam semoga selalu tercurahkan kepada Nabi Muhammad SAW, beserta keluarga, para sahabat dan kaum muslimin yang senantiasa istiqomah berjuang menegakkan Agama Allah dimuka bumi ini. Sehingga penulis dapat menyelesaikan penelitian yang berjudul "PENGKAJIAN PERBANDINGAN PERFORMANSI ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MACHINES (SVM) PADA ANALISIS SENTIMEN ULASAN APLIKASI GOBIS SUROBOYO BUS:

Laporan Skripsi ini ditulis dan disusun sebagai syarat kelulusan Program Sarjana (S1) Jurusan Informatika Fakultas Ilmu Komputer UPN "Veteran" Jawa Timur. Penelitian ini menjelaskan tentang analisis sentimen masyarakat terhadap aplikasi GOBIS Suroboyo Bus baik sentimen positif dan negatif serta membandingkan tingkat akurasi algoritma *Naive Bayes* dan *Support Vector Machines* (SVM).

Walaupun demikian, laporan ini tentu masih terdapat beberapa kekurangan di dalamnya dan akan sangat berarti jika pembaca dapat memberikan evaluasinya agar menjadi bahan perbaikan bagi penulis.

Surabaya, 11 Mei 2023

Penulis

UCAPAN TERIMAKASIH

Puji syukur kami panjatkan atas kehadiran Allah SWT atas segala limpahan berkat, rahmat, dan inayah-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Penulis menyadari tidak akan menyelesaikan skripsi ini dengan baik tanpa bimbingan, saran, motivasi, dan bantuan dari berbagai pihak. Pada kesempatan ini, penulis mengucapkan terimakasih yang sebesar – besarnya kepada:

1. Keluarga yang senantiasa mendukung dan mendoakan saya baik secara finansial maupun spiritual.
2. Prof. Dr. Ir. Akhmad Fauzi, M.MT. selaku Rektor Universitas Pembangunan Nasional Veteran Jawa Timur.
3. Dr. Ir. Novirina Hendrasarie, MT. selaku Dekan Fakultas Ilmu Komputer
4. Fetty Try Anggraeny, S.Kom, M.Kom Selaku Koordinator Program Studi Informatika
5. Yisti Vita Via, S.ST., M.Kom. selaku Dosen Pembimbing I
6. Made Hanindia Prami Swari, S.Kom., M.Cs. selaku Dosen Pembimbing II
7. Seluruh teman – teman yang saling membantu dan memberikan dukungan satu sama lain.

DAFTAR ISI

Abstrak	ii
KATA PENGANTAR	iii
UCAPAN TERIMAKASIH.....	iv
DAFTAR ISI.....	v
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xii
DAFTAR KODE.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Penelitian Terdahulu	6
2.2 GOBIS Suroboyo Bus.....	9
2.3 Analisis Sentimen	9
2.4 Analisis Sentimen Text Mining	10
2.5 Text Pre-processing	11

2.5.1	<i>Case Folding</i>	11
2.5.2	<i>Cleaning</i>	11
2.5.3	<i>Tokenization</i>	12
2.5.4	<i>Normalization</i>	12
2.5.5	<i>Stopword</i>	12
2.5.6	<i>Stemming</i>	13
2.6	Ekstraksi Fitur.....	13
2.7	<i>Naïve Bayes</i>	14
2.8	<i>Support Vector Machine (SVM)</i>	16
2.9	<i>Confusion Matrix</i>	18
2.10	<i>Google Play Store</i>	20
2.11	<i>Google Play Scrapper</i>	21
BAB III METODOLOGI PENELITIAN.....		22
3.1	Diagram Alur	22
3.2	Pengumpulan Data	23
3.3	<i>Preprocessing Data</i>	24
3.3.1	<i>Case Folding</i>	24
3.3.2	<i>Cleaning Data</i>	25
3.3.3	<i>Tokenization</i>	25
3.3.4	<i>Normalization</i>	26
3.3.5	<i>Stopword</i>	27

3.3.6	<i>Stemming</i>	27
3.4	<i>Labelling Data</i>	28
3.5	Ekstraksi Fitur.....	28
3.6	Klasifikasi Sentimen <i>Naïve Bayes</i>	31
3.7	Klasifikasi Sentimen <i>Support Vector Machines (SVM)</i>	35
3.8	Analisis Hasil Dan Uji Coba Pengujian Algoritma	44
BAB IV HASIL DAN PEMBAHASAN		46
4.1	Hasil Pengumpulan Data.....	46
4.2	Hasil <i>Pre Processing Data</i>	49
4.2.1	Hasil Implementasi <i>Case Folding</i>	49
4.2.2	Hasil Implementasi <i>Cleaning Data</i>	52
4.2.3	Hasil Implementasi <i>Tokenization Data</i>	56
4.2.4	Hasil Implementasi <i>Normalization Data</i>	58
4.2.5	Hasil Implementasi <i>Stopword</i>	63
4.2.6	Hasil Implementasi <i>Stemming</i>	66
4.3	Hasil Pelabelan Data	69
4.4	Hasil Implementasi Ekstraksi Fitur.....	72
4.5	Hasil Implementasi Algoritma Klasifikasi <i>Naïve Bayes</i>	73
4.6	Analisis Hasil Klasifikasi <i>Naïve Bayes</i>	79
4.7	Hasil Implementasi Algoritma Klasifikasi <i>Support Vector Machines</i> (<i>SVM</i>)	82

4.8	Analisis Hasil Klasifikasi <i>Support Vector Machines</i>	82
4.9	Hasil Perbandingan Uji Coba Klasifikasi menggunakan <i>Naïve Bayes</i> dan <i>Support Vector Machines (SVM)</i>	87
4.9.1.	Hasil Perbandingan Uji Coba Menggunakan 80% Data Latih	87
4.9.2.	Hasil Perbandingan Uji Coba Menggunakan 80% Data Latih dan 20% Data Uji	89
4.9.3.	Hasil Perbandingan Uji Coba Menggunakan 75% Data Latih dan 25% Data Uji	90
4.9.4.	Hasil Perbandingan Uji Coba Menggunakan 70% Data Latih dan 30% Data Uji	92
4.10	Analisis hasil Klasifikasi Per Tahun	97
4.10.1	Analisis Hasil Klasifikasi Tahun 2018	98
4.10.2	Analisis Hasil Klasifikasi Tahun 2019	101
4.10.3	Analisis Hasil Klasifikasi Tahun 2020	104
4.10.4	Analisis Hasil Klasifikasi Tahun 2021	107
4.10.5	Analisis Hasil Klasifikasi Tahun 2022	110
4.10.6	Analisis Hasil Klasifikasi April Tahun 2023	113
4.11	Analisis Perbandingan Klasifikasi Per Tahun.....	116
BAB V KESIMPULAN DAN SARAN.....		120
5.1	Kesimpulan	120
5.2	Saran	121

DAFTAR PUSTAKA 122

DAFTAR TABEL

Tabel 2. 1 Tabel Confusion matrix	18
Tabel 3. 1 Contoh Labelling Data	28
Tabel 3. 2 Contoh Frekuensi kata dalam kalimat.....	29
Tabel 3. 3 Hasil perhitungan DF	30
Tabel 3. 4 Contoh Ekstraksi Fitur	30
Tabel 3. 5 Contoh nilai TF IDF.....	31
Tabel 3. 6 Contoh frekuensi kemunculan kata.....	33
Tabel 3. 7 Contoh ekstraksi fitur pada dataset.....	38
Tabel 3. 8 Contoh pasangan kernel	39
Tabel 3. 9 Contoh hasil perhitungan kernelisasi	40
Tabel 3. 10 Contoh Perhitungan Matriks Hessian	40
Tabel 3. 11 Contoh hasil perhitungan nilai E_i pada iterasi ke – 1	41
Tabel 3. 12 Contoh hasil perhitungan nilai $\delta\alpha_i$ pada iterasi ke – 1	42
Tabel 3. 13 Contoh hasil perhitungan iterasi nilai E_i	42
Tabel 3. 14 Contoh hasil perhitungan iterasi nilai $\delta\alpha_i$	42
Tabel 3. 15 Contoh hasil perhitungan iterasi nilai α_i	43
Tabel 3. 16 Contoh hasil perhitungan nilai bias.....	43
Tabel 3. 17 Contoh hasil perhitungan kernel data uji dan data latih.....	44
Tabel 4. 1 Hasil crawling data aplikasi GOBIS	47
Tabel 4. 2 Hasil case folding data aplikasi GOBIS.....	50
Tabel 4. 3 Hasil cleaning data aplikasi GOBIS	54
Tabel 4. 4 Hasil tokezation data aplikasi GOBIS	57

Tabel 4. 5 Hasil normalization data aplikasi GOBIS.....	61
Tabel 4. 6 Hasil stopword data aplikasi GOBIS	64
Tabel 4. 7 Contoh hasil stemming data aplikasi GOBIS	68
Tabel 4. 8 Hasil pelabelan data aplikasi GOBIS.....	70
Tabel 4. 9 Hasil Confusion Matrix Naïve Bayes	79
Tabel 4. 10 Hasil Confusion Matrix Support Vector Machines (SVM)	84
Tabel 4. 11 Perbandingan Tingkat Akurasi Naive Bayes dan SVM Menggunakan Data Latih.....	89
Tabel 4. 12 Perbandingan Tingkat Akurasi Naive Bayes dan SVM.....	90
Tabel 4. 13 Perbandingan Tingkat Akurasi Naive Bayes dan SVM 75 : 25.....	92
Tabel 4. 14 Perbandingan Tingkat Akurasi Naive Bayes dan SVM 70 : 30.....	94
Tabel 4. 15 Hasil Confusion Matrix Data Tahun 2018.....	98
Tabel 4. 16 Hasil Confusion Matrix Data Tahun 2019.....	101
Tabel 4. 17 Hasil Confusion Matrix Data Tahun 2020.....	104
Tabel 4. 18 Hasil Confusion Matrix Data Tahun 2021	107
Tabel 4. 19 Hasil Confusion Matrix Data Tahun 2022	110
Tabel 4. 20 Hasil Confusion Matrix Data April Tahun 2023	113

DAFTAR GAMBAR

Gambar 2. 1 Proses Analisis Sentimen	10
Gambar 2. 2 Hyperplane pada SVM.....	16
Gambar 2. 3 Hasil klasifikasi menggunakan SVM.....	16
Gambar 3. 1 Diagram Alur.....	22
Gambar 3. 2 Tahap Pre-Processing Data	24
Gambar 3. 3 Hasil Case Folding	24
Gambar 3. 4 Hasil Cleaning Data	25
Gambar 3. 5 Hasil Tokenization	26
Gambar 3. 6 Hasil Normalization	26
Gambar 3. 7 Hasil Stopword.....	27
Gambar 3. 8 Hasil Stemming.....	27
Gambar 3. 9 Flowchart algoritma Naïve Bayes.....	32
Gambar 3. 10 Flowchart SVM.....	36
Gambar 4. 1 Output Implementasi Crawling Data.....	47
Gambar 4. 2 Output Implementasi Case Folding.....	49
Gambar 4. 3 Output Implementasi Cleaning Data	54
Gambar 4. 4 Output Implementasi Tokenization.....	56
Gambar 4. 5 Output Implementasi Normalization.....	60
Gambar 4. 6 Output Implementasi Stopword	63
Gambar 4. 7 Output Implementasi Stemming	67
Gambar 4. 8 Hasil Pelabelan Data	71
Gambar 4. 9 Output Implementasi Ekstraksi Fitur	73

Gambar 4. 10 Output Confusion Matrix Naive Bayes.....	77
Gambar 4. 11 Nilai Aktual Data Uji Naive Bayes	78
Gambar 4. 12 Nilai Prediksi Data Uji Naive Bayes.....	78
Gambar 4. 13 Clasification Report Naive Bayes	81
Gambar 4. 14 Output Confusion Matrix SVM	83
Gambar 4. 15 Nilai AKtual Data Uji Support Vector Machines	83
Gambar 4. 16 Nilai Prediksi Data Uji Support Vector Machines	84
Gambar 4. 17 Clasification Report SVM.....	86
Gambar 4. 18 Clasification Report Uji Coba Menggunakan Data Latih Naive Bayes	88
Gambar 4. 19 Clasification Report Uji Coba menggunakan data latih SVM.....	88
Gambar 4. 20 Clasification Report Uji Coba 75:25 Naive Bayes	90
Gambar 4. 21 Clasification Report Uji Coba 75:25 SVM.....	91
Gambar 4. 22 Clasification Report Uji Coba 70:30 Naive Bayes	93
Gambar 4. 23 Clasification Report Uji Coba 70:30 SVM.....	93
Gambar 4. 24 Diagram Perbandingan Uji Coba.....	95
Gambar 4. 25 Data Uji <i>Naive Bayes</i>	96
Gambar 4. 26 Data <i>SVM</i>	96
Gambar 4. 27 Clasification Report Tahun 2018	100
Gambar 4. 28 Clasification Report Tahun 2019	103
Gambar 4. 29 Clasification Report Tahun 2020	106
Gambar 4. 30 Clasification Report Tahun 2021	109
Gambar 4. 31 Clasification Report Tahun 2022	112
Gambar 4.32 Clasification Report Tahun 2023 Hingga Bulan April.....	115

Gambar 4. 33 Confusion Matrix Per Tahun.....	116
Gambar 4. 34 Perbandingan Sentimen Positif dan Negatif Per Tahun	117
Gambar 4. 35 Wordcloud Sentimen Positif	118
Gambar 4. 36 Wordcloud Sentimen Negatif.....	119

DAFTAR KODE

Kode 4. 1 Kode Program Crawling Data	46
Kode 4. 2 Kode Program Case Folding	49
Kode 4. 3 Kode Program Cleaning Data.....	53
Kode 4. 4 Kode Program Tokenization Data	56
Kode 4. 5 Kode Program Normalization Data	60
Kode 4. 6 Kode Program Stopword	63
Kode 4. 7 Kode Program Stemming	67
Kode 4. 8 Kode Program ekstraksi Fitur.....	72
Kode 4. 9 Kode Program Splitting Data	73
Kode 4. 10 Kode Program Klasifikasi Menggunakan Algoritma Naive Bayes....	77
Kode 4. 11 Kode Program SVM.....	82