

**ANALISIS SENTIMEN OPINI PENGGUNA MEDIA
SOSIAL TWITTER TERHADAP TRANSJAKARTA
MENGUNAKAN METODE *NAÏVE BAYES*
DAN *SUPPORT VECTOR MACHINE***

SKRIPSI

Oleh:
Dian Fitrianto Hidayat
201910225350



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2023**

LEMBAR PERSETUJUAN PEMBIMBING

Judul Proposal Tugas Akhir : Analisis Sentimen Opini Pengguna Media Sosial Twitter Terhadap Transjakarta Menggunakan Metode *Naïve Bayes* dan *Support Vector Machine*

Nama Mahasiswa : Dian Fitrianto Hidayat

Nomor Pokok Mahasiswa : 201910225350

Program Studi/Fakultas : Informatika / Ilmu Komputer

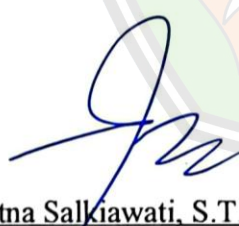
Tanggal Lulus Ujian Skripsi : 20 Juli 2023

Bekasi, 1 Agustus 2023

MENYETUJUI,

Pembimbing I

Pembimbing II


(Ratna Salkiawati, S.T., M.Kom)

NIDN. 0310038006


(Khairunnisa Fadhillah Ramdhania, M.Si)

NIDN. 0328039201

LEMBAR PENGESAHAN

Judul Tugas Akhir : Analisis Sentimen Opini Pengguna
Media Sosial Twitter Terhadap
Transjakarta Menggunakan Metode *Naïve*
Bayes dan *Support Vector Machine*

Nama Mahasiswa : Dian Fitrianto Hidayat

Nomor Pokok Mahasiswa : 201910225350

Program Studi/ Fakultas : Informatika/ Ilmu Komputer

Tanggal Lulus Ujian Tugas Akhir : 20/07/2023

Bekasi, 2 Agustus 2023

MENGESAHKAN,

Ketua Tim Penguji : Herlawati, S.Si., M.M., M.Kom.

NIDN : 0311097302

Penguji I : Andy Achmad Hendharsetiawan, S.T., M.T.I.

NIDN : 0317057204

Penguji II : Ratna Salkiawati, S.T., M.Kom.

NIDN : 0310038006

MENGETAHUI,

Ketua Program Studi Informatika



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

Dekan Fakultas Ilmu Komputer



Dr. Dra. Tyastuti Sri Lestari, MM.
NIP. 1408206



UNIVERSITAS BHAYANGKARA JAKARTA
RAYAFAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA

LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan di bawah ini :

Nama : Dian Fitrianto Hidayat
NPM : 201910225350
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Analisis Sentimen Opini Pengguna Media
Sosial Twitter Terhadap Transjakarta
Menggunakan Metode *Naïve Bayes* dan
Support Vector Machine

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudianhari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggung jawabkan 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, 04 Agustus 2023
Peneliti



Dian Fitrianto Hidayat

ABSTRAK

Dian Fitrianto Hidayat. 201910225350. Analisis Sentimen Opini Pengguna Media Sosial Twitter Terhadap Transjakarta Menggunakan Metode *Naïve Bayes* Dan *Support Vector Machine*. Bekasi: Fakultas Ilmu Komputer. Universitas Bhayangkara Jakarta Raya. 2023.

Twitter merupakan dari salah satu platform media sosial yang populer di Indonesia. Twitter dapat berguna untuk mengungkapkan perasaan bagi para penggunanya dan memberikan kritik atau saran terhadap suatu objek. Transjakarta merupakan salah satu transportasi umum yang memiliki jumlah penumpang harian terbanyak dibandingkan dengan transportasi umum lainnya. Hal tersebut membuat para pengguna Transjakarta dapat memberikan pendapat, kritik maupun saran dari pengalaman yang dialami. Salah satu permasalahan yang terjadi pada Transjakarta yaitu beberapa halte mengalami penutupan sementara waktu, selain itu marak terjadinya kasus asusila yang dialami oleh beberapa pengguna Transjakarta. Tujuan dari penelitian ini yaitu untuk menganalisis opini pengguna media sosial Twitter terhadap Transjakarta dengan menggunakan metode *Naïve Bayes* dan *Support Vector Machine*. Pada penelitian ini menggunakan *Lexicon Based* untuk pemberian label dari masing-masing kelas dan menggunakan dataset sebanyak 6736 *tweet*. Hasil yang didapatkan dari kelas positif berjumlah 2228 dan kelas negatif 2821. Berdasarkan metode *Naïve Bayes* didapatkan akurasi 76.43% dengan presisi 76.66%, *recall* 75.85% dan *f1-score* 76%. Sedangkan pada metode *Support Vector Machine* didapatkan akurasi sebesar 84.95% dengan presisi 83%, *recall* 83% dan *f1-score* 83%.

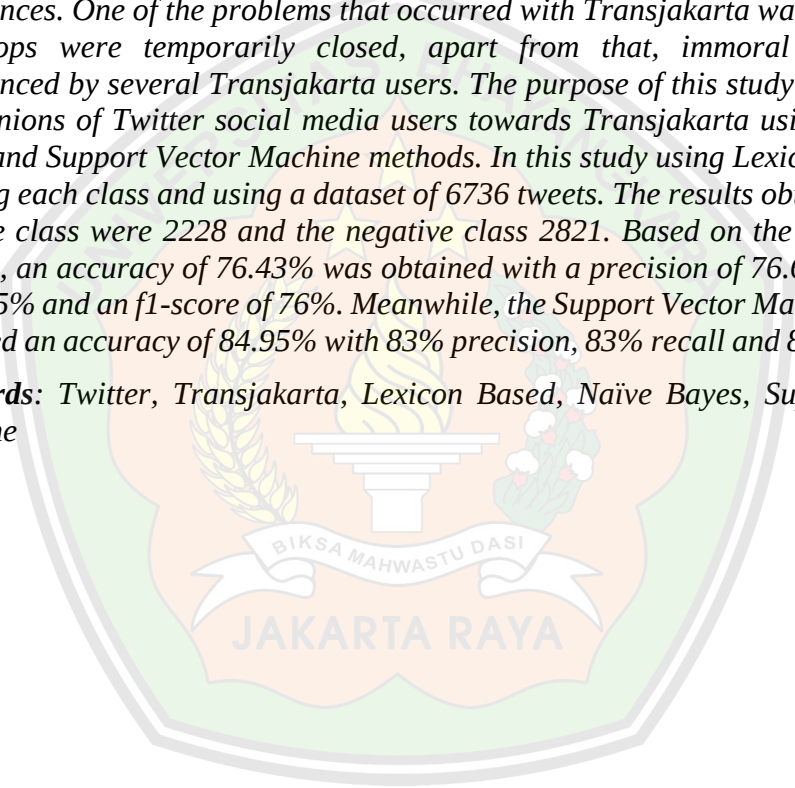
Kata Kunci: Twitter, Transjakarta, *Lexicon Based*, *Naïve Bayes*, *Support Vector Machine*

ABSTRACT

Dian Fitrianto Hidayat. 201910225350. *Sentiment Analysis of Twitter Social Media User Opinions on Transjakarta Using Naïve Bayes Method and Support Vector Machine.* Bekasi: Faculty of Computer Science. Jakarta Bhayangkara University. 2023.

Twitter is one of the most popular social media platforms in Indonesia. Twitter can be useful for expressing feelings for its users and providing criticism or suggestions for an object. Transjakarta is one of the public transportation that has the highest number of daily passengers compared to other public transportation. This allows Transjakarta users to provide opinions, criticisms and suggestions based on their experiences. One of the problems that occurred with Transjakarta was that several bus stops were temporarily closed, apart from that, immoral cases were experienced by several Transjakarta users. The purpose of this study is to analyze the opinions of Twitter social media users towards Transjakarta using the Naïve Bayes and Support Vector Machine methods. In this study using Lexicon Based for labeling each class and using a dataset of 6736 tweets. The results obtained for the positive class were 2228 and the negative class 2821. Based on the Naïve Bayes method, an accuracy of 76.43% was obtained with a precision of 76.66%, a recall of 75.85% and an f1-score of 76%. Meanwhile, the Support Vector Machine method obtained an accuracy of 84.95% with 83% precision, 83% recall and 83% f1-score.

Keywords: Twitter, Transjakarta, Lexicon Based, Naïve Bayes, Support Vector Machine



LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI

KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

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

Nama : Dian Fitrianto Hidayat

NPM : 201910225350

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:

ANALISIS SENTIMEN OPINI PENGGUNA MEDIA SOSIAL
TWITTER TERHADAP TRANSJAKARTA MENGGUNAKAN
METODE NAÏVE BAYES DAN SUPPORT VECTOR MACHINE

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalih medikan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin 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 : 04 Agustus 2023

Yang Menyatakan

Dian Fitrianto Hidayat

KATA PENGANTAR

Puji syukur kepada Allah SWT berkat Rahmat, Hidayah, dan Karunia-Nya kepada kita semua sehingga peneliti dapat menyelesaikan skripsi yang disusun sebagai salah satu syarat untuk mencapai gelar sarjana Strata Satu (S1) yang berjudul “Analisis Sentimen Opini Pengguna Media Sosial Twitter Terhadap Transportasi Umum Transjakarta Menggunakan Metode *Naïve Bayes* dan *Support Vector Machine*”. Dalam penyusunan skripsi ini peneliti mendapat banyak sekali bantuan, dukungan, dan bimbingan. Oleh karena itu, dalam kesempatan ini ingin menyampaikan terima kasih kepada:

1. Bapak Irjen Pol. (Purn). Prof. Dr. Drs. Bambang Karsono, S.H., M.M. Selaku Rektor Universitas Bhayangkara Jakarta Raya.
2. Ibu Dr. Dra. Tyastuti Sri Lestari, M.M. Selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
3. Bapak Ahmad Fathurrozi, S.E., M.M.S.I. Selaku kepala Program Studi Informatika Universitas Bhayangkara Jakarta Raya.
4. Ibu Ratna Salkiawati, S.T., M.Kom. Selaku pembimbing I yang telah memberikan arahan serta motivasi yang mudah dipahami saat sesi bimbingan.
5. Mrs Khairunnisa Fadhillah Ramdhanisa, S.Si., M.Si. Selaku pembimbing II yang telah pengertian terhadap waktu bimbingan.
6. Kedua orang tua peneliti. Yang sudah memberikan dukungan dari semua hal dan semangat untuk tidak mudah menyerah selama proses yang dilakukan oleh peneliti.

7. Pemilik NIM 11190930000104 dan seluruh teman seperjuangan atas bantuan, selalu memberi semangat, dukungan, serta kerja samanya dalam perkuliahan.

Peneliti menyadari bahwa dalam penyusunan skripsi ini masih jauh dari kata sempurna, serta masih banyak kekurangan baik dalam penelitian materi maupun dalam susunan bahasanya. Akhir kata, peneliti berharap agar skripsi ini dapat bermanfaat, bagi peneliti maupun orang lain yang membacanya.

Bekasi, 27 Juni 2023

Hormat saya,



Dian Fitrianto Hidayat

201910225350



DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN BUKAN PLAGIASI.....	iii
ABSTRAK	iv
ABSTRACT	v
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI	vi
KATA PENGANTAR	vi
DAFTAR ISI.....	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	5
1.3 Rumusan Masalah	5
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	6
1.6 Batasan Masalah.....	6
1.7 Sistematika Tugas Akhir	7
BAB II TINJAUAN PUSTAKA.....	9
2.1 Kajian Literatur	9
2.2 Analisis Sentimen	10
2.3 Twitter	11
2.4 Transjakarta.....	13
2.5 <i>Text Mining</i>	14
2.6 SEMMA	15
2.7 <i>Lexicon Based</i>	18

2.8	TF-IDF	18
2.9	<i>Naïve Bayes</i>	19
2.10	<i>Support Vector Machine</i>	21
2.11	<i>Grid Search Cross Validation</i>	23
2.12	<i>Confusion Matrix</i>	23
2.13	<i>Flowchart</i>	25
2.14	<i>Wordcloud</i>	25
2.15	<i>Python</i>	26
BAB III METODOLOGI PENELITIAN		27
3.1	Objek Penelitian	27
3.2	Kerangka Penelitian	29
3.3	Metode Pengolahan Data	31
3.4	Perangkat Penelitian	38
BAB IV HASIL DAN PEMBAHASAN		39
4.1	<i>Sample</i>	39
4.2	<i>Explore</i>	41
4.3	<i>Modify</i>	41
4.4	Model	48
4.4.1	<i>Lexicon Based</i>	48
4.4.2	Pembobotan Kata TF-IDF	51
4.4.3	<i>Naïve Bayes</i>	53
4.4.4	<i>Support Vector Machine</i>	56
4.5	Assess	62
4.5.1	<i>Confusion Matrix Naïve Bayes</i>	63
4.5.2	<i>Confusion Matrix Support Vector Machine</i>	64
BAB V PENUTUP		67
5.1	Kesimpulan	67
5.2	Saran	68
DAFTAR PUSTAKA		69
LAMPIRAN		73

DAFTAR TABEL

Tabel 2. 1 Penelitian Terkait	9
Tabel 2. 2 <i>Confusion Matrix</i>	24
Tabel 2. 3 Simbol-Simbol <i>Flowchart</i>	25
Tabel 3. 1 Spesifikasi <i>Hardware</i> dan <i>Software</i>	38
Tabel 4. 1 Hasil Proses <i>Case Folding</i>	42
Tabel 4. 2 Hasil Proses <i>Cleaning</i>	43
Tabel 4. 3 Hasil Proses <i>Tokenize</i>	44
Tabel 4. 4 Hasil <i>Normalize</i>	45
Tabel 4. 5 Hasil <i>Stopword Removal</i>	46
Tabel 4. 6 Hasil <i>Stemming</i>	48
Tabel 4. 7 Jumlah Aktual Data Set.....	49
Tabel 4. 8 Hasil Pelabelan Sentimen dengan Kamus <i>Lexicon Inset</i>	49
Tabel 4. 9 Total Jumlah Pada <i>Wordcloud</i> Positif	50
Tabel 4. 10 Total Jumlah pada <i>Wordcloud</i> Negatif.....	51
Tabel 4. 11 Pembobotan Kata TF-IDF	52
Tabel 4. 12 Pembobotan Manual TF-IDF	52
Tabel 4. 13 Perbandingan Rasio Data Latih dan Data Uji <i>Naïve Bayes</i>	53
Tabel 4. 14 Sampel Data Manual <i>Naïve Bayes</i>	54
Tabel 4. 15 Perbandingan Rasio Data Latih dan Data Uji <i>Support Vector Machine</i>	57
Tabel 4. 16 Kernel Dan Parameter Untuk Diuji	57
Tabel 4. 17 Pembobotan Kata TF-IDF Perhitungan Manual SVM.....	58
Tabel 4. 18 Fungsi Kernel	58
Tabel 4. 19 Hasil Perhitungan Fungsi Kernel	59
Tabel 4. 20 Hasil Perhitungan Matriks.....	59
Tabel 4. 21 Hasil Perhitungan <i>error</i>	60
Tabel 4. 22 Hasil Perhitungan <i>Delta Alpha</i>	60
Tabel 4. 23 Hasil Perhitungan <i>Alpha</i> Baru.....	61
Tabel 4. 24 Sampel Data Uji	61

Tabel 4. 25 Hasil Perhitungan <i>Dot Product</i> Data Uji Dengan Data Latih	62
Tabel 4. 26 Hasil Klasifikasi <i>Naïve Bayes</i>	63
Tabel 4. 27 <i>Confusion Matrix Naive Bayes</i>	63
Tabel 4. 28 Hasil <i>Grid Search Cross Validation</i>	65
Tabel 4. 29 Hasil Klasifikasi <i>Support Vector Machine</i>	65
Tabel 4. 30 Hasil <i>Confusion Matrix Support Vector Machine</i>	66



DAFTAR GAMBAR

Gambar 1. 1 Jumlah Penumpang Harian Transportasi Umum di Jakarta	2
Gambar 1. 2 <i>Tweet</i> Pengguna Transjakarta	3
Gambar 2. 1 <i>Text Analytics, Related Application Areas, and Enabling Disciplines</i>	15
Gambar 2. 2 Proses SEMMA <i>Data Mining</i>	16
Gambar 2. 3 Visualisasi <i>Wordcloud</i>	26
Gambar 3. 1 Struktur Organisasi Transjakarta	27
Gambar 3. 2 Kerangka Penelitian.....	30
Gambar 3. 3 Alur <i>Crawling</i> Dataset.....	32
Gambar 3. 4 Alur Tahapan <i>Modify</i>	34
Gambar 3. 5 Alur Pelabelan Data.....	35
Gambar 3. 6 Alur Klasifikasi <i>Naive Bayes</i>	36
Gambar 3. 7 Alur Klasifikasi <i>Support Vector Machine</i>	37
Gambar 4. 1 Instalasi <i>Library Snsrape</i> Untuk <i>Crawling</i> Data	40
Gambar 4. 2 Hasil <i>Crawling</i> Data Twitter	40
Gambar 4. 3 Hasil Pengambilan Dua Kolom Dataset	41
Gambar 4. 4 Proses <i>Case Folding</i>	42
Gambar 4. 5 Proses <i>Cleaning</i>	43
Gambar 4. 6 Proses <i>Tokenize</i>	44
Gambar 4. 7 Proses <i>Normalize</i>	45
Gambar 4. 8 Proses <i>Stopword Removal</i>	46
Gambar 4. 9 Proses <i>Stemming</i>	47
Gambar 4. 10 <i>Wordcloud</i> Kelas Positif	50
Gambar 4. 11 <i>Wordcloud</i> Kelas Negatif	51

DAFTAR LAMPIRAN

Lampiran 1: Hasil Plagiarism Checker Originality Report.....	74
Lampiran 2: Biodata Mahasiswa.....	75
Lampiran 3: Kartu Bimbingan Skripsi.....	76
Lampiran 4: Surat Rekomendasi Dari Pembimbing	79
Lampiran 5: Surat Izin Penelitian	80
Lampiran 6: Surat Pemberian Izin Penelitian	81
Lampiran 7: DataSet	82

