

**ANALISIS SENTIMEN SISTEM E-TILANG PADA TWITTER
MENGGUNAKAN METODE *MULTINOMIAL NAÏVE BAYES***



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER
2024**

TUGAS AKHIR

**ANALISIS SENTIMEN SISTEM E-TILANG PADA TWITTER
MENGGUNAKAN METODE MULTINOMIAL NAÏVE BAYES**

Diajukan untuk Memenuhi Persyaratan Guna Meraih Gelar Sarjana Komputer

Teknik Informatika Universitas Muhammadiyah Jember



PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MUHAMMADIYAH JEMBER

2024

HALAMAN PERSETUJUAN SIDANG TUGAS AKHIR

ANALISIS SENTIMEN SISTEM E-TILANG PADA TWITTER
MENGGUNAKAN METODE MULTINOMIAL NAÏVE BAYES

Oleh

Widianto Dwi Pangestu

1810651041

Telah disetujui bahwa Laporan Tugas Akhir ini untuk diajukan pada sidang
Tugas Akhir sebagai salah satu syarat kelulusan dan mendapatkan gelar
Sarjana Komputer (S.Kom)

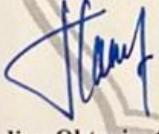
di

Universitas Muhammadiyah Jember

Menyetujui,

Pembimbing I

Pembimbing II

 Hardian Oktavianto, S.Si., M.Kom

NIDN. 0722108105

 Daryanto, S.Kom, M.Kom

NIDN. 0707077203

HALAMAN PENGESAHAN

ANALISIS SENTIMEN SISTEM E-TILANG PADA TWITTER
MENGGUNAKAN METODE MULTINOMIAL NAÏVE BAYES

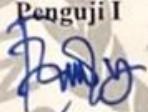
Oleh
Widianto Dwi Pangestu
18106651041

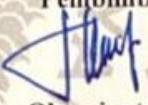
Telah mempertanggung jawabkan Laporan Tugas Akhirnya pada sidang Tugas Akhir Tanggal 19 Bulan Juli Tahun 2024 sebagai salah satu syarat kelulusan dan mendapatkan gelar Sarjana Komputer (S.Kom)

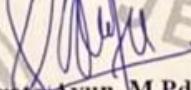
di

Universitas Muhammadiyah Jember

Menyetujui,

Pengaji I

Rosita Yanuarti, S.Kom., M.Cs
NIDN. 0629018601

Pembimbing I

Hardian Oktavianto, S.Si., M.Kom
NIDN. 0722108105

Pengaji II

Qurrota Ayun, M.Pd
NIDN. 0703069002

Pembimbing II

Darvanto, S.Kom, M.Kom
NIDN. 0707077203

Mengesahkan,
Dekan Fakultas Teknik

Dr. Ir. Muhtar, ST., MT., IPM
NIDN. 0010067301

Mengetahui,
Ketua Program Studi Teknik Informatika

Rosita Yanuarti, S.Kom., M.Cs
NIDN. 0629018601

HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertanda tangan dibawah ini :

Nama : Widianto Dwi Pangestu

Nim : 1810651041

Program Studi : Teknik Informatika

Perguruan Tinggi : Universitas Muhammadiyah Jember

Menyatakan bahwa Tugas Akhir yang berjudul "**Analisis Sentimen Sistem E-Tilang Pada Twitter Menggunakan Metode Multinomial Naïve Bayes**" bukan merupakan Tugas Akhir orang lain baik sebagian maupun keseluruhan kecuali dalam bentuk kutipan yang telah disebutkan sumbernya.

Demikian surat pernyataan ini dibuat dengan sebenar – benarnya dan apabila pernyataan ini tidak benar penulis bersedia mendapatkan sanksi akademik.

Jember, 19 Juli 2024

Yang Menyatakan,



Widianto Dwi Pangestu

NIM. 1810651041

ABSTRAK

E-tilang adalah digitalisasi proses tilang, teknologi ini diharapkan menjadikan seluruh proses tilang lebih efisien. Sistem e-tilang membantu pengendara menghindari banyaknya pelanggaran berlalu-lintas. E-tilang juga membantu penegak hukum tetap bersih dari pungutan liar, istilah damai, dan denda. Sistem e-tilang yang efisien dan efektif menimbulkan berbagai pendapat dari kalangan masyarakat. Perkembangan website 4.0 dan peningkatan penggunaan media sosial sendiri dapat dimanfaatkan untuk menjadi sumber informasi dan pengambilan keputusan ini dapat dicapai melalui *text mining* yang menggunakan komentar di media sosial atau website untuk melakukan analisis sentimen. Analisis sentimen dilakukan dengan penerapan metode *Multinomial Naive Bayes*. *Multinomial Naive Bayes* merupakan salah satu metode text mining pada proses pengklasifikasian teks dengan menggunakan nilai probabilitas suatu kelas dalam suatu dokumen. Penelitian ini menganalisis opini publik tentang sistem e-tilang berbasis data twitter menggunakan metode *Multinomial Naïve Bayes* dengan ekstraksi fitur TF-IDF. Yang bertujuan untuk mengetahui tingkat akurasi, presisi, dan *recall* menggunakan metode *Multinomial Naïve Bayes* dalam klasifikasi sentimen sistem e-tilang di twitter. Hasil pada penelitian ini menunjukkan nilai akurasi tertinggi diperoleh pada *fold* $k=10$ pada langkah uji 6 dengan hasil akurasi sebesar 91,11%, presisi 90,91%, dan *recall* 91%.

Kata Kunci : E-Tilang, Twitter, Analisis Sentimen, *Multinomial Naïve Bayes*, *TF-IDF*

ABSTRACT

E-ticketing is the digitization of the ticketing process, this technology is expected to make the entire ticketing process more efficient. The e-ticket system helps drivers avoid many traffic violations. E-tickets also help law enforcement stay clear of illegal levies, peace terms and fines. An efficient and effective e-ticket system raises various opinions from the public. The development of website 4.0 and the increased use of social media itself can be used as a source of information and decision making can be achieved through text mining which uses comments on social media or websites to carry out sentiment analysis. Sentiment analysis was carried out using the Multinomial Naive Bayes method. Multinomial Naive Bayes is a text mining method in the text classification process using the probability value of a class in a document. This research analyzes public opinion about the e-ticket system based on Twitter data using the Multinomial Naïve Bayes method with TF-IDF feature extraction. The aim is to determine the level of accuracy, precision and recall using the Multinomial Naïve Bayes method in classifying the sentiment of the e-ticket system on Twitter. The results of this study show that the highest accuracy value was obtained at fold k=10 in test step 6 with accuracy results of 91,11%, precision of 90,91%, and recall of 91%.

Keywords : *E-Tilang, Twitter, Sentiment Analysis, Multinomial Naïve Bayes, TF-IDF*

HALAMAN PERSEMPBAHAN

Puji syukur kepada Allah SWT atas segala nikmat, rahmat, dan hidayah-Nya yang senantiasa selalu dilimpahkan, sehingga penulisan Tugas Akhir ini dapat diselesaikan dengan baik. Penyusunan Tugas Akhir ini tidak akan terwujud tanpa adanya dukungan, bantuan, masukan, serta dorongan dari berbagai pihak. Maka dari itu, penulis sampaikan banyak terima kasih kepada :

1. Allah SWT, yang telah memberikan segala ridho, rahmat, dan hidayah-Nya sehingga semua proses penyusunan Tugas Akhir dapat diselesaikan dengan baik.
2. Bapak Dr. Ir. Muhtar, ST.,M.T.,IPM, selaku Dekan Fakultas Teknik Universitas Muhammadiyah Jember.
3. Ibu Rosita Yanuarti, S.Kom., M.Cs, selaku Ketua Program Studi Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember.
4. Bapak Hardian Oktavianto, S.Si., M.Kom, selaku Dosen Pembimbing I dan Bapak Daryanto, M.Kom, selaku Dosen Pembimbing II yang telah banyak memberikan waktu dan kesabarannya untuk membimbing dalam proses penyusunan Tugas Akhir ini, sehingga dapat terselesaikan dengan baik.
5. Ibu Rosita Yanuarti, S.Kom., M.Cs, selaku Dosen Penguji I dan Ibu Qurrota Ayun, M.Pd, selaku Dosen Penguji II yang telah memberikan saran dan masukan pada proses penyusunan Tugas Akhir ini.
6. Bapak dan Ibu Dosen Program Studi Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember, yang turut memberikan ilmu selama berada di bangku perkuliahan.
7. Kepada kedua orang tua tercinta saya, serta seluruh keluarga terima kasih karena telah memberi segala dukungan, semangat, pengorbanan, nasihat, dan do`a yang tidak pernah berhenti selama proses penyusunan Tugas Akhir ini.
8. Kepada kakak saya Galuh Rizki Eka Lestari, A.Md.Far. dan Amri Rivaldi S.kom yang menjadi salah satu sumber motivasi dan selalu memberikan dukungan, serta mengajarkan untuk terus berusaha tiada henti untuk menjadi

seseorang yang sukses, memberikan semangat dan contoh yang baik dalam menyelesaikan tanggung jawab yang telah diberikan.

9. Kepada adikku Dia Aulia dan Balqis Maryam Amalia yang selalu menghibur, memberikan semangat, memberikan motivasi dan selama ini sehingga terselesaikan skripsi ini.
10. Kepada partner spesial di hidup saya, Elfa Afkarina Herwanti, terima kasih karena sudah meluangkan waktu menemani, menjadi *support system* dan berkontribusi banyak dalam penulisan skripsi ini, memberikan dukungan dan semangat.
11. Kepada teman satu kontrakkan, Teguh, Dody, Dewa, Agil, Tito, Wildan, Rizal, dan Galang terima kasih telah banyak membantu, menemani dan turut mendukung dalam proses penyelesaian Tugas Akhir ini.
12. Kepada diri sendiri, terima kasih karena telah mampu bertahan dan berjuang sejauh ini. meski banyak rintangan, ketakutan, tidak percaya diri yang selalu menghantui setiap hari, terima kasih karena tetap tidak pernah menyerah sesulit apapun penyusunan skripsi ini dengan menyelesaikan sebaik dan semaksimal mungkin, ini merupakan pencapaian yang patut dibanggakan untuk diri sendiri.

Laporan Tugas Akhir ini masih jauh dari kata sempurna, tetapi proses pengerjaan Tugas Akhir ini telah diselesaikan dengan sebaik – baiknya. Maka dari itu, diharapkan laporan ini dapat membantu dan menjadi manfaat bagi kita semua. Akhir kata, terima kasih sebanyak – banyaknya kepada seluruh pihak yang terkait dan telah membantu. Semoga senantiasa Allah SWT selalu memberikan keberkahan untuk kita semua, Aamiin.

MOTTO

“Apabila sesuatu yang kau senangi tidak terjadi, maka senangilah apa yang terjadi”

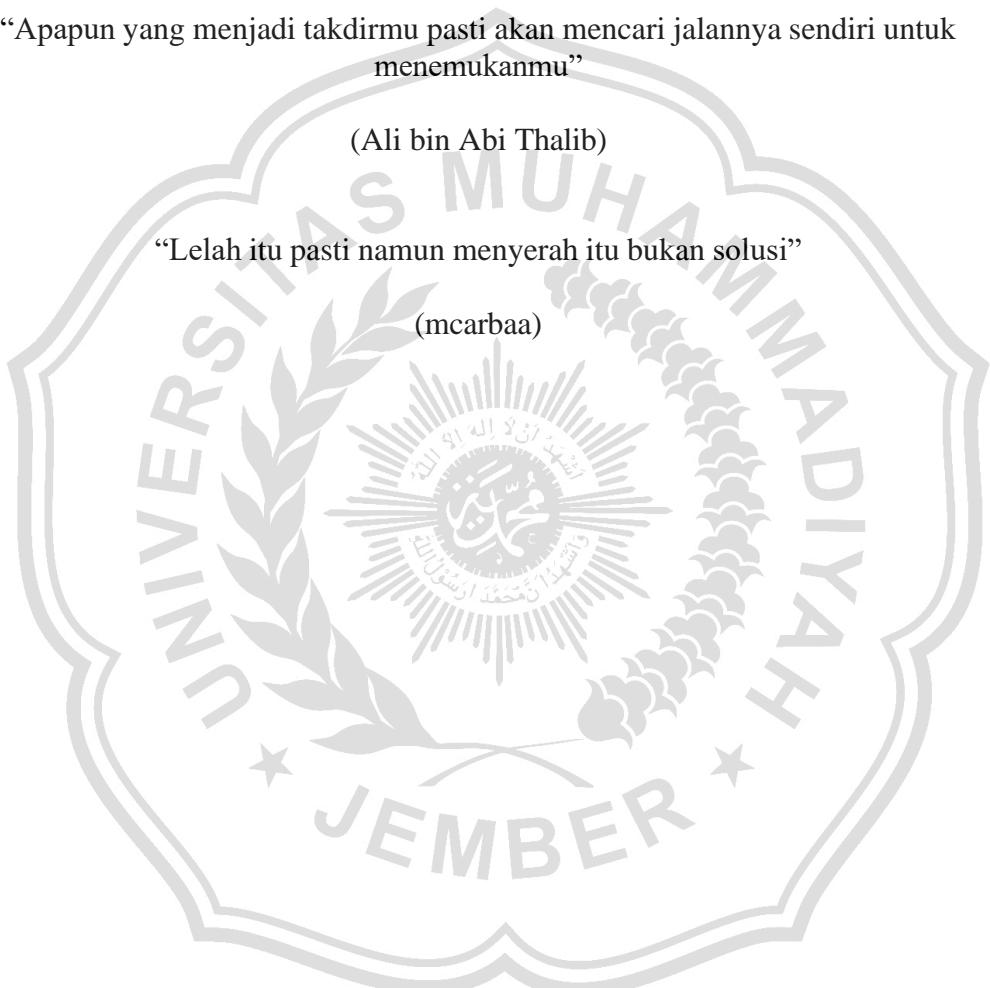
(Ali bin Abi Thalib)

“Apapun yang menjadi takdirmu pasti akan mencari jalannya sendiri untuk menemukanmu”

(Ali bin Abi Thalib)

“Lelah itu pasti namun menyerah itu bukan solusi”

(mcarbaa)



DAFTAR ISI

HALAMAN PERSETUJUAN SIDANG TUGAS AKHIR.....	i
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR	iii
ABSTRAK	iv
<i>ABSTRACT</i>	v
HALAMAN PERSEMBAHAN	vi
MOTTO	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
BAB I	14
PENDAHULUAN	14
12.1 Latar Belakang	14
12.2 Rumusan Masalah	16
12.3 Tujuan Penelitian.....	16
12.4 Manfaat Penelitian.....	16
12.5 Batasan Penelitian	16
BAB II	18
LANDASAN TEORI	18
2.1 E-tilang	18
2.2 Twitter	18
2.3 Analisis Sentimen.....	19
2.4 <i>Text Preprocessing</i>	20
2.5 Term Frequency-Inverse Document Frequency (TF-IDF).....	20
2.6 <i>Naïve Bayes</i>	21
2.7 <i>Multinomial Naïve Bayes</i>	22
2.8 Confusion Matrix.....	23
2.9 Penelitian terdahulu	24
BAB III.....	28
METODE PENELITIAN	28
3.1 Rancangan penelitian.....	28
3.2 Pengumpulan Data.....	28
3.3 <i>Text Preprocessing</i>	29
3.3.1 Pelabelan Data.....	30

3.3.2 <i>Cleansing</i>	30
3.3.3 <i>Case Folding</i>	31
3.3.4 <i>Stopword Removal</i>	31
3.3.5 <i>Stemming</i>	32
3.3.6 <i>Normalization</i>	32
3.3.7 <i>Tokenizing</i>	33
3.4 Implementasi	33
3.4.1 Pembobotan Kata (TF-IDF).....	33
3.4.2 Algoritma Klasifikasi <i>Multinomial Naïve Bayes</i>	36
3.5 Evaluasi dan Hasil	39
BAB IV	40
PEMBAHASAN	40
4.1 Implementasi Metode	40
4.1.1 Pengumpulan Data.....	40
4.1.2 Pelabelan Data	40
4.2 <i>Text Preprocessing</i>	42
4.2.1 <i>Cleansing</i>	42
4.2.2 <i>Case Folding</i>	43
4.2.3 <i>Stopword Removal</i>	43
4.2.4 <i>Stemming</i>	44
4.2.5 <i>Normalization</i>	45
4.2.6 <i>Tokenizing</i>	45
4.3 Pembobotan Kata.....	46
$TF-IDF$	46
4.4 Pembagian Data.....	47
4.5 Hasil Analisis	55
4.5.1 Hasil Nilai Akurasi, Presisi, da <i>Recall</i>	55
BAB V.....	58
KESIMPULAN DAN SARAN	58
5.1 Kesimpulan.....	58
5.2 Saran.....	58
DAFTAR PUSTAKA.....	59
LAMPIRAN – LAMPIRAN	61

DAFTAR GAMBAR

Gambar 3.1 Flowchart Tahapan Penelitian	28
Gambar 3.2 Flowchart Tahapan Text Preprocessing.....	27
Gambar 3.3 Flowchart Algoritma Multinomial Naive Bayes	33
Gambar 4.1 Data Hasil Crawling	40
Gambar 4.2 Data Label Positif dan Negatif	41
Gambar 4.3 Data Dengan Label Positif.....	41
Gambar 4.4 Data Dengan Label Negatif	41
Gambar 4.5 Hasil Akurasi, Presisi, Recall Pada 2-Fold.....	48
Gambar 4.6 Hasil Akurasi, Presisi, dan Recall Pada 4-Fold	50
Gambar 4.7 Hasil Akurasi, Presisi, dan Recall Pada 5-Fold	52
Gambar 4.8 Hasil Akurasi, Presisi, dan Recall Pada 10-Fold	55
Gambar 4.9 Visualisasi Word Cloud Komentar Positif.....	54
Gambar 4.9 Hasil Sentimen Positif dan Negatif	57

DAFTAR TABEL

Tabel 2.1 Confusion Matrix	23
Tabel 2.2 Penelitian Terdahulu.....	23
Tabel 3.1 Data Komentar Hasil Crawling	29
Tabel 3.2 Pelabelan Data Komentar	30
Tabel 3.3 Tahap Cleansing	30
Tabel 3.4 Tahap Case Folding.....	31
Tabel 3.5 Tahap Stopword Removal	31
Tabel 3.6 Tahap Stemming.....	32
Tabel 3.7 Tahap Normalization	32
Tabel 3.8 Tahap Tokenizing.....	33
Tabel 3.9 Data Pembobotan.....	33
Tabel 3.10 Hasil Perhitungan Term Frequency (TF)	34
Tabel 3.11 Hasil Perhitungan Nilai Inverse Document Frequency (IDF).....	35
Tabel 3.12 Hasil Perhitungan Pembobotan TF-IDF	35
Tabel 3.13 Data Uji	37
Tabel 3.14 Nilai Probabilitas Term Tweet Data Uji	37
Tabel 3.15 Hasil Sentimen.....	38
Tabel 3.16 Confusion Matrix.....	39
Tabel 3.17 Hasil Kriteria	39
Tabel 4.1 Hasil Label Data	42
Tabel 4.2 Proses Cleansing.....	42
Tabel 4.3 Proses Case Folding	43
Tabel 4.4 Proses Stopword Removal.....	43
Tabel 4.5 Proses Stemming	44
Tabel 4.6 Proses Normalization.....	45
Tabel 4.7 Proses Tokenizing	46
Tabel 4.8 Hasil Pembobotan Kata	46
Tabel 4.9 Confusion Matrix 2-Fold Cross Validation	47
Tabel 4.10 Confusion Matrix 4-Fold Cross Validation	49
Tabel 4.11 Confusion Matrix 5-Fold Cross Validation	51
Tabel 4.12 Confusion Matrix 10-Fold Cross Validation.....	54
Tabel 4.13 Hasil Keseluruhan Nilai Akurasi, Presisi, dan Recall	56