

**SKRIPSI**

**ANALISIS SENTIMEN DATA KOMENTAR SOSIAL MEDIA TWITTER  
TERHADAP PEMBERLAKUAN E-TILANG DI MASYARAKAT**



**Disusun oleh:**

**ANDRIANTIKA DEWI KUSUMA**

**Nomor Mahasiswa : 175410187**

**PROGRAM STUDI INFORMATIKA**

**PROGRAM SARJANA**

**FAKULTAS TEKNOLOGI INFORMASI**

**UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA**

**YOGYAKARTA**

**2024**

**SKRIPSI**

**ANALISIS SENTIMEN DATA KOMENTAR SOSIAL MEDIA TWITTER  
TERHADAP PEMBERLAKUAN E-TILANG DI MASYARAKAT**

Diajukan Sebagai Salah Satu Syarat Untuk Menyelesaikan Studi Jenjang Strata Satu

(S1)

Program Studi Informatika

Fakultas Teknologi Informasi

Universitas Teknologi Digital Indonesia

Yogyakarta

**Disusun Oleh :**

**ANDRIANTIKA DEWI KUSUMA**

**Nomor Mahasiswa : 175410187**

**PROGRAM STUDI INFORMATIKA**

**PROGRAM SARJANA**

**FAKULTAS TEKNOLOGI INFORMASI**

**UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA**

**YOGYAKARTA**

**2024**

**HALAMAN PERSETUJUAN**

**UJIAN SKRIPSI**

**Judul** : Analisis Sentimen Data Komentar Sosial Media Twitter Terhadap  
Pemberlakuan E-Tilang Di Masyarakat

**Nama** : Andriantika Dewi Kusuma

**NIM** : 175410187

**Jurusan** : Informatika

**Semester** : Genap (2023/2024)



**Telah diperiksa dan disetujui untuk diujikan di hadapan Dewan Penguji Skripsi**

**Yogyakarta, 28 Agustus 2024**

**Dosen Pembimbing,**

**Deborah Kurniawati, S.Kom., M.Cs**

**NIDN : 0511107301**

**HALAMAN PENGESAHAN**

**SKRIPSI**

**ANALISIS SENTIMEN DATA KOMENTAR SOSIAL MEDIA TWITTER  
TERHADAP PEMBERLAKUAN E-TILANG DI MASYARAKAT**

Telah Dipertahankan dan Diujikan di Depan Dewan Penguji Skripsi dan Dinyatakan  
Diterima Sebagai Syarat Guna Memperoleh Gelar Sarjana Komputer Universitas

Teknologi Digital Indonesia

Yogyakarta, 18 Juli 2024

Mengesahkan

Dewan Penguji

NIDN

1. Ariesta Damayanti, S.Kom., M.Cs. 0020047801
2. Dini Fakta Sari, S.T., M.T. 0507108401
3. Deborah Kurniawati, SKom., M.Cs. 0511107301

Tanda Tangan



Mengetahui

Ketua Program Studi Informatika



Dini Fakta Sari, S.T., M.T.

NIDN : 0507108401

## PERNYATAAN KEASLIAN SKRIPSI

Dengan ingin saya menyatakan bahwa naskah skripsi ini belum pernah diajukan untuk memperoleh gelar Sarjana Komputer di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara sah diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 28 Agustus 2024



Andriantika Dewi Kusuma

NIM : 175410187

## HALAMAN PERSEMBAHAN

Alhamdulillah robbil'alamin dengan mengucapkan syukur kepada Allah SWT.

Saya persembahkan skripsi ini untuk **kedua orangtua saya tercinta** yang telah membesarkan, mendidik, mendukung, mendoakan dan memberikan kasih sayangnya yang tak terhingga kepada saya.

Kepada ibu **Deborah Kurniawati, S.Kom., M.Cs.**, selaku dosen pembimbing saya ucapkan terimakasih banyak telah membantu dan memberikan bimbingan pengarahan serta semangat yang luar biasa dalam penyusunan tugas akhir saya.

Saya ucapkan terimakasih untuk **teman-teman seperjuangan, keluarga besar TI 4, teman-teman Taekwondo serta teman-teman angkatan 2017** yang tidak bisa saya sebutkan satu persatu telah banyak memberikan semangat dan membantu kelancaran pembuatan skripsi saya.

## **HALAMAN MOTTO**

“Barang siapa bertaqwa kepada Allah maka dia akan menjadikan jalan keluar baginya, dan memberinya rezeki dari jalan yang tidak ia sangka, dan barang siapa yang bertawakal kepada Allah niscaya Allah akan cukupkan (keperluan)nya, Sesungguhnya Allah melaksanakan kehendak-Nya, Allah telah menjadikan untuk setiap sesuatu kadarnya.” (Q.S At-Talaq : 2-3)

“Barang siapa yang mengerjakan kebaikan sekecil apapun, niscaya dia akan melihat (balasan)nya.” (Q.S Al-Zalzalah : 7)

## KATA PENGANTAR

Assalamualaikum warahmatullahi wabarokatuh.

Segala puji dan syukur saya panjatkan kepada Allah SWT dengan rahmat serta hidayahnya saya selaku penulis dapat menyelesaikan skripsi yang berjudul “ **Analisis Sentimen Data Komentar Sosial Media Twitter Terhadap Pemberlakuan E-Tilang Di Masyarakat** ”. Skripsi ini disusun sebagai syarat mendapatkan gelar Sarjana Komputer di Universitas Teknologi Digital Indonesia

Dalam proses penyusunan laporan skripsi ini saya selaku penulis menyadari bahwa saya sangat mengharapkan kritik dan saran dari semua pihak. Saya selaku penulis skripsi ini mendapatkan banyak bantuan, dukungan, serta bimbingan dari berbagai pihak, maka dari itu saya mengucapkan terimakasih yang tak terhingga kepada :

1. Sri Redjeki, S.Si., M.Kom., Ph.D., selaku ketua Universitas Teknologi Digital Indonesia.
2. Dr. Bambang Purnomosidi DP, S.E. Akt., S.Kom.,MMSI. Selaku Dekan Fakultas Teknologi Informasi.
3. Dini Fakta Sari, S.T., M.T. selaku ketua program studi strata 1 Informatika Universitas Teknologi Digital Indonesia.
4. Deborah Kurniawati, SKom., M.Cs , selaku dosen pembimbing skripsi yang telah membimbing dan memberikan arahan kepada saya selama mengerjakan skripsi.



5. Dosen program studi Informatika Universitas Teknologi Digital Indonesia yang telah memberikan bekal ilmu yang bermanfaat bagi saya.
6. Kedua orang tua saya tercinta yang selalu mendoakan, mendukung, memberikan semangat, motivasi kepada saya sehingga terciptanya keberhasilan skripsi saya.
7. Saudara yang selalu memberikan dukungan dan semangat yang luar biasa kepada saya.
8. Teman-teman yang selalu memberikan bantuan, semangat dan dukungan kepada saya.

Saya menyadari bahwa laporan skripsi ini jauh dari kata sempurna sehingga saya akan menerima semua jenis kritik dan saran yang bersifat membangun. Saya berharap skripsi ini dapat memberikan manfaat dan juga wawasan kepada pembaca.

Wassalamualaikum warahmatullahi wabarakatuh.

Yogyakarta, 28 Agustus 2024

Andriantika Dewi Kusuma

## DAFTAR ISI

<b>HALAMAN PERSETUJUAN.....</b>	<b>ii</b>
<b>HALAMAN PENGESAHAN.....</b>	<b>iii</b>
<b>PERNYATAAN KEASLIAN SKRIPSI.....</b>	<b>iv</b>
<b>HALAMAN PERSEMBAHAN.....</b>	<b>v</b>
<b>HALAMAN MOTTO .....</b>	<b>vi</b>
<b>KATA PENGANTAR.....</b>	<b>vii</b>
<b>DAFTAR ISI.....</b>	<b>ix</b>
<b>DAFTAR GAMBAR.....</b>	<b>xi</b>
<b>DAFTAR TABEL.....</b>	<b>xiii</b>
<b>INTISARI.....</b>	<b>xiv</b>
<b>ABSTRACT .....</b>	<b>xv</b>
<b>BAB 1 PENDAHULUAN .....</b>	<b>1</b>
<b>1.1 Latar Belakang Masalah .....</b>	<b>1</b>
<b>1.2 Rumusan Masalah .....</b>	<b>5</b>
<b>1.3 Ruang Lingkup.....</b>	<b>5</b>
<b>1.4 Tujuan Penelitian.....</b>	<b>6</b>
<b>1.5 Manfaat Penelitian.....</b>	<b>6</b>
<b>BAB 2 TINJAUAN PUSTAKA DAN DASAR TEORI .....</b>	<b>8</b>
<b>2.1 Tinjauan Pustaka .....</b>	<b>8</b>
<b>2.2 Dasar Teori .....</b>	<b>11</b>
<b>BAB 3 METODE PENELITIAN.....</b>	<b>21</b>

3.1	Bahan/Data .....	21
3.2	Peralatan .....	23
3.3	Analisis dan Perancangan Sistem .....	23
<b>BAB 4</b>	<b>IMPLEMENTASI DAN PEMBAHASAN SISTEM.....</b>	<b>40</b>
4.1	Implementasi dan Pembahasan Sistem.....	40
<b>BAB V</b>	<b>PENUTUP.....</b>	<b>61</b>
5.1	Kesimpulan .....	61
5.2	Saran .....	61
<b>DAFTAR PUSTAKA</b>	<b>.....</b>	<b>62</b>
<b>LAMPIRAN</b>	<b>.....</b>	<b>65</b>

## DAFTAR GAMBAR

Gambar 3. 1 Alur proses pengerjaan .....	24
Gambar 3. 2 Flowchart cleansing .....	25
Gambar 3. 3 Flowchart case folding.....	26
Gambar 3. 4 Flowchart tokenizing .....	27
Gambar 3. 5 Flowchart stopword removal.....	28
Gambar 3. 6 Flowchart stemming .....	29
Gambar 3. 7 Flowchart TF-IDF .....	32
Gambar 3. 8 Flowchart K-NN.....	35
Gambar 3. 10 Rancangan antarmuka data tweet .....	38
Gambar 3. 11 Rancangan antarmuka visualisasi.....	39
Gambar 4. 1 Potongan program scraping data.....	41
Gambar 4. 2 Hasil scraping data .....	41
Gambar 4. 3 Halaman Dashboard.....	42
Gambar 4. 4 Potongan program menu login .....	42
Gambar 4. 5 Tampilan Menu Tambah Data Admin .....	43
Gambar 4. 6 Potongan program menu simpan user.....	44
Gambar 4. 7 Tampilan Menu Tambah Data Latih.....	45
Gambar 4. 8 Tampilan halaman data latih .....	47
Gambar 4. 9 Potongan program pelabelan data .....	48
Gambar 4. 10 Potongan program cleansing .....	49
Gambar 4. 11 Potongan program case folding.....	49
Gambar 4. 12 Potongan program tokenizing .....	49
Gambar 4. 13 Potongan program stopword removal.....	50
Gambar 4. 14 Potongan program stemming .....	50
Gambar 4. 15 Hasil Preprocessing .....	51
Gambar 4. 16 Tampilan Menu Uji Data .....	52
Gambar 4. 17 Potongan program pembobotan TF-IDF .....	52

<b>Gambar 4. 18 Potongan program perhitungan nilai scalar .....</b>	<b>53</b>
<b>Gambar 4. 19 Potongan program perhitungan nilai vektor .....</b>	<b>54</b>
<b>Gambar 4. 20 Potongan program perhitungan nilai cosine similarity .....</b>	<b>55</b>
<b>Gambar 4. 21 Potongan program perhitungan nilai k. ....</b>	<b>55</b>
<b>Gambar 4. 22 Tampilan Menu Klasifikasi K-NN .....</b>	<b>56</b>
<b>Gambar 4. 23 Hasil pengujian sistem.....</b>	<b>57</b>
<b>Gambar 4. 24 Hasil confusion matrix .....</b>	<b>57</b>
<b>Gambar 4. 25 Presentase klasifikasi.....</b>	<b>59</b>
<b>Gambar 4. 26 Grafik klasifikasi .....</b>	<b>59</b>

## DAFTAR TABEL

Tabel 2. 1 Tinjauan Pustaka .....	8
Tabel 2. 2 Tabel confusion matrix .....	19
Tabel 3. 1 Tabel Hasil Cleansing. ....	25
Tabel 3. 2 Tabel Hasil Case Folding.....	26
Tabel 3. 3 Tabel Hasil Tokenizing .....	27
Tabel 3. 4 Tabel Hasil Stopword Removal.....	28
Tabel 3. 5 Tabel Hasil Stemming .....	30
Tabel 3. 6 Tabel contoh data latih .....	30
Tabel 3. 7 Tabel contoh data uji .....	31
Tabel 3. 8 Tabel menghitung TF.....	32
Tabel 3. 9 Tabel menghitung DF .....	33
Tabel 3. 10 Tabel menghitung IDF .....	33
Tabel 3. 11 Tabel menghitung TF-IDF .....	34
Tabel 3. 12 Tabel menghitung perkalian skalar ( $W*W_{di}$ ).....	35
Tabel 3. 13 Tabel menghitung panjang vector .....	36
Tabel 3. 14 Tabel Cosine Similarity.....	36
Tabel 4. 1 Pengujian sistem berdasarkan nilai k .....	58

## INTISARI

Twitter yang saat ini bisa disebut dengan aplikasi X adalah salah satu media sosial yang digunakan pengguna untuk media komunikasi dan informasi. Informasi yang didapat berupa komentar, pertanyaan, ataupun opini baik yang bersifat positif, netral, ataupun negative. Informasi yang disajikan salah satunya adalah pemberlakuan e-tilang. Informasi masyarakat dengan sentimen bisa dijadikan indicator untuk menentukan kualitas dari pemberlakuan e-tilang dari sisi opini masyarakat. Oleh karena itu, peneliti akan melakukan analisa tweet dari opini masyarakat mengenai pemberlakuan e-tilang di Indonesia.

Penelitian ini menggunakan metode K-Nearest Neighbor dengan beberapa tahapan. Tahap pre-processing yang terdiri atas cleansing, case folding, tokenizing, stop removal dan stemming. Tahap kedua yaitu perhitungan pembobotan dengan TF-IDF, dan yang ketiga adalah klasifikasi sentiment kedalam 3 kategori yaitu negative, positif dan netral. Data yang digunakan berupa opini masyarakat dengan jumlah data 1181 data tweet yang diperoleh dari hasil scraping menggunakan tweepy di python.

Hasil yang diperoleh dari pengujian analisis dengan menggunakan metode K-nearest Neighbor menghasilkan akurasi tertinggi dengan nilai  $k = 3$  yaitu 78%. Sedangkan nilai presisi tertinggi yaitu 72% dan nilai recall tertinggi yaitu 62%. Hasil analisis sentimen dari pemberlakuan system e-tilang di Indonesia mendapatkan respon positif sebesar 68%.

Kata kunci : *analisis sentiment, e-tilang, k-nearest neighbor, tf-idf, twitter*

## ABSTRACT

Twitter, which can currently be called the X application, is one of the social media used by users for communication and information media. The information obtained is in the form of comments, questions, or opinions, whether positive, neutral, or negative. One of the information presented is the implementation of e-tickets. Public information with sentiment can be used as an indicator to determine the quality of the implementation of e-ticketing from the side of public opinion. Therefore, the researcher will analyze tweets from public opinion regarding the implementation of e-ticketing in Indonesia.

This study uses the K-Nearest Neighbor method with several stages. The pre-processing stage consists of cleansing, case folding, tokenizing, stop removal and stemming. The second stage is the calculation of weighting with TF-IDF, and the third is the classification of sentiment into 3 categories, namely negative, positive and neutral. The data used was in the form of public opinion with a total of 1181 tweet data obtained from scraping results using tweepy in python.

The results obtained from the analysis test using the K-nearest Neighbor method produced the highest accuracy with a value of  $k = 3$  which was 78%. Meanwhile, the highest precision value is 72% and the highest recall value is 62%. The results of sentiment analysis from the implementation of the e-ticketing system in Indonesia received a positive response of 68%.

Keywords: *sentiment analysis, e-tilang, k-nearest neighbor, tf-idf, twitter*