

## BAB I

### PENDAHULUAN

#### 1.1 Latar Belakang

Sebagai warga negara Indonesia yang baik, pengemudi, pengguna jalan dan otoritas lainnya harus mematuhi peraturan yang berlaku. Aturan lalu lintas yang telah diperkenalkan harus menjadi peran penting bagi pengendara, dan tentu saja juga menghindari denda saat mengeluarkan tilang untuk alasan keamanan. Tingkat kesadaran masyarakat tentang kedisiplinan dalam berkendara di jalan raya sangatlah rendah dengan di buktikan angka kecelakaan yang tinggi yaitu mencapai 28-38 ribu pertahun meninggal di tanah air akibat kecelakaan lalu lintas [1].

Dalam usaha untuk mengurangi jumlah pelanggaran dan kecelakaan di jalan, Kepolisian Republik Indonesia menerapkan sistem penegakan hukum lalu lintas yang menggunakan bukti elektronik atau tilang elektronik yang dikenal sebagai *Electronic Traffic Law Enforcement* (E-TLE). Melalui penerapan kebijakan ini, tindakan akan diambil untuk menegur dan menindak individu-individu yang melanggar aturan dan kurang disiplin dalam mengemudi. Selain itu kebijakan ini untuk mengurangi tindak pidana suap kepada oknum-oknum yang tidak bertanggung jawab. Seiring berjalannya waktu, pihak kepolisian dan pemerintah telah memanfaatkan kemajuan teknologi dalam upaya meningkatkan kesadaran masyarakat akan pentingnya tertib berlalu lintas, dengan tujuan mengurangi jumlah korban jiwa dan pelanggaran di negara ini. Salah satu inovasi yang diterapkan adalah sistem tilang elektronik (E-Tilang), yang diharapkan dapat menurunkan tingkat pelanggaran yang terjadi [2].

Dengan adanya kejadian tersebut peneliti kemudian ingin mengetahui persepsi masyarakat tentang kebijakan penindakan pelanggaran lalu lintas dan bukti tilang elektronik (E-Tilang). Dalam studi ini, kami mengadopsi metode analisis sentimen untuk mengevaluasi opini masyarakat terkait penggunaan etilang di platform media sosial Twitter. Twitter, sebagai salah satu media sosial berorientasi teks yang populer dan sering digunakan, telah dipilih sebagai sumber data utama dalam penelitian ini [3].

Keberadaan layanan ini mengakibatkan kecenderungan masyarakat untuk lebih memilih mengungkapkan pendapat mereka melalui media sosial daripada melalui interaksi langsung. Ungkapan opini tersebut dapat dijadikan sebagai sumber informasi. Pendapat dan persepsi yang tertulis oleh individu di platform media sosial, seperti Twitter, memiliki variasi yang sangat beragam. Terdapat kalangan yang mendukung secara kuat, namun juga ada kalangan yang menolak dan mengkritik pemberlakuan tilang elektronik (E-Tilang). Respon masyarakat terhadap

implementasi tilang elektronik (E-Tilang) seharusnya menjadi pertimbangan bagi pemerintah, dengan membandingkan antara respons positif dan negatif yang disampaikan melalui Twitter. Salah satu cara untuk mengkategorikan opini atau persepsi positif dan negatif bisa dilakukan dengan cara analisis sentiment [4].

Analisis sentimen adalah bidang studi yang fokus pada menganalisis pendapat, evaluasi, dan sikap terhadap berbagai entitas, seperti produk, layanan, organisasi, individu, topik, dan atribut yang terkait. Tujuannya adalah mengklasifikasikan teks-teks berdasarkan opini atau sentimen yang terungkap di dalamnya [5]. Salah satu metode yang bisa digunakan dalam proses klasifikasi adalah *naive bayes*. Metode ini memiliki konsep yang mudah dimengerti dan memberikan hasil yang akurat dengan efisiensi tinggi.

Untuk melakukan penelitian analisis sentimen terhadap pemberlakuan kebijakan E-Tilang, peneliti ini menggunakan metode *Naive bayes* dan Seleksi Fitur *Query Expansion Ranking* (QER). Penggunaan Metode *Naive bayes* terbukti memiliki hasil yang sangat baik dalam banyak penelitian yang dilakukan dengan tingkat akurasi data besar. Data yang berupa komentar masyarakat yang ada dalam *twitter* akan diolah menggunakan *Naive bayes* dengan cara menghitung probabilitasnya. Tingkat akurasi *Naive bayes Classifier* yang tinggi adalah salah satu kelebihan algoritma *Naive bayes* apabila dibandingkan algoritma lain yang sejenis. Hal ini menjadikan algoritma *Naive bayes* cukup baik untuk digunakan dalam perhitungan yang memiliki jumlah data besar [6].

Penelitian ini menggunakan metode Naive Bayes sebagai algoritma untuk mengklasifikasikan sentimen pada tweet di platform media sosial Twitter. Naive Bayes adalah sebuah algoritma yang sederhana, yang menghitung kemungkinan kombinasi total data yang muncul serta nilai-nilai yang terdapat dalam kumpulan data. Data teks dari Twitter yang sebelumnya diproses terlebih dahulu kemudian dilakukan *pre-processing* dalam beberapa langkah *pre-processing* yang terdiri dari *Cleansing*, *Case folding*, *Tokenizing*, *Slangword*, *Stopword*, dan *Stemming*. Tujuan *pre-processing* ini adalah untuk membuat data lebih bersih dan terstruktur. Kemudian diekstraksi dan dibobot sebelum masuk ke proses probabilitas menggunakan *Naive bayes*. Salah satu metode ekstraksi dan pembobotan kata yang dapat digunakan adalah TF-IDF. TF-IDF adalah singkatan dari Term Frequency-Reverse Document Frequency, yang merupakan metode untuk menentukan nilai atau bobot kata dalam sebuah dokumen atau kalimat saat dokumen atau kalimat tersebut dimasukkan ke dalam data [7].

Secara mendasar, analisis sentimen melibatkan pengklasifikasian teks ke dalam dua kategori, yaitu positif dan negatif, dengan menghitung nilai dari fitur-fitur yang ada. Namun, salah

satu tantangan dalam analisis sentimen adalah dimensi yang tinggi dari fitur-fitur yang digunakan dalam proses klasifikasi. Dalam rangka mengurangi kompleksitas, beberapa penelitian telah menggunakan teknik seleksi fitur untuk mengurangi jumlah fitur yang digunakan dalam tahap klasifikasi [8].

Pemilihan fitur ini bertujuan untuk mengurangi jumlah fitur agar proses klasifikasi menjadi lebih efisien dan efektif. Dalam penelitian ini, digunakan metode Seleksi Fitur *Query Expansion Ranking* (QER), yang sebelumnya telah digunakan dalam penelitian dengan judul "*A New Feature Selection Method for Sentiment Analysis of Turkish Reviews*". Penelitian ini menerapkan teknik seleksi fitur yang disebut QER untuk melakukan analisis sentimen pada empat kategori yang berbeda. Selanjutnya, metode QER dibandingkan dengan teknik seleksi fitur lain, seperti Chi-Square dan Document Frequency Difference. Temuan dari penelitian ini menunjukkan bahwa QER memberikan hasil terbaik dengan tingkat akurasi mencapai 91% [9].

Pada penelitian ini, respon masyarakat di twitter dapat menentukan apakah hasil analisis sentimen adalah positif dan negatif sehingga berdasarkan data tersebut kita dapat mengetahui pendapat masyarakat indonesia secara general terhadap penerapan kebijakan E-Tilang. Pada penelitian sebelumnya dengan menggunakan metode seleksi fitur *query ekspansion ranking* dan *naive bayes* memiliki kelebihan dalam penggunaan data training yang kecil untuk proses pengklasifikasian sehingga menghasilkan akurasi tertinggi sebesar 86,6% pada *naive bayes* dan 75% pada seleksi fitur. Namun kekurangannya yaitu dengan lebih memperhatikan singkatan, gabungan dua kata atau lebih, kata-kata ambigu dan kalimat sarkastik supaya hasil *slangword* lebih sempurna. Kontribusi penelitian ini adalah penambahan *Query Expansion Ranking* bertujuan untuk mengurangi jumlah fitur sehingga dapat meningkatkan akurasi sehingga tingkat akurasi meningkat dari penelitian sebelumnya.

## **1.2 Perumusan Masalah**

### **1.2.1 Permasalahan**

Dengan adanya kebijakan E-Tilang yang baru di terapkan oleh pemerintah yang sosialisasinya sangat minim terhadap masyarakat, terdapat banyak respon positif dan negatif karena adanya beberapa opini tentang ketidak puasan dari masyarakat tentang adanya E-Tilang. Sehingga di butukan analisis sentimen, dimana dalam penelitian ini menggunakan Metode *Naive bayes*. Metode *Naive bayes* dalam proses klasifikasi membutuhkan waktu yang lama karena tergantung dari semakin banyaknya fitur. Sehingga membutuhkan seleksi fitur untuk memilih fitur yang penting untuk meningkatkan akurasinya. Metode yang digunakan dalam seleksi fitur adalah metode seleksi fitur *Query Expansion Ranking* (QER).

### **1.2.2 Metode Usulan**

Peneliti mengusulkan metode *naive bayes* dan seleksi fitur *Query Expansion Ranking*, dengan data yang dikumpulkan dari media sosial twitter tentang kebijakan E-Tilang. Dengan menggunakan metode *naive bayes* dan seleksi fitur *Query Expansion Ranking* peneliti dapat mengetahui tentang kebijakan E-Tilang apakah berdampak baik atau sebaliknya bagi masyarakat pengguna lalu lintas.

### **1.2.3 Pertanyaan Penelitian**

- a. Seberapa besar pengaruh Seleksi Fitur *Query Expansion Ranking* untuk meningkatkan akurasi dan komputasi dari Metode Klasifikasi *Naïve Bayes* ?

## **1.3 Tujuan dan Manfaat**

### **1.3.1 Tujuan**

- a. Untuk mengetahui seberapa besar pengaruh Seleksi Fitur *Query Expansion Ranking* untuk meningkatkan akurasi dan komputasi dari Metode Klasifikasi *Naïve Bayes*

### **1.3.2 Manfaat**

- a. Dapat mengetahui informasi terhadap opini atau persepsi masyarakat terhadap kebijakan E-Tilang berupa opini positif atau negatif guna dijadikan bahan evaluasi pemerintah.
- b. Sebagai media pendukung untuk mendapatkan *feedback* dari masyarakat dalam proses pengambilan keputusan.

## **1.4 Batasan-Batasan**

Penentuan Batasan masalah dilakukan agar tidak terjadi penyimpangan dalam pembahasan sehingga peneliti lebih terfokus dalam melakukan penelitian sebagai berikut:

- a. Data yang digunakan adalah hasil *tweet* dari opini masyarakat menggunakan media sosial twitter terhadap kebijakan E-Tilang di Indonesia selama periode tahun 2022 -2023 sebanyak 1172 data.
- b. Data yang akan disajikan berupa opini positif dan negatif di twitter terhadap kebijakan E-Tilang di Indonesia.
- c. Metode yang dipakai dalam penelitian ini menggunakan *Naive Bayes* dan *Seleksi Fitur Query Expansion Ranking*.

## **1.5 Sistematika Skripsi**

Berikut adalah pedoman yang harus diikuti saat menulis laporan berdasarkan format penulisan Skripsi Program Studi Sistem Informasi di Fakultas Teknik Universitas Trunojoyo Madura:

### **BAB I : PENDAHULUAN**

Pada bab ini membahas mengenai latar belakang, perumusan masalah, tujuan dan manfaat, batasan-batasan serta sistematika skripsi.

### **BAB II : KAJIAN PUSTAKA**

Pada bab ini membahas mengenai uraian tinjauan pustaka, serta teori penunjang lainnya sebagai bahan referensi.

### **BAB III : METODE USULAN**

Pada bab ini membahas tentang tahapan penelitian, metode yang diusulkan, arsitektur (Input, Output, Proses), metode penelitian, skenario pengujian serta data yang digunakan.

### **BAB IV : HASIL DAN ANALISIS**

Dalam bab ini membahas tentang hasil dari sistem dengan acuan pada model sistem yang sudah dibangun.

### **BAB V : PENUTUP**

Dalam bab ini membahas tentang kesimpulan skripsi yang dibuat secara keseluruhan serta kekurangan pada suatu sistem yang dibangun.