

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Kepolisian Indonesia adalah penegak hukum di negara Republik Indonesia yang bekerja untuk melaksanakan fungsi kepolisian, yang meliputi penegakan hukum, perlindungan, pengayoman, dan pelayanan masyarakat, serta berfungsi sebagai pemelihara keamanan dalam negeri.[1]. Keberadaan institusi kepolisian tengah berada dalam masa transisi yang masih dihadapkan dengan berbagai persoalan eksternal dan internal. Tingginya angka kejahatan dan kejahatan konvensional dan transnasional merupakan indikasi masalah eksternal, sedangkan kurangnya kapasitas sumber daya manusia (SDM) bagi aparat kepolisian merupakan masalah internal. Tuntutan dan kebutuhan masyarakat tersebut harus menentukan keberadaan dan fungsi polisi dalam masyarakat. Tugas polisi adalah untuk menjaga keamanan dan ketertiban masyarakat, untuk memastikan bahwa petugas penegak hukum bekerja menuju tujuan yang dimaksudkan, dan untuk mencegah kerugian terhadap individu, masyarakat, dan negara.[2].

Salah satu aspek yang menarik perhatian masyarakat adalah kinerja kepolisian. Kepercayaan masyarakat Indonesia terhadap polisi mulai berkurang dalam beberapa tahun terakhir, seperti beberapa kasus yang menimbulkan pertanyaan tentang kejujuran petugas polisi, seperti, yang dikemukakan oleh Pada 14 Oktober 2021, situs berita CNN Indonesia melaporkan bahwa tagar "Percuma melapor ke polisi" menjadi topik pembicaraan yang populer. Ketika seorang ibu di Sulawesi Selatan mengeluhkan ketiga anaknya yang dilecehkan secara seksual, tagar tersebut mendapat perhatian. Lantaran kasus dihentikan karena bukti tidak kuat, pengguna aktif media sosial pun ikut bersimpati. Selain kasus di atas, Menurut Asfinawati, Ketua Yayasan Lembaga Bantuan Hukum Indonesia (YLBHI), ada 68 kasus kekerasan polisi saat demonstrasi. Sebanyak 3.539 korban ditangkap secara sewenang-wenang, diikuti 3.539 korban ditahan sewenang-wenang dan 474 korban disiksa[3]. Terlepas dari kasus-kasus tersebut, masih terdapat beberapa dilema yang berakibat gambaran kepolisian umumnya masih

terlihat buram di mata publik dan media, yang seharusnya menjadi perhatian khusus yang perlu segera disikapi. Hal itu dilakukan untuk menyikapi dan menanggapi perasaan kurang baik yang masih ada dan berafiliasi dengan kepolisian saat ini. Menurut data yang terekam baik oleh media massa maupun *National Police Commission of Indonesia*, terdapat beberapa PR yang wajib segera diselesaikan menyangkut urusan internal institusi kepolisian sendiri [4].

Saat ini, media sosial banyak digunakan. Beriklan, membeli dan menjual, menyimpan data, dan sebagai sarana ekspresi diri pengguna adalah penggunaan umum media sosial. Penggunaan media sosial menjadi lebih mudah sebagai hasil dari kemajuan teknologi yang memungkinkan penggunaan media sosial melalui aplikasi seluler yang mudah digunakan kapan pun dan di mana pun Anda memiliki koneksi internet[5]. Menurut survei atau survei tahun 2017 oleh Asosiasi Penyelenggara Jasa Internet Indonesia (APJII), 143,26 juta dari total penduduk Indonesia atau 262 juta orang menggunakan Internet, dan 87,13% di antaranya menggunakan media sosial. Facebook adalah platform media sosial pertama yang digunakan secara luas pada tahun 2004. Twitter adalah platform media sosial yang populer pada tahun 2006 seperti halnya Facebook[6].

*Statista Research Department* tahun 2021 menyajikan statistik pengguna Twitter di seluruh dunia, dan Amerika Serikat memiliki jumlah pengguna Twitter terbesar. Indonesia memiliki jumlah pengguna Twitter terbanyak keenam di dunia dengan 17,5 juta orang. [7]. Twitter adalah media sosial tempat Anda dapat mengunduh tweet dengan istilah frase atau paragraf. Tweet sering digunakan untuk menyampaikan pendapat atau informasi tentang berbagai topik [8].

Analisis sentimen adalah metode yang digunakan untuk menentukan isi dari kumpulan data berupa teks negatif, positif atau netral. [9]–[11]. Metode analisis sentimen yang umum adalah *Naive Bayes Classifier* (NBC), *K-Nearest Neighbor* (KNN), *Support Vector Machine* (SVM). [12]. Karena kemudahan penggunaannya, *Naive Bayes Classifier* banyak digunakan[13]. Metode Naive Bayes memiliki keuntungan karena membutuhkan sedikit pelatihan data untuk memperkirakan parameter (mean dan varians variabel) yang diperlukan untuk klasifikasi[14]. Metode Naive Bayes sangat sederhana dan efektif, tetapi di sisi lain, Naive Bayes sangat sensitif terhadap pilihan fitur. Masalah terbesar dalam

klasifikasi teks adalah tingginya dimensi ruang fitur, yang sering terjadi pada teks dengan puluhan ribu fitur[14]. Sebagian besar fitur ini sepele dan tidak berguna untuk klasifikasi teks, dan bahkan dapat menurunkan akurasi, jadi memilih fitur yang tepat sangatlah penting.

Sebelumnya beberapa teknik telah diperoleh untuk memilih fitur dokumen, antara lain document density thresholding (DF), validasi informasi, mutual information (MI), term strength (TS), dan uji chi-square (X2). Penelitian sebelumnya yang menggunakan metode klasifikasi Naive Bayes dengan pemilihan fitur Chi-Square menunjukkan bahwa metode NBC mencapai akurasi 89%, presisi 89%, dan recall 100%. [15]. Tujuan dari penelitian ini adalah untuk menunjukkan bagaimana kinerja algoritma Naive Bayes Classifier dalam klasifikasi analisis sentimen dipengaruhi oleh pemilihan fitur Chi-Square. Menggunakan data dari Twitter dan media sosial lainnya, penelitian ini juga bertujuan untuk mengkaji bagaimana perasaan masyarakat Indonesia terhadap kinerja polisi.

## **1.2 Rumusan Masalah**

Perumusan masalah yang dihasilkan berdasarkan latar belakang diatas adalah seperti berikut:

### **1.2.1 Permasalahan**

Kepercayaan masyarakat terhadap polisi di Indonesia mulai menurun dalam beberapa tahun terakhir, dan sejumlah kejadian membuat masyarakat mempertanyakan kinerja polisi. Masih ada beberapa persoalan yang membuat citra polisi secara umum tidak jelas di mata publik dan media, yang seharusnya menjadi perhatian khusus dan membutuhkan penyelesaian segera.. Sehingga perlu dilakukan analisis terkait sentimen masyarakat tentang kinerja kepolisian.

Metode Naive Bayes sangat sederhana dan efektif, tetapi di sisi lain, *Naive Bayes* sangat sensitif terhadap pemilihan fitur. Masalah terbesar dalam klasifikasi teks adalah tingginya dimensi ruang fitur, yang sering terjadi pada teks dengan puluhan ribu fitur. Sebagian besar fitur ini sepele dan tidak berguna

untuk klasifikasi teks, dan bahkan dapat menurunkan akurasi, jadi memilih fitur yang tepat sangatlah penting.

### **1.2.2 Metode Usulan**

Penelitian ini dilakukan untuk membuktikan pengaruh seleksi fitur *Chi-Square* terhadap kinerja algoritma *Naïve Bayes Classifier* dalam menganalisis sentimen opini publik terhadap kinerja kepolisian. Menggunakan data dari Twitter dan media sosial lainnya, penelitian ini juga bertujuan untuk mengkaji bagaimana perasaan masyarakat Indonesia terhadap kinerja polisi.

### **1.2.3 Pertanyaan Penelitian**

1. Bagaimana performa yang dihasilkan oleh algoritma *Naïve Bayes Classifier* dan seleksi fitur *Chi-Square* dalam menganalisis sentimen opini publik terhadap kinerja kepolisian?
2. Apakah seleksi fitur *Chi-Square* berpengaruh dalam meningkatkan kinerja algoritma *Naïve Bayes* dan berapa nilai kritis pada taraf nyata  $\alpha$  yang optimal pada *Chi-Square*?

## **1.3 Tujuan dan Manfaat Penelitian**

Tujuan dan manfaat yang dihasilkan dari penelitian tugas akhir ini adalah sebagai berikut :

### **1.3.1 Tujuan Penelitian**

1. Mengetahui performa dan pengaruh yang dihasilkan oleh seleksi fitur *Chi-Square* terhadap kinerja algoritma *Naïve Bayes Classifier* dalam menganalisis sentimen opini publik terhadap kinerja kepolisian.
2. Mengetahui pengaruh dari seleksi fitur *Chi-Square* dalam meningkatkan kinerja algoritma *Naïve Bayes* dan nilai kritis pada taraf nyata  $\alpha$  yang optimal pada *Chi-Square*.

### 1.3.1 Manfaat Penelitian

1. Dapat memberikan wawasan dari penerapan seleksi fitur *Chi-Square* dan algoritma *Naïve Bayes*.
2. Mampu menghasilkan suatu sistem yang dapat menganalisa sentimen masyarakat dengan menggunakan data dari media sosial Twitter.
3. Menjadi referensi bagi penelitian selanjutnya.

### 1.4 Batasan Masalah

Ada batasan-batasan yang diterapkan pada penelitian tugas akhir ini, diantaranya yaitu:

1. Kata kunci yang digunakan untuk mengambil data *tweet* berdasarkan tagar #percumalaporpholisi, #polisiindo, dan #oknumpolisi dalam kurun waktu tahun 2021 - 2022.
2. Data hasil *crawling* akan diberi label secara manual oleh 3 pakar.
3. Sistem yang dibuat tidak memperhatikan kesalahan ejaan pada dokumen yang diuji.
4. Peneliti menggunakan nilai kritis pada taraf nyata  $\alpha$  0.25, 0.1, 0.05 pada seleksi fitur dengan menggunakan metode *Chi-Square*.

### 1.5 Sistematika Penulisan

Pada penelitian tugas akhir ini, digunakan sistematika penulisan sebagai berikut:

## BAB I PENDAHULUAN

Bab pertama penelitian ini memuat pendahuluan yang memberikan gambaran umum tentang penyusunannya. Penulis merangkum isi dari masing-masing lima bab tersebut, yang dirinci menjadi latar belakang masalah, cara mendefinisikannya, tujuan dan manfaat penelitian, masalah, dan sistematika penulisan.

## **BAB II KAJIAN PUSTAKA**

Pada bab ini memuat kajian penelitian terdahulu yang melandasi penulisan penelitian ini, dan kajian pustaka yang memuat landasan teori seperti Analisis Sentimen, *Text Mining*, Twitter, *Data Crawling*, *Text Preprocessing*, Seleksi Fitur *Chi-Square*, *Naïve Bayes Classifier*, *Confusion Matrix* dan Penelitian Terkait.

## **BAB III METODE USULAN**

Bab ketiga adalah metode atau metode penelitian yang diusulkan, yang meliputi jenis dan sumber data. Bab ini juga menjelaskan tentang teknik pengolahan data untuk tahapan penelitian dan evaluasi hasil analisis klasifikasi..

## **BAB IV HASIL DAN PEMBAHASAN**

Bab ini berisi tentang hasil implementasi sistem dengan menerapkan metode seleksi fitur *Chi-Square* dan *Naive Bayes* didalamnya. Hasil implementasi terdiri dari lingkungan uji coba, skenario uji coba yang dilakukan, tahapan analisis data, hasil uji coba, dan analisis hasil uji coba dari hasil klasifikasi.

## **BAB V PENUTUP**

Bab ini membahas temuan penelitian yang disesuaikan dengan pertanyaan penelitian dan rekomendasi untuk pengembangan aplikasi lebih lanjut.