

BAB I

PENDAHULUAN

1.1. Latar Belakang

Seiring berkembangnya teknologi informasi jejaring sosial dan platform ulasan daring seperti *Twitter*, *Instagram*, dan *Google Review*, dan platform e-commerce seperti *Shopee* dan *Tokopedia*, telah menjadi sarana penting bagi konsumen untuk berbagi pendapat, ulasan, dan pengalaman mereka tentang produk atau layanan. Ulasan ini mengandung banyak pendapat baik positif maupun negatif, yang dapat mempengaruhi persepsi pelanggan lain untuk membeli sesuatu.

Seiring dengan meningkatnya jumlah ulasan yang diterima setiap hari, perusahaan menghadapi tantangan dalam mengolah data secara manual. Proses analisis sentimen secara manual memerlukan waktu serta sumber daya yang besar, terutama jika data yang harus dianalisis sangat banyak. Oleh karena itu, dibutuhkan sistem otomatis yang dapat memproses dan menganalisis sentimen pelanggan secara efisien. Analisis sentimen, yang juga dikenal sebagai *opinion mining*, merupakan teknik dalam pemrosesan bahasa alami (NLP) yang bertujuan untuk mengidentifikasi opini yang terkandung dalam teks[1]. Melalui analisis sentimen, perusahaan dapat memahami pola kepuasan pelanggan dan mengambil langkah strategis untuk meningkatkan kualitas layanan atau produk mereka.

Penelitian ini dilakukan karena desain *jersey* Tim Nasional (Timnas) Indonesia dinilai tidak sesuai dengan harapan para penggemar. Selain itu, desainer yang bertanggung jawab atas pembuatan *jersey* tersebut tidak menerima kritik atau masukan dari berbagai pihak, yang berujung pada meningkatnya ketidakpuasan. Informasi mengenai ketidakpuasan ini bersumber dari berita yang diterbitkan oleh Kompas, yang melaporkan adanya keluhan dan kritik dari para penggemar terkait desain *jersey* yang dianggap tidak memenuhi ekspektasi mereka. Situasi ini menimbulkan berbagai permasalahan, seperti menurunnya minat terhadap *jersey* Timnas dan potensi penurunan angka penjualan. Lebih dari itu, ketidakpuasan terhadap desain *jersey* ini mencerminkan permasalahan yang lebih mendalam terkait bagaimana aspirasi dan harapan penggemar terhadap tim nasional Indonesia belum terpenuhi. Topik ini dipilih karena memiliki relevansi dengan isu-isu yang

sedang dihadapi oleh Timnas Indonesia dan para penggemarnya. Penelitian ini juga menyoroti peluncuran *jersey* terbaru yang dilakukan pada 18 Maret 2024 oleh Ersopo, yang memicu kontroversi di kalangan penggemar. Dengan menggunakan platform Twitter, salah satu media sosial yang paling populer di Indonesia.

Analisis sentimen merupakan proses mengolah data dari berbagai sumber, seperti internet dan juga media sosial, untuk mengidentifikasi pandangan seseorang terhadap isu tertentu, seperti kepuasan layanan atau kebijakan. Proses ini mengubah informasi tidak terstruktur menjadi data terstruktur yang lebih mudah dianalisis[2]. Dengan kata lain, analisis ini bertujuan untuk memahami bagaimana perasaan atau pandangan seseorang terhadap suatu topik yang diekspresikan melalui kata-kata dalam bentuk tulisan. Dari hasil analisis sentimen ini memberikan gambaran persentase mengenai sentimen positif maupun negatif terhadap subjek tertentu. Dalam penerapannya, analisis ini bermanfaat sebagai evaluasi kinerja produk di pasar hingga reputasi suatu entitas di mata publik[3].

Selain itu dengan banyaknya penggunaan bahasa informal dalam ulasan pelanggan, penggunaan kata-kata slang, singkatan, dan ejaan yang tidak sesuai dengan kaidah baku sering kali ditemukan, sehingga menyulitkan algoritma untuk mengklasifikasikan sentimen secara akurat. Kondisi ini menciptakan peluang bagi penelitian lebih lanjut untuk mengembangkan metode *Naïve Bayes* yang lebih efektif dalam mengolah teks berbahasa Indonesia dengan struktur yang lebih kompleks.

Penelitian terdahulu yang berkaitan dengan analisis sentimen yang dilakukan oleh[4]. Diantaranya adalah penggunaan *Naïve Bayes* untuk menganalisis pendapat masyarakat terhadap vaksinasi covid-19 di Indonesia, pada penelitian tersebut data yang digunakan berjumlah 2278 data. Dari penelitian tersebut menyatakan bahwa penggunaan algoritma *Naïve Bayes* sudah sangat baik dengan hasil akurasi 93%. Penelitian pernah dilakukan oleh Nur Amalia dkk., 2023 dengan menggunakan metode *Naïve Bayes* dan *Decision Tree* dalam menganalisis sentimen pengguna *twitter* terhadap pelaksanaan kurikulum MBKM. Data yang digunakan berjumlah 1275 data *tweet*, Dari penelitian tersebut *Naïve Bayes* memperoleh hasil akurasi yang lebih tinggi dibandingkan dengan *Decision Tree*.

Akurasi *Naïve Bayes* sebesar 81,15%. Sedangkan akurasi yang diperoleh *Decision Tree* sebesar 68,19%[5]. Penelitian selanjutnya tentang analisis sentimen terhadap layanan aplikasi Grab yang dilakukan oleh[6]. Penelitian tersebut menggunakan metode *Naïve Bayes*. Data yang digunakan didapatkan melalui proses *crawling* yaitu sebanyak 2000 data. Dari penelitian tersebut, metode *Naïve Bayes* memperoleh nilai akurasi dengan tingkat yang baik sebesar 84,36%.

Penelitian ini bertujuan untuk menganalisis opini dan pandangan masyarakat terhadap *jersey* baru timnas Indonesia dengan menggunakan analisis sentimen berbasis model klasifikasi *Naive Bayes* pada platform *Twitter*. Metode *Naive Bayes* dipilih karena memiliki sejumlah keunggulan, seperti kesederhanaan, kecepatan, dan tingkat akurasi yang tinggi[7]. Dalam penelitian ini, *twitter* dimanfaatkan sebagai sumber utama pengumpulan data. Data yang diperoleh akan diklasifikasikan ke dalam kategori *positif* dan *negatif*. Pengumpulan data dilakukan melalui teknik *crawling* untuk pengambilan data secara *real-time* dari *Twitter*.

1.2. Rumusan Masalah

1.2.1. Permasalahan

Berdasarkan latar belakang yang telah dijelaskan, pada awal peluncuran *jersey* timnas Indonesia, momen tersebut langsung menjadi viral dan trending di salah satu platform media sosial *Twitter*. Namun, perhatian yang diperoleh justru mengarah ke komentar negatif, dengan banyak pengguna yang menyoroti berbagai aspek dari *jersey* tersebut. Isu ini memicu diskusi hangat, yang sebagian besar berisi kritik terhadap desain, material, atau konsep yang dianggap kurang memenuhi ekspektasi.

1.2.2. Metode Usulan

Berdasarkan permasalahan tersebut maka dilakukan penelitian mengenai analisis sentimen kepuasan pelanggan terhadap produk *jersey* timnas Indonesia menggunakan metode *Naive Bayes*

1.2.3. Pertanyaan Penelitian

1. Berapa nilai akurasi, *precision*, *recall* dan *f1-score* yang dihasilkan metode *Naive Bayes* untuk klasifikasi sentimen positif dan negatif pada data ulasan *Twitter* mengenai produk jersey timnas Indonesia?

1.3. Tujuan dan Manfaat Penelitian

1.3.1. Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut :

1. Menerapkan metode *Naive Bayes* untuk mengklasifikasikan sentimen ulasan pelanggan ke dalam kategori positif dan negatif.
2. Menganalisis hasil metode *Naive Bayes* dalam analisis sentimen pada ulasan pelanggan, termasuk nilai akurasi, *precision*, *recall* dan *f1-score*.

1.3.2. Manfaat Penelitian

Manfaat penelitian ini adalah sebagai berikut :

1. Manfaat Akademis

- Menambah wawasan dalam bidang analisis sentimen, khususnya yang menggunakan metode *Naive Bayes*.
- Menjadi referensi bagi penelitian selanjutnya yang terkait dengan analisis teks dan pengolahan bahasa alami (*Natural Language Processing*).

2. Manfaat Praktis

- Membantu perusahaan dalam memahami kepuasan pelanggan melalui analisis sentimen ulasan pelanggan.
- Memberikan informasi yang relevan bagi perusahaan untuk meningkatkan kualitas produk atau layanan sesuai dengan kebutuhan pelanggan.

1.4. Batasan Masalah

Agar penelitian lebih terarah, terdapat beberapa batasan masalah dalam penelitian ini, yaitu :

1. Data ulasan yang digunakan hanya diambil dari satu platform yaitu *Twitter*, data diperoleh pada tanggal 09 April 2024 hingga tanggal 01 November 2024 sebanyak 1500 data.
2. Bahasa yang dianalisis dalam ulasan adalah bahasa Indonesia.
3. Metode analisis yang digunakan hanya sebatas pada *Naïve Bayes*, tanpa membandingkan dengan metode lain.
4. Analisis hanya fokus pada klasifikasi sentimen menjadi dua kategori utama: positif dan negatif.

1.5. Sistematika Proposal

Sistematika proposal ini disusun dengan sistematika sebagai berikut:

BAB I: PENDAHULUAN

Bab ini berisi latar belakang masalah, rumusan masalah, permasalahan, metode usulan, pertanyaan penelitian, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika proposal.

BAB II: TINJAUAN PUSTAKA

Berisi kajian teori mengenai analisis sentimen, metode *Naïve Bayes*, pengolahan bahasa alami, serta penelitian-penelitian terkait.

BAB III: METODOLOGI PENELITIAN

Menjelaskan metode penelitian yang digunakan, termasuk teknik pengumpulan data, perancangan alur *sentiment analysis*, tahapan penelitian, arsitektur sistem, data penelitian, *pre-processing data*, dijelaskan tentang TF-IDF, tahapan klasifikasi data menggunakan metode *Naïve Bayes*, visualisasi dan evaluasi model.

BAB IV: HASIL DAN PEMBAHASAN

Bab ini membahas tentang implementasi sistem yang dilakukan berdasarkan

rancangan yang telah disusun pada bab sebelumnya. Dalam bab ini, dijelaskan berbagai tahapan, termasuk skenario pengujian, hasil dari proses *crawling* data pada *Twitter*, hasil pelabelan manual, serta hasil dari tahap *preprocessing* data yang mencakup *case folding*, *cleaning*, *tokenizing*, *filtering*, dan *stemming*. Hasil pembobotan menggunakan TF-IDF, hasil klasifikasi menggunakan *Naive Bayes*, serta evaluasi model menggunakan *confusion matrix*. Terakhir hasil dari uji skenario.

BAB V: KESIMPULAN DAN PENUTUP

Bab ini memaparkan kesimpulan dari hasil penelitian yang mencakup visualisasi kata, tingkat akurasi, serta nilai precision, recall, dan F1-score yang diperoleh. Selain itu, bab ini juga memberikan saran untuk penelitian selanjutnya yang akan membahas topik serupa, yaitu analisis sentimen.