

# BAB I PENDAHULUAN

## 1.1 Latar Belakang

Perkembangan teknologi informasi dan komunikasi (TIK) yang pesat dalam dua dekade terakhir telah membawa transformasi besar dalam berbagai aspek kehidupan manusia. Salah satu sektor yang mengalami perubahan signifikan adalah perdagangan, terutama melalui kehadiran platform *e-commerce*. Digitalisasi dalam perdagangan memungkinkan konsumen untuk melakukan transaksi dengan lebih mudah, cepat, dan fleksibel tanpa terikat oleh batasan geografis dan waktu. Indonesia, sebagai negara dengan populasi terbesar di Asia Tenggara, telah menjadi pasar yang menjanjikan bagi perkembangan *e-commerce*, yang ditunjukkan oleh tingginya penetrasi internet dan penggunaan perangkat mobile yang semakin meluas (Badan Pusat Statistik, 2023).

Di antara berbagai platform *e-commerce* yang ada, Shopee menempati posisi penting sebagai salah satu pemain utama dalam ekosistem digital di Indonesia. Shopee tidak hanya menyediakan kemudahan dalam proses belanja daring, akan tetapi juga menawarkan berbagai fitur interaktif yang mendorong keterlibatan pengguna secara aktif. Salah satu fitur penting adalah sistem ulasan dan penilaian produk oleh pembeli. Fitur ini memberikan kesempatan bagi pengguna untuk berbagi pengalaman mereka dalam menggunakan produk atau layanan, yang pada gilirannya menjadi referensi penting bagi calon pembeli lainnya dalam membuat keputusan pembelian. Dengan demikian, ulasan pelanggan telah menjadi komponen krusial dalam proses bisnis *e-commerce*, baik dari sudut pandang konsumen maupun pelaku usaha.

Namun, seiring dengan meningkatnya jumlah transaksi dan aktivitas pengguna di Shopee, volume data ulasan aplikasi di *Google Play Store* juga mengalami lonjakan yang signifikan setiap harinya. Ulasan-ulasan ini sangat bervariasi dalam gaya bahasa, panjang teks, dan konteks penyampaian. Hal ini menciptakan tantangan besar dalam pengelolaan data tersebut agar dapat diolah menjadi informasi yang bermakna dan dimanfaatkan secara optimal oleh pihak-pihak terkait, seperti pengelola platform, pengembang aplikasi, maupun peneliti. Oleh karena itu, dibutuhkan pendekatan yang mampu mengolah data teks dalam skala besar secara efisien dan akurat. Salah satu metode yang umum digunakan dalam konteks ini adalah analisis sentimen, yang bertujuan untuk mengidentifikasi opini atau persepsi pengguna terhadap aplikasi melalui ulasan yang mereka berikan.

Analisis sentimen adalah proses otomatis dalam pemrosesan bahasa alami yang bertujuan untuk mendeteksi dan mengklasifikasikan opini atau emosi pengguna seperti positif, negatif, atau netral terhadap suatu entitas berdasarkan ulasan yang mereka unggah (Kumar et al., 2023). Hasil analisis sentimen dapat memberikan wawasan mendalam mengenai kepuasan pelanggan, persepsi merek, serta isu-isu yang sering muncul dalam interaksi pengguna dengan produk atau layanan tertentu. Berdasarkan informasi tersebut, pengelola platform dapat melakukan evaluasi dan perbaikan sistem secara berkelanjutan, sementara penjual dapat merumuskan strategi pemasaran dan peningkatan kualitas produk yang lebih tepat sasaran.

Salah satu metode yang populer dalam implementasi analisis sentimen adalah algoritma *Naïve Bayes*, yang merupakan algoritma klasifikasi berbasis probabilitas. Algoritma ini dikenal karena kesederhanaan model, efisiensi dalam komputasi, dan

performa yang cukup baik dalam pengolahan teks skala besar. Meskipun algoritma ini mengasumsikan bahwa fitur-fitur dalam dokumen bersifat independen satu sama lain asumsi yang jarang terpenuhi secara sempurna dalam data nyata berbagai penelitian menunjukkan bahwa *Naïve Bayes* tetap mampu memberikan hasil klasifikasi yang kompetitif, terutama dalam domain analisis teks.

Penelitian oleh (Agustina et al., 2022) menunjukkan keberhasilan penerapan algoritma *Naïve Bayes* dalam analisis sentimen pada ulasan pengguna Shopee, di mana algoritma ini mampu mengklasifikasikan sentimen dengan akurasi yang cukup tinggi. Meskipun demikian, penelitian tersebut menekankan pentingnya tahapan preprocessing yang cermat, seperti tokenisasi, stemming, dan penghapusan kata umum, serta pemilihan fitur yang relevan dan proses pengolahan data yang optimal untuk meningkatkan kinerja klasifikasi.

Penelitian yang dilakukan oleh (Rahel Lina Simanjuntak et al., 2023) juga membandingkan performa *Naïve Bayes* dengan algoritma lain, dan hasilnya menunjukkan bahwa *Naïve Bayes* tetap menjadi pilihan utama karena keunggulannya dalam efisiensi dan kemudahan implementasi. Selain itu, cakupan analisis sentimen tidak terbatas pada ulasan produk saja. (Wandani, 2021) memperluas penerapan analisis sentimen pada data media sosial khususnya Twitter, selama *event flash sale* yang diadakan oleh Shopee. Mereka menunjukkan bahwa analisis sentimen mampu menangkap persepsi publik secara real-time, yang sangat berguna bagi manajemen dalam merancang strategi promosi dan penyesuaian layanan.

Sementara itu, penelitian yang dilakukan oleh (Rahel Lina Simanjuntak et al., 2023) mengembangkan metode query expansion ranking dalam klasifikasi

multinomial *Naïve Bayes* pada data Shopee saat Hari Belanja Online Nasional. Inovasi tersebut menunjukkan pentingnya pengembangan teknik-teknik lanjutan dalam meningkatkan akurasi dan relevansi hasil klasifikasi sentimen, terutama dalam konteks event besar yang menghasilkan data dalam jumlah sangat besar dan kompleks.

Penelitian yang dilakukan oleh (Siniwi et al., 2021) juga memberikan kontribusi penting dalam konteks ini. Mereka menegaskan bahwa hasil analisis sentimen dapat dijadikan umpan balik yang konstruktif bagi para penjual dalam meningkatkan kualitas produk dan pelayanan mereka. Selain itu, penelitian oleh (Sihombing et al., 2021) serta (Darmawan et al., 2020) juga memperlihatkan manfaat langsung dari analisis sentimen bagi toko-toko online, terutama dalam memahami dan merespons harapan pelanggan. Sementara itu (Rani & Candra, 2023) menekankan bahwa hasil analisis sentimen dapat menjadi indikator penting dalam menilai kualitas produk maupun pelayanan yang diberikan oleh penjual, serta menjadi bahan evaluasi yang berkelanjutan untuk menjaga kualitas ekosistem *e-commerce* secara keseluruhan.

Dari berbagai hasil penelitian tersebut, jelas bahwa penerapan analisis sentimen dengan algoritma *Naïve Bayes* dalam konteks *e-commerce*, khususnya platform Shopee, merupakan topik yang sangat relevan dan strategis. Pada satu sisi, kondisi ideal yang ingin dicapai adalah sistem otomatis yang mampu secara akurat mengklasifikasikan sentimen pengguna secara *real-time*. Namun, di sisi lain, kondisi faktual menunjukkan adanya tantangan-tantangan kompleks seperti kualitas data

ulasan yang tidak seragam, kebutuhan akan *preprocessing* yang tepat, serta kebutuhan untuk mengembangkan metode yang adaptif terhadap dinamika data.

Oleh karena itu, penelitian ini menjadi penting karena bertujuan tidak hanya untuk mengimplementasikan algoritma *Naïve Bayes* dalam klasifikasi sentimen ulasan pengguna Shopee, tetapi juga untuk mengkaji efektivitasnya dalam konteks data ulasan yang kompleks. Kebaruan dalam penelitian ini terletak pada pendekatan *pre-processing* yang ditingkatkan serta analisis performa klasifikasi berdasarkan variasi jenis ulasan dan konteks penggunaannya. Berdasarkan uraian di atas, penulis tertarik untuk mengangkat sebuah penelitian dengan judul “Analisis Sentimen pada Platform Shopee menggunakan Algoritma *Naïve Bayes*”.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah:

1. Bagaimana proses *pre-processing* data dilakukan terhadap ulasan pengguna Shopee di *Google Play Store* sebelum diterapkan algoritma *Naïve Bayes*
2. Bagaimana kinerja algoritma *Naïve Bayes* dalam mengklasifikasikan sentimen ulasan pengguna Shopee di *Google Play Store* ke dalam kategori positif, negatif, dan netral
3. Bagaimana hasil klasifikasi sentimen tersebut dapat menggambarkan persepsi dan tingkat kepuasan pengguna terhadap aplikasi Shopee di *Google Play Store*.

### 1.3 Batasan Masalah

Agar penelitian ini lebih terfokus pada pembahasan dan tetap terarah serta menghindari pembahasan yang terlalu luas, maka dalam penelitian ini ditetapkan beberapa batasan masalah sebagai berikut:

1. Data yang digunakan dalam penelitian ini hanya berupa ulasan pengguna aplikasi Shopee yang diperoleh dari *Google Play Store*, tidak mencakup data dari platform lain seperti media sosial atau situs *e-commerce* lainnya.
2. Analisis sentimen yang dilakukan dibatasi pada tiga kategori, yaitu positif, negatif, dan netral, tanpa mempertimbangkan tingkat intensitas atau emosi yang lebih kompleks.
3. Metode yang digunakan dalam analisis sentimen hanya difokuskan pada algoritma *Naïve Bayes*, tanpa melakukan perbandingan dengan algoritma machine learning lainnya.
4. Data yang diambil hanya pada periode 1 Agustus – 9 Agustus 2025 untuk mendapatkan representasi ulasan terbaru dari pengguna aplikasi Shopee di *Google Play Store*.

### 1.4 Tujuan Penelitian

Penelitian ini memiliki beberapa tujuan utama yang hendak dicapai sebagai bentuk kontribusi terhadap pemahaman dan penerapan analisis sentimen pada data ulasan pengguna aplikasi Shopee. Tujuan-tujuan tersebut antara lain:

1. Menganalisis dan menjelaskan proses pre-processing data yang dilakukan terhadap ulasan pengguna Shopee di *Google Play Store* sebelum diterapkan algoritma *Naïve Bayes*.
2. Mengevaluasi kinerja algoritma *Naïve Bayes* dalam mengklasifikasikan sentimen ulasan pengguna Shopee di *Google Play Store* ke dalam kategori positif, negatif, dan netral.
3. Menginterpretasikan hasil klasifikasi sentimen untuk menggambarkan persepsi dan tingkat kepuasan pengguna terhadap aplikasi Shopee di *Google Play Store*.

### **1.5 Manfaat Penelitian**

Berdasarkan latar belakang yang telah di jabarkan diatas maka manfaat dari penelitian yang dilakukan adalah:

#### **1. Manfaat Teoritis**

Penelitian ini diharapkan memberikan kontribusi bagi pengembangan ilmu komputer, khususnya pada bidang analisis sentimen dan penerapan algoritma machine learning, melalui:

1. Penyajian deskripsi menyeluruh mengenai karakteristik data ulasan pengguna pada aplikasi *e-commerce*.
2. Penjabaran proses prapengolahan data teks sebelum dilakukan analisis sentimen.
3. Evaluasi kinerja algoritma *Naïve Bayes* dalam mengklasifikasikan sentimen pada ulasan pengguna.

4. Penyediaan referensi empiris yang dapat digunakan dalam penelitian lanjutan terkait integrasi teknik data mining dan pengolahan bahasa alami.
5. Pemberian kontribusi dalam memperluas kajian teoritis tentang penerapan kecerdasan buatan dalam memahami persepsi pengguna secara otomatis.

## **2. Manfaat Praktis**

Hasil dari penelitian ini juga memiliki manfaat praktis yang dapat dirasakan oleh berbagai pihak, antara lain: .

1. Menjadi media penerapan ilmu yang telah diperoleh penulis selama masa studi serta memenuhi salah satu syarat penyelesaian tugas akhir jenjang S1.
2. Memberikan dasar analisis bagi pengelola aplikasi Shopee dalam memahami opini dan kepuasan pengguna.
3. Menjadi bahan pertimbangan dalam pengambilan keputusan strategis untuk meningkatkan kualitas produk dan layanan.
4. Memberikan wawasan berbasis data kepada pelaku bisnis *e-commerce* dalam merancang strategi bisnis yang lebih responsif terhadap kebutuhan pasar.
5. Menyediakan panduan awal bagi praktisi dalam mengembangkan solusi teknologi berbasis analisis sentimen secara efisien.