

BAB II

LANDASAN TEORI

2.1. Analisis Sentimen

Analisis sentimen merupakan teknik pengolahan data teks yang bertujuan untuk mengidentifikasi, mengklasifikasikan, serta memahami opini atau perasaan yang terkandung dalam suatu teks [1]. Proses ini biasanya digunakan untuk mengetahui apakah suatu ulasan, komentar, atau opini bersifat positif, negatif, atau netral. Dalam perkembangannya, analisis sentimen banyak dimanfaatkan pada berbagai bidang seperti pemasaran digital, evaluasi layanan, serta pemantauan opini publik di media sosial [2]. Dengan mengubah data teks yang bersifat subjektif menjadi informasi terstruktur, analisis sentimen memberikan peluang besar bagi organisasi maupun perusahaan dalam memahami persepsi masyarakat secara lebih sistematis [3].

Analisis sentimen sangat relevan untuk diterapkan pada data ulasan pengguna aplikasi Grab di Play Store yang jumlahnya terus bertambah setiap hari [4]. Ulasan tersebut memuat berbagai persepsi pengguna, mulai dari kepuasan terhadap fitur hingga keluhan atas layanan yang kurang optimal [5]. Dengan melakukan klasifikasi menggunakan metode *Naive Bayes*, ulasan dapat dikelompokkan ke dalam kategori positif, negatif, atau netral sehingga menghasilkan gambaran yang lebih jelas mengenai kecenderungan opini masyarakat [6]. Informasi yang diperoleh dari hasil klasifikasi ini dapat digunakan sebagai dasar evaluasi dalam peningkatan kualitas layanan Grab sekaligus memperlihatkan efektivitas metode *Naive Bayes* dalam menganalisis opini pengguna aplikasi transportasi online [7].

2.2. Aplikasi Grab

Grab adalah salah satu aplikasi berbasis layanan transportasi online yang berkembang pesat di Asia Tenggara, termasuk Indonesia. Aplikasi ini menyediakan berbagai layanan mulai dari transportasi roda dua (*GrabBike*) dan roda empat (*GrabCar*), layanan pesan-antar makanan (*GrabFood*), pengiriman barang (*GrabExpress*), hingga layanan pembayaran digital melalui *OVO*. Dengan keberagaman fitur yang ditawarkan, Grab tidak hanya menjadi sarana transportasi, tetapi juga bagian dari ekosistem digital yang memudahkan aktivitas sehari-hari masyarakat. Popularitasnya membuat aplikasi ini memperoleh jutaan unduhan di Play Store serta menjadi salah satu pilihan utama bagi pengguna layanan digital di Indonesia.

Ulasan pengguna di Play Store memberikan gambaran nyata mengenai kepuasan maupun keluhan terhadap layanan Grab, sehingga opini tersebut dapat diolah menjadi informasi yang bermanfaat. Melalui analisis sentimen, opini tersebut dapat dipetakan menjadi kategori positif, negatif, atau netral untuk mengetahui kecenderungan persepsi masyarakat. Penerapan metode klasifikasi seperti *Naive Bayes* mampu membantu mengidentifikasi pola sentimen secara sistematis, sehingga ulasan yang semula hanya berupa teks bebas dapat menjadi sumber data yang terstruktur. Dengan cara ini, perusahaan dapat memperoleh wawasan yang lebih mendalam tentang kualitas layanan berdasarkan suara langsung dari para penggunanya.

2.3. Aplikasi PlayStore

Google Play Store merupakan platform distribusi digital resmi yang dikembangkan oleh Google untuk perangkat berbasis Android. Melalui Play Store,

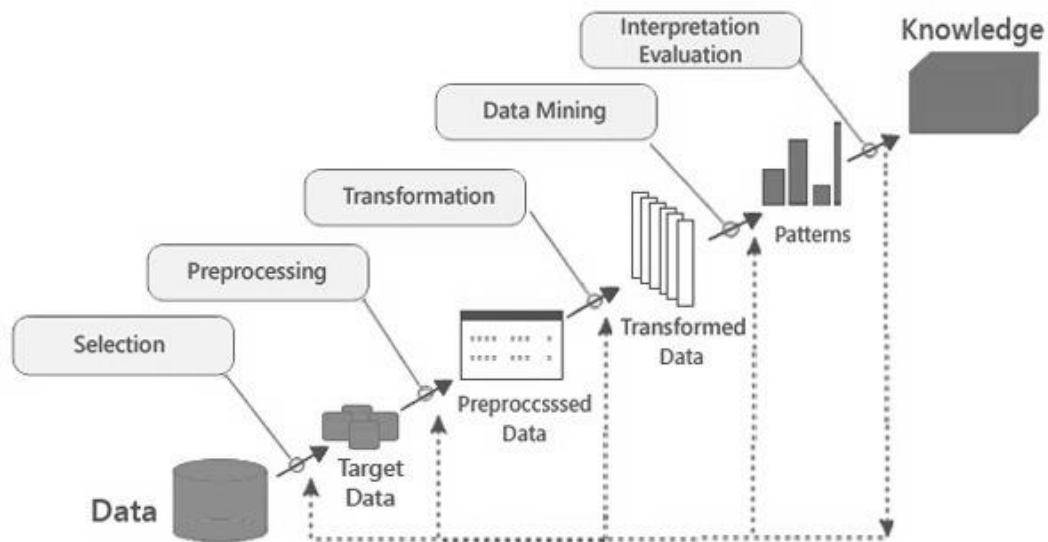
pengguna dapat mengunduh berbagai aplikasi, game, buku, film, hingga musik secara mudah dan terpusat. Selain menjadi tempat distribusi aplikasi, Play Store juga menyediakan sistem penilaian berupa rating dan ulasan pengguna. Rating digunakan untuk memberikan skor terhadap kualitas aplikasi, sedangkan ulasan berfungsi sebagai sarana berbagi pengalaman, baik positif maupun negatif. Fitur ulasan ini menjadi salah satu indikator penting dalam menilai kualitas aplikasi sekaligus membantu pengguna lain dalam menentukan pilihan sebelum mengunduh.

Ulasan yang terdapat di Play Store memiliki nilai penting karena merepresentasikan pengalaman nyata pengguna terhadap suatu aplikasi, termasuk aplikasi Grab yang menjadi salah satu layanan transportasi online populer. Data ulasan tersebut mencerminkan persepsi masyarakat yang beragam, mulai dari kepuasan hingga keluhan terhadap layanan. Informasi yang terkandung di dalamnya sangat relevan untuk dikaji lebih lanjut melalui analisis sentimen dengan tujuan mengungkap kecenderungan opini publik. Dengan demikian, Play Store bukan hanya berfungsi sebagai tempat distribusi aplikasi, tetapi juga sebagai sumber data yang kaya untuk memahami pengalaman pengguna terhadap layanan digital.

2.4. Knowledge Discovery in Database

Knowledge Discovery in Database (KDD) merupakan proses identifikasi pola, informasi, dan pengetahuan yang berguna dari kumpulan data yang besar dan kompleks [8]. Proses KDD melibatkan beberapa tahapan, mulai dari pemilihan data, pembersihan data, transformasi, hingga penerapan teknik analisis seperti data

mining untuk menemukan pola tersembunyi [9]. Tujuan utama KDD adalah mengubah data mentah menjadi informasi yang bermakna dan dapat digunakan untuk pengambilan keputusan, prediksi, maupun evaluasi performa suatu sistem atau layanan [10].



Gambar 2. 1. Knowledge Discovery in Database

Dalam konteks penelitian ini, prinsip KDD dapat diterapkan untuk memahami pola opini pengguna aplikasi Grab melalui ulasan di Play Store [10]. Dengan pendekatan klasifikasi sentimen menggunakan metode *Naive Bayes*, data ulasan yang awalnya tidak terstruktur dapat diolah menjadi informasi yang jelas tentang kecenderungan sentimen, baik positif, negatif, maupun netral. Proses ini mencerminkan esensi KDD, yakni mengubah data teks mentah menjadi pengetahuan yang bermanfaat bagi pengambilan keputusan terkait peningkatan layanan aplikasi.

2.4.1. Data Mining

Data mining adalah proses pengolahan dan analisis data dalam jumlah besar untuk menemukan pola, tren, dan informasi yang berguna secara otomatis atau

semi-otomatis [11]. Proses ini memanfaatkan teknik statistik, machine learning, dan algoritma tertentu untuk mengekstrak pengetahuan dari data mentah yang tersebar. Data mining sering diterapkan dalam berbagai bidang seperti bisnis, kesehatan, keuangan, dan teknologi informasi untuk mendukung pengambilan keputusan yang lebih tepat dan berbasis data, termasuk dalam pengelolaan ulasan, transaksi, dan perilaku pengguna [12].

Dalam konteks penelitian ini, data mining berperan sebagai dasar pemahaman bagaimana ulasan pengguna Grab di Play Store dapat diolah menjadi informasi yang terstruktur [13]. Melalui prinsip-prinsip data mining, data teks dari ulasan dapat dianalisis untuk mengetahui kecenderungan opini pengguna, sehingga klasifikasi sentimen menggunakan *Naive Bayes* dapat diterapkan secara efektif [14]. Dengan demikian, data mining membantu mengidentifikasi pola sentimen positif, negatif, maupun netral yang muncul dari ulasan pengguna, sehingga memberikan gambaran yang jelas mengenai persepsi masyarakat terhadap layanan Grab.

2.4.2. Database dan Data Processing

Database merupakan kumpulan data yang tersimpan secara terstruktur dan sistematis sehingga memudahkan akses, pengelolaan, dan pemrosesan informasi. Dalam konteks analisis sentimen, database berfungsi untuk menampung ulasan atau komentar pengguna yang akan dianalisis. Data yang tersimpan biasanya mencakup informasi teks ulasan, tanggal, rating, dan metadata lain yang relevan. Proses *data processing* mencakup pengorganisasian, pembersihan, dan transformasi data agar siap digunakan dalam analisis. Tahapan ini penting untuk memastikan data yang

dianalisis akurat, konsisten, dan mudah diproses oleh algoritma klasifikasi seperti *Naive Bayes*.

Dalam konteks penelitian ini, database memuat ulasan pengguna aplikasi Grab yang diambil dari Play Store, yang menjadi sumber utama untuk analisis sentimen. Data processing dilakukan untuk menyiapkan ulasan agar dapat diklasifikasikan menjadi kategori positif, negatif, atau netral. Dengan adanya pengolahan data yang sistematis, metode *Naive Bayes* dapat bekerja lebih efektif dalam menilai kecenderungan opini pengguna. Hal ini memungkinkan penelitian untuk menghasilkan informasi yang merefleksikan persepsi pengguna terhadap layanan Grab secara lebih akurat dan terstruktur.

2.4.3. Visualization

Visualisasi merupakan teknik untuk menyajikan data dalam bentuk grafis atau gambar sehingga informasi dapat dipahami dengan lebih cepat dan jelas. Dengan bantuan visualisasi, pola, tren, atau distribusi data menjadi lebih mudah dikenali dibandingkan hanya melihat data mentah. Berbagai jenis visualisasi dapat digunakan, seperti diagram batang, diagram lingkaran, atau grafik garis, yang masing-masing memiliki kegunaan tersendiri tergantung sifat data dan tujuan analisis. Selain itu, visualisasi juga membantu dalam mengkomunikasikan hasil analisis kepada pihak yang bukan ahli data, sehingga informasi yang kompleks dapat diterima dengan lebih mudah.

Dalam konteks penelitian analisis sentimen pada layanan aplikasi Grab di Play Store, visualisasi dapat digunakan untuk menampilkan proporsi ulasan positif, negatif, dan netral secara jelas. Dengan menyajikan hasil klasifikasi dalam bentuk

grafik atau diagram, tren persepsi pengguna terhadap layanan Grab dapat segera terlihat, misalnya apakah mayoritas pengguna puas atau tidak. Hal ini memudahkan pemangku kepentingan, seperti pengembang aplikasi atau analis, untuk memahami kecenderungan opini masyarakat dengan cepat dan mengambil keputusan yang tepat berdasarkan data yang terstruktur.

2.4.4. Statistik

Statistik merupakan cabang ilmu yang mempelajari pengumpulan, pengolahan, analisis, dan interpretasi data untuk mendapatkan informasi yang berguna. Dalam praktiknya, statistik digunakan untuk menggambarkan pola, tren, maupun hubungan antarvariabel dalam bentuk angka, tabel, atau grafik. Dengan teknik statistik, data mentah yang bersifat acak dan kompleks dapat diubah menjadi informasi yang lebih terstruktur dan mudah dipahami. Selain itu, statistik juga memungkinkan pengambilan keputusan yang lebih objektif karena didasarkan pada data yang terukur dan terverifikasi, bukan sekadar asumsi atau opini subjektif.

Dalam konteks penelitian analisis sentimen pada ulasan aplikasi Grab, statistik berperan penting dalam menggambarkan distribusi sentimen pengguna, seperti proporsi ulasan positif, negatif, dan netral. Statistik membantu menilai pola opini masyarakat terhadap layanan aplikasi, misalnya melihat kategori sentimen mana yang mendominasi, serta memberikan gambaran kuantitatif tentang kepuasan atau ketidakpuasan pengguna. Dengan pendekatan statistik, hasil klasifikasi menggunakan metode *Naive Bayes* dapat dianalisis secara objektif sehingga informasi yang diperoleh lebih mudah dipahami dan dijadikan dasar evaluasi layanan.

2.4.5. Pattern Recognition

Pattern recognition atau pengenalan pola merupakan cabang dari ilmu komputer dan kecerdasan buatan yang memfokuskan pada identifikasi pola atau karakteristik tertentu dalam data. Proses ini melibatkan pengelompokan, klasifikasi, dan pemodelan data berdasarkan ciri-ciri yang terlihat atau terukur. Metode pattern recognition banyak diterapkan pada berbagai bidang, mulai dari pengolahan citra, pengenalan suara, hingga analisis teks, dengan tujuan utama untuk mengekstraksi informasi penting dari data yang kompleks dan menghasilkan keputusan yang akurat berdasarkan pola yang ditemukan.

Dalam konteks penelitian ini, konsep pattern recognition menjadi relevan karena analisis sentimen pada ulasan pengguna Grab pada dasarnya merupakan proses pengenalan pola dalam data teks. Setiap ulasan dapat dianggap sebagai pola yang mengandung informasi subjektif, baik berupa opini positif, negatif, maupun netral. Dengan pendekatan klasifikasi, metode seperti *Naive Bayes* berfungsi sebagai alat untuk mengenali pola-pola tersebut dan mengelompokkan ulasan ke dalam kategori yang sesuai, sehingga memudahkan pemahaman terhadap kecenderungan persepsi pengguna terhadap layanan aplikasi.

2.5. Metode *Naïve Bayes*

Metode *Naive Bayes* adalah salah satu algoritma klasifikasi dalam machine learning yang didasarkan pada Teorema Bayes dengan asumsi independensi antar fitur. Algoritma ini bekerja dengan menghitung probabilitas suatu data termasuk ke dalam kelas tertentu berdasarkan frekuensi kemunculan kata atau fitur dalam data pelatihan. Meskipun sederhana, *Naive Bayes* terbukti efektif untuk tugas klasifikasi teks seperti spam detection, analisis sentimen, dan rekomendasi, karena

kemampuannya dalam memproses data dalam jumlah besar dengan cepat dan menghasilkan prediksi yang relatif akurat.

Dalam konteks penelitian ini, *Naive Bayes* digunakan untuk mengklasifikasikan ulasan pengguna aplikasi Grab di Play Store ke dalam kategori sentimen positif, negatif, dan netral. Dengan pendekatan probabilistik, algoritma ini memungkinkan pengolahan opini pengguna yang beragam menjadi informasi yang lebih terstruktur. Hasil klasifikasi membantu menggambarkan pola persepsi masyarakat terhadap layanan Grab, sehingga pihak pengembang dapat memahami kecenderungan opini secara umum dan mengambil keputusan berbasis data mengenai evaluasi kualitas layanan.

2.6. Model Klasifikasi

Metode model klasifikasi merupakan salah satu pendekatan dalam *machine learning* yang bertujuan untuk mengelompokkan data ke dalam kategori atau kelas tertentu berdasarkan pola yang dimiliki [15]. Dalam konteks analisis teks, model klasifikasi memanfaatkan fitur-fitur dari data teks untuk menentukan kategori, misalnya positif, negatif, atau netral [16]. Salah satu algoritma populer yang sering digunakan adalah *Naive Bayes*, yang bekerja berdasarkan prinsip probabilitas dan asumsi independensi antarfitur. Metode ini relatif sederhana, cepat, dan efektif untuk data teks yang bersifat besar dan bervariasi, sehingga sering diaplikasikan dalam analisis sentimen, spam detection, maupun sistem rekomendasi [17].

Dalam konteks penelitian ini, model klasifikasi digunakan untuk menganalisis ulasan pengguna aplikasi Grab di Play Store [18]. Data ulasan yang beragam dikategorikan ke dalam sentimen positif, negatif, atau netral menggunakan

pendekatan probabilistik dari *Naive Bayes*. Dengan cara ini, pola opini pengguna dapat dipetakan secara sistematis, sehingga memudahkan pemahaman terhadap persepsi masyarakat mengenai kualitas layanan Grab. Model klasifikasi menjadi alat utama dalam penelitian ini untuk menyajikan informasi yang lebih terstruktur dan terukur dari sekumpulan ulasan yang sebelumnya bersifat subjektif dan tersebar.

2.7. Alat Bantu Program Aplikasi Orange

Orange adalah sebuah aplikasi open-source berbasis *data mining* dan *machine learning* yang menyediakan antarmuka visual untuk analisis data tanpa memerlukan banyak penulisan kode. Aplikasi ini dilengkapi dengan berbagai modul untuk pra-pemrosesan data, klasifikasi, regresi, klastering, dan visualisasi hasil analisis. Salah satu keunggulan Orange adalah kemampuannya menampilkan alur proses analisis dalam bentuk *workflow* yang intuitif, sehingga pengguna dapat menghubungkan berbagai komponen, seperti input data, algoritma klasifikasi, dan evaluasi model, secara langsung. Selain itu, Orange mendukung integrasi dengan skrip Python, memungkinkan fleksibilitas bagi pengguna yang ingin menyesuaikan analisis lebih lanjut.

Dalam konteks penelitian ini, Orange dapat digunakan sebagai alat bantu untuk memvisualisasikan dan mengklasifikasikan data ulasan pengguna aplikasi Grab. Dengan antarmuka *drag-and-drop*, proses penerapan metode *Naive Bayes* untuk menganalisis sentimen menjadi lebih mudah dan terstruktur. Peneliti dapat memasukkan dataset ulasan, menentukan kolom atribut yang relevan, dan langsung memperoleh hasil klasifikasi serta evaluasi akurasi model. Hal ini membuat Orange

menjadi sarana praktis untuk memahami kecenderungan opini pengguna, sekaligus mempermudah interpretasi hasil analisis secara visual bagi pihak yang tidak terbiasa dengan pemrograman kompleks.

2.8. Kelebihan dan Kekurangan Naive Bayes

Metode *Naive Bayes* adalah algoritma klasifikasi berbasis probabilitas yang sederhana namun efektif, yang bekerja dengan asumsi bahwa setiap fitur pada data bersifat independen satu sama lain [19]. Salah satu kelebihan utamanya adalah kemampuannya untuk menangani dataset besar dengan cepat serta memberikan hasil klasifikasi yang cukup akurat meskipun modelnya sederhana [20]. Selain itu, metode ini tidak memerlukan banyak parameter dan relatif mudah diimplementasikan [20]. Namun, *Naive Bayes* juga memiliki keterbatasan, terutama ketika fitur-fitur dalam data ternyata memiliki korelasi satu sama lain, karena asumsi independensi bisa mengurangi akurasi prediksi. Metode ini juga sensitif terhadap kata-kata yang jarang muncul dalam teks, sehingga bisa mempengaruhi hasil klasifikasi jika distribusi data tidak merata.

Dalam konteks penelitian ini, *Naive Bayes* dipilih karena kemampuannya yang cepat dan efektif dalam mengklasifikasikan sentimen pengguna dari ulasan di Play Store. Dengan asumsi independensi kata-kata dalam ulasan, metode ini dapat membedakan ulasan positif, negatif, dan netral dengan cara yang sistematis. Kelebihan ini membuat proses analisis lebih efisien, terutama ketika jumlah ulasan cukup banyak. Meski demikian, keterbatasan seperti sensitivitas terhadap distribusi kata yang tidak merata menjadi perhatian saat menafsirkan hasil klasifikasi,

sehingga diperlukan pemahaman konteks ulasan agar interpretasi sentimen tetap akurat.

2.9. Evaluasi Model Naïve Bayes

Metode evaluasi model *Naive Bayes* umumnya dilakukan dengan mengukur kinerja klasifikasi menggunakan beberapa metrik standar, seperti akurasi, presisi, recall, dan F1-score. Akurasi menunjukkan sejauh mana model mampu memprediksi label yang benar dari keseluruhan data, sedangkan presisi menilai proporsi prediksi positif yang benar-benar relevan. Recall mengukur kemampuan model dalam menemukan seluruh data relevan, dan F1-score merupakan kombinasi presisi dan recall yang memberikan gambaran keseimbangan antara kedua metrik tersebut. Evaluasi ini penting untuk memastikan bahwa model *Naive Bayes* bekerja secara konsisten dan dapat diandalkan dalam membedakan kategori teks, terutama pada data yang bersifat subjektif seperti ulasan pengguna.

Dalam konteks penelitian ini, evaluasi model *Naive Bayes* dilakukan untuk menilai kemampuan klasifikasi sentimen pada ulasan pengguna aplikasi Grab di Play Store. Metrik-metrik seperti akurasi, presisi, recall, dan F1-score digunakan untuk mengetahui sejauh mana model dapat membedakan ulasan positif, negatif, dan netral dengan benar. Hasil evaluasi ini memberikan gambaran tentang efektivitas metode *Naive Bayes* dalam menangkap opini pengguna dan membantu menilai kualitas layanan yang diberikan oleh aplikasi Grab berdasarkan kecenderungan sentimen yang muncul.

2.10. Kelebihan Penelitian

Analisis sentimen dengan metode *Naive Bayes* memiliki kelebihan utama dalam kesederhanaan implementasi dan kecepatan proses klasifikasi. Metode ini

mampu menangani data teks dalam jumlah besar secara efisien tanpa memerlukan sumber daya komputasi yang tinggi. Selain itu, *Naive Bayes* juga dikenal cukup akurat dalam memprediksi kategori sentimen, terutama ketika data telah dipersiapkan secara terstruktur. Kelebihan lain adalah metode ini mudah dipahami dan diterapkan, sehingga dapat menjadi pilihan praktis bagi penelitian yang ingin menganalisis opini masyarakat secara cepat dan sistematis.

Dalam konteks penelitian ini, kelebihan tersebut sangat relevan karena data ulasan pengguna Grab di Play Store sangat banyak dan bervariasi. Penggunaan *Naive Bayes* memungkinkan penelitian untuk mengklasifikasikan opini pengguna ke dalam kategori positif, negatif, dan netral secara efisien. Dengan pendekatan ini, penelitian dapat memberikan gambaran yang jelas mengenai kecenderungan sentimen pengguna terhadap layanan Grab, sekaligus memberikan dasar bagi pengembang untuk memahami persepsi masyarakat dan meningkatkan kualitas layanan secara lebih terfokus.