

BAB I

PENDAHULUAN

1.1 Latar Belakang

Teknologi telah menjadi bagian penting dalam kehidupan masyarakat modern. Dengan adanya teknologi yang terus berkembang, semakin banyak kegiatan yang dapat dilakukan secara virtual. Gaya hidup masyarakat pun turut mengalami perubahan yang cukup signifikan, di mana mereka menginginkan segala sesuatu dilakukan dengan lebih cepat dan mudah daripada sebelumnya. Perubahan ini terjadi hampir di seluruh aktivitas kehidupan, termasuk perilaku masyarakat dalam melakukan perjalanan (*traveling*). Saat ini, sebagian besar masyarakat tidak tertarik untuk pergi ke agen perjalanan konvensional. Mereka telah beralih ke layanan pemesanan tiket transportasi dan akomodasi secara *online* melalui *website* agen perjalanan *online* (*online travel agency*). Kehadiran agen perjalanan *online* sebagai pelaku bisnis baru dalam industri pariwisata berhasil menggeser posisi agen perjalanan konvensional (Piambudi et al., 2019).

Seiring dengan perkembangan teknologi *smartphone* yang semakin luas dan canggih, agen perjalanan *online* mulai mengubah model bisnis mereka dari layanan berbasis *website* ke aplikasi *mobile* (Dwikesumasari & Ervianty, 2017). Keputusan ini dipengaruhi oleh fakta bahwa aplikasi *mobile* lebih mudah diunduh dan digunakan oleh pengguna *smartphone*. Akibatnya, popularitas aplikasi perjalanan *online* semakin meningkat dan semakin banyak masyarakat yang memilih untuk menggunakan aplikasi daripada mengakses *website*. Fenomena ini membuat persaingan di pasar aplikasi perjalanan *online* semakin ketat. Hasil survei bertajuk “*End of Year Vacation Report 2022*” yang dirilis oleh Populix pada bulan Desember 2022 mengungkapkan bahwa Traveloka menempati urutan pertama dalam daftar jajaran aplikasi perjalanan *online* yang paling populer di kalangan masyarakat Indonesia (Populix, 2022).

Sebagai aplikasi perjalanan *online* terpopuler di Indonesia, Traveloka tentu menerima beragam kritik dan saran dari para penggunanya. Hal ini dapat terlihat dari jumlah ulasan aplikasi Traveloka di Google Play Store yang mencapai 1,64 juta pada bulan Februari 2023. Ulasan di Google Play Store terdiri dari ulasan bintang

dan ulasan teks. Namun, ulasan teks jauh lebih bermanfaat karena dapat memberikan detail tentang pengalaman pengguna dalam menggunakan aplikasi (Wahyudi & Kusumawardana, 2021). Ulasan pengguna seringkali dijadikan referensi bagi pengguna lain sebelum mereka memutuskan untuk mengunduh dan menggunakan suatu aplikasi. Di sisi lain, ulasan tersebut juga berguna bagi pengembang aplikasi (*developer*) untuk meningkatkan kualitas dan performa aplikasi. Maka dari itu, diperlukan analisis sentimen terhadap ulasan aplikasi Traveloka untuk mengetahui respons pengguna terhadap aplikasi tersebut.

Analisis sentimen adalah studi komputasional yang menganalisis pendapat, sikap, dan emosi seseorang terhadap suatu entitas, seperti individu, acara, atau topik (Chaudhari, 2021). Tujuannya adalah untuk memahami sentimen yang terkandung dalam sebuah teks, seperti ulasan pengguna. Dengan demikian, dapat diketahui apakah ulasan tersebut termasuk ke dalam sentimen positif atau negatif. Analisis sentimen dapat dilakukan melalui pendekatan *Natural Language Processing* (NLP) dan *machine learning*. Pendekatan NLP seperti *text preprocessing* digunakan untuk mengidentifikasi dan mengekstrak pola serta fitur-fitur linguistik dalam teks yang menunjukkan sentimen positif atau negatif. Selanjutnya, pendekatan *machine learning* dilakukan untuk membuat model yang dapat mempelajari pola dan fitur-fitur dalam teks sehingga dapat mengklasifikasikan sentimen dari teks baru yang belum pernah dilihat oleh model sebelumnya.

Ada beberapa algoritma *machine learning* yang dapat digunakan untuk mengklasifikasikan sentimen, salah satunya adalah algoritma *Naïve Bayes* (Medhat et al., 2014). *Naïve Bayes* adalah sebuah algoritma klasifikasi yang menggunakan perhitungan probabilitas untuk memprediksi kelas data. Salah satu varian dari algoritma *Naïve Bayes* adalah *Multinomial Naïve Bayes* (MNB), yang didesain untuk bekerja dengan tipe data diskrit dan berdasarkan pada konsep frekuensi kata dalam dokumen (Yuyun et al., 2021). Namun, *Multinomial Naïve Bayes* memiliki kelemahan dalam mengatasi data yang tidak seimbang. *Complement Naïve Bayes* (CNB) adalah varian yang dirancang untuk memperbaiki asumsi berat yang digunakan pada *Multinomial Naïve Bayes* agar lebih sesuai dengan data yang tidak seimbang (Hasibuan & Heriyanto, 2022). Penelitian terkait *Multinomial Naïve Bayes* dan *Complement Naïve Bayes* sudah pernah dilakukan sebelumnya.

Penelitian yang dilakukan oleh (Abd et al., 2020) membahas mengenai perbandingan metode ekstraksi fitur *Term Frequency* (TF) dan *Term Frequency-Inverse Document Frequency* (TF-IDF) dengan algoritma klasifikasi *Naïve Bayes* (*Gaussian*, *Multinomial*, *Complement*, dan *Bernoulli*). Dataset yang digunakan dalam penelitian ini berjumlah 208 artikel politik yang diberi label reformis, konservatif, dan revolusioner. Hasil penelitian menunjukkan bahwa *Complement Naïve Bayes* memiliki akurasi terbaik sebesar 96,77% dengan menggunakan TF atau TF-IDF. *Multinomial Naïve Bayes* juga memiliki akurasi yang baik, yaitu 93,54% dengan TF dan 95,16% dengan TF-IDF. Namun, *Gaussian* dan *Bernoulli Naïve Bayes* memiliki akurasi yang lebih rendah dibandingkan dengan *Complement* dan *Multinomial Naïve Bayes*, baik menggunakan TF atau TF-IDF.

Penelitian lain yang juga menggunakan algoritma *Multinomial Naïve Bayes* juga dilakukan oleh (Hadaina & Budiyo, 2022) yang membahas tentang analisis sentimen ulasan aplikasi CoLearn di Google Play Store. Dataset yang digunakan dalam penelitian ini terdiri dari 536 ulasan yang kemudian dikelompokkan menjadi ulasan positif dan negatif. Dari hasil pengelompokan yang dilakukan, didapatkan 439 ulasan positif dan 97 ulasan negatif. Beberapa tahapan *preprocessing*, seperti *case folding*, *cleaning*, *tokenizing*, penggantian *slang word*, penghapusan *stop word*, dan *stemming*. Untuk menguji performa dari metode klasifikasi yang digunakan, dataset dibagi menjadi dua bagian, yaitu data training sebesar 90% dan data testing sebesar 10%. Berdasarkan hasil pengujian, metode *Multinomial Naïve Bayes* berhasil mencapai tingkat akurasi sebesar 88,89%.

Berdasarkan penelitian-penelitian sebelumnya diketahui bahwa algoritma *Multinomial* dan *Complement Naïve Bayes* memiliki akurasi yang cukup tinggi. Oleh karena itu, penulis tertarik untuk melakukan penelitian yang berjudul “Perbandingan Kinerja Algoritma Multinomial Dan Complement Naïve Bayes Untuk Analisis Sentimen Ulasan Aplikasi Traveloka Di Google Play Store”.

1.2 Rumusan Masalah

Berdasarkan penjelasan latar belakang yang telah disampaikan sebelumnya, dapat dirumuskan suatu permasalahan sebagai berikut.

1. Bagaimana mengimplementasikan algoritma *Multinomial* dan *Complement Naïve Bayes* dalam analisis sentimen ulasan aplikasi Traveloka di Google Play Store?
2. Mana yang lebih unggul, algoritma *Multinomial* atau *Complement Naïve Bayes* dalam melakukan analisis sentimen ulasan aplikasi Traveloka di Google Play Store?

1.3 Tujuan

Berdasarkan rumusan masalah yang telah dipaparkan, maka tujuan dari penelitian ini adalah sebagai berikut.

1. Mengimplementasikan algoritma *Multinomial* dan *Complement Naïve Bayes* dalam analisis sentimen ulasan aplikasi Traveloka di Google Play Store.
2. Membandingkan performa algoritma *Multinomial* dan *Complement Naïve Bayes* dalam analisis sentimen ulasan aplikasi Traveloka di Google Play Store untuk menentukan algoritma mana yang lebih unggul.

1.4 Manfaat

Adapun manfaat dari penelitian ini adalah sebagai berikut.

1. Memberikan pemahaman tentang konsep analisis sentimen dan algoritma *Naive Bayes*.
2. Memberikan informasi mengenai perbandingan kinerja antara algoritma *Multinomial Naive Bayes* dan *Complement Naive Bayes* pada analisis sentimen.
3. Menjadi bahan referensi untuk penelitian lebih lanjut di bidang analisis sentimen.

1.5 Batasan Masalah

Adapun batasan masalah yang ditetapkan dalam penelitian ini adalah sebagai berikut.

1. Data yang digunakan dalam penelitian ini adalah ulasan pengguna aplikasi Traveloka yang didapatkan dari Google Play Store.

2. Ulasan yang digunakan adalah ulasan berbahasa Indonesia yang diurutkan menurut ulasan yang paling relevan.
3. Ulasan akan diklasifikasikan ke dalam dua kategori, yaitu kategori sentimen positif dan negatif.
4. Implementasi dilakukan dengan pengkodean program menggunakan bahasa pemrograman Python.