

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Dalam era digital ini, ulasan pengguna di *Google Play Store* tidak hanya sekadar menyampaikan kesan individual, tetapi juga menjadi sumber informasi berharga bagi calon pengguna untuk memilih aplikasi yang sesuai dengan kebutuhan mereka [1].

Google Play Store menyediakan konten-konten digital seperti game, aplikasi, film, musik, dan buku dengan kategori yang beragam. Pada halaman *Google Play Store*, ulasan yang diberikan yaitu berupa rating bintang dari satu sampai lima ditambah dengan ulasan teks. Aplikasi yang memiliki rating bintang tinggi akan direkomendasikan oleh Google dan muncul pada halaman depan atau muncul pada aplikasi terbaik jika diurutkan melalui rating. Calon pengguna pada umumnya ingin mengetahui pendapat atau pengalaman dari pengguna lain terkait dengan aplikasi yang akan digunakan dan akan mencari aplikasi berdasarkan rating terbaik [2].

Penelitian terkait mengenai penerapan Algoritma *Naïve Bayes* dalam klasifikasi Aplikasi *Google Play Store* telah banyak dilakukan untuk mengatasi masalah pada ulasan pengguna. Salah satunya yaitu studi pada aplikasi Shopee menunjukkan bahwa algoritma *Naïve Bayes* dapat mengklasifikasikan sentiment ulasan menjadi positif dan negatif dengan

akurasi tinggi setelah melalui tahapan praproses data seperti pembersihan, *stemming*, dan *normalisasi* [3]. Penelitian lain pada aplikasi PeduliLindungi menggunakan *Naïve Bayes Multinomial* untuk menganalisis ulasan dengan fokus pada peningkatan akurasi menggunakan teknik seleksi fitur, yang menghasilkan performa klasifikasi yang signifikan [4]. Selain itu, pada implementasi metode *Naive Bayes* dalam menganalisis sentimen terhadap aplikasi McDonald's di *Google Play Store* dengan menggunakan ulasan pengguna sebagai sumber data. Proses analisis melibatkan langkah-langkah preprocessing, seperti tokenisasi, eliminasi stopwords, dan vektorisasi menggunakan metode *Term Frequency-Inverse Document Frequency (TF-IDF)*. Dengan model *Naive Bayes*, penelitian mencapai tingkat akurasi 88%, menunjukkan keahlian model dalam mengklasifikasikan ulasan ke dalam kategori sentimen positif, negatif, atau netral dengan tingkat presisi yang tinggi [5].

Pengguna sering menghadapi kesulitan dalam menemukan aplikasi yang relevan dan sesuai dengan kebutuhan mereka di *Aplikasi Google Play Store*. Hal ini disebabkan karena banyaknya aplikasi yang tersedia dan kurangnya personalisasi dalam rekomendasi. Akibatnya, pengguna sering kali harus mencoba beberapa aplikasi sebelum menemukan yang paling sesuai, yang tidak hanya membuang waktu tetapi juga menurunkan pengalaman pengguna secara keseluruhan.

Untuk penelitian ini penulis metode yaitu machine learning dengan menitik beratkan pada penggunaan metode *Naïve Bayes Multinomial*

dengan seleksi fitur index, dengan alasan utama yang berdasar pada penelitian penulis ingin mencoba apakah metode ini dapat diimplementasikan dengan sempurna pada Dataset Stem Review supaya Akurasinya bisa mencapai 98.66%. Berdasarkan penelitian dapat dilihat bermacam-macam metode untuk meningkatkan akurasi dari metode naïve bayes dan untuk metode yang terbaik adalah menggunakan seleksi fitur, dalam penelitian menjelaskan bahwa metode *Naïve Bayes* sangat sensitif terhadap seleksi fitur sehingga jumlah fitur yang digunakan akan sangat berpengaruh pada metode ini [6].

Adapun algoritma yang akan digunakan pada penelitian ini ialah Algoritma *Naïve Bayes*. *Naïve Bayes Classifier* merupakan algoritma untuk pengklasifikasian dengan metode probabilitas dan statistic untuk memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya [7]. Beberapa penelitian yang menerapkan *Naïve Bayes Classifier* sudah banyak diteliti. Seperti pada penelitian yang diimplementasikan oleh Muhamad, yang mengemukakan *Naïve Bayes* mempunyai banyak keunggulan, diantaranya memiliki akurasi yang tinggi dengan bentuk algoritma yang sederhana namun cepat menyelesaikan masalah, namun validasi pengujian yang dilakukan dalam penelitian tersebut hanya satu kali. Mekanisme pengujian dilakukan dengan menggunakan 123 dataset dengan parameter akurasi saja, peneliti akan melakukan pengujian dengan mekanisme membagi 303 data kedalam 5 subset yang akan

divalidasi dengan 5-fold cross validation dengan parameter pengujian akurasi, presisi dan recall [8].

Dari uraian di atas, maka penelitian akan berfokus pada pengembangan algoritma *Naïve Bayes*. Dengan judul penelitian **“PENERAPAN METODE ALGORITMA *NAÏVE BAYES* PADA KLASIFIKASI APLIKASI *GOOGLE PLAY STORE* DENGAN MENGGUNAKAN PLATFORM *MACHINE LEARNING*”**.

1.2 RUMUSAN MASALAH

Berdasarkan latar belakang masalah, maka didapatkan rumusan masalah penelitian ini adalah “Bagaimana algoritma *Naïve Bayes* dapat diterapkan untuk mengklasifikasi aplikasi di *Google Play Store* berdasarkan kategori, rating aplikasi yang paling banyak diminati “

1.3 BATASAN MASALAH

Batasan pada penelitian ini adalah sebagai berikut :

1. Bahasa pemrograman yang digunakan adalah *Python* dengan menggunakan *Google Colab*.
2. Menggunakan metode *Naïve Bayes*.
3. Dataset yang digunakan berjumlah 10842 aplikasi yang terdapat pada *Google Play Store*.
4. Aplikasi berdasarkan rating tertinggi.

1.4 TUJUAN DAN MANFAAT PENELITIAN

1.4.1 Tujuan Penelitian

Tujuan dari penelitian ini adalah :

1. Menerapkan Algoritma *Naïve Bayes* untuk mengklasifikasi aplikasi di *Google Play Store* berdasarkan kategori dan rating tertinggi.
2. Menggunakan platform *Machine Learning* untuk mempermudah proses pengolahan data dan evaluasi kinerja algoritma *Naïve Bayes* dalam klasifikasi

1.4.2 Manfaat Penelitian

Manfaat dari penelitian ini adalah :

1. Membantu pengguna Aplikasi *Google Play Store* untuk menemukan aplikasi yang relevan dengan cepat berdasarkan preferensi kategori atau rating tertinggi
2. Meningkatkan efisiensi dalam pengelompokan data yang tersedia di Aplikasi *Google Play Store* secara cepat dan akurat

1.5 SISTEMATIKA PENULISAN

Sistematika ini menggambarkan tentang pembahasan yang penulis buat pada tugas akhir untuk memudahkan dalam memahami penulisan laporan ini. Adapun sistematika penulisan dalam penelitian ini disusun sebagai berikut :

BAB I : PENDAHULUAN

Dalam bab pendahuluan penulis membahas tentang latar belakang masalah, perumusan masalah, batasan masalah, tujuan dan manfaat penelitian serta sistematika penulisan.

BAB II : LANDASAN TEORI

Dalam bab ini penulis menguraikan teori-teori yang mendasari pembahasan secara khusus dan berisikan definisi-definisi yang melandasi penelitian seperti Machine Learning, Naïve Bayes dan Penelitian Sejenis.

BAB III : METODOLOGI PENELITIAN

Dalam bab ini penulis menjelaskan metodologi yang digunakan dalam penelitian termasuk Teknik pengumpulan data dan metode analisis data.

BAB IV : ANALISIS DAN HASIL

Dalam bab ini penulis memaparkan analisi data yang diperoleh selama penelitian dan hasil dari analisis tersebut.

Bab ini juga mencakup interpretasi hasil serta diskusi mengenai temuan penelitian.

BAB V : PENUTUP

Dalam bab ini penulis memberikan Kesimpulan dari hasil penelitian dan menyampaikan saran-saran yang relavan untuk penelitian lebih lanjut

