

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terkait

Beberapa penelitian sebelumnya telah menggunakan algoritma *Naive Bayes* dalam analisis sentimen pengguna terhadap aplikasi digital. Salah satunya menganalisis ulasan pengguna aplikasi perbankan digital di Indonesia dari *Google Play Store*, menggunakan pendekatan CRISP-DM. Ulasan diklasifikasikan menjadi sentimen positif (46%) dan negatif (54%) dengan akurasi model mencapai 89% pada pembagian data 70:30. Hasil ini menunjukkan efektivitas *Naive Bayes* dalam analisis sentimen di sektor perbankan digital [14].

Penelitian lain telah membandingkan kinerja algoritma *Naive Bayes* dan SVM dalam analisis sentimen ulasan pengguna *Mobile Legends: Bang Bang* dari *Google Play Store*. Dari 796 data, ditemukan ketidakseimbangan dengan dominasi sentimen negatif (73%). Setelah upsampling menggunakan SMOTE, *Naive Bayes* mencapai akurasi 82%, sementara SVM memperoleh akurasi lebih tinggi, yaitu 91%. Temuan ini menunjukkan efektivitas *Naive Bayes* dalam menangani data yang tidak seimbang. [15].

Selain itu, riset yang telah dilaksanakan untuk mengevaluasi sentimen para pengguna aplikasi DANA di *Google Play Store* Indonesia memanfaatkan algoritma *Naive Bayes*. Data ulasan diperoleh dari Kaggle.com dan diproses menggunakan *RapidMiner* yang melibatkan tahapan tokenisasi, transformasi huruf kecil, penghapusan *stopword*, penyaringan panjang token, dan *stemming*. Penelitian ini membagi data 70% sebagai data latih dan 30% sebagai data uji, metode *Naive Bayes* yang dibangun berhasil mencapai akurasi sebesar 88,58%. Penelitian ini menggarisbawahi efektivitas *Naive Bayes* dalam analisis sentimen pada aplikasi keuangan digital [16].

2.2 Landasan Teori

2.2.1 Analisis Sentimen

Analisis sentimen adalah domain penelitian yang menekankan pemahaman pendapat, sentimen, evaluasi, sikap, dan emosi yang disampaikan dalam data tekstual. Analisis ini melibatkan pemeriksaan komputasi proses kognitif individu, keadaan afektif, *respons* emosional, dan pola perilaku. Selain itu, analisis sentimen memiliki kemampuan untuk meneliti repositori dokumen online, termasuk artikel berita digital, ulasan konsumen, dan tanggapan publik, kemudian mengklasifikasikannya ke dalam kategori positif, negatif, dan netral. [17]. Definisi lain menyatakan bahwa analisis sentimen merupakan upaya untuk memahami dan menafsirkan informasi tekstual secara mandiri untuk mendapatkan wawasan berharga [18].

2.2.2 Aplikasi TikTok

TikTok adalah aplikasi jejaring sosial yang memungkinkan penggunanya untuk produksi dan membagikan konten video singkat. TikTok adalah aplikasi inovatif untuk pembuatan video, didirikan pada September 2016 oleh sebuah perusahaan yang dikenal sebagai *ByteDance*, yang berasal dari China. TikTok mengalami lonjakan popularitas yang signifikan pada tahun 2018 di Indonesia dan selanjutnya diakui sebagai aplikasi utama di *Google Play Store*. [19].

2.2.3 *Google Play Store*

Google Play Store dikenal sebagai platform digital yang berperan sebagai pusat utama aplikasi pada perangkat dengan sistem operasi *Android* [20]. Dalam *Google Play Store*, ada bagian penilaian serta ulasan dari para pengguna untuk aplikasi yang sudah tersedia. Selain itu, *Google Play Store* juga dapat diartikan sebagai sebuah *platform* populer dari Google yang digunakan untuk mengunduh aplikasi, dengan fitur yang memungkinkan pengguna memberikan ulasan terhadap aplikasi yang telah mereka gunakan. Menurut portal data statistik, sekitar bulan Desember 2023, ada sekitar 2.430.000 aplikasi *mobile* yang tersedia di *Google Play Store* [21]. Salah satu fitur menarik dari *Google Play Store* adalah adanya kemampuan untuk menampilkan ulasan dari penggunanya.

2.2.4 *Algoritma Naive Bayes*

Naive Bayes adalah algoritma pembelajaran mesin yang sering digunakan untuk mengatasi masalah klasifikasi, terutama dalam pengelompokan teks terkait dengan data pelatihan yang berdimensi tinggi. [22]. Definisi lainnya menyebutkan bahwa *Naive Bayes* merupakan sebuah metode yang digunakan untuk klasifikasi data dengan berdasarkan pada probabilitas yang mungkin akan terjadi di masa depan [23]. Metode yang sering disebut sebagai *Naive Bayes Classifier* ini berfungsi dalam klasifikasi terawasi dengan cara menetapkan label kelas pada *instance* atau catatan di masa depan berdasarkan penggunaan probabilitas bersyarat. Algoritma ini memiliki kelebihan dalam hal kecepatan, kesederhanaan, dan kinerja yang baik pada data berukuran besar. Meskipun asumsi independensi antar fitur sering kali tidak sepenuhnya terpenuhi, yang dapat menjadi kelemahannya, *Naive Bayes* tetap menjadi pilihan yang andal dan efisien dalam berbagai tugas klasifikasi.