BAB I
PENDAHULUAN

1.1 Latar Belakang Masalah

Aplikasi digital berkembang dengan cepat, terutama dalam dunia perbankan dan transaksi elektronik. Flip, sebuah aplikasi yang menawarkan layanan transfer antar bank tanpa biaya administrasi telah berhasil menarik perhatian banyak pengguna [1]. Namun, terlepas dari manfaat ini, beberapa ulasan pengguna pada *platform* seperti *Google Play Store* menunjukkan berbagai pengalaman, baik yang positif maupun negatif yang dapat mempengaruhi citra aplikasi [2].

Ulasan negatif yang sering muncul dapat berdampak buruk pada persepsi pengguna. Sebaliknya, ulasan positif dapat membantu sebuah aplikasi meningkatkan citranya dan menarik lebih banyak konsumen. Hal ini menunjukkan bahwa *brand awareness* dan *perceived quality* berperan penting dalam membangun *brand associations* dan *brand loyalty*, yang pada akhirnya mempengaruhi *purchase intention* konsumen [3].

Analisis sentimen dari ulasan pengguna adalah metode yang sangat baik untuk mengevaluasi perspektif pengguna secara menyeluruh. Perusahaan dapat mengidentifikasi kekuatan dan kekurangan layanannya dengan memeriksa ulasan yang ada. Namun, salah satu masalah yang paling sering ditemui dalam penelitian ini adalah ketidakseimbangan data, yang terjadi ketika jumlah ulasan positif lebih besar daripada jumlah ulasan negatif. Hal ini berpotensi untuk mengubah hasil analisis serta keakuratan model prediksi sentimen yang digunakan [4].

Seiring dengan perkembangan fintech di Indonesia, terutama selama pandemi COVID-19, aplikasi Flip muncul sebagai inovasi menarik dengan layanan transfer bank tanpa biaya administrasi. Namun, distribusi data ulasan pengguna di Google Play Store menunjukkan persepsi yang beragam, baik positif maupun negative [5]. Untuk mengatasi ketidakseimbangan data ulasan tersebut, penelitian ini menggunakan metode Naive Bayes serta teknik balancing data dalam mengevaluasi distribusi data dan meningkatkan kualitas analisis sentimen.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang dipaparkan, rumusan masalah dalam penelitian ini adalah sebagai berikut.

1. Bagaimana kinerja algoritma Naive Bayes dalam mengklasifikasikan sentimen ulasan pengguna aplikasi Flip berdasarkan data yang tidak seimbang?
2. Metode *oversampling* atau *undersampling* apa yang paling efektif untuk meningkatkan performa model Naive Bayes saat menangani data sentimen ulasan aplikasi Flip yang tidak seimbang?

1.3 Batasan Masalah

Berdasarkan rumusan masalah tersebut, maka ruang lingkup dan batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Data penelitian ini hanya mencakup ulasan aplikasi Flip dari *Google Play Store*.
2. Metode Naive Bayes adalah satu-satunya metode yang digunakan dalam penelitian ini untuk klasifikasi sentimen.
3. Representasi teks dilakukan menggunakan TF-IDF.
4. Variabel yang dianalisis terbatas pada *content* (teks ulasan) dan *score* (nilai rating) dari dataset.
5. Sentimen ulasan diasumsikan sepenuhnya berdasarkan nilai *score*: nilai ≤3 dikategorikan negatif, sementara nilai >3 dikategorikan positif.
6. Dataset diasumsikan representatif meskipun hanya mencakup 1000 ulasan dari keseluruhan populasi ulasan aplikasi Flip.
7. Penelitian hanya membandingkan beberapa metode *oversampling* dan *undersampling* yang dipilih, seperti ADASYN, K-Means SMOTE, *Random Over Sampler*, *Random Under Sampler*, *Repeated Nearest Neighbour*, dan Tomek Links.
8. Evaluasi model hanya dilakukan menggunakan metrik *precision*, *recall*, *F1-score*, dan *accuracy*.

1.4 Tujuan Penelitian

Tujuan penelitian yang ingin dicapai oleh penulis adalah sebagai berikut:

1. Mengevaluasi kemampuan model Naive Bayes dalam mengklasifikasikan sentimen ulasan pengguna aplikasi Flip berdasarkan data yang tidak seimbang.
2. Membandingkan pengaruh metode *oversampling* dan *undersampling* ADASYN, K-Means SMOTE, *Random Over Sampler*, *Random Under Sampler*, *Repeated Nearest Neighbour*, dan Tomek Links untuk meningkatkan performa model Naive Bayes dalam menangani ketidakseimbangan data ulasan aplikasi Flip.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat bagi berbagai pihak sebagai berikut:

1. Penelitian ini dapat menjadi sumber wawasan dan referensi dalam bidang analisis sentimen, khususnya dalam *review* aplikasi digital dengan menggunakan metode Naive Bayes, serta menjadi salah satu syarat untuk memperoleh gelar sarjana.
2. Berkontribusi pada pengembangan strategi untuk menangani data yang tidak seimbang dengan menggunakan berberapa teknik *oversampling* dan *undersampling*.
3. Menjadi referensi dalam studi yang serupa, terutama yang menggunakan data ulasan aplikasi atau metode Naive Bayes, dan membantu peneliti lain dalam menyelidiki lebih lanjut teknik-teknik untuk menangani data yang tidak seimbang.

1.6 Sistematika Penulisan

Sistematika penulisan penelitian ini dibagi menjadi 5 (lima) bab, dengan masing-masing bab diuraikan sebagai berikut:

BAB I PENDAHULUAN, menguraikan latar belakang masalah, rumusan masalah, batasan masalah, tujuan, manfaat penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA, berisi tinjauan pustaka, dasar-dasar teori yang digunakan.

BAB III METODE PENELITIAN, berisi objek penelitian, alur penelitian.

BAB IV HASIL DAN PEMBAHASAN, menyajikan hasil dan pembahasan berdasarkan penelitian yang dilakukan sesuai dengan metodologi yang digunakan.

BAB V PENUTUP, mengakhiri penelitian dengan menarik kesimpulan dari uraian bab sebelumnya dan memberikan saran untuk penelitian selanjutnya.