

BAB I

PENDAHULUAN

1.1 Latar Belakang

Di era digital, layanan pinjaman tunai berbasis aplikasi semakin berkembang pesat, memberikan kemudahan bagi masyarakat dalam memperoleh akses ke pinjaman tunai tanpa perlu melalui proses yang sulit seperti di lembaga keuangan tradisional [1]. Tren pinjaman online semakin tinggi seiring dengan perubahan gaya hidup dan kebutuhan finansial masyarakat yang menginginkan solusi kredit yang cepat, mudah, dan fleksibel [2]. Kredivo merupakan aplikasi pinjaman online yang populer di Indonesia yang menyediakan layanan pinjaman tunai dengan proses cepat dan persyaratan yang relatif mudah dibandingkan layanan perbankan konvensional. Namun, banyak pengguna yang memberikan ulasan beragam terkait pengalaman mereka dalam menggunakan layanan pinjaman tunai dari aplikasi ini [3]. Beberapa pengguna mengeluhkan aspek tertentu seperti proses verifikasi yang lama, transparansi biaya dan bunga, sistem penagihan, serta kecepatan pencairan dana [4]. Sebaliknya, ada pula pengguna yang merasa puas dengan fleksibilitas dan kemudahan akses pinjaman [5].

Beberapa penelitian terdahulu yang dilakukan untuk mendukung perbandingan algoritma *machine learning* dalam analisis sentimen pada aplikasi layanan keuangan menunjukkan beberapa tren yang konsisten. Sejumlah studi pada aplikasi seperti BRImo dan BCA Mobile secara konsisten menemukan bahwa algoritma *Support Vector Machine* (SVM) memiliki performa yang lebih unggul dibandingkan Naïve Bayes, dengan akurasi masing-masing mencapai 97,69% dan 85% [3][4]. Hasil serupa juga ditemukan pada analisis spesifik aplikasi Kredivo di mana SVM terbukti lebih superior[5]. Di sisi lain, ketika algoritma *ensemble* seperti Random Forest disertakan, hasilnya menunjukkan performa yang sangat kompetitif. Penelitian pada aplikasi Halo BCA menemukan bahwa Random Forest justru mencapai akurasi tertinggi sebesar 91,28%, melampaui SVM dan Naïve Bayes[6]. Meskipun studi-studi tersebut berhasil membandingkan performa akurasi, sebuah celah penelitian yang

signifikan dan konsisten teridentifikasi: masalah ketidakseimbangan kelas data. Banyak dari penelitian tersebut, seperti yang tertera pada analisis sentimen Kredivo menggunakan *Support Vector Machine* dan *Naïve Bayes* secara eksplisit menyebutkan bahwa jumlah ulasan positif yang jauh mendominasi menjadi kelemahan yang menyebabkan model cenderung bias dan lemah dalam mendeteksi sentimen minoritas (negatif)[7]. Kurangnya eksplorasi solusi untuk masalah fundamental inilah yang menjadi justifikasi utama dilakukannya penelitian ini, di mana dampak teknik penyeimbangan data SMOTE akan diuji secara spesifik.

Sebagian besar penelitian yang ada cenderung fokus pada pencapaian akurasi keseluruhan dan seringkali kurang mendalam dalam menangani masalah ketidakseimbangan data. Masalah ini sangat umum terjadi pada data ulasan, di mana sentimen positif jauh lebih dominan. Permasalahan tersebut, yang bahkan diakui sebagai keterbatasan oleh penelitian sebelumnya, menyebabkan model yang dihasilkan cenderung lemah dalam mengidentifikasi sentimen minoritas (negatif) [1].

Untuk mengisi celah penelitian inilah, studi ini diusulkan. Penelitian ini tidak hanya akan membandingkan performa tiga algoritma utama (*SVM*, *Naïve Bayes*, dan *Random Forest*), tetapi secara spesifik akan menguji dan menganalisis dampak dari penerapan teknik *oversampling SMOTE* sebagai solusi untuk masalah ketidakseimbangan data. Melalui pendekatan ini, penelitian ini diharapkan dapat memberikan pemahaman yang lebih komprehensif mengenai cara membangun model analisis sentimen yang lebih seimbang dan andal untuk aplikasi *fintech*.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang telah disajikan sebelumnya, masalah-masalah yang ada dapat dijabarkan menjadi beberapa poin sebagai berikut:

1. Bagaimana perbandingan performa dari algoritma *Support Vector Machine* (*SVM*), *Naïve Bayes*, dan *Random Forest* dalam

mengklasifikasikan sentimen pada dataset ulasan aplikasi Kredivo yang tidak seimbang?

2. Bagaimana penerapan teknik *oversampling* SMOTE (*Synthetic Minority Over-sampling Technique*) dapat memengaruhi dan meningkatkan performa model dalam mengatasi masalah ketidakseimbangan data, khususnya dalam mendeteksi kelas sentimen negatif?
3. Bagaimana sebuah *dashboard* interaktif dapat dirancang untuk menyajikan wawasan (*insight*) utama dari hasil analisis sentimen secara ringkas?

1.3 Batasan Masalah

Dalam penelitian ini, terdapat beberapa batasan masalah pada perusahaan yang dianalisis. Batasan-batasan tersebut adalah sebagai berikut:

1. Penelitian ini hanya menggunakan data ulasan aplikasi Kredivo yang tersedia di Google Play Store sebesar 10000 ulasan.
2. Kategorisasi sentimen terbatas pada dua kategori yaitu positif dan negatif dengan menggunakan rating, 1-2 adalah negatif dan 4-5 adalah positif.
3. Algoritma yang digunakan hanya meliputi Support Vector Machine, Naïve Bayes, dan Random Forest tanpa membandingkan dengan model deep learning.
4. Data ulasan diambil dalam rentang waktu 2025.
5. Tidak membahas faktor eksternal yang dapat memengaruhi sentimen pengguna, seperti contohnya promosi atau perubahan kebijakan Kredivo.
6. Penelitian tidak menganalisis potensi bias yang mungkin timbul dari karakteristik pengguna yang memilih untuk memberikan ulasan di Google Play Store

7. Proses pemodelan algoritma machine learning tidak melibatkan optimasi hyperparameter, parameter yang digunakan adalah parameter standart dari library scikit-learn

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Membandingkan performa klasifikasi sentimen dari algoritma *Support Vector Machine* (SVM), *Naïve Bayes*, dan *Random Forest* pada dataset ulasan aplikasi Kredivo yang tidak seimbang untuk menentukan model dengan performa dasar terbaik.
2. Menganalisis dampak dan mengukur efektivitas penerapan teknik *oversampling* SMOTE dalam meningkatkan kemampuan model untuk mengklasifikasikan kelas sentimen minoritas (negatif).
3. Membangun sebuah *dashboard* interaktif sebagai *proof-of-concept* yang mampu menyajikan wawasan (*insight*) utama dari hasil analisis sentimen secara visual, seperti tema-tema yang sering muncul pada keluhan negatif.

1.4.2 Manfaat Penelitian

Manfaat pada penelitian ini adalah sebagai berikut:

A. Manfaat Teoritis

1. Hasil penelitian ini dapat menjadi referensi dan memperkaya literatur akademis di bidang analisis sentimen, khususnya mengenai perbandingan performa algoritma *machine learning* pada studi kasus ulasan aplikasi fintech di Indonesia.
2. Memberikan kontribusi pengetahuan mengenai dampak dan efektivitas penerapan teknik *oversampling* SMOTE dalam mengatasi masalah ketidakseimbangan data pada model klasifikasi teks yang berbeda (SVM, Naïve Bayes, dan

Random Forest), yang dapat menjadi acuan bagi penelitian sejenis di masa depan.

B. Manfaat Praktis

1. Bagi Perusahaan, Penelitian ini memberikan rekomendasi model terbaik yang dapat diimplementasikan untuk sistem monitoring opini publik secara otomatis, serta menghasilkan *insight* berbasis data untuk mendukung pengambilan keputusan dalam rangka peningkatan kualitas layanan dan pengalaman pengguna.
2. Bagi pengguna aplikasi, penelitian ini berguna untuk membantu pengguna untuk memahami aspek-aspek positif dan negatif dari layanan Kredivo berdasarkan hasil analisis sentimen.

1.5 Sistematika Penulisan

1) BAB I: Pendahuluan

Bab ini menjelaskan latar belakang penelitian, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan. Bab ini memberikan dasar-dasar mengapa penelitian ini dilakukan dan apa yang ingin dicapai.

2) BAB II: Landasan Teori

Bab ini memuat teori-teori yang mendukung penelitian, termasuk penjelasan tentang Sentiment Analysis, algoritma SVM, algoritma Naïve Bayes, algoritma Random Forest dan teori terkait lainnya. Selain itu, juga disajikan kajian terhadap penelitian-penelitian terdahulu yang relevan dengan topik ini.

3) BAB III: Metodologi Penelitian

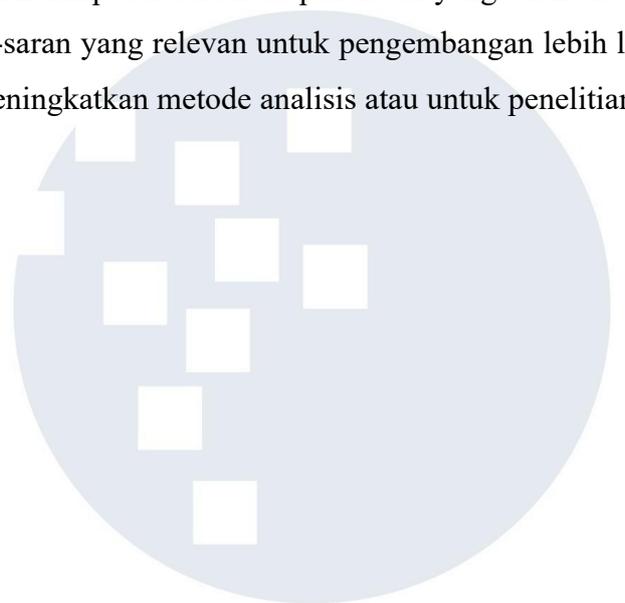
Bab ini menjelaskan metode penelitian yang digunakan, mencakup teknik pengumpulan data, metode pengolahan data (pre-processing), serta metode analisis sentimen yang diterapkan. Proses pemodelan dan evaluasi metode juga dijelaskan di bagian ini.

4) BAB IV: Analisis dan Hasil Penelitian

Pada bab ini, hasil penelitian akan dibahas secara detail. Mulai dari pengumpulan data, proses pre-processing data, analisis sentimen, pemodelan dengan algoritma yang digunakan, hingga evaluasi hasil. Bab ini merupakan inti dari penelitian karena berisi temuan utama.

5) BAB V: Simpulan dan Saran

Bab ini menyajikan kesimpulan dari hasil penelitian yang telah dilakukan, serta memberikan saran-saran yang relevan untuk pengembangan lebih lanjut. Saran ditujukan untuk meningkatkan metode analisis atau untuk penelitian lanjutan di bidang yang sama



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA