

BAB I

PENDAHULUAN

1.1 Latar Belakang

Menurut data yang diberikan oleh Asosiasi Kartu Kredit Indonesia (*Asosiasi Kartu Kredit Indonesia*, n.d.), pengguna kartu kredit mengalami peningkatan. Pada tahun 2015, jumlah pengguna kartu kredit yang tercatat mencapai angka 16.863.842. Selanjutnya, jumlah ini mengalami peningkatan signifikan, mencapai 17.693.269 pada tahun 2023. Peningkatan ini sejalan dengan tingkat kejahatan di sektor perbankan, khususnya dalam penipuan transaksi. Data statistik menunjukkan bahwa Indonesia menempati peringkat kedua di antara negara-negara dengan risiko penipuan transaksi kartu kredit tertinggi, mencapai 18,3% (Goyal & Manjhvar, 2020). Penipuan transaksi kartu kredit dapat terjadi ketika penipu menggunakan informasi palsu untuk melakukan transaksi atau ketika kartu kredit yang sah digunakan untuk kegiatan penipuan.

Algoritma klasifikasi pada *machine learning* dapat diimplementasikan untuk mengklasifikasikan penipuan transaksi kartu kredit. Namun, implementasi model *machine learning* dalam klasifikasi penipuan transaksi kartu kredit juga menghadapi banyak tantangan, terutama terkait dengan ketidakseimbangan kelas. Ketidakseimbangan kelas adalah masalah umum dalam data penipuan kartu kredit, di mana jumlah transaksi penipuan (kelas minoritas) biasanya jauh lebih sedikit daripada jumlah transaksi yang sah (kelas mayoritas). Fenomena ini dapat menyebabkan model *machine learning* cenderung memihak pada kelas mayoritas, menghasilkan kinerja yang rendah dalam mendeteksi kasus penipuan (Ningsih et al., 2022). Ketidakseimbangan kelas ini dapat menyebabkan model menjadi tidak sensitif terhadap kasus penipuan yang sebenarnya, karena mayoritas model akan cenderung memprediksi transaksi sebagai transaksi yang sah.

Salah satu pendekatan untuk menangani masalah ketidakseimbangan kelas adalah dengan menggunakan teknik sampling. Teknik sampling mengubah distribusi data antara kelas mayoritas dan kelas minoritas dalam dataset pelatihan

untuk mencapai keseimbangan jumlah data di setiap kelas (Sutoyo & Fadlurrahman, 2020). Salah satu teknik sampling adalah *oversampling* yang merupakan teknik yang bertujuan untuk meningkatkan jumlah kelas minoritas agar seimbang atau hampir seimbang dengan kelas mayoritas (Ningsih et al., 2022; Purohit & Vishwakarma, 2021).

Beberapa penelitian telah mengimplementasikan teknik *oversampling* untuk mengatasi masalah ketidakseimbangan kelas seperti pada penelitian oleh (Ningsih et al., 2022), dalam penelitian tersebut didapatkan bahwa teknik *oversampling SMOTE* berhasil menaikkan *recall* sebesar 10% pada model random forest. Adapun penelitian lainnya dari (Hidayat et al., 2021), dalam penelitian tersebut didapatkan bahwa teknik *oversampling SMOTE* memberikan performa yang lebih bagus dari *ADASYN* ketika dipasangkan dengan algoritma *SVM* untuk klasifikasi dataset Airbnb.

Sesuai dengan latar belakang di atas maka penulis mengambil judul “PERBANDINGAN TEKNIK OVERSAMPLING DALAM MENGATASI PERMASALAHAN KETIDAKSEIMBANGAN KELAS PADA KASUS PENIPUAN KARTU KREDIT”. Penelitian ini bertujuan untuk mengevaluasi sejauh mana teknik *oversampling* dapat meningkatkan performa model *machine learning* untuk klasifikasi kasus penipuan transaksi kartu kredit pada dataset dengan kelas tidak seimbang. Dengan demikian, penelitian ini diharapkan dapat memberikan wawasan yang lebih mendalam tentang solusi yang paling efektif dalam menangani masalah ketidakseimbangan kelas pada klasifikasi penipuan kartu kredit.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang di atas, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana kinerja teknik *oversampling* dalam meningkatkan keseimbangan kelas?

2. Bagaimana pengaruh penerapan teknik *oversampling* terhadap performa model *machine learning* yang dibangun?
3. Bagaimana perbandingan antara teknik *oversampling* *SMOTE*, *Borderline SMOTE*, dan *ADASYN*?

1.3 Ruang Lingkup Penelitian

Ruang Lingkup masalah pada penelitian ini adalah:

1. Dataset yang digunakan adalah dataset *credit card fraud detection* yang diambil dari Kaggle.com.
2. Algoritma *oversampling* yang digunakan adalah *SMOTE*, *Borderline SMOTE* dan *ADASYN*
3. Algoritma klasifikasi yang digunakan adalah *Decision Tree* dan *Logistic Regression*.
4. Metrik untuk mengevaluasi model yang digunakan adalah *precision*, *recall*, *f1 score*, *confusion matrix*, *area under receiver operating characteristic curve* dan *area under precision-recall curve*.

1.4 Tujuan

Sesuai dengan rumusan masalah yang telah diuraikan sebelumnya, maka tujuan dari penelitian ini adalah sebagai berikut:

1. Mengukur kinerja teknik *oversampling* dalam meningkatkan keseimbangan kelas.
2. Mengukur kinerja teknik *oversampling* dalam meningkatkan performa model *machine learning* yang dibangun untuk deteksi penipuan kartu kredit.
3. Membandingkan teknik *oversampling* *SMOTE*, *Borderline SMOTE*, dan *ADASYN*.

1.5 Manfaat Penelitian

Penelitian ini akan memberikan kontribusi baru terhadap literatur ilmiah dalam bidang deteksi penipuan kartu kredit dengan membandingkan efektivitas teknik *oversampling*. Hasil penelitian ini dapat memberikan wawasan yang berharga bagi praktisi industri perbankan dan keamanan finansial untuk memilih teknik yang paling efektif dalam mengatasi masalah ketidakseimbangan kelas dalam deteksi penipuan transaksi kartu kredit. Selain itu, temuan dari penelitian ini juga dapat memberikan landasan untuk pengembangan lebih lanjut dalam penggunaan teknologi *machine learning* dalam meningkatkan keamanan transaksi kartu kredit.

1.6 Sistematika Penulisan

Sistematika penulisan skripsi ini disusun untuk memberikan gambaran umum tentang penelitian yang dilakukan. Sistematika penulisan skripsi ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini menguraikan tentang latar belakang melakukan penelitian, rumusan masalah, ruang lingkup penelitian, manfaat penelitian, dan sistematika penulisan yang akan diterapkan.

BAB II KAJIAN TEORI

Bab ini mencakup referensi yang relevan dengan topik penelitian untuk mendukung penulisan skripsi. Bab ini membahas berbagai konsep dasar dan teori-teori yang terkait dengan topik penelitian yang sedang dilakukan.

BAB III METODE PENELITIAN

Bab ini berisi tentang proses perancangan berdasarkan analisis permasalahan sesuai dengan metode penelitian yang dipilih dalam penelitian yang dilakukan.

BAB IV HASIL DAN PEMBAHASAN

Bab ini akan memuat penjelasan mengenai tahap implementasi secara bertahap. Selain itu, hasil evaluasi implementasi juga akan disampaikan.

BAB V PENUTUP

Bab ini berisi kesimpulan yang didapatkan dari hasil penelitian yang telah dilakukan juga saran yang bertujuan untuk peningkatan topik skripsi yang dapat digunakan di masa mendatang.