

# **BAB I**

## **PENDAHULUAN**

### **1.1. Latar Belakang**

Industri keuangan saat ini telah mengalami proses perubahan, dimana transaksi online telah menjadi bagian integral dari kehidupan modern. Hal yang mendorong kemajuan industri keuangan dibuktikan dengan keberadaan platform *e-commerce*, perbankan digital dan layanan pembayaran online. Dan salah satu aspek yang mendorong pertumbuhan platform *e-commerce* adalah kenyamanan yang ditawarkan kepada konsumen. Dimana kenyamanan yang diberikan dalam kemudahan dan kenyamanan konsumen ialah ketika melakukan pembayaran yang dilakukan secara online. Dengan beberapa klik, individu dapat menjelajahi dan membeli produk atau layanan dari kenyamanan rumah mereka sendiri. Selain itu, berbagai inovasi dalam teknologi keuangan (*fintech*) telah memberikan kemampuan untuk melakukan transaksi dengan cepat, aman, dan efisien.

Namun, seiring dengan keuntungan dan kenyamanan yang didapatkan dari ekosistem transaksi online, hal ini juga memunculkan tantangan serius, terutama dalam kecurangan transaksi online. Kecurangan transaksi online adalah salah satu bentuk penipuan yang dilakukan sengaja dan menyebabkan kerugian dengan trik implisit maupun eksplisit yang dilakukan oleh pihak tertentu dengan tujuan untuk mendapatkan keuntungan finansial yang signifikan. Kecurangan transaksi online dapat melibatkan berbagai tindakan yang tidak sah, contohnya seperti mencuri identitas, menggunakan kartu kredit palsu, dan aktivitas penipuan lainnya.

Deteksi penipuan transaksi online merupakan bentuk upaya yang harus dilakukan, karena kecurangan transaksi online sangat merugikan konsumen. Dengan mendeteksi kecurangan transaksi online dapat memberikan perlindungan konsumen dari kehilangan dana atau identitas pribadi yang dapat disalahgunakan (Faried Zamachsari dan Niken Puspitasari, 2021). Tak hanya berdampak pada konsumen, bentuk kecurangan transaksi online mempengaruhi integritas sistem pembayaran. Kecurangan dalam hal ini dapat merusak reputasi lembaga keuangan dan perusahaan penyedia layanan. Dengan demikian, deteksi kecurangan transaksi online tidak hanya menjadi elemen penting dalam melindungi individu dan

perusahaan dari risiko finansial, tetapi juga dalam menjaga integritas dan kepercayaan dalam ekosistem bisnis online secara keseluruhan. Adapun cara untuk mendeteksi kecurangan transaksi online dapat dilihat dari beberapa faktor data misalnya jenis transaksi, saldo sebelum melakukan transaksi dan saldo sesudah transaksi.

Selanjutnya untuk dapat mengklasifikasikan data kecurangan transaksi online perlu model atau oleh data yang digunakan untuk mendeteksi kecurangan transaksi online. Sehingga untuk melakukan oleh data tersebut menggunakan proses Data Mining. Dalam *Knowledge Discovery in Database (KDD)*, data mining memegang peran kunci sebagai elemen yang sangat signifikan dalam mengekstrak pola atau model dari data dengan penerapan algoritma tertentu. Klasifikasi, sebagai salah satu model dalam data mining, adalah teknik untuk memprediksi data dengan membuat estimasi nilai dari suatu data yang diperoleh dari sumber data yang berbeda.

Namun salah satu permasalahan umum algoritma klasifikasi tidak memperhitungkan ketidakseimbangan dalam dataset, yang dapat mengakibatkan penurunan kinerja prediksi dan peningkatan bias pada kelas mayoritas. Dua pendekatan umum untuk menangani ketidakseimbangan data adalah oversampling dan undersampling. Oversampling bertujuan untuk menyamakan distribusi kelas minoritas dengan menduplikasi data kelas minoritas secara identik. Sementara itu, undersampling berusaha untuk mengimbangi kelas minoritas dengan menghapus sebagian data dari kelas mayoritas hingga distribusi menjadi seimbang, namun ini dapat mengakibatkan kehilangan informasi yang signifikan (Wijayanti, Kencana, & Sumarjana, 2021). Oleh karena itu, peneliti memilih pendekatan oversampling untuk menyeimbangkan data. Berbagai metode oversampling yang umum digunakan termasuk ROS (Random Over Sampling), SMOTE, dan ADASYN. Penelitian ini menggunakan metode SMOTE, yang meningkatkan jumlah data kelas minoritas dengan mereplikasi data secara acak menggunakan  $k$  tetangga terdekat. Dengan demikian, metode SMOTE dapat mengurangi risiko overfitting dan tetap mempertahankan integritas data asli (Suryana, Pratiwi, & Prasetyo, 2021).

Adapun penelitian sebelumnya yang terkait, adalah penelitian yang

dilakukan oleh Yerik Afrianto dengan judul Penerapan Metode CRISP-DM dalam Klasifikasi Data Ulasan Pengunjung Destinasi Danau Toba Menggunakan Algoritma *Naïve Bayes Classifier* (NBC) dan *Decision Tree* (DT). Dari penelitian ini menyimpulkan bahwa hasil perbandingan nilai algoritma NBC dan DT menunjukkan bahwa model dengan performa terbaik ialah DT menggunakan operator SMOTE UPsampling dengan nilai akurasi (98,27%), presisi (98,83%), recall (97,71%), f-measure (98,26%), dan nilai AUC (0,982). Selanjutnya, penelitian dari Afrilio Fransenda dengan judul Integrasi Metode Decision Tree dan SMOTE untuk Klasifikasi Data Kecelakaan Lalu Lintas. Dari penelitian ini menyimpulkan bahwa Pengujian menggunakan Decision Tree dan SMOTE Split Data diperoleh akurasi 71.12% dengan perbandingan 1:9 dan juga diperoleh pula presisi 89.71% (3:7) dan area under curve (AUC) sebesar 0.773 (1:9). Sehingga dengan hasil tersebut maka penelitian ini masuk pada kategori *fair classification* (cukup). Dari kedua penelitian tersebut, akurasi yang dihasilkan oleh metode Decision Tree cukup baik dengan menghasilkan akurasi yang cukup.

Berkaitan dengan hal itu, maka peneliti akan melakukan klasifikasi dengan metode *Decision Tree*. Penggunaan metode ini memberikan kemudahan dalam interpretasi dan kecepatan komputasi. Dengan pertimbangan ini, penggunaan Decision Tree merupakan strategi yang tepat sesuai dengan kebutuhan analisis data dan tujuan proyek yang dihadapi saat ini. Dan dalam penelitian ini juga akan membandingkan hasil dari model klasifikasi decision tree dengan proses algoritma SMOTE dan tanpa proses algoritma SMOTE dalam studi kasus ketidakseimbangan data kecurangan transaksi online.

## **1.2. Rumusan Masalah**

Berdasarkan dengan latar belakang diatas, maka peneliti mengambil rumusan masalah sebagai berikut:

1. Bagaimana penerapan metode SMOTE dalam menangani ketidakseimbangan data?
2. Bagaimana penerapan klasifikasi Decision Tree dalam mendeteksi kecurangan transaksi online?

3. Bagaimana hasil kinerja klasifikasi decision tree dengan dan tanpa pengaruh penerapan algoritma SMOTE mendeteksi kecurangan transaksi online?

### **1.3. Tujuan**

Adapun tujuan dari penelitian ialah analisis kinerja sistem klasifikasi dalam mendeteksi kecurangan transaksi online dengan menggunakan algoritma SMOTE dan Decision Tree.

### **1.4. Manfaat**

1. Memberikan solusi untuk mengatasi masalah kecurangan transaksi online.
2. Meningkatkan keamanan transaksi online.
3. Membantu *e-commerce* dan lembaga keuangan untuk melindungi diri dari kerugian finansial.
4. Memberikan wawasan dan referensi tentang ilmu data mining, serta memperoleh gambaran dari langkah proses klasifikasi *Decision Tree* menggunakan algoritma SMOTE.

### **1.5. Batasan Masalah**

Untuk mencapai agar penelitian dapat dilakukan lebih terarah, maka penulis membatasi penelitiannya sebagai berikut:

1. Data yang digunakan dalam penelitian adalah data “*Online Payment*” diambil dari Kaggle dengan berupa data numerik. Data diakses pada tanggal 2 November 2023, link yang diakses:  
<https://www.kaggle.com/datasets/rupakroy/online-payments-fraud-detection-dataset>
2. Menggunakan metode *decision tree* untuk melakukan klasifikasi
3. Bahasa pemrograman yang digunakan dalam penelitian adalah bahasa Python, dengan menggunakan tools *Jupyter Notebook*.
4. Skenario yang dijalankan sesuai dengan skenario pengujian yang sudah ditentukan.

5. Evaluasi model yang digunakan dalam penelitian ialah *k-fold cross validation*.