

BAB I

PENDAHULUAN

1.1 Latar Belakang

Financial services merupakan sebuah sektor yang bergerak dalam bidang penyedia layanan keuangan dengan tujuan supaya memudahkan masyarakat dalam mengelola aset finansialnya. Layanan tersebut dapat berupa peminjaman/pembiayaan, investasi, simpanan, pembayaran (Amalia, 2022). Salah satu layanan pembiayaan dapat digunakan oleh pelanggan untuk kredit rumah, kendaraan bermotor, mobil, alat berat. Melalui layanan tersebut, pelanggan mendapatkan pinjaman uang dengan jumlah sesuai dengan kredit yang akan dilakukan. Dalam melakukan peminjaman uang tersebut, perusahaan tentu akan melakukan seleksi yang ketat supaya dapat memberikan pembiayaan kepada pelanggan dengan tepat. Namun pada praktiknya, masih banyak aksi penipuan yang dilakukan oleh pelanggan ketika melakukan peminjaman uang di perusahaan layanan keuangan (Alwadain, Ali and Muneer, 2023).

Data menunjukkan bahwa kejahatan di sektor finansial mengalami kenaikan, mulai dari tahun 2017 sejumlah 370.916 kasus hingga tahun 2021 sejumlah 1.686.121 kasus. Dari sejumlah kasus tersebut, pada tahun 2021 terdapat 197.967 kasus penipuan (*fraud*) pada perusahaan layanan peminjaman uang dan sewa (Daly and Caporal, 2023). Salah satu penyebab terjadinya penipuan tersebut adalah pemalsuan data (*data falsification*) yang dilakukan oleh pelanggan pada saat melakukan peminjaman uang dan tidak diketahui sepenuhnya oleh perusahaan

khususnya pada bagian *financial risk management* (Rahayu, 2014). Penyebab yang lain adalah *fraud* yang terjadi pada saat transaksi (*transaction fraud*) (Diadiushkin, Sandkuhl and Maiatin, 2019). Misalnya, ditemukan sebuah transaksi yang mencurigakan sehingga mengakibatkan penipuan dan berdampak pada pelanggan maupun perusahaan. Kedua hal tersebut tentu menjadi perhatian perusahaan layanan keuangan karena dapat mengakibatkan kerugian yang besar bagi perusahaan maupun pelanggan.

Cara yang dapat dilakukan untuk mengatasi penipuan pada perusahaan layanan keuangan adalah melakukan deteksi, baik itu deteksi data diri (profil) pelanggan maupun deteksi transaksi yang dilakukan oleh pelanggan. Deteksi dapat dilakukan dengan cara melihat dan mencocokkan kembali satu persatu data diri pelanggan yang akan melakukan peminjaman uang. Melalui data tersebut, perusahaan akan tahu mana pelanggan yang benar-benar memberikan data diri yang asli, dan mana yang palsu. Selain itu, dilakukan juga pengecekan berkala untuk setiap transaksi yang dilakukan oleh pelanggan. Apabila terdapat transaksi yang mencurigakan, maka dari perusahaan akan memberikan tanda pada transaksi tersebut dan memasukkan dalam kategori *fraud*. Namun, cara yang dilakukan tersebut masih manual dan konvensional, serta rawan terjadi *human error* ketika melakukan pengecekan (Ali et al., 2022). Selain itu, data pada perusahaan semakin lama semakin mengalami kenaikan. Tentu cara tersebut tidak efektif karena akan menyulitkan perusahaan dalam pengecekan dan pencocokkan data pelanggan.

Saat ini, terdapat cara lain yang lebih efektif digunakan untuk mendeteksi kemungkinan terjadi *fraud* pada layanan keuangan. Salah satu cara atau metode

yang digunakan adalah dengan menggunakan pendekatan *machine learning* (Zhao and Bai, 2022). Melalui *machine learning*, seluruh pengecekan dan validasi data dilakukan sepenuhnya oleh komputer. Perusahaan tidak perlu lagi melakukan pengecekan satu per satu data secara manual karena sudah sepenuhnya dikerjakan oleh model *machine learning* secara otomatis. Model pada algoritma *machine learning* dapat digunakan untuk mendeteksi *fraud* yang terjadi pada data diri pelanggan (*customer fraud*) (Ali et al., 2022) maupun deteksi *fraud* yang terjadi pada transaksi pelanggan (*transaction fraud*) (Liu, Chen and Yu, 2021). Algoritma yang digunakan pada setiap deteksi tentu berbeda dan menghasilkan akurasi yang berbeda juga.

Fokus pada penelitian ini adalah deteksi *fraud* yang terjadi pada transaksi di sebuah perusahaan layanan keuangan. Pada penelitian sebelumnya, deteksi *fraud* pada transaksi layanan keuangan telah dilakukan menggunakan algoritma *machine learning*, *deep learning*, dan kombinasi beberapa algoritma. Hasilnya dapat dikatakan baik untuk mendeteksi *fraud* pada transaksi keuangan. Akan tetapi, pada kasus *fraud*, pasti terjadi masalah *imbalanced dataset* (ketidakseimbangan kelas pada suatu dataset) dimana terjadi ketimpangan jumlah yang sangat besar antara kelas yang satu (*majority class*) dengan kelas yang lainnya (*minority class*). Hal tersebut pada *machine learning* dapat mengakibatkan *bias* (akurasi cenderung tinggi untuk *majority class*). Pada penelitian ini, peneliti akan membuat sebuah model untuk mendeteksi *fraud* pada transaksi pelanggan (*transaction fraud*) dimana pada kasus tersebut sering terjadi masalah *imbalanced dataset*. Model yang akan dibuat menggunakan teknik *Ensemble Learning*. *Ensemble learning* merupakan

sebuah teknik untuk mengkombinasikan *multiple models* pada *machine learning* untuk menghasilkan akurasi prediksi yang lebih tinggi daripada menggunakan *single model*. Kombinasi beberapa algoritma dapat meningkatkan akurasi dibandingkan dengan hanya menggunakan satu algoritma saja. Pada penelitian ini, peneliti akan mengkombinasikan tiga algoritma dasar *machine learning*, yaitu *Support Vector Machine (SVM)*, *k-Nearest-Neighbour (KNN)*, dan *Logistic Regression (LR)*. Dan untuk meta model, peneliti menggunakan algoritma *Random Forest (RF)*. Teknik yang digunakan untuk proses *ensemble learning* adalah *single stacking*.

1.2 Rumusan Masalah

Masalah *fraud* pada perusahaan layanan keuangan disebabkan oleh dua faktor. Faktor pertama adalah pemalsuan data diri pelanggan ketika akan melakukan peminjaman uang dengan cara pelanggan memberikan *fake document* kepada perusahaan. Faktor kedua adalah terjadinya transaksi yang tidak wajar dan mencurigakan sehingga itu diduga sebagai penipuan. Kedua masalah tersebut sama-sama memiliki potensi untuk merugikan perusahaan. Perusahaan yang tidak rutin melakukan pengecekan transaksi pelanggan akan mudah kecolongan apabila suatu saat terjadi *fraud* pada transaksi. Pengecekan transaksi secara manual menjadi cara yang kurang efektif untuk dilakukan oleh perusahaan. Selain itu, masalah *imbalanced dataset* juga berpengaruh terhadap model *machine learning* yang dibuat karena dapat mengakibatkan bias.

1.3 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Penelitian ini hanya berfokus pada deteksi *fraud* untuk transaksi pelanggan.
2. Data transaksi diambil dari sumber sekunder melalui sumber online.

1.4 Keaslian Penelitian

Keaslian penelitian ini adalah pada penggunaan model *ensemble learning*. Pada penelitian sebelumnya, model *ensemble learning* yang sering digunakan adalah *bagging* dan *boosting*. Pada kedua model tersebut terbatas hanya menggunakan satu metode pembelajaran yang sama (homogen). Sedangkan pada penelitian ini menggunakan *stacking* dimana metode pembelajaran yang digunakan bisa berbeda-beda (heterogen). Pemilihan kombinasi metode pembelajaran juga menjadi keaslian pada penelitian ini, yaitu menggunakan algoritma SVM, kNN, LR sebagai model dasar (*base model*) dan RF sebagai *meta model*. Penelitian ini juga membuat variasi *resampling* metode SMOTE dengan jumlah yang berbeda-beda sehingga dapat dibuat perbandingan mana jumlah sampling yang paling optimal untuk digunakan.

1.5 Manfaat Penelitian

Manfaat pada penelitian ini adalah sebagai berikut:

1. Bagi perusahaan layanan keuangan, penelitian ini memberi manfaat untuk membantu dalam proses deteksi transaksi setiap pelanggan dan menentukan transaksi mana yang *fraud* dan mana yang tidak *fraud*.
2. Pengembangan teori tentang *fraud detection*. Teori yang dikembangkan yaitu kombinasi algoritma *machine learning* yang nantinya akan diproses melalui metode *ensemble learning*. Hal tersebut dapat menjadi referensi pada penelitian

selanjutnya apabila ingin membangun sebuah sistem *fraud detection* berkaitan dengan pemilihan algoritma yang akan digunakan untuk mendeteksi *fraud*.

3. Pengembangan teori untuk mengatasi masalah *imbalanced dataset*.

1.6 Tujuan Penelitian

Penelitian ini bertujuan untuk mengatasi masalah *fraud* yang terjadi pada saat proses transaksi di sebuah perusahaan layanan keuangan. Cara mengatasinya adalah dengan menggunakan teknik *ensemble learning*. Melalui model yang telah dibuat, perusahaan layanan keuangan tidak perlu lagi secara manual melakukan pengecekan setiap transaksi pelanggan. Penelitian ini juga bertujuan untuk menghasilkan sebuah analisis berupa performa algoritma yang akan digunakan untuk memprediksi *fraud* pada transaksi. Selain itu, peneliti juga akan membahas terkait *imbalanced dataset* yang sering terjadi pada masalah *fraud detection*.

1.7 Sistematika Penulisan

Sistematika penulisan pada penelitian ini mengacu pada tata cara penulisan yang telah ditetapkan oleh Pascasarjana Universitas Atma Jaya Yogyakarta yang terdiri dari:

BAB I PENDAHULUAN

Pada bab ini akan menjelaskan masalah umum tentang laporan tesis yang meliputi latar belakang, rumusan masalah, batasan masalah, keaslian penelitian, manfaat penelitian, tujuan penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Pada bab ini akan menjelaskan tentang penelitian-penelitian terdahulu yang terkait dengan penelitian ini. Tinjauan pustaka diperoleh dari publikasi jurnal nasional, jurnal internasional, dan literasi ilmiah lainnya.

BAB III LANDASAN TEORI

Pada bab ini akan menjelaskan tentang seluruh teori yang akan digunakan pada penelitian ini. Sumber referensi berasal dari literasi ilmiah seperti jurnal, buku, karya ilmiah, serta literasi tambahan pada sumber online di internet.

BAB IV METODOLOGI PENELITIAN

Pada bab ini akan menjelaskan tentang metode dan langkah-langkah yang digunakan pada penelitian. Langkah-langkah tersebut dimulai dari tahap pengumpulan data, pengolahan data, dan pengujian model.

BAB V HASIL PENELITIAN DAN PEMBAHASAN

Pada bab ini akan menjelaskan detail tentang pengolahan data, pembuatan model, dan pengujian / evaluasi model. Setelah itu akan dilakukan analisis dan pembahasan terhadap model yang telah dibuat.

BAB VI KESIMPULAN DAN SARAN

Pada bab ini akan menjelaskan tentang kesimpulan yang diperoleh dari hasil penelitian ini. Selain itu akan dijelaskan juga saran untuk mengembangkan penelitian ini.