

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Ada berbagai jenis penipuan: penipuan asuransi, penipuan kartu kredit, penipuan pernyataan, penipuan sekuritas, dan lain-lain. Dari semuanya, penipuan kartu kredit adalah jenis yang paling umum. Ini didefinisikan sebagai penggunaan akun kartu kredit yang tidak sah. Ini terjadi ketika pemegang kartu dan penerbit kartu tidak mengetahui bahwa kartu tersebut sedang digunakan oleh pihak ketiga (Makki, 2019). Dalam beberapa tahun terakhir, volume transaksi kartu kredit telah meningkat secara dramatis karena mempopulerkan kartu kredit dan pesatnya perkembangan *e-services*, termasuk *e-commerce*, *e-finance*, dan pembayaran seluler (Zhang, David, Foroosh, & Gong, 2020).

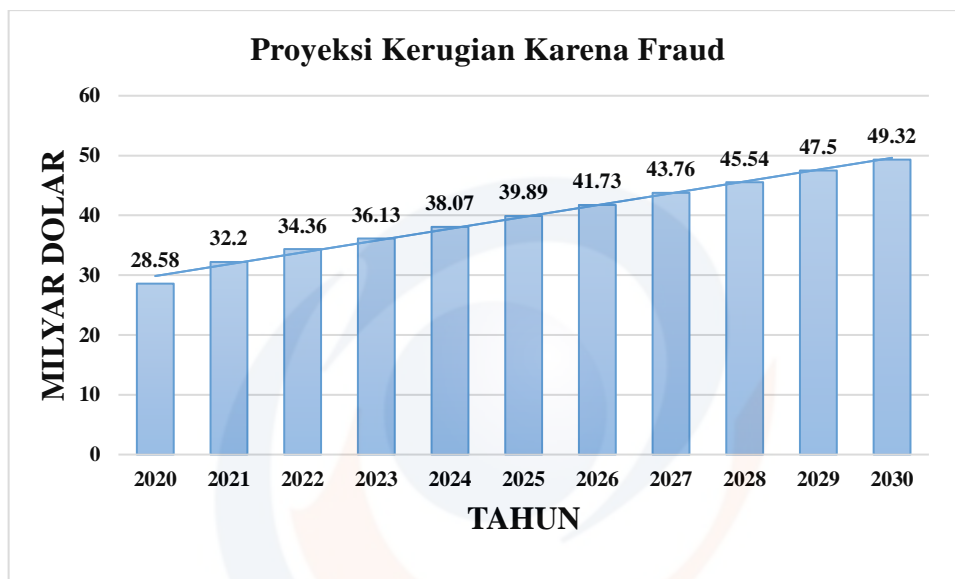
Banyak perusahaan yang mengembangkan bisnis menggunakan media online. Perusahaan ini menyediakan fasilitas untuk berbelanja online bagi pelanggan. Sebagian besar pelanggan menggunakan kartu kredit untuk membeli barang secara online. Dengan cara ini, beberapa pelanggan dapat menjadi pencuri yang telah mencuri kartu seseorang untuk melakukan transaksi online. Hal ini dinilai sebagai penipuan kartu kredit yang harus dideteksi (Paruchuri, 2017).

Kerugian yang ditimbulkan dari tindakan fraud ini bisa mencapai milyaran dolar. Menurut *The Nilson Report*, 28,65 miliar dolar hilang karena penipuan kartu pembayaran di seluruh dunia pada tahun 2019, meningkat 19,5% dari 23,97 miliar dolar pada tahun 2017. Kerugian diproyeksikan meningkat menjadi 35,67 miliar dolar dalam lima tahun dan 40,63 miliar dolar dalam 10 tahun (Itri, 2021).

Menurut Federal Trade Commission (FTC), pencurian identitas adalah jenis penipuan paling umum yang dilaporkan oleh konsumen, dengan penipuan kartu kredit menyumbang 33% dari semua keluhan pencurian identitas pada tahun 2020 (FTC, 2021). FTC lebih lanjut melaporkan bahwa penipuan kartu kredit telah meningkat sebesar 104% dalam dekade terakhir, dengan 1,4 juta laporan penipuan diajukan pada tahun 2020 saja (FTC, 2021). Lonjakan penipuan kartu kredit ini

telah menyebabkan lembaga keuangan menerapkan sistem deteksi penipuan yang lebih kuat.

Penipuan kartu kredit adalah masalah signifikan yang mempengaruhi konsumen, pedagang, dan lembaga keuangan secara global. Perusahaan kartu kredit dan lembaga keuangan telah menerapkan berbagai teknik deteksi penipuan untuk mencegah aktivitas penipuan, termasuk model penilaian kredit, pembelajaran mesin, dan teknik analitik data.



Gambar 1. Proyeksi kerugian karena fraud pada kartu kredit

Sistem deteksi penipuan memainkan peran penting dalam mengamankan lembaga keuangan dan meminimalkan risiko kerugian finansial. Apa pun sistem yang diterapkan, penipu akan terus mencari celah baru (Itoo, 2021). Dengan demikian, terus meningkatkan dan berinvestasi di sistem deteksi penipuan adalah suatu keharusan, itulah tantangan bagi semua lembaga keuangan.

Salah satu isu krusial yang perlu dipelajari lebih lanjut adalah masalah ketidakseimbangan kelas dataset. Dataset dianggap tidak seimbang ketika kelas tidak terwakili secara merata dalam dataset. Dalam dataset kartu kredit, jumlah transaksi penipuan jauh lebih sedikit daripada jumlah transaksi normal (N. Boutaher, 2020). Ini adalah masalah umum dalam proses deteksi penipuan karena kasus umum adalah transaksi yang sah dan yang jarang terjadi adalah transaksi penipuan. Kinerja teknik pembelajaran mesin sangat bergantung pada kualitas data pelatihan (Sanober, et al., 2021).

Karena masalah kelas yang tidak seimbang dalam dataset kartu kredit dunia nyata, peneliti menggunakan berbagai jenis teknik augmentasi data sebelum melatih pengklasifikasi untuk meningkatkan efisiensi model pembelajaran (Moreno-Barea, 2020). Banyak algoritma mesin pembelajaran *deep learning* diperkenalkan untuk menangani masalah klasifikasi yang tidak seimbang seperti yang dilakukan oleh (Al Olaimat, 2020), (Langevin, 2021) dan (Tamtama, 2022). Augmentasi data adalah proses menghasilkan data pelatihan baru dari data asli. Ini adalah teknik penting untuk meningkatkan keandalan dan kinerja model *deep learning*, membuatnya sangat efisien dalam aplikasi dunia nyata. Dengan menggunakan metode augmentasi data, kita dapat meningkatkan ukuran dan keragaman data pelatihan alih-alih mengumpulkan lebih banyak data (Strelcenia E, 2023).

Salah satu alasan utama untuk menghasilkan data buatan adalah *oversampling* kelas minoritas (Cordón, 2018). *Oversampling* dilakukan untuk mempelajari dataset yang tidak seimbang. Dalam banyak dataset dunia nyata, masalah ketidakseimbangan tetap ada. Dalam domain penipuan kartu kredit, jumlah kegiatan penipuan minimal jika dibandingkan dengan kegiatan hukum, sehingga menciptakan masalah kelas yang tidak seimbang (Benchaji, Douzi, & El Ouahidi, 2019). Dalam skenario ini, menjadi sulit bagi algoritma klasifikasi untuk mengidentifikasi kelas minoritas (transaksi penipuan) (Cai, Wang, Zhou, Xu, & Jing, 2019). Augmentasi data kelas minoritas adalah cara untuk mengatasi masalah kelas yang tidak seimbang. Ini dimungkinkan dengan menghasilkan contoh baru dengan sifat yang mirip dengan data asli. Augmentasi membantu menghindari masalah *underrepresentation* dari kelas minoritas, sekaligus menghindari *overfitting* juga (Sayed, Soliman, & Hassanien, 2021).

Model generatif, yang merupakan metode *unsupervised*, menjadi sangat populer dalam literatur terbaru dan telah diterapkan dalam berbagai sistem deteksi penipuan untuk penipuan kartu kredit dan asuransi, dan bahkan anti pencucian uang. Model-model ini dapat mempelajari representasi fitur mentah yang lebih dalam dan lebih kompleks dari ruang laten set pelatihan (Hilal, Gadsden, & Yawney, 2022). Di antara metode generatif, yang paling populer adalah *Generative Adversarial Networks* (GANs), dan *Variational Autoencoders* (VAEs). Namun,

penerapan model generatif mendalam untuk oversampling dataset yang tidak seimbang masih terbatas (Fajardo, et al., 2021).

1.2 Identifikasi Masalah

Berdasarkan latar belakang masalah dapat dirumuskan identifikasi masalah sebagai berikut :

- 1) Ketidakmampuan metode tradisional: Metode deteksi fraud yang umum digunakan mungkin tidak mampu mengidentifikasi pola fraud yang kompleks dan tidak terstruktur pada data transaksi kartu kredit. Masalah ini dapat menyebabkan banyak transaksi fraud yang tidak terdeteksi, menyebabkan kerugian bagi perusahaan dan pemegang kartu.
- 2) Overfitting dan Underfitting: Algoritma deteksi fraud tradisional mungkin menghadapi masalah overfitting atau underfitting. Overfitting dapat menyebabkan model terlalu kaku dan tidak dapat umumkan, sementara underfitting dapat menghasilkan model yang tidak cukup sensitif terhadap pola fraud.
- 3) Ketimpangan dataset: Dataset transaksi kartu kredit cenderung memiliki ketimpangan kelas, di mana sebagian besar transaksi adalah non-fraud dan hanya sebagian kecil yang fraud. Ketimpangan ini dapat menyebabkan kesulitan dalam mengenali dan mengklasifikasikan transaksi fraud, sehingga mengurangi performa deteksi.

1.3 Perumusan Masalah

Berdasarkan latar belakang dan identifikasi masalah yang telah diuraikan, dapat diambil beberapa rumusan masalah atau *Research Question* (RQ) sebagai berikut:

RQ1. Bagaimana cara menggunakan metode VAEs dan GANs dengan menggunakan pendekatan untuk mendeteksi kasus fraud pada transaksi kartu kredit?

RQ2. Bagaimana mengatasi kekurangan data fraud pada data transaksi kartu kredit untuk membangun model analisis data yang akurat?

RQ3. Bagaimana performa deteksi fraud menggunakan metode VAEs dan GANs ?

1.4 Batasan Penelitian

Untuk menghindari meluasnya pokok bahasan, penelitian ini dibatasi untuk hal-hal berikut :

- a. Lingkungan sistem (*environment*) untuk pembentukan model ML menggunakan bahasa pemrograman Python di Kaggle.com dengan 2 Core CPU, 13 GB RAM, Tesla P100 GPU.
- b. Dataset yang digunakan dalam penelitian ini digunakan dataset *Credit Card Fraud Detection*, yang dapat diunduh dari Kaggle . Himpunan data ini berisi transaksi, terjadi dalam dua hari, yang dilakukan pada September 2013 oleh pemegang kartu di Eropa. Dataset berisi 284.807 transaksi di mana 492 transaksi adalah penipuan dan sisanya asli. Mempertimbangkan angka-angkanya, kita dapat melihat bahwa dataset ini sangat tidak seimbang, di mana hanya 0,173% transaksi yang dicap sebagai penipuan

1.5 Tujuan Penelitian

Tujuan dari penelitian ini sebagai berikut:

- 1) Menerapkan metode VAEs dan GANs dengan pendekatan Domain Adaptation untuk mendeteksi kasus fraud pada data transaksi kartu kredit.
- 2) Membangun model analisis data yang akurat untuk deteksi fraud pada data transaksi kartu kredit.
- 3) Memberikan kontribusi untuk pengembangan metode deteksi fraud pada transaksi kartu kredit yang lebih efektif dan efisien.

1.6 Manfaat Penelitian

Adapun manfaat dari penelitian adalah sebagai berikut :

- 1) Meningkatkan efektivitas deteksi fraud: Penelitian ini dapat meningkatkan efektivitas deteksi fraud pada transaksi kartu kredit, sehingga dapat membantu institusi keuangan dan pengguna kartu kredit dalam mencegah kerugian finansial yang disebabkan oleh penipuan
- 2) Meningkatkan efisiensi pengolahan data: Dalam penelitian ini, digunakan metode deep learning yang mampu membuat data sintetis untuk pengolahan data yang lebih baik

- 3) Memberikan pandangan baru dalam deteksi fraud: Dalam penelitian ini, digunakan metode VAEs dan GANs dalam deteksi fraud pada transaksi kartu kredit, yang memberikan pandangan baru dalam deteksi fraud pada industri keuangan.

1.7 Sistematika Penulisan

Secara garis besar, Proposal Tesis ini terdiri 3 (tiga) BAB sebagai berikut:

BAB 1 – Pendahuluan

Menjelaskan latar belakang penelitian, identifikasi masalah penelitian, rumusan masalah penelitian, tujuan penelitian dan manfaat penelitian.

BAB 2 – Landasan Teori

Menjelaskan teori-teori yang relevan dengan penelitian ini, metode dan hasil dari para peneliti sebelumnya.

BAB 3 – Metodologi Penelitian

Menjelaskan tentang alur penelitian, kerangka konseptual penelitian, teknik pengumpulan data, pembentukan model dan kerangka kerja yang diusulkan, implementasi model dan kerangka kerja, dan analisis dan evaluasi data.