

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Fraud atau penipuan merupakan tindakan ilegal yang dapat merugikan suatu organisasi atau perusahaan. *Fraud* adalah suatu perbuatan melawan atau melanggar hukum yang dilakukan oleh orang dari dalam atau dari luar organisasi, dengan maksud untuk memperkaya atau mendapatkan keuntungan diri sendiri, orang lain, atau badan hukum lain yang secara langsung atau tidak langsung merugikan pihak lain [1]. *Fraud* dapat mengancam keberlangsungan perekonomian suatu negara. Laporan [2] dari *Association of Certified Fraud Examiners (ACFE)* tahun 2018 menunjukkan bahwa kerugian yang dialami suatu organisasi karena *fraud* sekitar 5% dari pendapatan kotor suatu organisasi.

Tindakan *fraud* sering terjadi pada berbagai jenis transaksi dan praktik akuntansi di bidang perbankan. Perkembangan ekonomi yang begitu pesat menuntut terjadinya transaksi keuangan setiap saat sehingga pembayaran tunai dinilai telah ketinggalan zaman. Oleh sebab itu bermunculan berbagai macam metode pembayaran transaksi keuangan non tunai salah satunya menjadi pilihan adalah kartu kredit. Kartu kredit saat ini dianggap sebagai salah satu bagian dari gaya hidup masyarakat modern [3].

Data yang dirilis oleh Bank Indonesia sampai akhir tahun 2018 menunjukkan bahwa penggunaan kartu kredit dalam transaksi non tunai menduduki peringkat ketiga setelah uang elektronik dan debit. Menurut data

dari AKKI (2020), jumlah transaksi menggunakan kartu kredit dari tahun 2009 sampai 2019 selalu mengalami peningkatan. Pada tahun 2019, jumlah transaksi menggunakan kartu kredit menembus angka 340 juta rupiah.

Semakin meningkatnya penggunaan kartu kredit sebagai media transaksi menyebabkan meningkatnya kasus *fraud*. Tindakan *fraud* berakibat pada kerugian keuangan serta penurunan tingkat kepercayaan nasabah terhadap pihak bank. Penanganan tindakan *fraud* dapat dilakukan dengan audit kecurangan (*Fraud Audit*) yang bertujuan untuk menemukan kecurangan [1].

Dibutuhkan teknologi yang dapat mengklasifikasikan transaksi *fraud* dan tidak *fraud* dengan cepat, akurat, dan efisien karena teknologi tersebut akan mengklasifikasikan berdasarkan model yang sudah dilatih sebelumnya. Salah satunya yaitu dengan menerapkan pendekatan *deep learning* dengan algoritme *Artificial Neural Network* (ANN). Algoritme ANN dipilih merujuk pada dataset yang digunakan berupa data transaksi berbentuk data tabular.

Merujuk pada penelitian yang telah dilakukan sebelumnya, algoritme ANN terbukti mampu mengklasifikasikan suatu masalah dengan baik. Penelitian [4] membuktikan bahwa algoritme ANN menunjukkan performa yang baik untuk klasifikasi peminjaman nasabah bank. Penelitian [5] menerapkan algoritme ANN dengan metode *Backpropagation* untuk mendeteksi penipuan kartu kredit. Hasil penelitian menunjukkan hasil yang memuaskan dengan nilai akurasi sebesar 99.92%. Penelitian lain terkait

deteksi penipuan kartu kredit dilakukan pada penelitian [6] dimana algoritme ANN memiliki nilai akurasi yang paling tinggi di antara algoritme *Support Vector Machine* (SVM) dan *k-Nearest Neighbors* (kNN). Berdasarkan latar belakang yang telah diuraikan, penulis akan melakukan penelitian dengan judul “Penerapan Algoritme *Artificial Neural Network* (ANN) untuk *Fraud Detection* pada Transaksi Kartu Kredit”.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, maka didapat rumusan masalah sebagai berikut:

1. Atribut apa saja yang digunakan untuk melakukan klasifikasi?
2. Dapatkah pendekatan *deep learning* dengan algoritme *Artificial Neural Network* (ANN), mengklasifikasikan transaksi kartu kredit yang *fraud* dan tidak *fraud*?
3. Sejauh mana performa yang bisa dihasilkan?

1.3 Batasan Penelitian

Agar dapat memaksimalkan kinerja berdasarkan hasil dari rumusan masalah, maka diberikan batasan-batasan sebagai berikut:

1. Mengklasifikasikan transaksi kartu kredit ke dalam label yang telah ditentukan yaitu *fraud* dan tidak *fraud* (normal).
2. Menggunakan dataset dari salah satu bank di Indonesia yang didapat dari kompetisi FinHack tahun 2018 tanpa menggunakan dataset pembanding.

3. Algoritme yang digunakan adalah *Artificial Neural Network (ANN)* tanpa menggunakan algoritme pembandingan.

1.4 Maksud dan Tujuan Penelitian

Maksud dan tujuan penelitian ini dilakukan adalah sebagai berikut:

1. Mengetahui atribut apa saja yang mempengaruhi suatu transaksi dikatakan *fraud* atau tidak *fraud*.
2. Menerapkan pendekatan *deep learning* dengan algoritme *Artificial Neural Network (ANN)* untuk mengklasifikasikan transaksi kartu kredit secara cepat, akurat, dan efisien.
3. Mengetahui performa yang dihasilkan ANN dalam mengklasifikasikan transaksi kartu kredit.

1.5 Manfaat Penelitian

Adapun manfaat yang dapat diambil dari penelitian ini adalah sebagai berikut:

1. Membantu pihak bank dalam mengklasifikasikan transaksi kartu kredit ke dalam kategori *fraud* dan tidak *fraud* sehingga dapat menjadi bahan pertimbangan untuk menyetujui atau menolak transaksi.
2. Penelitian ini diharapkan memberikan gambaran untuk mendeteksi penipuan transaksi kartu kredit menggunakan algoritme ANN.
3. Mengetahui performa pendekatan *deep learning* dengan algoritme *Artificial Neural Network (ANN)* dalam mengklasifikasikan transaksi kartu kredit.

1.6 Metode Penelitian

1.6.1 Metode Pengumpulan Data

Pengumpulan data pada penelitian ini didapat dari dataset transaksi salah satu bank di Indonesia. Dataset yang digunakan memiliki jumlah transaksi sebanyak 13.125 transaksi. Dari data tersebut terdapat 27 atribut dan 1 kelas yang menandakan transaksi *fraud* atau tidak *fraud*.

1.6.2 Metode Analisis

Pada tahap analisis dilakukan beberapa tahapan yang mengacu pada konsep KDD (*Knowledge Discovery in Databases*). Tahapan tersebut yaitu sebagai berikut:

1. *Data Selection*

Pada tahap ini, dilakukan pemilihan himpunan data atau memfokuskan pada subset variabel atau sampel data yang akan digunakan untuk penelitian dan menghapus data yang memiliki nilai konstan/tetap.

2. *Pre-processing*

Pada tahap ini, dilakukan beberapa proses untuk menangani data kosong (*handling missing value*). Data yang kosong akan menyebabkan hasil analisa tidak akurat. Untuk menangani hal ini, data kosong pada suatu atribut akan diisi menggunakan median dari atribut tersebut.

3. *Transformation*

Proses *transformation* dilakukan dengan mengubah data menjadi bentuk yang sesuai untuk proses *data mining*. Rentang nilai yang tidak seimbang pada setiap atribut, akan berpengaruh terhadap hasil klasifikasi. Sehingga diperlukan proses normalisasi agar tidak terjadi perbedaan nilai yang jauh antar atribut. Proses normalisasi akan mengubah data ke dalam skala tertentu yang lebih kecil. Pada penelitian ini, normalisasi dilakukan dengan menerapkan metode *Z Score Normalization*.

Proses *balancing data* juga dilakukan pada tahap ini. Pada dataset yang tersedia, terdapat 1 kelas yang berisi data transaksi *fraud* sejumlah 12.215 data dan transaksi tidak *fraud* sejumlah 910 data. Untuk mengatasi data yang tidak seimbang, digunakan metode SMOTE untuk *resampling data*.

4. *Data Mining*

Pada penelitian ini, proses *data mining* dilakukan dengan melakukan klasifikasi transaksi menggunakan pendekatan *deep learning* dengan algoritme *Artificial Neural Network* (ANN). Dataset yang telah melewati proses *pre-processing* dan *transformation* akan dilatih menggunakan algoritme ANN sehingga akan menghasilkan sebuah model klasifikasi yang siap digunakan untuk mendeteksi penipuan transaksi kartu kredit.

5. Evaluation

Pada tahap ini, model yang dihasilkan dari proses *data mining* akan dilakukan evaluasi menggunakan metode *confusion matrix*, *learning curve*, dan kurva ROC yang akan dijelaskan lebih lengkap pada Metode Pengujian.

1.6.3 Metode Perancangan

Metode perancangan yang diterapkan pada penelitian ini yaitu menggunakan diagram alir atau *flowchart* untuk menunjukkan langkah-langkah yang dilakukan untuk membuat suatu model klasifikasi menggunakan algoritme *Artificial Neural Network (ANN)*.

1.6.4 Metode Implementasi

Pada tahap ini dilaksanakan implementasi metode dan algoritme yang sudah direncanakan. Implementasi berupa pembuatan program berbasis web untuk mendeteksi penipuan transaksi kartu kredit menggunakan bahasa pemrograman Python dengan *library* pendukung.

1.6.5 Metode Pengujian

Proses pengujian dilakukan dengan membagi data menjadi 2 bagian yaitu data latih (*data training*) dan data uji (*data testing*) dengan komposisi 80% data latih dan 20% data uji. Model yang

dihasilkan pada proses pelatihan akan diuji menggunakan set data uji menggunakan metode *confusion matrix* dengan menghitung nilai *accuracy*, *error rate*, *precision*, *recall*, *specificity*, dan *f1 score*. Selain *confusion matrix*, *learning curve* dan kurva ROC juga akan diterapkan pada penelitian ini.

1. Accuracy

Accuracy digunakan untuk menghitung berapa persen *classifier* dapat mengklasifikasi dengan benar antara transaksi *fraud* dan transaksi tidak *fraud*. Nilai *accuracy* ditunjukkan oleh Persamaan 1.1.

$$Accuracy = \frac{TN + TP}{TN + TP + FP + FN} \quad \dots\dots\dots (1.1)$$

2. Error Rate

Error rate digunakan untuk menghitung berapa persen tingkat kesalahan atau kekeliruan *classifier* dalam melakukan klasifikasi. Nilai *error rate* ditunjukkan oleh Persamaan 1.2.

$$Error\ rate = \frac{FP + FN}{P + N} \quad \dots\dots\dots (1.2)$$

3. Precision

Precision akan menghitung rasio benar transaksi *fraud* dibandingkan dengan keseluruhan transaksi yang diklasifikasikan *fraud*. Nilai *precision* ditunjukkan oleh Persamaan 1.3.

$$Precision = \frac{TP}{TP + FP} \dots\dots\dots (1.3)$$

4. *Recall* atau *Sensitivity*

Recall digunakan untuk menghitung rasio klasifikasi transaksi *fraud* dibandingkan dengan keseluruhan transaksi yang benar *fraud*. Nilai *recall* ditunjukkan oleh Persamaan 1.4.

$$Recall = \frac{TP}{TP + FN} \dots\dots\dots (1.4)$$

5. *Specificity*

Specificity digunakan untuk menghitung nilai kebenaran *classifier* dalam mengklasifikasi transaksi tidak *fraud* dibandingkan dengan keseluruhan data transaksi tidak *fraud*. Nilai *specificity* ditunjukkan oleh Persamaan 1.5.

$$Specificity = \frac{TN}{TN + FP} \dots\dots\dots (1.5)$$

6. *F1 Score*

F1 score digunakan untuk membandingkan rata-rata *precision* dengan *recall* yang dibobotkan atau dengan kata lain, *f1 score* berfungsi untuk mengukur keseimbangan antara *precision* dengan *recall*. Nilai *f1 score* ditunjukkan oleh Persamaan 1.6.

$$F1\ Score = \frac{2 * recall * precision}{recall + precision} \dots\dots\dots (1.6)$$

7. *Learning Curve*

Learning curve atau kurva pembelajaran merupakan *plot* kinerja pembelajaran model selama proses *training*. Kurva ini digunakan untuk mendiagnosis masalah dalam proses *training*. Pada penelitian ini, akan diterapkan *learning curve* untuk meninjau nilai *loss* dan *accuracy* selama proses *training*.

8. Kurva ROC

Receiver Operating Characteristic atau ROC merupakan kurva yang membandingkan antara *true positive rate* dan *false positive rate*. Kurva ROC menggambarkan visualisasi untuk semua *threshold* klasifikasi yang mungkin. Area di bawah kurva ROC disebut dengan AUC. AUC digunakan sebagai standar pengukuran untuk kinerja klasifikasi.

Keterangan istilah:

1. *True Positive (TP)* merupakan transaksi *fraud* yang diklasifikasikan dengan benar sebagai transaksi *fraud*.
2. *True Negative (TN)* merupakan transaksi tidak *fraud* yang diklasifikasikan dengan benar sebagai transaksi tidak *fraud*.
3. *False Positive (FP) – Type 1 error* merupakan transaksi tidak *fraud* yang diklasifikasikan sebagai transaksi *fraud*.
4. *False Negative (FN) – Type 2 error* merupakan transaksi *fraud* yang diklasifikasikan sebagai transaksi tidak *fraud*.

1.7 Sistematika Penulisan

Adapun sistematika penulisan yang digunakan pada penelitian ini adalah sebagai berikut:

BAB I : PENDAHULUAN

Bab ini berisi penjelasan umum mengenai penelitian yang akan dilakukan, meliputi latar belakang, rumusan masalah, batasan masalah, maksud dan tujuan penelitian, metode penelitian, dan sistematika penulisan.

BAB II : LANDASAN TEORI

Bab ini berisi tinjauan pustaka terkait penelitian yang akan dilakukan serta menguraikan secara terperinci teori-teori yang berhubungan dengan pokok permasalahan yang dijadikan sebagai acuan dan landasan dalam penelitian ini.

BAB III : ANALISIS DAN PERANCANGAN

Bab ini akan menguraikan secara detail analisis dan rancangan penelitian yang akan diterapkan.

BAB IV : HASIL DAN PEMBAHASAN

Bab ini berisi pembahasan dan pemaparan hasil dari penelitian yang dilakukan.

BAB V : PENUTUP

Bab ini berisi kesimpulan yang didapat dari pembahasan pada bab-bab sebelumnya serta saran agar penelitian selanjutnya bisa menghasilkan hasil yang lebih baik.