Progresif: Jurnal Ilmiah Komputer
https://ojs.stmik-banjarbaru.ac.id/index.php/progresif/index
Jl. Ahmad Yani, K.M. 33,5 - Kampus STMIK Banjarbaru
Loktabat – Banjarbaru (Tlp. 0511 4782881), e-mail: puslit.stmikbjb@gmail.com
e-ISSN: 2685-0877

Penerapan Algoritma *Decision Tree* Dalam Deteksi *Fraud* Transaksi Kartu Kredit

DOI: http://dx.doi.org/10.35889/progresif.v21i2.2751

Gladisya Devina Agustine[1], Irwansyah[2*]
Teknik Informatika, Universitas Muhammadiyah Prof. DR. Hamka, Jakarta, Indonesia
*e-mail *Corresponding Author:* irwansyah@uhamka.ac.id

Abstract

Credit card fraud poses a serious threat in digital financial systems. Manual detection of suspicious transactions has become ineffective due to detecting fraudulent transactions using the Decision Tree algorithm. The dataset used was obtained from Kaggle and underwent preprocessing and attribute selection. The model was tested under four data split scenarios: 90:10, 80:20, 70:30, and 60:40. Performance evaluation was conducted using a confusion matrix with accuracy, precision, and recall metrics. The results show that the 60:40 data split yielded the best performance, with an accuracy of 97,47%, precision of 86,34%, and recall of 78,67%. These findings indicate that the Decision Tree algorithm can produce highly accurate classification results even without applying data balancing techniques.
Kata kunci: *Credit Card; Fraud Detection; Decision Tree; Data Mining.*

Abstrak

Penipuan dalam transaksi kartu kredit merupakan ancaman serius dalam sistem keuangan digital. Deteksi secara manual terhadap transaksi yang mencurigakan menjadi tidak efektif seiring dengan meningkatnya volume data. Penelitian ini bertujuan untuk mengembangkan model klasifikasi untuk mendeteksi transaksi fraud menggunakan algoritma *Decision Tree* C4.5. Dataset yang digunakan diperoleh dari Kaggle dan telah melalui proses praproses dan seleksi atribut. Pengujian dilakukan dengan empat skenario pembagian data *training* dan data *testing*, yaitu 90:10, 80:20, 70:30, dan 60:40. Evaluasi performa dilakukan menggunakan *confusion matrix* dengan metrik akurasi, presisi, dan *recall*. Hasil menunjukkan bahwa pembagian data 60:40 memberikan performa terbaik dengan nilai akurasi sebesar 97,47%, presisi 86,34%, dan recall 78,67%. Model ini menunjukkan bahwa algoritma *Decision Tree* mampu memberikan hasil klasifikasi yang sangat baik bahkan tanpa teknik penyeimbangan data.
Kata kunci: *Kartu Kredit; Deteksi Penipuan; Decision Tree; Data Mining.*

1. Pendahuluan

Perkembangan teknologi digital yang sangat cepat di sektor keuangan telah membawa kemudahan bagi masyarakat dalam melakukan berbagai macam transaksi, termasuk penggunaan kartu kredit. Seiring dengan kemunculan *platform* jual-beli dan transaksi daring, pola konsumsi individu maupun bisnis menjadi lebih efisien dan instan [1]. Meski begitu, kemajuan ini juga disertai tantangan besar, yaitu meningkatnya risiko terjadinya penipuan kartu kredit [2]. Penipuan semacam ini merupakan kejahatan yang tak hanya menimbulkan kerugian bagi pemilik kartu, tetapi juga mengancam stabilitas dan citra lembaga keuangan [3]. Oleh sebab itu, dibutuhkan sistem yang mampu mendeteksi transaksi penipuan secara tepat dan seketika di era digital saat ini.

Kejahatan penipuan kini semakin kompleks dan sulit diidentifikasi dengan cara manual, sehingga diperlukan sistem pendeteksian otomatis yang dapat mengenali pola transaksi mencurigakan dengan efisiensi tinggi [4]. Berbagai teknik deteksi telah dikembangkan, namun masih banyak menghadapi permasalahan seperti akurasi prediksi yang rendah dan ketidakseimbangan dalam distribusi data [5].

Sebagai pendekatan alternatif, algoritma *Decision Tree* dipilih dalam penelitian ini karena memiliki sejumlah kelebihan dibanding metode *data mining* lainnya [6]. Kelebihan tersebut antara lain adalah kemudahan dalam pemahaman, penerapan yang sederhana, tidak memerlukan pengetahuan teknis yang mendalam, serta kemampuannya untuk menangani data numerik maupun kategorikal dalam skala besar, termasuk data yang tidak seimbang [7]. *Decision Tree* mengklasifikasikan data berdasarkan atribut-atribut tertentu melalui pembentukan struktur pohon keputusan yang sistematis dan logis [8]. Dengan membangun pohon keputusan berdasarkan aktivitas transaksi pengguna, metode ini diharapkan mampu mengidentifikasi potensi penipuan dengan lebih akurat dan efisien.

Penelitian ini bertujuan untuk membangun dan mengevaluasi model klasifikasi berbasis algoritma *Decision Tree* dalam mendeteksi transaksi penipuan pada kartu kredit. Manfaat dari penelitian ini adalah memberikan solusi klasifikasi yang mudah dipahami serta mendukung proses pengambilan keputusan secara cepat, khususnya bagi pengguna non-teknis.

2. Tinjauan Pustaka

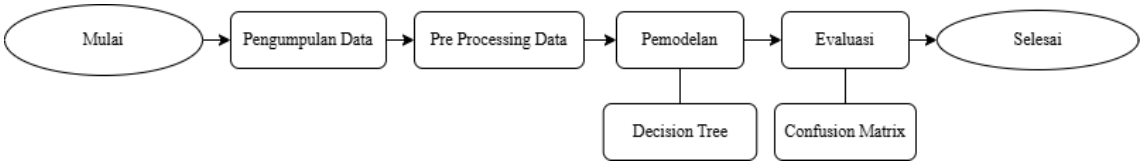
Beberapa penelitian sebelumnya telah menunjukkan keberhasilan penggunaan algoritma *Decision Tree* dalam mendeteksi penipuan transaksi kartu kredit. Penelitian oleh [6] mencatat bahwa *Decision Tree* mampu mencapai akurasi hingga 98% dalam mendeteksi transaksi ilegal, namun penelitian tersebut belum membahas struktur pohon keputusan secara rinci dan keterkaitannya dengan interpretasi hasil klasifikasi. Sementara itu, penelitian [9] menggabungkan *Decision Tree* C4.5 dengan teknik *SMOTE* untuk menyeimbangkan data, tetapi menyebabkan pohon yang dihasilkan menjadi sangat kompleks dan sulit dipahami secara visual. Studi oleh [10] menggunakan *Random Forest* dan *Neural Networks*, memberikan akurasi tinggi namun tidak memberikan gambaran struktur klasifikasi yang bisa dijelaskan secara visual. Selanjutnya, [11] menerapkan pendekatan *hybrid machine learning*, namun kurang fokus pada model yang mudah dijelaskan seperti *Decision Tree*. Dalam penelitian [12] menggunakan *Random Forest* dan *Decision Tree*, namun tidak mendalami pengaruh pemilihan atribut dan interpretasi logika pohon terhadap hasil klasifikasi. Penelitian oleh [13] juga menggunakan *Decision Tree* dalam mendeteksi penipuan kartu kredit. Mereka menunjukkan bahwa akurasi dapat mencapai 99,05% sebelum parameter tuning dan turun menjadi 74,76% sesudahnya. Hal ini mengindikasikan bahwa pengaturan parameter dan pemilihan atribut memiliki peran penting dalam performa model.

Penelitian-penelitian tersebut umumnya memiliki fokus pada peningkatan sesuai akurasi dan penggunaan teknik *balancing* data. Namun, sebagian besar tidak akan menekankan aspek transparansi hasil klasifikasi atau kemudahan pemahaman bagi pengguna non-teknis. Berdasarkan studi-studi terdahulu, dapat disimpulkan bahwa *Decision Tree* merupakan salah satu algoritma yang banyak digunakan untuk mendeteksi penipuan pada transaksi kartu kredit karena kemampuannya dalam menghasilkan model klasifikasi yang transparan dan mudah dipahami. Meskipun beberapa penelitian mencapai akurasi tinggi, namun interpretasi terhadap logika pohon dan keterkaitannya dengan hasil klasifikasi belum banyak dikaji secara mendalam.

Penelitian ini menghadirkan kebaruan dengan menekankan pada aspek interpretabilitas hasil klasifikasi menggunakan *Decision Tree* dalam *RapidMiner*, serta mengevaluasi performa model tanpa menerapkan teknik *balancing* data. Dengan pendekatan ini, diharapkan model yang dihasilkan lebih mudah dianalisis secara visual, serta tetap memberikan performa klasifikasi yang sangat baik.

3. Metodologi

Alur penelitian disajikan seperti pada Gambar 1.



Gambar 1. Alur Penelitian

Dalam penelitian ini, digunakan pendekatan data mining dengan fokus pada teknik klasifikasi menggunakan algoritma *Decision Tree*. Proses penelitian dilakukan secara bertahap, dimulai dari pengumpulan data, dilanjutkan dengan tahapan *pre-processing*, kemudian implementasi algoritma *Decision Tree* menggunakan aplikasi *RapidMiner* [14]. Setelah model terbentuk, kinerjanya dievaluasi menggunakan *Confusion Matrix* untuk mengukur tingkat akurasi, presisi, dan *recall*.

3.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini bersumber dari situs *Kaggle* dengan format Excel dan mencakup sebanyak 10.000 transaksi kartu kredit. Dari seluruh data tersebut, terdapat 7.195 transaksi yang teridentifikasi sebagai *fraud*, sementara sisanya sebanyak 9.805 transaksi tergolong sebagai transaksi normal. Seluruh data kemudian diimpor ke dalam aplikasi *RapidMiner* untuk proses analisis lebih lanjut. Dataset ini dapat diakses secara publik melalui https://www.kaggle.com/datasets/anurag629/credit-card-fraud-transaction-data

Credit Card Fraud Transaction Data

Data Card | Code (7) | Discussion (1) | Suggestions (0)

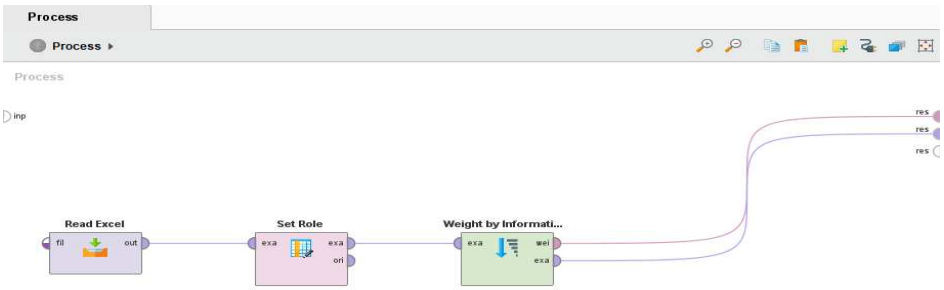
Transaction ID	Date	Day of Week	Time	
95680 unique values	14-Oct-20 50% 13-Oct-20 50% Other (2) 0%	Wednesday 50% Tuesday 50% Other (26) 0%	13 6% 16 6% Other (88405) 88%	Visa Mas
#3577 209	14-Oct-20	Wednesday	19	Vis
#3039 221	14-Oct-20	Wednesday	17	Mas
#2694 780	14-Oct-20	Wednesday	14	Vis
#2640 960	13-Oct-20	Tuesday	14	Vis
#2771 031	13-Oct-20	Tuesday	23	Vis
#3446 698	13-Oct-20	Tuesday	20	Mas
#3652 191	13-Oct-20	Tuesday	18	Vis
#3161 927	13-Oct-20	Tuesday	18	Mas
#3025 809	13-Oct-20	Tuesday	23	Mas
#3413 696	14-Oct-20	Wednesday	23	Mas
#2667 502	13-Oct-20	Tuesday	11	Vis
#3474 192	14-Oct-20	Wednesday	1	Mas

Gambar 2. Contoh Data pda Dataset Kaggle

3.2 Pre-processing Data

Tahapan *pre-processing* dilakukan untuk memastikan bahwa data yang digunakan dalam kondisi bersih dan siap digunakan dalam proses klasifikasi. Beberapa tahapan *pre-processing* yang dilakukan sebagai berikut:

1) Seleksi Data



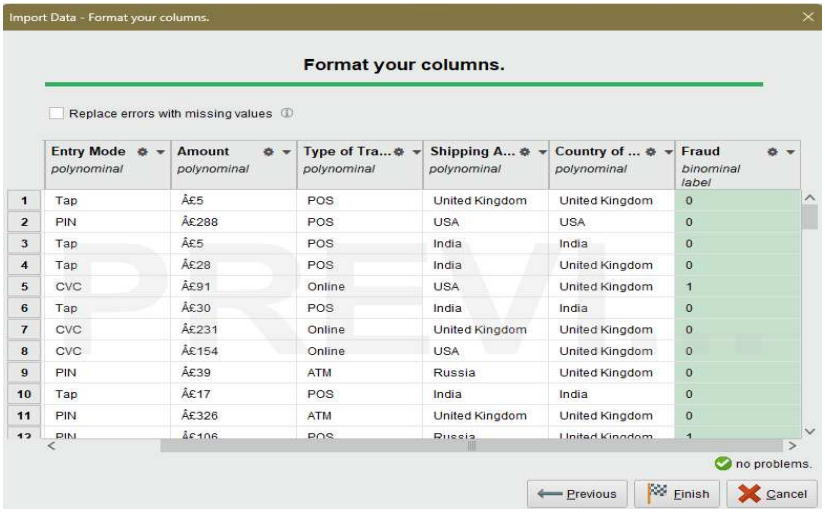
Gambar 3. Proses Weight by Information Gain

Dalam penelitian ini, dilakukan proses seleksi atribut sebagai bagian dari tahap *pre-processing* data. Dari total 14 atribut yang terdapat pada dataset asli hasil unduhan dari *Kaggle*. Untuk menentukan atribut yang paling relevan, digunakan metode *Weight by Information Gain* yang tersedia pada perangkat lunak *RapidMiner*. Metode ini menghitung bobot setiap atribut terhadap target klasifikasi, yaitu *Fraud*, berdasarkan seberapa besar informasi yang diberikan oleh atribut tersebut dalam membedakan antara kelas *fraud* dan *non-fraud* [15]. Hasil seleksi disajikan seperti pada Gambar 3. Berdasarkan hasil perhitungan (Gambar 3) tersebut, 7 atribut dengan nilai informasi tertinggi dipilih untuk dilanjutkan ke tahap klasifikasi, yaitu:

Tabel 1. Atribut Data Seleksi

Atribut Data Sebelum Seleksi	Atribut Data Setelah Seleksi
Transaction ID	Time
Date	Entry Mode
Day of Week	Shipping Address
Time	Country of Residence
Type of Card	Amount
Entry Mode	Country of Transaction
Amount	Fraud
Type of Transaction	
Merchant Group	
Country of Residence	
Gender	
Age	
Bank	
Fraud	

2) Transformasi Data



Gambar 4. Transformasi Data

Transformasi Data (Gambar 4) dilakukan untuk memastikan bahwa data siap digunakan dalam proses analisis. Data di ubah ke dalam format yang sesuai untuk pemrosesan di *RapidMiner*. Konversi tipe data dilakukan, dari integer (*real*) menjadi kategori (binominal) untuk atribut "*Fraud*" yang membutuhkan klasifikasi. Sebagai bagian dari proses ini, kolom label ditambahkan untuk membedakan antara transaksi penipuan dan tidak penipuan.

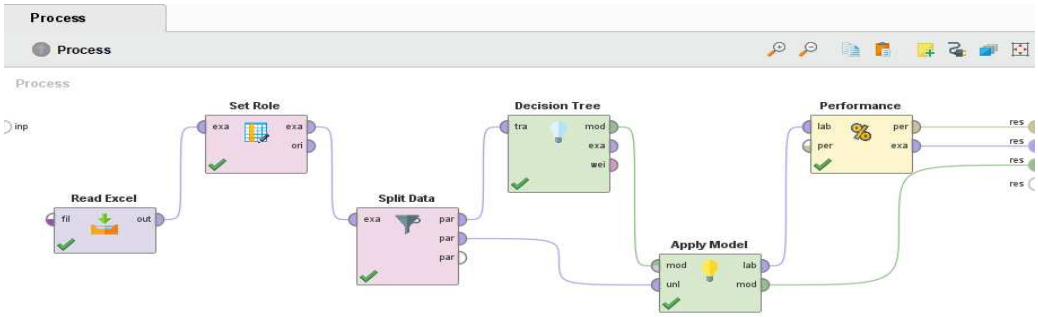
3) Data Siap Olah

Tabel 2. Data Siap Olah

Time	Entry Mode	Amount	Type of Transaction	Shipping Address	Country of Residence	Fraud
19	Tap	Â£5	POS	United Kingdom	United Kingdom	0
17	PIN	Â£288	POS	USA	USA	0
14	Tap	Â£5	POS	India	India	0
14	Tap	Â£28	POS	India	United Kingdom	0
23	CVC	Â£91	Online	USA	United Kingdom	1
20	Tap	Â£30	POS	India	India	0
18	CVC	Â£231	Online	United Kingdom	United Kingdom	0
18	CVC	Â£154	Online	USA	United Kingdom	0
23	PIN	Â£39	ATM	Russia	United Kingdom	0
23	Tap	Â£17	POS	India	India	0
11	PIN	Â£326	ATM	United Kingdom	United Kingdom	0
1	PIN	Â£106	POS	Russia	United Kingdom	1
21	PIN	Â£21	ATM	United Kingdom	United Kingdom	0
20	PIN	Â£211	ATM	United Kingdom	United Kingdom	0
…	…	…	…	…	…	…
15	Tap	Â£351	POS	India	India	0

3.3 Proses RapidMiner

Setelah data siap, proses klasifikasi dilakukan menggunakan algoritma *Decision Tree* yang tersedia dalam RapidMiner. Algoritma ini bekerja dengan membangun pohon keputusan berdasrkan atribut-atribut pada dataset untuk memisahkan data ke dalam kelas "*fraud*" dan "*non-fraud*". Model yang terbentuk kemudian digunakan untuk memprediksi data uji menggunakan operator *Apply Model* dan operator *Performance* untuk mengukur hasil akurasi.



Gambar 5. Proses RapidMiner

Proses klasifikasi diatas dimulai dengan mengimpor data menggunakan operator *Read Excel*, di mana dataset yang telah dibersihkan dan dipilih atributnya dimasukkan ke dalam lingkungan kerja *RapidMiner*. Setelah data dimuat, digunakan operator *Set Role* untuk menetapkan peran masing-masing atribut. Pada tahap ini, atribut *Fraud* diatur sebagai label (kelas target) yang akan diprediksi, sementara atribut lain digunakan sebagai atribut *input*.

Selanjutnya, data dibagi menjadi data latih dan data uji menggunakan operator *Split Data*. Proses pembagian ini menggunakan beberapa skenario rasio seperti 90:10, 80:20, 70:20, dan 60:40 guna menguji pengaruh proporsi data terhadap performa model. Setelah data dibagi, bagian data latih digunakan dalam operator *Decision Tree* untuk membentuk model klasifikasi berdasarkan algoritma Decision Tree. Model ini bekerja dengan membangun struktur pohon keputusan dari atribut-atribut yang paling informatif dalam membedakan kelas *fraud* dan *non-fraud*.

Model yang telah terbentuk kemudian diterapkan ke data uji menggunakan operator *Apply Model*. Proses ini menghasilkan prediksi terhadap data uji berdasarkan pola yang telah dipelajari oleh model dari data latih. Terakhir, performa model dievaluasi seperti akurasi, presisi, dan *recall*, berdasarkan perbandingan antara hasil prediksi dan data aktual pada data uji.

3.4 Evaluasi Model

Evaluasi model dilakukan berdasarkan *confusion matrix*, yang menunjukkan nilai *True Positive* (TP), yaitu jumlah prediksi yang benar bahwa suatu data termasuk ke dalam kelas positif, *False Positive* (FP), yaitu jumlah prediksi yang salah bahwa suatu data termasuk ke dalam kelas positif, *False Negative* (FN), yaitu jumlah prediksi yang salah bahwa suatu data termasuk ke dalam kelas negatif, dan *True Negative* (TN), yaitu jumlah prediksi yang benar bahwa suatu data termasuk ke dalam kelas negatif [16]. Nilai-nilai ini digunakan untuk menghitung semua metrik evaluasi model. Rumus-rumus yang digunakan sebagai berikut [17]:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

Klasifikasi performa model klasifikasi dapat ditentukan berdasarkan rentang nilai akurasi yang diperoleh. Sebagaimana Tabel 1. *Klasifikasi Performa Berdasarkan Nilai Akurasi* berikut [18].

Tabel 3. Klasifikasi Performa Berdasarkan Nilai

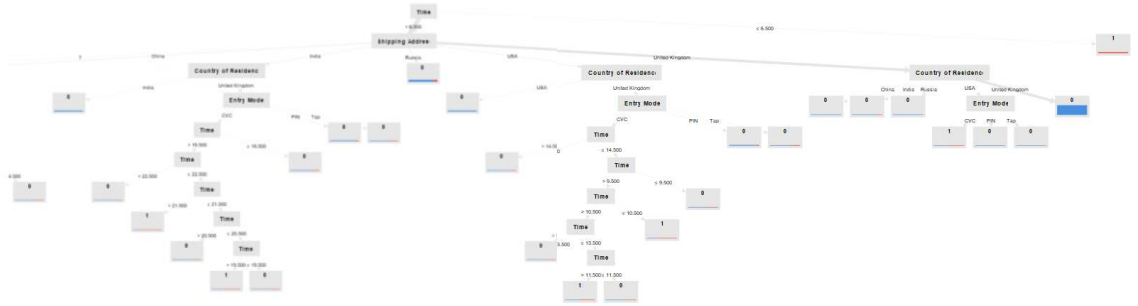
Rentang Nilai Akurasi (%)	Klasifikasi Performa
90-100	Sangat Baik
80-90	Baik
70-80	Cukup
60-70	Buruk
<= 60	Sangat Buruk

4. Hasil dan Pembahasan

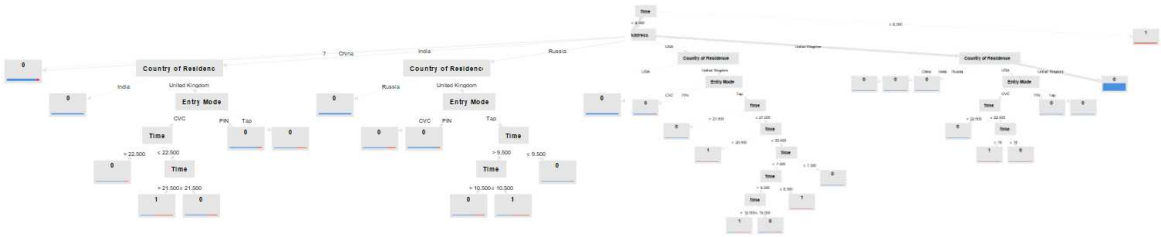
4.1 Hasil Klasifikasi

Proses klasifikasi pada penelitian ini dilakukan menggunakan algoritma *Decision Tree*, dengan empat skenario pembagian data, yaitu 90:10, 80:20, 70:30, dan 60:40. Masing-masing skenario menghasilkan struktur pohon keputusan yang berbeda sesuai dengan jumlah data pelatihan dan pola atribut yang dipelajari oleh model. Visualisasi pohon keputusan dari masing-masing skenario ditampilkan pada Gambar 6 hingga Gambar 9.

Gambar 6. Pohon Keputusan Skenario 60:40



Gambar 7. Pohon Keputusan Skenario 70:30



Gambar 8. Pohon Keputusan Skenario 80:20

Gambar 9. Pohon Keputusan Skenario 90:10

Secara umum, atribut *Time*, *Entry Mode*, dan *Shipping Address* secara konsisten muncul sebagai simpul awal maupun simpul internal dalam pohon, menandakan perannya yang signifikan dalam membedakan transaksi *fraud* dan *non-fraud*. Meski struktur dan kedalaman pohon sedikit bervariasi antar scenario, pola klasifikasi yang dihasilkan tetap menunjukkan konsistensi logis dan keterbacaan yang baik.

Dari keempat skenario, struktur pohon pada skenario 60:40 dipilih sebagai fokus utama untuk evaluasi lebih lanjut karena menghasilkan performa terbaik berdasarkan pengukuran akurasi, presisi, dan *recall*.

4.2 Hasil Evaluasi

Untuk menilai kinerja model yang dibangun, dilakukan pengujian dengan empat variasi rasio pembagian antara data pelatihan dan data pengujian, yaitu 90:10, 80:20, 70:30, dan 60:40. Proses evaluasi menggunakan *Confusion Matrix* guna memperoleh hasil pengukuran berupa akurasi, presisi, dan *recall*.

Tabel 4. Hasil Evaluasi Confusion Matrix berikut menunjukkan hasil evaluasi kinerja model klasifikasi berdasarkan nilai akurasi, presisi, dan *recall* dari masing-masing skenario pembagian data.

Tabel 4. Hasil Evaluasi Confusion Matrix

Rasio (%)	Akurasi	Presisi	Recall
60:40	**97,47%**	**86,34%**	**78,67%**
70:30	97,44%	88,44%	75,81%
80:20	97,38%	89,63%	73,50%
90:10	97,21%	88,27%	72,44%

Berdasarkan hasil evaluasi diatas, model terbaik diperoleh pada skenario pembagian data 60:40 dengan nilai akurasi sebesar 97,47%, presisi 86,34%, dan *recall* 78,67%. Jika diklasifikasikan berdasarkan **Tabel 3. Klasifikasi Performa Berdasarkan Nilai Akurasi**, maka model ini berada pada kategori "Sangat Baik" karena berada dalam rentang 90%-100%. Perhitungan nilai akurasi, presisi, dan *recall* pada skenario 60:40 dapat dijelaskan menggunakan *confusion matrix* yang dihasilkan oleh *RapidMiner* sebagai berikut:

1) *True Positive* (TP) sebanyak 354 transaksi *fraud* yang berhasil terdeteksi secara benar sebagai penipuan.
2) *False Positive* (FN) yaitu sebanyak 56 transaksi *non-fraud* yang secara keliru diprediksi sebagai penipuan.
3) *True Negative* (TN) yaitu sebanyak 5494 transaksi *non-fraud* yang berhasil diklasifikasikan dengan tepat sebagai transaksi valid.
4) *False Negative* (FN) yaitu sebanyak 96 transaksi *fraud* yang tidak berhasil terdeteksi, sehingga diklasifikasikan sebagai transaksi valid.

Perhitungan menggunakan rumus:

$$Accuracy = \frac{354 + 5494}{354 + 5494 + 56 + 96} \times 100\% = 97{,}47\%$$

$$Precision = \frac{354}{354 + 56} \times 100\% = 86{,}34\%$$

$$Recall = \frac{354}{354 + 96} \times 100\% = 78{,}67\%$$

Hasil perhitungan secara manual tersebut konsisten dengan hasil evaluasi dari *RapidMiner*, menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam mendeteksi transaksi *fraud*.

4.3 Pembahasan

Berdasarkan hasil evaluasi pada Tabel 4. Hasil Evaluasi Confusion Matrix, dapat disimpulkan bahwa pembagian data dengan rasio 60:40 menghasilkan performa terbaik, dengan akurasi 97,47%, presisi 86,34%, dan *recall* 78,67%. Nilai akurasi yang tinggi menunjukkan bahwa sebagian besar data berhasil diklasifikasikan dengan tepat oleh model. Namun, nilai *recall* yang relatif lebih rendah dibandingkan akurasi mengindikasikan bahwa masih terdapat beberapa data *fraud* yang tidak terdeteksi, yang dalam konteks keamanan finansial dapat menjadi perhatian penting.

Algoritma C4.5 menyusun pohon keputusan dengan memprioritaskan atribut paling informatif. Perhitungan *Information Gain* mengidentifikasi variabel *Time*, *Shipping Address, dan Country of Residence* sebagai faktor yang paling berpengaruh dalam pemisahan kelas.

Jika dibandingkan dengan temuan pada penelitian-penelitian sebelumnya yang juga menunjukkan efektivitas algoritma *Decision Tree* dan *Random Forest* dalam mendeteksi transaksi *fraud*. Misalnya, [17] mencatat akurasi sebesar 96% menggunakan *Random Forest,* dan [9] melaporkan akurasi 95% dengan *Decision Tree* dan SMOTE.

Model pada penelitian ini tidak menggunakan tekik *balancing* tambahan seperti *SMOTE*, namun tetap menghasilkan akurasi yang sangat baik. Hal ini menunjukkan bahwa data asli memiliki pola yang cukup kuat untuk diklasifikasikan oleh *Decision Tree* secara efektif.

5. Simpulan

Berdasarkan hasil pengujian dengan empat skenario pembagian data (90:10, 80:20, 70:30, dan 60:40), diperoleh bahwa skenario pembagian data 60:40 menghasilkan performa terbaik dengan nilai akurasi sebesar 97,47%, presisi 86,34%, dan recall 78,67%. Hasil ini menunjukkan bahwa algoritma *Decision Tree* mampu mengklasifikasikan transaksi *fraud* dan *non-fraud* dengan tingkat keakuratan yang sangat baik untuk kasus klasifikasi transaksi data tidak seimbang, bahkan tanpa menerapkan teknik *oversampling*. Selain itu, pemilihan atribut

menggunakan metode *Weight by Information Gain* terbukti efektif dalam meningkatkan kualitas klasifikasi dengan memprioritaskan atribut yang paling informatif.

Dengan demikian, model klasifikasi yang dibangun dalam penelitian ini dapat digunakan sebagai pendekatan alternatif dalam mendeteksi transaksi penipuan kartu kredit secara otomatis dan efisien. Untuk pengembangan lebih lanjut, disarankan agar penelitian mendatang mengeksplorasi penggunaan algoritma lain dan mempertimbangkan teknik *balancing* data serta *tuning* parameter untuk meningkatkan performa *recall.*

Referensi:

[1] Ismawati, "Penggunaan Data Mining Untuk Mendeteksi Penipuan Transaksi Kartu Kredit Algoritma Decision Tree," *JAMASTIKA*, vol. 4, no.1, pp. 95–101, Apr. 2025, doi: https://doi.org/10.35473/jamastika.v4i1.3632.

[2] D. Hendarsyah, S. Tinggi, I. Ekonomi, and S. Bengkalis, "Analisis Perilaku Konsumen Dan Keamanan Kartu Kredit Perbankan," *Jurnal Perbankan Syariah*, vol. 1, no. 1, pp. 85–96, Apr. 2020, doi: https://doi.org/10.46367/jps.v1i1.204.

[3] W. Nugraha, D. Risdiansyah, D. Purwaningtias, T. Hidayatulloh, and S. Suhada, "Kombinasi Tomek-Link Dan Smote Untuk Mengatasi Ketidakseimbangan Kelas Pada Credit Card Fraud," *Jurnal Larik: Ladang Artikel Ilmu Komputer*, vol. 2, No.2, no. 2, pp. 32–40, Dec. 2022, doi: 10.31294/larik.v2i2.1789.

[4] A. Joshi, "Decision Tree Algorithm for Credit Card Fraud Detection," *Webology*, vol. 18, No. 4, pp. 2055–2061, 2021, doi: 10.29121/web/v18i4/103.

[5] T. Kumar, "Comparison of Logistic Regression and Decision Tree method for Credit Card Fraud Detection," *Int J Res Appl Sci Eng Technol*, vol. 9, no. 5, pp. 680–683, May 2021, doi: 10.22214/ijraset.2021.34241.

[6] I. Werdiningsih *et al.*, "Identifying Credit Card Fraud in Illegal Transactions Using Random Forest and Decision Tree Algorithms," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 12, no. 3, pp. 477–484, Nov. 2023, doi: 10.32736/sisfokom.v12i3.1730.

[7] P. T. S. Ningsih, M. Gusvarizon, and R. Hermawan, "Analisis Sistem Pendeteksi Penipuan Transaksi Kartu Kredit dengan Algoritma Machine Learning," *Jurnal Teknologi Informatika dan Komputer*, vol. 8, no. 2, pp. 386–401, Sep. 2022, doi: 10.37012/jtik.v8i2.1306.

[8] Hammed and S. Jumoke, "An implementation of decision tree algorithm augmented with regression analysis for fraud detection in credit card," *International Journal of Computer Science and Information Security*, vol. 18, no. 2, pp. 79–88, Feb. 2020, Accessed: Mar 29, 2025. [Online]. Available: https://sites.google.com/site/ijcsis/

[9] L. Darell Perwara and F. A. Bachtiar, "Penerapan Algoritma Decision Tree C4.5 Untuk Deteksi Fraud Pada Kartu Kredit dengan Oversampling Synthetic Minority Technique (SMOTE)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 4, no. 8, pp. 2664–2669, Aug. 2020, Accessed: Mar 29, 2025. [Online]. Available: http://j-ptiik.ub.ac.id

[10] S. Wijaya, W. Wesly, K. Ginting, and A. Dharma, "Analysis of Credit Card Fraud Detection Performance Using Random Forest Classifier & Neural Networks Model," *Engineering and Technology Journal*, vol. 9, no. 02, pp. 3516–3520, Feb. 2024, doi: 10.47191/etj/v9i02.11.

[11] E. F. Malik, K. W. Khaw, B. Belaton, W. P. Wong, and X. Chew, "Credit Card Fraud Detection Using a New Hybrid Machine Learning Architecture," *Mathematics*, vol. 10, no. 9, pp. 1–16, May 2022, doi: 10.3390/math10091480.

[12] A. Kurniawan and Y. Yulianingsih, "Pendugaan Fraud Detection pada kartu kredit dengan Machine Learning," *KILAT*, vol. 10, Number 2, no. 2, pp. 320–325, Oct. 2021, doi: 10.33322/kilat.v10i2.1482.

[13] D. Shah and L. Kumar Sharma, "Credit Card Fraud Detection using Decision Tree and Random Forest," *ITM Web of Conferences*, vol. 53, no. ICDSIA-2023, pp. 2–12, 2023, doi: 10.1051/itmconf/20235302012.

[14] D. Elisa Sinaga *et al.*, "KLIK: Kajian Ilmiah Informatika dan Komputer Analisis Data Mining Algoritma Decision Tree Pada Prediksi Persediaan Obat (Studi Kasus : Apotek Franch Farma)," *Kajian Ilmiah Informatika dan Komputer*, vol. 2, no. 4, pp. 123–131, Feb. 2022, Accessed: May 25, 2025. [Online]. Available: https://djournals.com/klik

[15] P. Tiwari, S. Mehta, N. Sakhuja, J. Kumar, and A. K. Singh, “Credit Card Fraud Detection using Machine Learning: A Study,” *Techincal Report*, pp. 2–10, Aug. 2021, Accessed: May 25, 2025. [Online]. Available: http://arxiv.org/abs/2108.10005

[16] Ainurrohmah, “Akurasi Algoritma Klasifikasi pada Software Rapidminer dan Weka,” *Prosiding Seminar Nasional Matematika*, vol. 4, No. 1, pp. 493–499, 2021, doi: https://journal.unnes.ac.id/sju/index.php/prisma/.

[17] T. S. Lestari and D. A. N. Sirodj, “Klasifikasi Penipuan Transaksi Kartu Kredit Menggunakan Metode Random Forest,” *Jurnal Riset Statistika*, vol. 1, no. 2, pp. 160–167, Feb. 2022, doi: 10.29313/jrs.v1i2.525.

[18] A. I. Sang, E. Sutoyo, and I. Darmawan, “Analisis Data Mining Untuk Klasifikasi Data Kualitas Udara Dki Jakarta Menggunakan Algoritma Decision Tree Dan Support Vector Machine Data Minning”, *e-Proceeding of Engineering*, vol. 8, no. 5, pp. 8954–8963, Oct. 2021, Accessed: May 29, 2025. [Online]. Available: https://openlibrary.telkomuniversity.ac.id/pustaka/170546/analisis-data-mining-untuk-klasifikasi-data-kualitas-udara-dki-jakarta-menggunakan-algoritma-decision-tree-dan-support-vector-machine.html