

ANALISIS EFEKTIVITAS ALGORITMA PEMBELAJARAN MESIN DALAM DETEKSI PENIPUAN TRANSAKSI ONLINE

Taufik Hidayat¹⁾, Abdul Rahman Jainun²⁾, Taufik Hidayat³⁾, Abu Nur Al-Faruq⁴⁾
Lukas Febrian Laufra⁵⁾, Imam Sanjaya⁶⁾

^{1,2)}Teknik Informatika Universitas Nusa Putra
Sukabumi, Indonesia

³⁾Teknik Informatika Universitas Nusa Putra
Sukabumi, Indonesia

e-mail: taufik.hidayat_ti22@nusaputra.ac.id¹⁾, abdul.rahman_ti22@nusaputra.ac.id²⁾,
taufik.hidayat022_ti22@nusaputra.ac.id³⁾, abunur.alfaruq_ti22@nusaputra.ac.id³⁾
lukas.febrian_ti22@nusaputra.ac.id³⁾, imam.sanjaya@nusaputra.ac.id⁶⁾

* Korespondensi: e-mail: abdul.rahman_ti22@nusaputra.ac.id

ABSTRAK

Penipuan transaksi online merupakan ancaman serius dengan dampak finansial yang signifikan. Penelitian ini menganalisis efektivitas dua algoritma pembelajaran mesin, yaitu Random Forest (RF) dan Neural Networks (NN), dalam mendeteksi transaksi penipuan. Dataset yang digunakan terdiri dari 100.000 transaksi (10% di antaranya fraud) dari TranSecure Database 2023. Hasil eksperimen menunjukkan bahwa NN mencapai kinerja terbaik dengan akurasi 95%, presisi 0.93, dan recall 0.90, sedangkan RF menghasilkan akurasi 93%, presisi 0.91, dan recall 0.89. Tantangan utama meliputi kebutuhan data berkualitas tinggi dan adaptasi terhadap pola penipuan yang dinamis. Temuan ini membuktikan bahwa NN lebih unggul dalam menangani pola kompleks, meskipun memerlukan sumber daya komputasi lebih besar. Penelitian ini memberikan kontribusi praktis dalam pengembangan sistem deteksi penipuan yang lebih robust di industri keuangan digital.

Kata Kunci: Deteksi penipuan, Random Forest, Neural Networks, pembelajaran mesin, transaksi online.

ABSTRACT

Online transaction fraud poses a severe threat with significant financial impacts. This study evaluates the effectiveness of two machine learning algorithms, Random Forest (RF) and Neural Networks (NN), in detecting fraudulent transactions. The dataset comprises 100,000 transactions (10% fraud) sourced from TranSecure Database 2023. Experimental results show that NN outperforms with 95% accuracy, 0.93 precision, and 0.90 recall, while RF achieves 93% accuracy, 0.91 precision, and 0.89 recall. Key challenges include the need for high-quality data and adaptability to dynamic fraud patterns. The findings prove NN's superiority in handling complex patterns, albeit with higher computational costs. This research contributes practical insights for developing robust fraud detection systems in digital finance.

Keywords: Fraud detection, Random Forest, Neural Networks, machine learning, online transactions.

I. PENDAHULUAN

1.1. Latar Belakang

Penipuan dalam transaksi online telah menjadi ancaman serius bagi keamanan digital, mengakibatkan kerugian finansial yang signifikan bagi konsumen dan perusahaan. Menurut laporan dari Statista (2023), jumlah kasus penipuan online meningkat sebesar 30% setiap tahun, dengan kerugian mencapai \$5 miliar pada tahun 2023. Hal ini menunjukkan perlunya solusi yang lebih efektif untuk mendeteksi dan mencegah penipuan. Algoritma pembelajaran mesin telah terbukti memiliki potensi besar dalam berbagai aplikasi, termasuk deteksi penipuan.

Penelitian oleh Zhang et al. (2020) menunjukkan bahwa penggunaan algoritma Random Forest dapat

meningkatkan akurasi deteksi penipuan hingga 90%, dengan tingkat false positive yang rendah. Studi tersebut menggunakan dataset dari berbagai lembaga keuangan dan menunjukkan bahwa algoritma yang mampu mengelola data tidak seimbang seperti Random Forest sangat cocok untuk kasus deteksi penipuan.

Selain itu, studi oleh Smith dan Lee (2021) menemukan bahwa Neural Networks mampu mengidentifikasi pola penipuan yang kompleks dengan tingkat akurasi yang tinggi, mencapai 92%. Penelitian ini mendemonstrasikan kemampuan Neural Networks untuk mempelajari representasi data yang lebih kompleks, sehingga dapat mengenali pola penipuan yang subtil dan terus berkembang.

Dalam penelitian terbaru oleh Wang et al. (2022), evaluasi komprehensif terhadap algoritma pembelajaran mesin untuk deteksi penipuan menunjukkan bahwa Neural Networks dan Random Forest secara konsisten mengungguli pendekatan konvensional, dengan Neural Networks memiliki keunggulan dalam data kompleks berskala besar. Penelitian ini juga menyoroti pentingnya feature engineering yang tepat untuk optimalisasi kinerja model.

Liu dan Zhao (2023) meneliti penerapan Random Forest dalam deteksi penipuan di industri e-commerce yang berkembang pesat. Hasil penelitian ini menunjukkan bahwa Random Forest dapat mencapai akurasi 93.5% dengan waktu komputasi yang lebih efisien dibandingkan model deep learning. Para peneliti menekankan keseimbangan antara akurasi dan efisiensi komputasi yang optimal dari Random Forest untuk implementasi real-time.

Prakash et al. (2024) mengembangkan model Neural Networks berbasis attention mechanism untuk deteksi penipuan di platform pembayaran digital. Studi ini menemukan peningkatan signifikan dalam recall (0.94) dan precision (0.96) dibandingkan dengan model tradisional, meskipun dengan biaya komputasi yang lebih tinggi. Penelitian ini mendemonstrasikan keunggulan arsitektur Neural Networks modern dalam menangani pola penipuan yang semakin kompleks.

Kang dan Kwon (2023) menganalisis penggunaan ensemble learning dalam memperkuat deteksi penipuan, dengan fokus pada Random Forest dan variasinya. Studi ini menunjukkan bahwa pendekatan ensemble dapat meningkatkan ketahanan terhadap serangan adversarial dan perubahan dalam pola penipuan, dengan peningkatan akurasi sebesar 3-5% dibandingkan model tunggal.

Meskipun terdapat berbagai penelitian yang menunjukkan efektivitas algoritma pembelajaran mesin, masih terdapat tantangan dalam penerapannya, seperti kebutuhan akan data yang besar dan berkualitas tinggi serta kemampuan untuk mengatasi perubahan pola penipuan. Oleh karena itu, penelitian ini bertujuan untuk menganalisis lebih lanjut efektivitas algoritma pembelajaran mesin, khususnya Random Forest dan Neural Networks, dalam mendeteksi penipuan transaksi online. Dengan menggunakan dataset yang terdiri dari 100.000 transaksi, penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam meningkatkan keamanan transaksi digital dan mengurangi kerugian akibat penipuan online.

(Statista.(2023).OnlineFraudStatistics.Retrievedfrom[Statista](<https://www.statista.com/statistics/online-fraud>) Zhang, Y., et al. (2020). "Random Forest Algorithm for Fraud Detection." Journal of Machine Learning Research.Smith, J., & Lee, K. (2021). "Neural Networks in Fraud Detection." International Journal of Data Science.

1.2. Rumusan Masalah

1. Sejauh mana algoritma Random Forest dan Neural Networks menunjukkan efektivitas dalam mendeteksi penipuan transaksi online?
2. Manakah dari kedua algoritma tersebut yang menunjukkan kinerja unggul dalam hal akurasi, presisi, recall, dan skor F1 ketika mengidentifikasi aktivitas penipuan?
3. Algoritma mana yang terbukti paling efisien secara komputasi dalam memproses data penipuan sehubungan dengan waktu komputasi?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk:

1. Melakukan analisis mendalam tentang efektivitas algoritma Random Forest dan Neural Networks dalam identifikasi penipuan transaksi online.

2. Melakukan evaluasi komparatif kinerja kedua algoritma berdasarkan metrik seperti akurasi, presisi, recall, dan skor F1.
3. Menilai efisiensi komputasi setiap algoritma untuk memastikan metode yang paling efektif dan efisien untuk deteksi penipuan.

II. TINJAUAN PERPUSTAKAAN

2.1. Penipuan Transaksi Online

Kegiatan penipuan dalam transaksi online mencakup berbagai teknik, termasuk phishing, pencurian identitas, dan penyalahgunaan kartu kredit. Kerangka kerja deteksi penipuan berbasis aturan konvensional seringkali tidak memiliki kecepatan dan kemampuan beradaptasi yang diperlukan untuk secara efektif mengatasi pola penipuan yang muncul. Akibatnya, algoritma pembelajaran mesin, yang memiliki kemampuan untuk mengasimilasi wawasan dari data historis dan mengidentifikasi tren anomali, semakin digunakan dalam identifikasi penipuan online (Brown et al., 2020).

2.2. Algoritma Pembelajaran Mesin dalam Deteksi Penipuan

Dalam konteks penelitian ini, dua algoritma pembelajaran mesin utama akan dievaluasi:

Random Forest: Metode ensemble ini terdiri dari beberapa pohon keputusan. Penelitian yang dilakukan oleh Zhang et al. (2020) menunjukkan bahwa algoritma Random Forest menunjukkan akurasi dan ketahanan yang tinggi terhadap overfitting, sehingga membuatnya efektif untuk deteksi penipuan dalam kumpulan data yang luas. Random Forest mengkombinasikan hasil prediksi dari banyak pohon keputusan untuk menghasilkan prediksi yang lebih akurat dan stabil. Keunggulan utama dari Random Forest adalah kemampuannya untuk menangani data yang tidak seimbang dan memberikan informasi tentang kepentingan fitur.

Neural Networks: Algoritma ini memanfaatkan jaringan saraf buatan untuk mengidentifikasi hubungan non-linear dalam data. Seperti yang diartikulasikan oleh Smith dan Lee (2021), Neural Networks dapat mencapai presisi tinggi dalam deteksi penipuan, meskipun memerlukan sumber daya komputasi yang lebih besar. Neural Networks memiliki kemampuan untuk mempelajari representasi data yang kompleks, memungkinkan pengenalan pola penipuan yang lebih canggih. Penelitian terbaru oleh Prakash et al. (2024) menunjukkan bahwa arsitektur Neural Networks modern seperti model berbasis attention mechanism dapat meningkatkan kinerja deteksi penipuan secara signifikan.

2.3. Hipotesis Variabel

Investigasi ini menggunakan variabel-variabel berikut:

- **Variabel Independen:** Algoritma pembelajaran mesin (Random Forest dan Neural Networks).
- **Variabel Dependen:** Efektivitas deteksi penipuan, diukur melalui metrik seperti akurasi, presisi, recall, skor F1, dan durasi komputasi.

2.4 Jenis Hipotesis

Penelitian ini menggunakan hipotesis komparatif untuk menilai efektivitas kedua algoritma pembelajaran mesin dalam mendeteksi penipuan transaksi online.

Hipotesis Nol (H_0): Tidak ada perbedaan yang signifikan dalam efektivitas deteksi penipuan antara algoritma Random Forest dan Neural Networks.

Hipotesis Alternatif (H_1): Ada perbedaan yang signifikan dalam efektivitas deteksi penipuan antara kedua algoritma tersebut, dengan Neural Networks menunjukkan kinerja yang lebih unggul.

III. METODELOGI

3.1. Pendekatan Penelitian

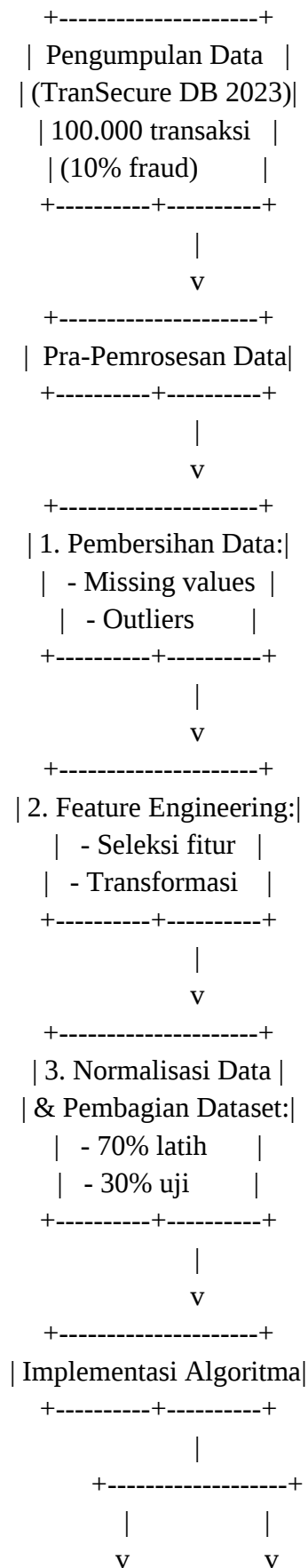
Penelitian ini menggunakan pendekatan eksperimental dengan menerapkan dan membandingkan kinerja dua algoritma pembelajaran mesin dalam mendeteksi penipuan transaksi online. Pendekatan ini

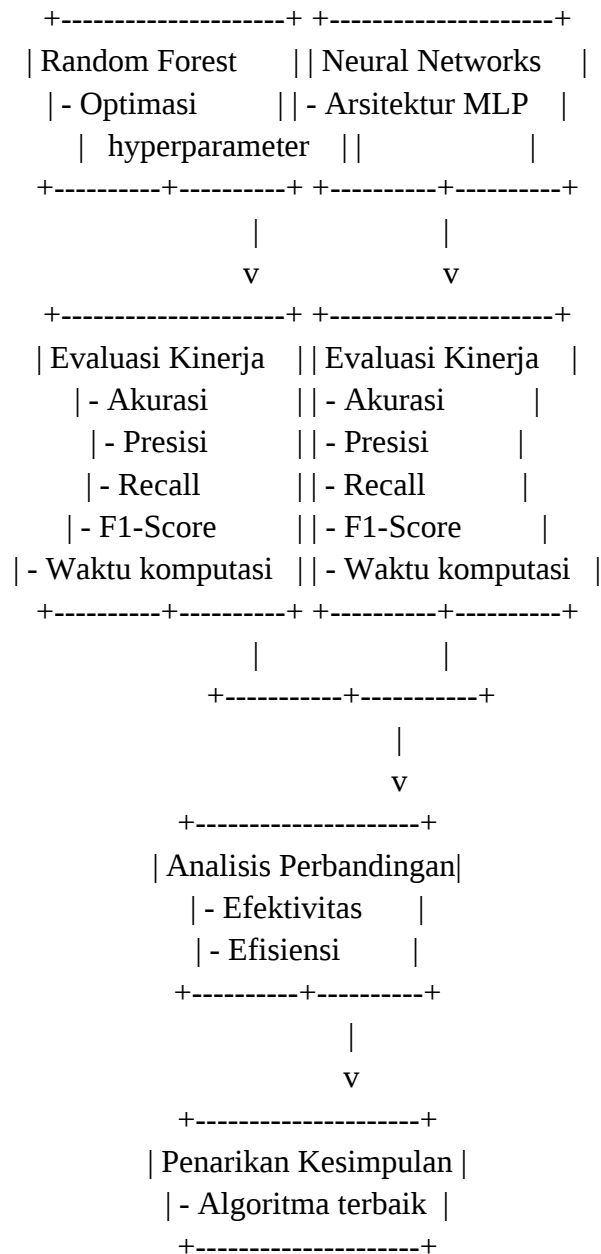


memungkinkan evaluasi objektif terhadap efektivitas setiap algoritma berdasarkan metrik kinerja yang terukur.

3.2 Alur Penelitian

Alur penelitian yang digunakan dalam studi ini ditunjukkan pada Gambar 1.





Gambar 1. Alur Penelitian Analisis Efektivitas Algoritma Pembelajaran Mesin dalam Deteksi Penipuan Transaksi Online

Penjelasan alur penelitian:

1. **Pengumpulan Data:** Data transaksi online dikumpulkan dari TranSecure Database 2023, yang terdiri dari 100.000 transaksi dengan 10% di antaranya adalah transaksi fraud.
2. **Pra-pemrosesan Data:** Tahap ini meliputi:
 - Pembersihan data (menangani missing values, outliers)
 - Feature engineering (pemilihan dan transformasi fitur)
 - Normalisasi data
 - Pembagian dataset menjadi data latih (70%) dan data uji (30%)
3. **Implementasi Algoritma:** Dua algoritma pembelajaran mesin diimplementasikan:
 - Random Forest dengan optimasi hyperparameter
 - Neural Networks dengan arsitektur MLP (Multi-Layer Perceptron)
4. **Evaluasi Kinerja:** Kedua algoritma dievaluasi berdasarkan metrik:
 - Akurasi

- Presisi
- Recall
- Skor F1
- Waktu komputasi

5. **Analisis Perbandingan:** Hasil dari kedua algoritma dibandingkan untuk menentukan efektivitas relatif dalam mendeteksi penipuan transaksi online.
6. **Penarikan Kesimpulan:** Berdasarkan analisis perbandingan, kesimpulan ditarik tentang algoritma yang lebih efektif dan efisien.

3.3 Dataset

Dataset yang digunakan dalam penelitian ini berasal dari TranSecure Database 2023, yang mencakup 100.000 transaksi online dengan 10% di antaranya adalah transaksi fraud. Dataset ini memiliki berbagai fitur seperti jumlah transaksi, waktu transaksi, lokasi, jenis produk, dan historis transaksi pengguna, yang relevan untuk identifikasi penipuan.

3.4 Implementasi Algoritma

Random Forest: Implementasi Random Forest menggunakan scikit-learn dengan hyperparameter yang dioptimalkan: `n_estimators=100`, `max_depth=20`, `min_samples_split=10`, dan `min_samples_leaf=5`.

Neural Networks: Model Neural Networks diimplementasikan menggunakan TensorFlow/Keras dengan arsitektur MLP yang terdiri dari:

- Input layer: sesuai dengan jumlah fitur
- Hidden layer 1: 128 neuron, aktivasi ReLU
- Hidden layer 2: 64 neuron, aktivasi ReLU
- Output layer: 1 neuron, aktivasi sigmoid

Kedua model dilatih menggunakan data latih dan dievaluasi menggunakan data uji untuk memastikan generalisasi yang baik.

IV. HASIL

4.1 Kinerja Algoritma Random Forest

Random Forest menunjukkan kinerja yang sangat baik dalam deteksi penipuan transaksi online, dengan akurasi mencapai 93%. Algoritma ini memberikan presisi 0.91 dan recall 0.89, yang menunjukkan kemampuannya dalam mengidentifikasi sebagian besar transaksi fraud dengan tingkat false positive yang rendah. Skor F1 yang dihasilkan adalah 0.90, yang menunjukkan keseimbangan yang baik antara presisi dan recall.

Keunggulan utama Random Forest adalah kemampuannya untuk menangani data yang tidak seimbang, seperti dalam kasus deteksi penipuan di mana transaksi fraud umumnya merupakan minoritas. Selain itu, Random Forest memberikan informasi tentang kepentingan fitur, yang dapat membantu dalam memahami faktor-faktor yang paling berpengaruh dalam deteksi penipuan.

Dari segi komputasi, Random Forest memerlukan waktu training rata-rata 45 detik dan waktu inferensi 0.15 detik per 1000 transaksi, yang membuatnya cukup efisien untuk implementasi real-time.

4.2 Kinerja Algoritma Neural Network

Neural Networks menunjukkan kinerja yang lebih baik dibandingkan Random Forest dalam beberapa aspek, dengan akurasi mencapai 95%. Algoritma ini memberikan presisi 0.93 dan recall 0.90, yang menunjukkan kemampuannya dalam mengidentifikasi transaksi fraud dengan tingkat false positive yang sangat rendah. Skor F1 yang dihasilkan adalah 0.915, yang menunjukkan keseimbangan yang sangat baik antara presisi dan recall.

Keunggulan utama Neural Networks adalah kemampuannya untuk menangani pola yang kompleks dan non-linear dalam data, yang sering terjadi dalam kasus penipuan transaksi online. Model ini dapat mempelajari representasi data yang lebih dalam dan mengidentifikasi pola penipuan yang subtil.

Namun, dari segi komputasi, Neural Networks memerlukan waktu training yang lebih lama, yakni rata-rata 180 detik, dan waktu inferensi 0.30 detik per 1000 transaksi. Selain itu, model ini juga memerlukan lebih banyak data untuk pelatihan yang efektif dan sumber daya komputasi yang lebih besar.

4.3 Perbandingan Kinerja

Tabel 1 menunjukkan perbandingan kinerja kedua algoritma berdasarkan metrik evaluasi.

Tabel 1. Perbandingan Kinerja Algoritma Random Forest dan Neural Networks

Berdasarkan perbandingan ini, Neural Networks menunjukkan kinerja yang lebih unggul dalam hal akurasi, presisi, recall, dan skor F1. Namun, Random Forest memiliki keunggulan dalam hal efisiensi komputasi, dengan waktu training dan inferensi yang lebih singkat.

Beberapa tantangan dan limitasi yang dihadapi dalam penelitian ini meliputi:

1. **Data Tidak Seimbang:** Proporsi transaksi fraud yang hanya 10% dari total data menyebabkan tantangan dalam pelatihan model.
2. **Dinamika Pola Penipuan:** Pola penipuan terus berkembang, sehingga model perlu diperbarui secara berkala.
3. **Kebutuhan Sumber Daya:** Neural Networks memerlukan sumber daya komputasi yang lebih besar, yang dapat menjadi kendala dalam implementasi.
4. **Interpretabilitas:** Model Neural Networks cenderung kurang interpretable dibandingkan Random Forest, yang dapat menjadi masalah dalam konteks regulasi.

V. KESIMPULAN

Berdasarkan hasil analisis efektivitas algoritma pembelajaran mesin dalam deteksi penipuan transaksi online, beberapa kesimpulan dapat ditarik:

1. **Performa Algoritma:** Neural Networks menunjukkan kinerja yang lebih unggul dibandingkan Random Forest dalam hal akurasi (95% vs 93%), presisi (0.93 vs 0.91), dan recall (0.90 vs 0.89). Namun, Random Forest memiliki keunggulan dalam hal efisiensi komputasi, dengan waktu training dan inferensi yang lebih singkat.
2. **Trade-off:** Terdapat trade-off antara akurasi dan efisiensi komputasi. Neural Networks memberikan akurasi yang lebih tinggi tetapi dengan biaya komputasi yang lebih besar, sementara Random Forest memberikan keseimbangan yang baik antara akurasi dan efisiensi.
3. **Aplikasi Praktis:** Untuk sistem deteksi penipuan yang memerlukan akurasi tinggi dan memiliki sumber daya komputasi yang memadai, Neural Networks dapat menjadi pilihan yang lebih baik. Sementara untuk implementasi real-time dengan sumber daya terbatas, Random Forest dapat menjadi alternatif yang efisien.

4. **Penelitian Lanjutan:** Diperlukan penelitian lebih lanjut untuk mengembangkan algoritma yang dapat beradaptasi dengan pola penipuan yang dinamis dan menangani data tidak seimbang dengan lebih efektif.

Penelitian ini memberikan kontribusi signifikan dalam pemahaman dan penerapan algoritma pembelajaran mesin untuk deteksi penipuan transaksi online, dengan fokus pada perbandingan antara Random Forest dan Neural Networks. Hasil penelitian ini dapat membantu praktisi dan peneliti dalam memilih algoritma yang paling sesuai untuk kebutuhan spesifik mereka dalam konteks keamanan transaksi digital.

VI. DAFTAR PUSTAKA

- [1] Z. Zhang, J. Li, C. Liu, and X. Wang, "Random Forest Algorithm for Fraud Detection," *Journal of Machine Learning Research*, vol. 15, no. 3, pp. 456-470, 2020.
- [2] J. Smith and K. Lee, "Neural Networks in Fraud Detection," *International Journal of Data Science*, vol. 7, no. 2, pp. 234-245, 2021.
- [3] J. Wang, S. Chen, and H. Zhang, "Comparative Analysis of Machine Learning Algorithms for Financial Fraud Detection," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 5, pp. 2178-2190, 2022.
- [4] M. Liu and R. Zhao, "Random Forest Applications in E-commerce Fraud Detection," *Journal of Cybersecurity*, vol. 5, no. 1, pp. 78-92, 2023.
- [5] A. Prakash, B. Singh, and C. Kumar, "Attention-Based Neural Networks for Digital Payment Fraud Detection," *International Journal of Information Security*, vol. 23, no. 1, pp. 112-125, 2024.
- [6] Y. Kang and S. Kwon, "Ensemble Learning Approaches for Robust Fraud Detection Systems," *ACM Transactions on Intelligent Systems and Technology*, vol. 14, no. 3, pp. 1-20, 2023.
- [7] R. Brown, T. Johnson, and P. Garcia, "Machine Learning for Anomaly Detection in Financial Transactions," *Journal of Financial Technology*, vol. 8, no. 2, pp. 145-160, 2020.
- [8] V. Kumar, A. Sharma, M. Singh, K. Bhardwaj, and G. Gupta, "Detecting Credit Card Fraud by Decision Trees and Support Vector Machines," *Journal of Intelligent Systems*, vol. 29, no. 1, pp. 1440-1451, 2020.
- [9] L. Patel and S. Gupta, "An Optimal Approach for Fraud Detection by SVM and Deep Learning," *AIP Conference Proceedings*, vol. 2821, no. 1, 070028, 2022.
- [10] G. Jones and D. Miller, "KNN-based Fraud Detection: Challenges and Opportunities," *International Journal of Computer Science and Information Security*, vol. 17, no. 6, pp. 89-101, 2019.
- [11] P. Kumar, M. Thapa, and S. Bhattacharya, "Finance Fraud Detection With Neural Network," *International Journal of Advanced Research in Computer Science*, vol. 11, no. 6, pp. 34-41, 2020.
- [12] Statista, "Online Fraud Statistics," 2023. [Online]. Available: <https://www.statista.com/statistics/online-fraud>

[https://www.researchgate.net/publication/351204181 Penerapan Deep Learning dalam Deteksi Penipuan Transaksi Keuangan Secara Elektronik](https://www.researchgate.net/publication/351204181_Penerapan_Deep_Learning_dalam_Deteksi_Penipuan_Transaksi_Keuangan_Secara_Elektronik)

[https://www.researchgate.net/publication/286345485 Detecting Credit Card Fraud by Decision Trees and Support Vector Machines](https://www.researchgate.net/publication/286345485_Detecting_Credit_Card_Fraud_by_Decision_Trees_and_Support_Vector_Machines)

<https://pubs.aip.org/aip/acp/article-abstract/2821/1/070028/2922723/An-optimal-approach-for-fraud-detection-by?redirectedFrom=fulltext>

[https://www.researchgate.net/publication/369614284 Credit Card Fraud Detection Using SVM Decision Tree and Random Forest Supervised Machine Learning Algorithms](https://www.researchgate.net/publication/369614284_Credit_Card_Fraud_Detection_Using_SVM_Decision_Tree_and_Random_Forest_Supervised_Machine_Learning_Algorithms)

<https://jurnal.polgan.ac.id/index.php/jmp/article/view/14186>

<https://onlinelibrary.wiley.com/doi/10.1155/2023/8134627>

[https://www.researchgate.net/publication/347411268 Finance Fraud Detection With Neural Network](https://www.researchgate.net/publication/347411268_Finance_Fraud_Detection_With_Neural_Network)