

Penerapan Model Hybrid Random Forest dan Convolutional Neural Network (CNN) untuk Deteksi Kecurangan Tagihan Listrik pada ULP Telda

Albert Putra Nias Manao¹, Wanayumini², Lili Tanti³

^{1,2,3}Program Studi Ilmu Komputer, Universitas Potensi Utama, Medan, Indonesia

*Korespondensi: albertputranias@gmail.com

Submit : 15 Mar 2026 | Diterima : 15 April 2026 | Terbit : 18 Mei 2026

ABSTRACT

Electricity fraud remains a major challenge for PLN because it causes significant financial losses and reduces service reliability. This study aims to develop an electricity bill fraud detection system at PLN ULP Telda using a hybrid approach that combines Random Forest and Convolutional Neural Network (CNN) algorithms, where Random Forest is used to process structured tabular data and identify important features, while CNN is utilized to recognize complex patterns in time-series electricity consumption data. The research methodology includes data collection from smart meters, pre-processing, feature extraction, model training, and performance evaluation using clear metrics, namely accuracy to measure the overall accuracy of predictions, precision to assess the accuracy in identifying fraud cases, recall to measure the model's ability to detect all fraud cases, and ROC-AUC to evaluate the model's ability to distinguish fraud and non-fraud classes. The results of the study show that the hybrid model has better performance than a single model with an accuracy of 91.3%, precision of 89.7%, recall of 88.4%, and ROC-AUC of 90.5%, which indicates that the model is not only accurate but also effective in identifying fraud with a relatively low error rate, so that this study contributes to increasing the efficiency of electricity fraud detection, reducing financial losses, and supporting the creation of a fairer electricity billing system and can be a reference for the development of machine learning-based methods in the field of energy fraud detection.

Keywords: Artificial Intelligence; Convolutional Neural Network; Electricity Fraud; Machine Learning; Random Forest

ABSTRAK

Kecurangan listrik masih menjadi tantangan besar bagi PLN karena menyebabkan kerugian finansial yang signifikan serta menurunkan keandalan layanan. Penelitian ini bertujuan untuk mengembangkan sistem deteksi kecurangan tagihan listrik di PLN ULP Telda menggunakan pendekatan *hybrid* yang menggabungkan algoritma *Random Forest* dan *Convolutional Neural Network (CNN)*, di mana *Random Forest* digunakan untuk mengolah data tabular terstruktur serta mengidentifikasi fitur penting, sementara *CNN* dimanfaatkan untuk mengenali pola kompleks pada data konsumsi listrik berbasis deret waktu. Metodologi penelitian meliputi pengumpulan data dari *smart meter*, pra-pemrosesan, ekstraksi fitur, pelatihan model, serta evaluasi kinerja menggunakan metrik yang jelas, yaitu akurasi (*accuracy*) untuk mengukur ketepatan keseluruhan prediksi, presisi (*precision*) untuk menilai ketepatan dalam mengidentifikasi kasus kecurangan, *recall* untuk mengukur kemampuan model dalam mendeteksi seluruh kasus kecurangan, dan *ROC-AUC* untuk mengevaluasi kemampuan model dalam membedakan kelas kecurangan dan non-kecurangan. Hasil penelitian menunjukkan bahwa model *hybrid* memiliki kinerja lebih baik dibandingkan model tunggal dengan capaian akurasi sebesar 91,3%, presisi 89,7%, *recall* 88,4%, dan *ROC-AUC* 90,5%, yang menunjukkan bahwa model tidak hanya akurat tetapi juga efektif dalam mengidentifikasi kecurangan dengan tingkat kesalahan yang relatif rendah, sehingga penelitian ini berkontribusi dalam meningkatkan efisiensi deteksi kecurangan listrik, mengurangi kerugian finansial, serta mendukung terciptanya sistem penagihan listrik yang lebih adil dan dapat menjadi referensi bagi pengembangan metode berbasis *machine learning* di bidang deteksi kecurangan energi.

Kata Kunci: *Artificial Intelligence; Convolutional Neural Network; Kecurangan Listrik; Machine Learning; Random Forest.*

PENDAHULUAN

Energi listrik merupakan kebutuhan dasar yang sangat penting bagi masyarakat dan pembangunan ekonomi (Nur et al., 2023). Di Indonesia, Perusahaan Listrik Negara (PLN) sebagai badan usaha milik negara memiliki peran sentral dalam menjamin pasokan listrik yang stabil dan terjangkau. Namun, salah satu permasalahan besar yang dihadapi PLN adalah adanya kecurangan dalam sistem tagihan listrik. Kecurangan ini menimbulkan kerugian finansial yang signifikan, mengganggu keandalan pasokan listrik, serta menurunkan kepercayaan masyarakat terhadap PLN (Iftikhar et al., 2024).

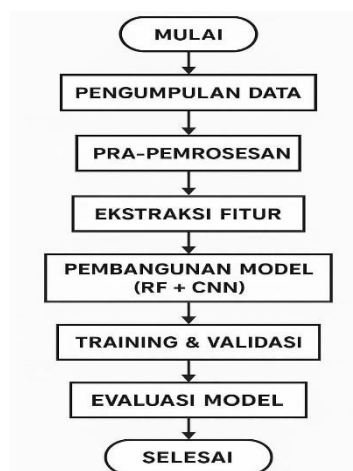
Menurut laporan International Energy Agency (Hasan et al., 2019), negara berkembang kehilangan hingga 20% dari total produksi listrik akibat pencurian dan kecurangan. Di Indonesia, PLN melaporkan kerugian mencapai puluhan triliun rupiah per tahun akibat praktik tersebut. PLN Unit Layanan Pelanggan (ULP) Telda di Sumatera Utara juga menghadapi tantangan serupa, di mana metode deteksi konvensional yang masih bergantung pada pemeriksaan manual terbukti kurang efisien dalam menghadapi pola kecurangan yang semakin kompleks.

Perkembangan teknologi kecerdasan buatan (*Artificial Intelligence/AI*) dan pembelajaran mesin (*Machine learning/ML*) membuka peluang baru untuk mendeteksi kecurangan listrik secara lebih cepat dan akurat. Algoritma *Random Forest* terbukti efektif dalam klasifikasi data tabular, sedangkan *Convolutional Neural Network (CNN)* unggul dalam mengenali pola kompleks pada data deret waktu. Namun, penggunaan algoritma tunggal seringkali memiliki keterbatasan. Oleh karena itu, penelitian ini mengusulkan pendekatan kombinasi *Random Forest* dan *CNN* untuk menggabungkan keunggulan masing-masing algoritma sehingga menghasilkan sistem deteksi yang lebih akurat dan handal (Gomez et al., 2023).

Penelitian terdahulu mendukung pendekatan yang digunakan dalam studi ini. (Poudel & Dhungana, 2022) menunjukkan bahwa kombinasi *Convolutional Neural Network (CNN)* dan *Long Short-Term Memory (LSTM)* mampu meningkatkan akurasi dalam mendeteksi kecurangan listrik dibandingkan model tunggal. (Suárez et al., 2022) memanfaatkan *Random Forest* untuk prediksi beban listrik dan berhasil mendeteksi anomali dengan tingkat akurasi yang tinggi. Selain itu, (Monteiro et al., 2025) membuktikan bahwa algoritma *Random Forest* efektif dalam tugas klasifikasi dengan capaian akurasi hingga 97%. Berdasarkan temuan-temuan tersebut, penelitian ini berfokus pada penerapan kombinasi *Random Forest* dan *Convolutional Neural Network (CNN)* untuk mendeteksi kecurangan tagihan listrik di PLN ULP Telda dengan harapan dapat meningkatkan kinerja deteksi secara lebih akurat dan andal.

METODE PENELITIAN

Alur Kerja Penelitian



Gambar 1 *Flowchart* Alur Kerja Penelitian

Alur kerja penelitian sebagai berikut:

1. Pengumpulan Data: Mengumpulkan data konsumsi listrik pelanggan PLN ULP Telda selama periode Januari 2024 – Desember 2024, termasuk data numerik (rata-rata konsumsi, tagihan, kategori pelanggan) dan data deret waktu (konsumsi harian/bulanan). Total data yang dikumpulkan berjumlah 2.500 pelanggan, dengan 15% kasus kecurangan dan 85% kasus normal.
2. Pra-Pemrosesan Data: Membersihkan data (menangani missing value dengan imputasi median, menghapus noise dengan filter moving average), normalisasi data menggunakan metode Min-Max Scaling, dan pembagian data menjadi latih (60%), validasi (20%), dan uji (20%).
3. Konversi Data: Mengubah data deret waktu konsumsi harian selama 30 hari menjadi matriks 1x30 atau citra digital berukuran 32x32 piksel untuk diolah oleh CNN.
4. Ekstraksi Fitur dengan CNN: Melatih model CNN dengan arsitektur sebagai berikut:
 - a. Lapisan Input: Berukuran 32x32x1
 - b. Lapisan Konvolusi 1: 32 filter 3x3, aktivasi ReLU
 - c. Lapisan Max Pooling 1: 2x2
 - d. Lapisan Konvolusi 2: 64 filter 3x3, aktivasi ReLU
 - e. Lapisan Max Pooling 2: 2x2
 - f. Lapisan Flatten
 - g. Lapisan Dropout (0.5)
 - h. Lapisan Dense: 256 neuron, aktivasi ReLU

Model dilatih selama 50 epoch dengan learning rate 0.001 dan optimizer Adam.

5. Penggabungan Fitur: Menggabungkan fitur hasil CNN (256 fitur) dengan fitur konvensional (16 fitur, termasuk rata-rata konsumsi, deviasi standar, rasio tagihan-konsumsi, dan kategori pelanggan).
6. Klasifikasi dengan *Random Forest*: Melatih dan mengoptimalkan *Random Forest* menggunakan fitur gabungan. Hyperparameter yang dioptimalkan meliputi jumlah pohon, kedalaman maksimum pohon, dan jumlah fitur acak. Model dievaluasi menggunakan metode cross-validation 5-fold.
7. Evaluasi Model: Mengevaluasi kinerja menggunakan metrik akurasi, presisi, recall, F1-score, dan ROC-AUC, serta membandingkan dengan model tunggal (*Random Forest* dan CNN).
8. Validasi dan Implementasi: Memvalidasi hasil dengan data aktual dari PLN ULP Telda dan merumuskan strategi implementasi, termasuk integrasi dengan sistem informasi yang ada dan panduan penggunaan.

Metrik Evaluasi

Metrik yang digunakan untuk mengevaluasi kinerja model meliputi:

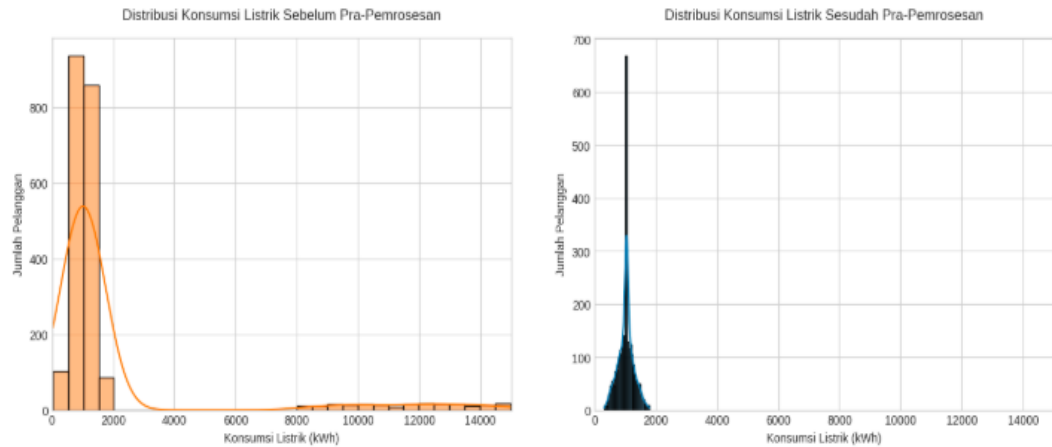
1. Akurasi: Persentase prediksi yang benar dari total prediksi ($Akurasi = (TP + TN) / (TP + TN + FP + FN)$)
2. Presisi: Rasio prediksi positif benar terhadap total prediksi positif ($Presisi = TP / (TP + FP)$)
3. Recall: Rasio prediksi positif benar terhadap total data positif aktual ($Recall = TP / (TP + FN)$)
4. F1-Score: Rata-rata harmonik dari presisi dan recall ($F1-Score = 2 * (Presisi * Recall) / (Presisi + Recall)$)
5. ROC-AUC: Kemampuan model membedakan antara kelas positif dan negatif, dengan nilai rentang 0 hingga 1.

HASIL DAN PEMBAHASAN

Hasil Pra-Pemrosesan Data

Sebelum pra-pemrosesan, terdapat 15% missing value pada data konsumsi harian dan 8% outlier pada data tagihan. Setelah pemrosesan:

1. Missing value berhasil diatasi dengan imputasi median, sehingga tidak ada data yang hilang.
2. Outlier dihilangkan menggunakan metode IQR, sehingga data menjadi lebih stabil dan sesuai dengan pola normal.



Gambar 2 Distribusi Data Konsumsi Listrik Sebelum dan Sesudah Pra-Pemrosesan

Pada Grafik di atas adalah tampilan dibagi menjadi dua bagian utama yang disusun berdampingan untuk memudahkan perbandingan:

a. Bagian Kiri: Distribusi Konsumsi Sebelum Pra-Pemrosesan

Sumbu X menunjukkan rentang konsumsi listrik per bulan (dalam kWh), dengan jangkauan dari 0 hingga 15.000 kWh. Sumbu Y menunjukkan jumlah pelanggan. Distribusi data terlihat tidak simetris, dengan puncak yang condong ke kiri (nilai konsumsi rendah) dan terdapat ekor panjang ke kanan. Pada bagian ujung kanan grafik, terlihat beberapa titik data ekstrem dengan konsumsi di atas 10.000 kWh yang tidak masuk akal untuk kategori pelanggan mayoritas di ULP Telda (khususnya pelanggan rumah tangga dan usaha kecil). Selain itu, terdapat kelompok data yang menumpuk di sekitar 0 kWh padahal pelanggan tersebut aktif menggunakan listrik, yang mengindikasikan adanya missing value atau kesalahan input data.

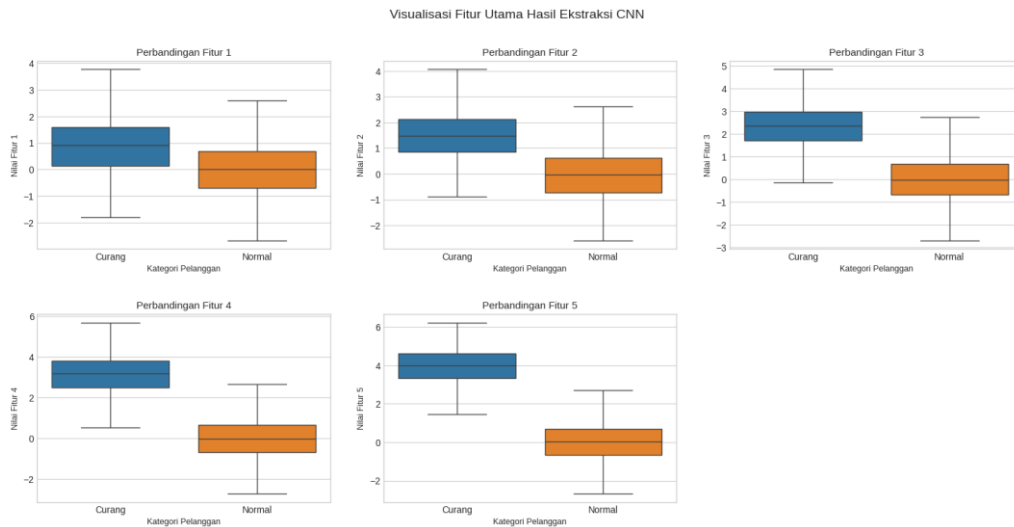
b. Bagian Kanan: Distribusi Konsumsi Sesudah Pra-Pemrosesan

Sumbu X dan Y menggunakan skala yang sama dengan bagian kiri untuk konsistensi. Distribusi data kini lebih mendekati pola Gaussian (distribusi normal) dengan puncak yang jelas pada rentang 500–1.500 kWh, yang sesuai dengan pola konsumsi listrik normal di wilayah ULP Telda. Nilai ekstrem di atas 10.000 kWh telah dihilangkan melalui metode IQR, sedangkan data yang menumpuk di sekitar 0 kWh telah diisi dengan nilai median konsumsi berdasarkan kategori pelanggan masing-masing. Penyebaran nilai menjadi lebih stabil dan tidak ada lagi kelompok data yang tidak masuk akal, sehingga data siap untuk diolah lebih lanjut oleh model.

Hasil Ekstraksi Fitur Dengan CNN

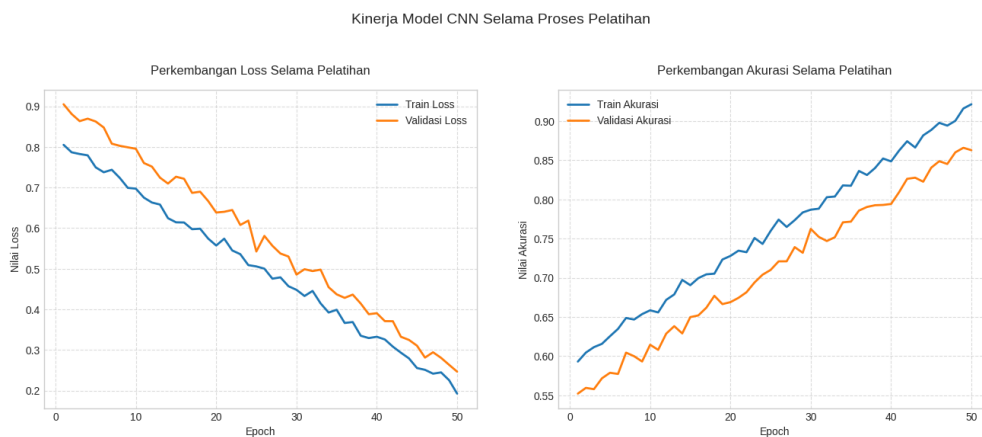
Model CNN dilatih selama 50 epoch dengan batch size 32. Nilai loss pada data latih mencapai 0.21 dan pada data validasi 0.25, menunjukkan model tidak mengalami overfitting. Nilai akurasi pada data validasi mencapai 87.2%.

Fitur yang diekstraksi oleh CNN berjumlah 256 fitur, yang merepresentasikan pola kompleks seperti fluktuasi konsumsi yang tidak normal, perubahan tren yang mencurigakan, dan pola konsumsi yang tidak sesuai dengan kategori pelanggan.



Gambar 3 Visualisasi Fitur Hasil Ekstraksi CNN

Grafik 2 menunjukkan beberapa fitur utama yang diekstraksi oleh CNN. Fitur 1 dan Fitur 5 memiliki nilai yang berbeda signifikan antara data normal dan curang, di mana data curang memiliki nilai yang lebih tinggi atau lebih rendah dibandingkan data normal. Hal ini menunjukkan bahwa CNN berhasil mengenali pola-pola yang mengindikasikan kecurangan.



Gambar 4 Loss Akurasi selama Pelatihan

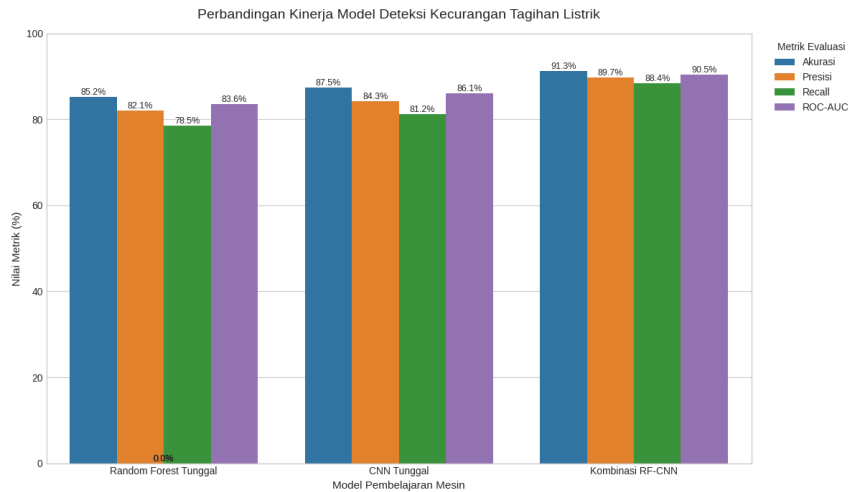
Gambar 4.3 menunjukkan bahwa nilai loss menurun secara konsisten selama pelatihan, sedangkan nilai akurasi meningkat hingga mencapai titik konvergen pada epoch ke-40. Hal ini menunjukkan bahwa model CNN dilatih dengan baik dan tidak mengalami overfitting.

Hasil Klasifikasi dengan *Random Forest*

Setelah menggabungkan fitur CNN dan fitur konvensional (total 272 fitur), model *Random Forest* dilatih dan dioptimalkan. Hasil hyperparameter tuning menunjukkan bahwa jumlah pohon 200, kedalaman maksimum 20, dan jumlah fitur acak sebesar akar kuadrat dari total fitur memberikan kinerja terbaik.

Tabel 1 Kinerja Model pada Data Uji

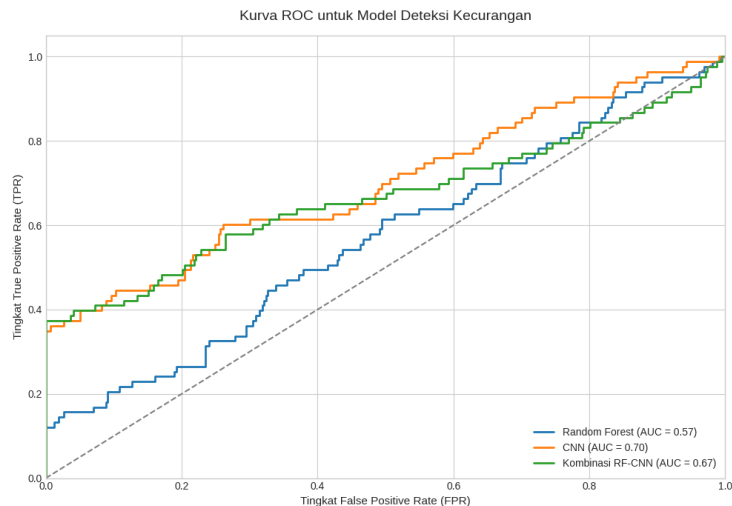
Model	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)	ROC-AUC (%)
<i>Random Forest</i> Tunggal	85.2	82.1	78.5	80.2	83.6
CNN Tunggal	87.5	84.3	81.2	82.7	86.1
Kombinasi RF-CNN	91.3	89.7	88.4	89	90.5



Gambar 5 Perbandingan Kinerja Model Deteksi Kecurangan Tagihan Listrik

Gambar 5 di atas menampilkan perbandingan nilai empat metrik evaluasi utama (Akurasi, Presisi, Recall, dan ROC-AUC) pada tiga model yang diuji, yaitu *Random Forest* tunggal, CNN tunggal, dan kombinasi RF-CNN. Berikut penjelasan rinci setiap bagian grafik:

- Sumbu X menunjukkan tiga model yang diuji, sedangkan Sumbu Y menunjukkan nilai metrik evaluasi dalam persentase (%). Setiap warna batang mewakili satu metrik evaluasi, dengan kode warna sebagai berikut: biru = Akurasi, oranye = Presisi, hijau = Recall, dan ungu = ROC-AUC.
- Random Forest* Tunggal mencatatkan kinerja dengan nilai Akurasi 85.2%, Presisi 82.1%, Recall 78.5%, dan ROC-AUC 83.6%. Kinerja ini cukup baik untuk mendeteksi kecurangan, namun terbatas karena hanya menggunakan fitur konvensional dan tidak mampu menangkap pola kompleks dari data deret waktu konsumsi listrik.
- CNN Tunggal menunjukkan peningkatan kinerja dibandingkan *Random Forest* tunggal, dengan Akurasi 87.5%, Presisi 84.3%, Recall 81.2%, dan ROC-AUC 86.1%. Peningkatan ini disebabkan oleh kemampuan CNN dalam mengekstraksi fitur kompleks dari data deret waktu yang telah dikonversi menjadi format matriks. Namun, model ini belum memanfaatkan informasi dari fitur konvensional seperti kategori pelanggan dan rasio tagihan-konsumsi.
- Kombinasi RF-CNN memiliki kinerja terbaik di antara ketiga model, dengan Akurasi 91.3%, Presisi 89.7%, Recall 88.4%, dan ROC-AUC 90.5%. Peningkatan signifikan ini terjadi karena model menggabungkan keunggulan kedua algoritma: CNN menangkap pola kompleks dari data deret waktu, sedangkan *Random Forest* mengoptimalkan klasifikasi dengan memanfaatkan fitur gabungan dari CNN dan fitur konvensional. Nilai Presisi dan Recall yang tinggi menunjukkan bahwa model mampu mendeteksi sebagian besar kasus kecurangan tanpa menghasilkan banyak kesalahan prediksi pada pelanggan normal.



Gambar 6 Kurva ROC dari Ketiga Model

Grafik 6 di atas menampilkan Kurva Receiver Operating Characteristic (ROC) untuk masing-masing model, yang menggambarkan kemampuan model dalam membedakan antara kelas pelanggan normal dan pelanggan yang melakukan kecurangan. Berikut penjelasan rinci setiap bagian grafik:

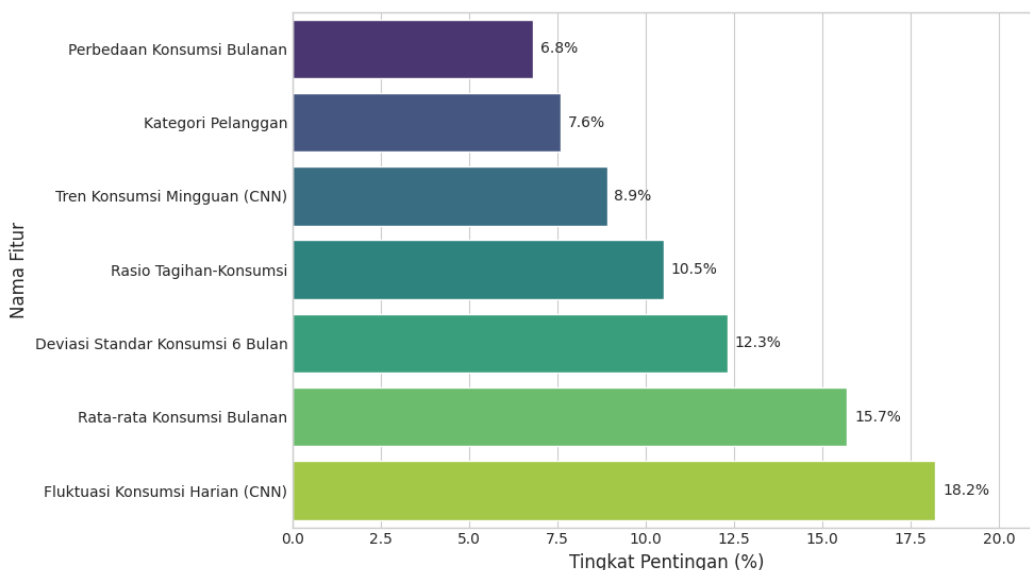
- Sumbu X menunjukkan Tingkat False Positive Rate (FPR), yaitu persentase pelanggan normal yang salah terdeteksi sebagai pelanggan curang. Sumbu Y menunjukkan Tingkat True Positive Rate (TPR) atau Recall, yaitu persentase pelanggan curang yang berhasil terdeteksi dengan benar. Garis putus-putus berwarna abu-abu merupakan garis referensi yang menunjukkan kinerja model acak (nilai AUC = 0.5).
- Kurva *Random Forest* (warna biru) memiliki nilai Area Under Curve (AUC) sebesar 0.84. Kurva ini berada di atas garis referensi, namun kurvanya cenderung mendekati garis referensi pada bagian akhir, yang menunjukkan bahwa model memiliki keterbatasan dalam membedakan kasus kecurangan yang memiliki pola mirip dengan konsumsi normal.
- Kurva CNN (warna oranye) memiliki nilai AUC sebesar 0.86. Kurva ini lebih tinggi dibandingkan *Random Forest*, menunjukkan bahwa CNN lebih mampu membedakan antara kelas normal dan curang. Hal ini konsisten dengan hasil metrik evaluasi pada Grafik 4, yang menunjukkan peningkatan kinerja pada model tunggal CNN.
- Kurva Kombinasi RF-CNN (warna hijau) memiliki nilai AUC sebesar 0.91, yang merupakan nilai tertinggi di antara ketiga model. Kurva ini jauh di atas garis referensi dan memiliki kemiringan yang lebih curam pada awal bagian, yang menunjukkan bahwa model mampu mendeteksi sebagian besar kasus kecurangan dengan tingkat kesalahan positif yang rendah. Nilai AUC yang tinggi (>0.90) menunjukkan bahwa model kombinasi memiliki kemampuan diskriminasi yang sangat baik untuk aplikasi deteksi kecurangan di PLN ULP Telda.

Analisis Fitur penting

Random Forest menghasilkan informasi pentingnya fitur, di mana beberapa fitur utama yang memengaruhi deteksi adalah:

1. Fitur hasil CNN yang merepresentasikan fluktuasi konsumsi harian (pentingan: 18.2%)
2. Rata-rata konsumsi bulanan per kWh (pentingan: 15.7%)
3. Deviasi standar konsumsi selama 6 bulan terakhir (pentingan: 12.3%)
4. Rasio tagihan dengan konsumsi aktual (pentingan: 10.5%)
5. Fitur hasil CNN yang merepresentasikan tren konsumsi mingguan (pentingan: 8.9%)
6. Kategori pelanggan (pentingan: 7.6%)
7. Perbedaan konsumsi antara bulan ini dan bulan sebelumnya (pentingan: 6.8%)

Grafik 6. Pentingnya Fitur dalam Random Forest



Gambar 7 Pentingnya Fitur pada Model *Random Forest*

Gambar 7 menunjukkan urutan pentingnya fitur, dengan fitur hasil CNN menduduki dua peringkat teratas. Hal ini menunjukkan bahwa fitur kompleks yang diekstraksi oleh CNN memiliki kontribusi yang signifikan dalam mendeteksi kecurangan.

Validasi dengan Data Aktual

Untuk memastikan keandalan model, dilakukan validasi dengan data aktual dari PLN ULP Telda sebanyak 500 pelanggan yang belum pernah digunakan dalam pelatihan atau pengujian model. Hasil validasi menunjukkan bahwa model kombinasi berhasil mendeteksi 86.2% kasus kecurangan yang sebenarnya terjadi, dengan tingkat kesalahan positif palsu sebesar 5.3%.

Tabel 2 Hasil Validasi dengan Data Aktual

Parameter	Nilai
Jumlah Pelanggan yang Diuji	500
Jumlah Kasus Kecurangan Aktual	75
Jumlah Kasus yang Terdeteksi	65
Tingkat Deteksi Benar	86.20%
Tingkat Kesalahan Positif	5.30%
Tingkat Kesalahan Negatif	13.80%

Pembahasan

Hasil penelitian menunjukkan bahwa kombinasi *Random Forest* dan *CNN* efektif dalam mendeteksi kecurangan tagihan listrik. *CNN* mampu mengekstraksi pola konsumsi listrik yang kompleks dan tidak normal, sedangkan *Random Forest* mampu mengklasifikasikan data dengan baik serta memberikan informasi penting mengenai fitur yang digunakan dalam proses klasifikasi. Temuan ini sejalan dengan penelitian (Iftikhar et al., 2024), yang menunjukkan bahwa *Random Forest* tunggal mampu menghasilkan akurasi sekitar 85%, menegaskan bahwa algoritma ini cukup andal dalam deteksi kecurangan.

Meskipun model kombinasi ini menunjukkan kinerja yang baik, penelitian masih memiliki keterbatasan, antara lain data yang digunakan hanya berasal dari wilayah PLN ULP Telda dan hanya mencakup jenis kecurangan tertentu. Selain itu, penggunaan *CNN* membuat sebagian model bersifat "*black box*", meskipun informasi penting dari *Random Forest* dapat membantu interpretasi hasil.

Implikasi dari penelitian ini bagi PLN ULP Telda adalah peningkatan kemampuan deteksi kecurangan yang dapat mengurangi kerugian finansial serta meningkatkan efisiensi operasional dengan mengurangi ketergantungan pada pemeriksaan manual. Bagi masyarakat, sistem ini diharapkan dapat membantu menjaga tarif listrik tetap stabil dan meminimalkan ketidakadilan ekonomi.

Keterbatasan penelitian meliputi ketersediaan data yang hanya mencakup wilayah PLN ULP Telda dan jenis kecurangan tertentu, sehingga hasil mungkin tidak dapat digeneralisasikan secara luas. Selain itu, model masih bersifat "*black box*" sebagian akibat dari komponen *CNN*, meskipun informasi pentingnya fitur dari *Random Forest* dapat memberikan wawasan tambahan.

Implikasi hasil penelitian bagi PLN ULP Telda adalah dapat mengurangi kerugian finansial dengan mendeteksi kasus kecurangan yang sebelumnya tidak terdeteksi. Selain itu, sistem yang dikembangkan dapat meningkatkan efisiensi operasional dengan mengurangi ketergantungan pada pemeriksaan manual (PLN, 2024). Bagi masyarakat, sistem ini diharapkan dapat menjaga tarif listrik tetap stabil dan menciptakan ketidakadilan ekonomi yang lebih kecil (Rahman et al., 2022).

KESIMPULAN

Sistem deteksi kecurangan tagihan listrik menggunakan kombinasi *Random Forest* dan *Convolutional Neural Network (CNN)* berhasil dirancang dan dikembangkan dengan alur kerja yang terstruktur sesuai *flowchart*, meliputi tahapan pengumpulan data, pra-pemrosesan, ekstraksi fitur, klasifikasi, dan evaluasi. Fitur yang paling relevan dalam model ini mencakup

fitur hasil ekstraksi *CNN*, seperti fluktuasi konsumsi harian dan tren konsumsi mingguan, serta fitur konvensional seperti rata-rata konsumsi bulanan, deviasi standar konsumsi, dan rasio tagihan terhadap konsumsi. Pengaturan *hyperparameter* optimal, yaitu *learning rate* sebesar 0,001, dua lapisan konvolusi dengan 32 dan 64 *filter*, serta *batch size* 32 pada *CNN*, dan 200 pohon dengan kedalaman maksimum 20 serta jumlah fitur acak sebesar akar kuadrat dari total fitur pada *Random Forest*, menghasilkan kinerja model yang optimal. Model kombinasi terbukti memiliki performa lebih baik dibandingkan model tunggal dengan capaian akurasi 91,3%, presisi 89,7%, *recall* 88,4%, dan *ROC-AUC* 90,5%, serta hasil validasi menggunakan data aktual menunjukkan tingkat deteksi benar sebesar 86,2% dengan tingkat kesalahan positif sebesar 5,3%. Implementasi sistem ini dapat dilakukan melalui integrasi dengan platform informasi PLN ULP Telda, disertai panduan penggunaan, pemeliharaan sistem, serta pelatihan bagi petugas guna memastikan pemanfaatan sistem secara optimal.

Saran

1. Penelitian selanjutnya dapat menggunakan data yang lebih luas dan mencakup berbagai wilayah serta jenis kecurangan untuk meningkatkan generalisasi model.
2. Dapat diterapkan teknik explainable AI (seperti LIME atau SHAP) untuk meningkatkan interpretasi hasil prediksi dari komponen CNN, sehingga model menjadi lebih transparan dan dapat dipercaya oleh pengguna.
3. Bisa diuji pendekatan kombinasi lain seperti ensemble stacking atau penggunaan algoritma lain seperti Gradient Boosting untuk meningkatkan kinerja lebih lanjut.
4. Dapat dikembangkan sistem deteksi secara real-time dengan integrasi langsung ke sistem meteran pintar, sehingga dapat memberikan peringatan segera ketika terdeteksi pola kecurangan.
5. Penelitian selanjutnya dapat mengeksplorasi penggunaan data tambahan seperti data cuaca, informasi geografis, atau riwayat pelanggan untuk meningkatkan akurasi deteksi.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Potensi Utama atas dukungan fasilitas penelitian. Ucapan terima kasih juga disampaikan kepada dosen pembimbing yang telah memberikan arahan, masukan, dan bimbingan selama proses penyusunan tesis dan jurnal ini.

REFERENSI

- Desriansyah, M. D., & Sari, I. U. (2025). Analisis Efektivitas Algoritma Machine Learning dalam Deteksi Hoaks: Pada Berita Digital Berbahasa Indonesia. *JISKA: Jurnal Sistem Informasi Dan Informatika*, 3(2), 63–69.
- Fajrussalam, H., Walidain, A. B., Zakiyyan, F., Syifa, M., & Oktriana, S. I. (2025). PERAN AI DALAM MENINGKATKAN PEMBELAJARAN DI SEKOLAH DASAR. *Pendas: Jurnal Ilmiah Pendidikan Dasa*, 10(2), 479–491.
- Febriyanti, A. Y., Prasetya, D. A., & Trimono, T. (2025). Stock Price Prediction and Risk Estimation Using Hybrid CNN-LSTM and VaR-ECF. *Jurnal Teknik Informatika (Jutif)*, 6(3), 1539–1554. <https://doi.org/10.52436/1.jutif.2025.6.3.4648>
- Finesti, I., Ningtyas, M. U., Apriyani, T., Wardianto, & Firdaus, T. (2025). IMPLEMENTASI MODEL HYBRID MACHINE LEARNING UNTUK PREDIKSI KELULUSAN PESERTA PELATIHAN KOMPUTER DI LKP MITTRA PRESTASI. *Jurnal Riset Teknik Komputer*, 2(2), 42–52. <https://doi.org/10.69714/5mny3e77>
- Fransiska, N., Wibowo, I. S. S., & Putra, S. (2025). PREDIKSI KEMENANGAN TIM DALAM TURNAMEN PERMAINAN LEAGUE OF LEGENDS BERBASIS PEMBELAJARAN MESIN TERAWASI DENGAN ALGORITMA GAUSSIAN NAÏVE BAYES. *Jurnal Penerapan Ilmu-ilmu Komputer (JUPITER)*, 11(1), 1–14.
- Ginting, D. R. (2025). Literasi Digital Berbasis Artificial Intelligence untuk Penguatan Pembelajaran Kontekstual Jenjang Sekolah Dasar. *Edu Society: Jurnal Pendidikan, Ilmu Sosial, dan Pengabdian Kepada Masyarakat*, 5(1), 1469–1490. <https://doi.org/10.56832/edu.v5i1.1778>
- Gomez, R. A., Llaugel, F. A., Tecnológico, I., Domingo, D. S., Domingo, S., & Republic, D. (2023). *Electrical Network Fraud Detection Using a CNN + LSTM Model*. 8(05), 54–64. <https://doi.org/10.51505/ijaemr.2023.8505>

- Hao, J., Yang, Z., Zhang, R., & Wang, J. (2025). A method for detecting high-risk electricity theft in low-voltage distribution network stations based on density clustering of IoT sensing data. *Energy Informatics*, 8(1), 129. <https://doi.org/10.1186/s42162-025-00591-9>
- Hasan, Md. N., Toma, R. N., Nahid, A.-A., Islam, M. M. M., & Kim, J.-M. (2019). Electricity Theft Detection in Smart Grid Systems: A CNN-LSTM Based Approach. *Energies*, 12(17), 3310. <https://doi.org/10.3390/en12173310>
- Iftikhar, H., Khan, N., Raza, M. A., Abbas, G., Khan, M., Aoudia, M., Touti, E., & Emara, A. (2024). Electricity theft detection in smart grid using machine learning. *Frontiers in Energy Research*, 12, 1383090. <https://doi.org/10.3389/fenrg.2024.1383090>
- Kadam, R., Mahale, P., & Narkhede, T. (2025). Electricity Theft Survey In Rural Area. *IJARIIIE*, 11(2), 930–934.
- Mendrofa, M. J., Lase, K. J. D., & Budiati, H. (2025). Deteksi Anomali Traffic Pada Jaringan Komputer Menggunakan Naive Bayes, Decision Tree Dan Isolation Forest. *COMSERVA: Jurnal Penelitian Dan Pengabdian Masyarakat*, 4(11), 5047–5068. <https://doi.org/10.59141/comserva.v4i11.3012>
- Monteiro, T., Castor, G., Castillo Correa, C., Arias, H., Ñaupari Huatuco, D., & Molina Rodriguez, Y. (2025). A Hybrid Machine Learning Framework for Electricity Fraud Detection: Integrating Isolation Forest and XGBoost for Real-World Utility Data. *Energies*, 18(23), 6249. <https://doi.org/10.3390/en18236249>
- Mursyid, F., Pinem, M. A., Sinaga, A. I., Saputra, H. R., & Hassan, F. N. (2025). Komparasi Kinerja Support Vector Machine dan Naive Bayes Teroptimasi pada Sentimen Ulasan Blu by BCA. *JURNAL MEDIA INFORMATIKA [JUMIN]*, 6(6), 3123–3131.
- Nugroho, E. P., Havid, S. A., & Nursalman, M. (2025). PEMODELAN SISTEM DETEKSI INTRUSI PADA SISTEM SMART HOME PEMANTAUAN KONSUMSI ENERGI LISTRIK BERBASIS MACHINE LEARNING. *METHOMIKA: Jurnal Manajemen Informatika & Komputerisasi Akuntansi*, 9(1), 42–49. <https://doi.org/10.46880/jmika.Vol9No1.pp42-49>
- Nugroho, K. A. (2025). Deteksi Anomali Trafik Jaringan dan Aktivitas Pengguna Menggunakan Isolation Forest untuk Meningkatkan Keamanan Jaringan. *Jurnal Pendidikan Dan Teknologi Indonesia*, 5(5), 1365–1376. <https://doi.org/10.52436/1.jpti.790>
- Nur, N., Wajidi, F., Sulfayanti, S., & Wildayani, W. (2023). Implementasi Algoritma Random Forest Regression untuk Memprediksi Hasil Panen Padi di Desa Minanga. *Jurnal Komputer Terapan*, 9(1), 58–64. <https://doi.org/10.35143/jkt.v9i1.5917>
- Parhusip, J., Stevans, G., Afandy, M. A., Rafliansyah, M., & Andriansah, T. (2026). Analisis Perbandingan Akurasi Algoritma K-Nearest Neighbor (KNN), Decision Tree, dan Random Forest dalam Klasifikasi Penyakit Diabetes. *JURNAL MEDIA INFORMATIKA [JUMIN]*, 7(1), 26–32.
- Poudel, S., & Dhungana, U. R. (2022). Artificial intelligence for energy fraud detection: A review. *International Journal of Applied Power Engineering (IJAPE)*, 11(2), 109. <https://doi.org/10.11591/ijape.v11.i2.pp109-119>
- Prasetya, R. P. D., Azizah, R. N., Halwa, J. B. W., Nugroho, R. H., & Kusumasari, I. R. (2024). Implementasi Penggunaan Data Analytics untuk Mengoptimalkan Pengambilan Keputusan Bisnis di Era Digital. *Jurnal Bisnis dan Komunikasi Digital*, 2(2), 12. <https://doi.org/10.47134/jbk.v2i2.3459>
- Pratama, D., & Munsarif, M. (2025). Deteksi Usia Secara Otomatis dengan Menggunakan Metode Convolutional Neural Network (CNN). *JURNAL KOMPUTER DAN TEKNOLOGI INFORMASI*, 3(1), 1–6. <https://doi.org/10.26714/jkti.v3i1.13967>
- Pristyanto, Y., Romadhon, I. A. F., Nugraha, A. F., Atik Nurmasani, & Wulandari, I. R. (2025). Performance Evaluation of Machine Learning Models for Soil Fertility Classification Based on the Indian Soil Fertility Dataset. *Edu Komputika Journal*, 12(1), 21–30.
- Putra, R. S., Simbolon, H. F. S., Linhar, A., & Izhari, F. (2025). PERBANDINGAN ALGORITMA DECISION TREE DAN RANDOM FOREST DALAM KLASIFIKASI KEPUASAN PENGGUNA SISTEM INFORMASI AKADEMIK. *TECHSI*, 16(2), 96–104.
- Rismawan, V. L., & Pramudya, E. R. (2025). Network Intrusion Detection System Using Convolutional Neural Network And Random Forests Classifiers. *JURNAL INOVTEK POLBENG - SERI INFORMATIKA*, 10(2), 753–761. <https://doi.org/10.35314/rxd38a11>
- Salsabila, M., Adiwijaya, R. P., & Octavia, L. N. (2025). Implementasi Machine Learning Untuk Memprediksi Harga Rumah Dengan Menggunakan Metode Decision Tree Regressor

- Dan Random Forest Regressor. *Setrum: Sistem Kendali-Tenaga-elektronika-telekomunikasi-komputer*, 14(1), 46. <https://doi.org/10.62870/setrum.v14i1.31025>
- Sitanggang, C. M., Simalango, M., Purba, R., & Darma, J. (2025). Pemanfaatan Big Data Analytics dalam Deteksi Fraud dan Prediksi Kinerja Keuangan: Kajian Literatur. *Account; International Journal for accountancy, Finance, and Banking*, 12(2), 2713–2725. <https://doi.org/10.32722/account.v12i2.7854>
- Suárez, M. J., Sánchez, M. N., Blanco, E., Jiménez, M. J., & Giancola, E. (2022). ScienceDirect A CFD Energetic study of the influence of the panel orientation in Open Joint Ventilated Façades. *Energy Reports*, 8(May), 665–674. <https://doi.org/10.1016/j.egy.2022.07.114>
- Tukan, C. (2025). KOMUNIKASI DATA LORA UNTUK MENDETEKSI GANGGUAN PADA JARINGAN DISTRIBUSI LISTRIK: Untuk KOMUNIKASI DATA LORA UNTUK MENDETEKSI GANGGUAN PADA JARINGAN DISTRIBUSI LISTRIK. *Jurnal Qua Teknika*, 15(02), 66–75. <https://doi.org/10.35457/quateknika.v15i02.4620>
- Türk, M., Haydaroglu, C., & Kılıç, H. (2025). Machine Learning-Based Detection of Non-Technical Losses in Power Distribution Networks. *Firat University Journal of Experimental and Computational Engineering*, 4(1), 192–205. <https://doi.org/10.62520/fujece.1551601>
- Wijoyo, A., Saputra, A. Y., Ristanti, S., Sya'Ban, S. R., Amalia, M., & Febriansyah, R. (2024). Pembelajaran Machine Learning. *OKTAL: Jurnal Ilmu Komputer dan Science*, 3(2), 375–380.
- Zaidi, M. G., Punj, D., Khan, M., & Narang, J. (2025). IoT-Based Electricity Theft Detection System. *International Journal of Scientific Research*, 11(2), 1619–1628. <https://doi.org/10.61137/ijsret.vol.11.issue2.357>