

UAS DEEP LEARNING

Pengembangan Sistem Deteksi Anomali pada Data Sensor IoT Industri

Menggunakan Arsitektur Hybrid CNN–LSTM



disusun oleh:

M. Febri Ardiansyah

G1A022049

Baim Mudrik Aziz

G1A022071

Mezi

G1A022077

Dosen Pengampu:

Ir. Arie Vatesia, S.T., M.T.I., P.hD., IPP.

PROGRAM STUDI INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS BENGKULU

2025

ABSTRAK

Internet of Things (IoT) menghasilkan data sensor industri dalam jumlah besar yang bersifat deret waktu (time series). Data tersebut rentan mengandung anomali akibat gangguan operasional atau kegagalan komponen mesin. Penelitian ini bertujuan untuk mengembangkan sistem deteksi anomali pada data sensor IoT industri menggunakan arsitektur hybrid Convolutional Neural Network–Long Short-Term Memory (CNN–LSTM). Dataset yang digunakan adalah MetroPT-3 Air Compressor Dataset yang diperoleh dari UCI Machine Learning Repository, yang merekam berbagai parameter operasional kompresor udara.

Tahapan penelitian meliputi pemuatan dan eksplorasi data, pelabelan anomali berbasis aturan tekanan dan suhu, normalisasi menggunakan Min-Max Scaler, serta segmentasi data menggunakan metode sliding window. Dataset dibagi menjadi data latih, validasi, dan uji dengan mempertahankan urutan temporal. Model CNN–LSTM dilatih menggunakan class weighting dan early stopping untuk mengatasi ketidakseimbangan kelas. Selain itu, dilakukan eksperimen variasi arsitektur dan hyperparameter tuning menggunakan Keras Tuner. Evaluasi performa dilakukan menggunakan metrik Accuracy, Precision, Recall, F1-Score, dan ROC-AUC, serta dibandingkan dengan model baseline Random Forest dan Support Vector Machine (SVM).

Hasil eksperimen menunjukkan bahwa model hybrid CNN–LSTM mampu mencapai performa sangat baik dengan akurasi tinggi dan kemampuan deteksi anomali yang lebih seimbang dibandingkan model baseline. Model terbaik selanjutnya dikonversi ke format TensorFlow Lite (TFLite) sehingga siap diimplementasikan pada perangkat edge IoT.

Kata kunci: CNN-LSTM, deteksi anomali, IoT, time series, TensorFlow Lite

Keyword: CNN-LSTM, deteksi anomali, IoT, time series, TensorFlow Lite

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan Internet of Things (IoT) telah membawa transformasi besar pada sistem otomasi industri (Industrial IoT / IIoT) dengan pemanfaatan sensor-sensor yang tersebar dalam rangka memantau kondisi operasional mesin secara real-time. Sensor-sensor tersebut menghasilkan data deret waktu (time series) secara kontinu, yang mencerminkan dinamika variabel penting seperti tekanan, suhu, getaran, dan konsumsi energi dalam proses industri. Namun demikian, data sensor sering dipengaruhi oleh noise, kesalahan pengukuran, serta anomali akibat keausan komponen dan kegagalan sistem, sehingga deteksi dini terhadap kondisi anomali merupakan aspek krusial untuk menjaga keandalan dan efektivitas operasional industri (Chatterjee & Ahmed, 2022).

Salah satu tantangan penting dalam sistem IoT industri adalah kehadiran anomali-anomali pada data sensor, yang dapat muncul akibat keausan komponen, kesalahan pengukuran, gangguan lingkungan, atau kegagalan sistem itu- sendiri. Anomali ini mencerminkan penyimpangan dari perilaku normal dan sering kali menjadi indikasi awal terjadinya kerusakan atau gangguan operasional yang serius. Deteksi dini terhadap anomali ini menjadi kunci untuk meningkatkan keandalan dan keamanan sistem industri melalui tindakan preventif, seperti perbaikan atau pemeliharaan sebelum terjadi kegagalan total. Metode tradisional untuk deteksi anomali, seperti berbasis aturan (rule-based) dan teknik statistik, memiliki keterbatasan signifikan dalam menangani data sensor IoT yang bersifat multivariat, tidak linear, dan berirama temporal kompleks. Metode statistik cenderung memerlukan asumsi distribusi data yang sering tidak realistis dalam praktik, serta kurang fleksibel dalam menangkap hubungan fitur jarak jauh dalam data deret waktu sensor. Akibatnya, pendekatan konvensional sering gagal ketika diterapkan pada skenario industri yang nyata (Singh et al., 2024).

Untuk mengatasi keterbatasan tersebut, pendekatan deep learning semakin banyak digunakan dalam deteksi anomali karena kemampuannya untuk mempelajari pola kompleks secara otomatis dari data mentah, tanpa memerlukan feature engineering manual yang intensif. Teknik deep learning seperti Convolutional Neural Network (CNN)

dan Long Short-Term Memory (LSTM) telah terbukti efektif dalam menangkap pola lokal dan dependensi temporal dalam data sensor. CNN dikembangkan untuk mengekstraksi fitur lokal dan pola struktural dalam sekuens data, sementara LSTM dirancang untuk menangkap ketergantungan jangka panjang dalam deret waktu yang sangat relevan dalam analisis data sensor. Dalam penelitian ini, fokus utama adalah mengembangkan sistem deteksi anomali pada data sensor kompresor udara industri berbasis IoT dengan menggunakan arsitektur CNN–LSTM dan membandingkan performanya dengan model pembelajaran mesin konvensional. Kompresor udara merupakan salah satu mesin penting dalam lini produksi yang perlu dipantau secara terus-menerus untuk menghindari kerusakan yang tidak terdeteksi. Pemilihan dataset MetroPT-3 sebagai studi kasus memberikan konteks nyata pada sistem industri berbasis sensor karena dataset ini mencerminkan karakteristik data sensor operasi nyata seperti variasi normal dan kejadian abnormal operasional mesin (link dataset dapat dicantumkan jika tersedia). Dengan pendekatan ini, diharapkan sistem dapat memberikan peringatan dini terhadap kondisi abnormal yang terjadi berdasarkan pola spasial-temporal data sensor, sehingga meningkatkan efektivitas pemeliharaan prediktif dan keandalan operasional keseluruhan (Gopali, 2022).

Salah satu tantangan penting dalam sistem IoT industri adalah kehadiran anomali pada data sensor, yang dapat muncul akibat keausan komponen, gangguan lingkungan, kesalahan pengukuran, atau kegagalan subsistem. Anomali ini mencerminkan penyimpangan dari perilaku normal dan sering kali menjadi indikasi awal terjadinya kerusakan atau gangguan operasional yang serius. Deteksi dini terhadap anomali ini menjadi kunci untuk meningkatkan keandalan dan keamanan sistem industri melalui tindakan preventif, seperti pemeliharaan prediktif sebelum terjadinya kegagalan total (*predictive maintenance*). Metode tradisional untuk deteksi anomali, seperti berbasis aturan (*rule-based*) dan teknik statistik, memiliki keterbatasan signifikan dalam menangani data sensor IoT yang bersifat multivariat, tidak linear, dan berirama temporal kompleks. Metode statistik cenderung memerlukan asumsi distribusi data yang sering tidak realistis dalam praktik, serta kurang fleksibel dalam menangkap hubungan fitur jarak jauh dalam data deret waktu sensor. Akibatnya, pendekatan konvensional sering gagal ketika diterapkan pada skenario industri nyata (Duraj et al., 2025).

Kajian komprehensif juga menunjukkan berbagai metode deteksi anomali dalam sistem IoT yang mengkombinasikan teknik Machine Learning klasik (mis. Random Forest, SVM) dengan pendekatan *deep learning* untuk mengatasi tantangan data besar, non-linear, dan *imbalanced* (Krzyszto, 2024). Selain itu, riset terbaru bahkan mengemukakan taksonomi lengkap teknik deteksi anomali *multivariate time-series*, mencakup metode statistik hingga modern termasuk deep learning, Transformer, dan metrik evaluasi serta dataset umum yang digunakan dalam penelitian (Wang et al., 2025).

Penelitian lain memainkan peran penting dalam mengkaji kinerja berbagai algoritma deteksi anomali pada jaringan berbasis IoT, menunjukkan bahwa model deep learning seperti LSTM dan CNN mampu menangkap dinamika data sensor lebih efektif dibanding model machine learning tradisional seperti Random Forest atau SVM (Li et al., 2024). Selain itu, penelitian yang membandingkan metode deep learning menunjukkan bahwa beberapa varian LSTM dan CNN memiliki kelebihan masing-masing di konteks tertentu, misalnya LSTM unggul dalam menangkap pola temporal sementara CNN menawarkan efisiensi komputasi pada ekstraksi fitur lokal pada deret waktu (Wu, 2024)

Pemilihan dataset **MetroPT-3 Air Compressor** sebagai studi kasus memberikan konteks yang realistis bagi sistem industri berbasis sensor karena dataset ini mencerminkan karakteristik operasi nyata seperti variasi normal dan kejadian abnormal pada kompresor udara. Penerapan pendekatan CNN–LSTM diharapkan mampu memberikan peringatan dini terhadap kondisi abnormal berdasarkan pola spasial-temporal data sensor, sehingga meningkatkan efektivitas pemeliharaan prediktif dan keandalan operasional keseluruhan (Gopali, 2022)

Dengan demikian, penelitian ini berupaya mengembangkan sistem deteksi anomali berbasis CNN–LSTM pada data sensor kompresor udara industri dan membandingkan kinerjanya dengan model pembelajaran mesin konvensional (mis. Random Forest & SVM), guna memberikan kontribusi terhadap penelitian deteksi anomali multivariate time-series pada IoT industri.

BAB II

METODOLOGI

2.1 Identifikasi Masalah dan Dataset

Permasalahan yang diangkat adalah mendeteksi anomali pada data sensor IoT industri yang bersifat multivariat dan time series. Dataset yang digunakan adalah **MetroPT-3 Air Compressor Dataset** dari UCI Machine Learning Repository. Dataset ini merekam berbagai parameter operasional kompresor udara, seperti tekanan, suhu, kelembaban, dan arus motor. Sebanyak 100.000 baris data pertama digunakan untuk eksperimen guna menjaga efisiensi komputasi.

2.2 Pra-pemrosesan dan Eksplorasi Data

Tahap pra-pemrosesan meliputi konversi kolom waktu menjadi format timestamp dan penetapan timestamp sebagai indeks data. Selanjutnya dilakukan pelabelan anomali berbasis aturan (rule-based), yaitu:

- Data dikategorikan sebagai **anomali** apabila tekanan TP2 < 5 bar saat arus motor tinggi, atau suhu oli > 80°C.
- Label 0 menunjukkan kondisi normal, dan label 1 menunjukkan kondisi anomali.

Distribusi kelas divisualisasikan untuk memahami tingkat ketidakseimbangan data.

2.3 Normalisasi Data

Data sensor dinormalisasi menggunakan **Min-Max Scaling** ke rentang [0,1]. Normalisasi ini bertujuan untuk menyamakan skala fitur agar proses pelatihan model CNN-LSTM lebih stabil dan konvergen.

2.4 Segmentasi Data Menggunakan Sliding Window

Data deret waktu disegmentasi menggunakan metode sliding window dengan ukuran window sebesar 50 timestep. Setiap window merepresentasikan satu sampel input model, sedangkan label ditentukan berdasarkan kondisi pada timestep berikutnya.

2.5 Pembagian Dataset

Dataset hasil windowing dibagi menjadi:

- 70% data latih
- 15% data validasi
- 15% data uji

Pembagian dilakukan secara berurutan tanpa pengacakan untuk menjaga karakteristik temporal data.

2.6 Perancangan dan Implementasi Model CNN–LSTM

Model CNN–LSTM terdiri dari lapisan Conv1D dan MaxPooling sebagai feature extractor, diikuti lapisan LSTM untuk mempelajari ketergantungan temporal. Lapisan Dropout digunakan untuk mencegah overfitting, dan lapisan Dense sigmoid digunakan untuk klasifikasi biner.

2.7 Pelatihan, Optimasi, dan Deployment

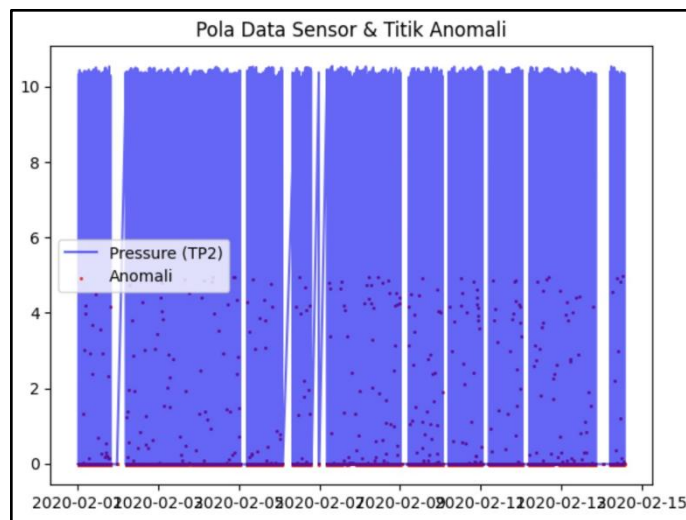
Model dilatih menggunakan optimizer Adam dan loss binary cross-entropy dengan class weighting. Early stopping digunakan untuk mencegah overfitting. Optimasi hyperparameter dilakukan menggunakan **Keras Tuner**. Model terbaik kemudian dikonversi ke format **TensorFlow Lite (TFLite)** untuk mendukung deployment pada perangkat edge.

BAB III

HASIL DAN PEMBAHASAN

3.1 Eksplorasi dan Pra-pemrosesan Data

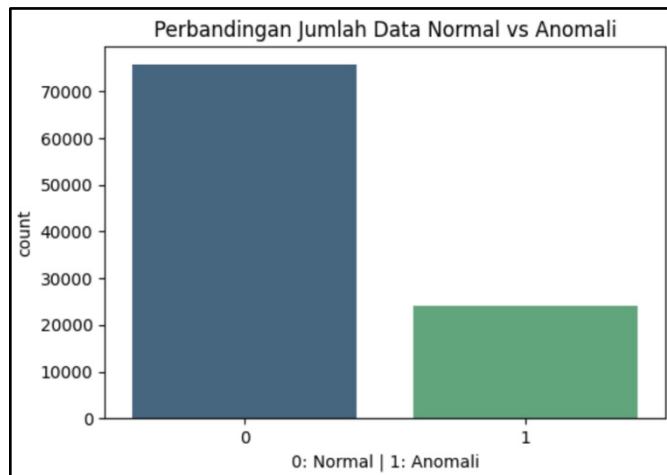
Hasil eksplorasi menunjukkan bahwa data sensor bersifat multivariat dan tidak seimbang, dengan jumlah data normal jauh lebih banyak dibandingkan data anomali. Kondisi ini menegaskan bahwa permasalahan yang dihadapi merupakan kasus imbalanced classification.



Gambar 3.1 Pola Tekanan Sensor TP2 dan Titik Anomali

Penjelasan:

Gambar ini menunjukkan pola deret waktu sensor tekanan TP2 pada kompresor udara, di mana garis biru merepresentasikan kondisi normal dan titik merah menunjukkan data yang dilabeli sebagai anomali berdasarkan aturan tekanan dan suhu. Visualisasi ini memperlihatkan bahwa anomali muncul sebagai penyimpangan signifikan dari pola tekanan normal, sehingga menegaskan pentingnya deteksi dini terhadap kondisi operasional tidak normal pada sistem IoT industri.



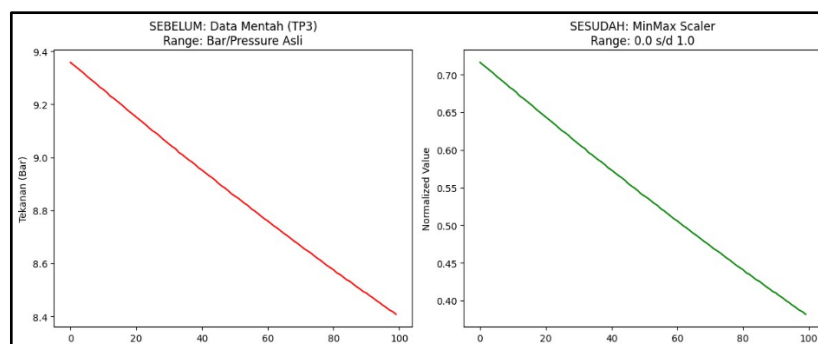
Gambar 3.2 Distribusi Data Normal dan Anomali

Penjelasan:

Gambar ini menampilkan perbandingan jumlah data normal dan data anomali dalam dataset MetroPT-3. Terlihat bahwa data normal jauh lebih dominan dibandingkan data anomali, yang menunjukkan bahwa permasalahan deteksi anomali pada penelitian ini termasuk kasus ketidakseimbangan kelas (imbalanced data), sehingga diperlukan teknik class weighting dalam pelatihan model.

3.2 Normalisasi dan Segmentasi Data

Normalisasi menggunakan Min-Max Scaling berhasil mempertahankan pola temporal data tanpa mengubah bentuk gelombang sinyal sensor. Proses sliding window menghasilkan data tiga dimensi dengan format (samples, time steps, features) yang sesuai untuk model CNN–LSTM.



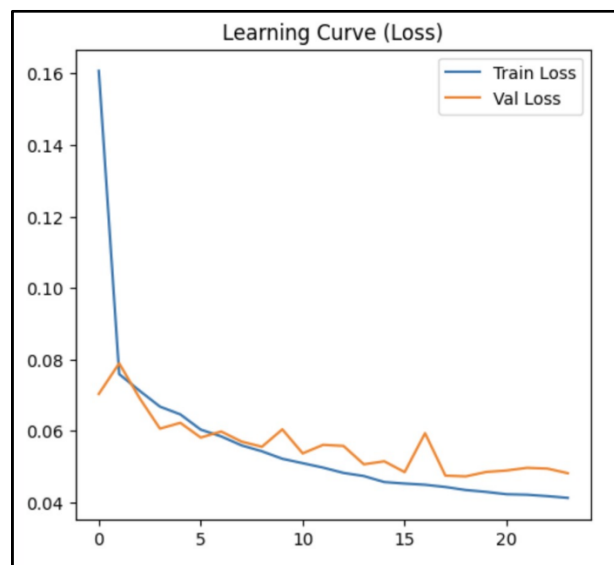
Gambar 3.3 Perbandingan Data Sensor Sebelum dan Sesudah Normalisasi

Penjelasan:

Gambar ini memperlihatkan sinyal sensor TP3 sebelum dan sesudah dilakukan normalisasi menggunakan Min-Max Scaling. Hasil visualisasi menunjukkan bahwa normalisasi tidak mengubah pola temporal data, melainkan hanya menyamakan skala nilai ke rentang 0–1 agar proses pelatihan model CNN–LSTM menjadi lebih stabil dan efektif.

3.3 Evaluasi Model CNN–LSTM

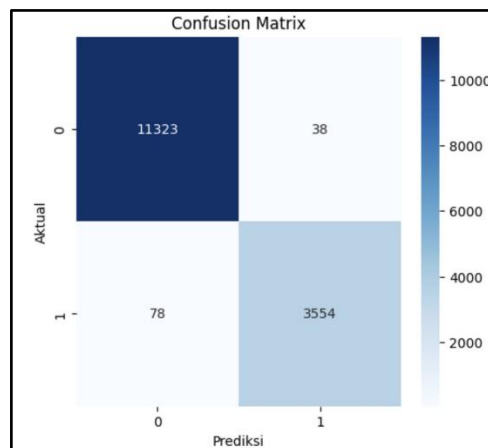
Model CNN–LSTM menunjukkan performa yang sangat baik dengan nilai accuracy, precision, recall, dan F1-score yang tinggi. Nilai ROC-AUC mendekati 1,0 menandakan kemampuan pemisahan kelas yang sangat baik.



Gambar 3.4 Kurva Pembelajaran (Loss Training dan Validasi)

Penjelasan:

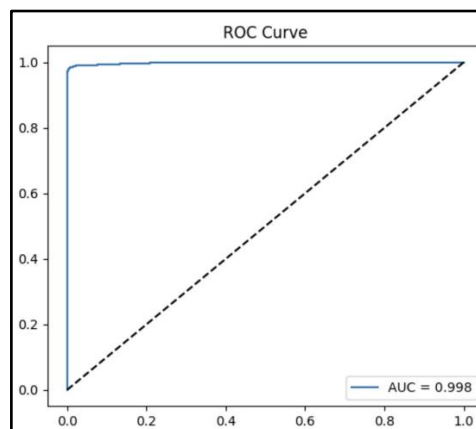
Gambar ini menunjukkan kurva loss pada data latih dan data validasi selama proses pelatihan model CNN–LSTM. Penurunan loss yang konsisten dan stabil menandakan bahwa model mampu mempelajari pola data dengan baik serta tidak mengalami overfitting yang signifikan berkat penerapan early stopping.



Gambar 3.5 Confusion Matrix Model CNN–LSTM

Penjelasan:

Gambar ini menampilkan confusion matrix hasil prediksi model CNN–LSTM pada data uji. Mayoritas data, baik kelas normal maupun anomali, berhasil diklasifikasikan dengan benar, yang menunjukkan bahwa model memiliki kemampuan deteksi anomali yang baik pada data sensor IoT industri.



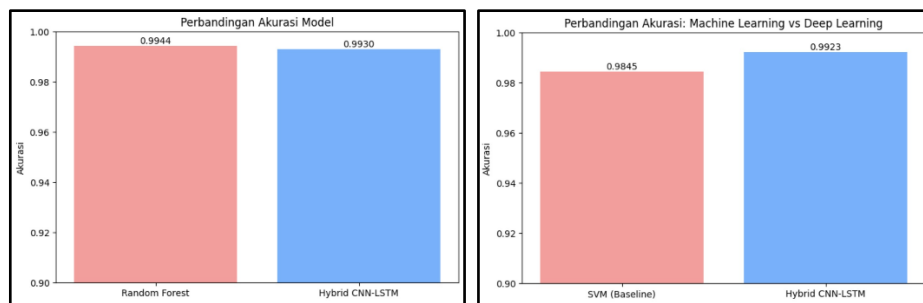
Gambar 3.6 Kurva ROC Model CNN–LSTM

Penjelasan:

Gambar ini memperlihatkan kurva ROC yang menggambarkan hubungan antara True Positive Rate dan False Positive Rate pada berbagai threshold. Nilai AUC yang mendekati 1 menunjukkan bahwa model CNN–LSTM memiliki kemampuan pemisahan kelas yang sangat baik antara data normal dan data anomali.

3.4 Perbandingan dengan Model Baseline

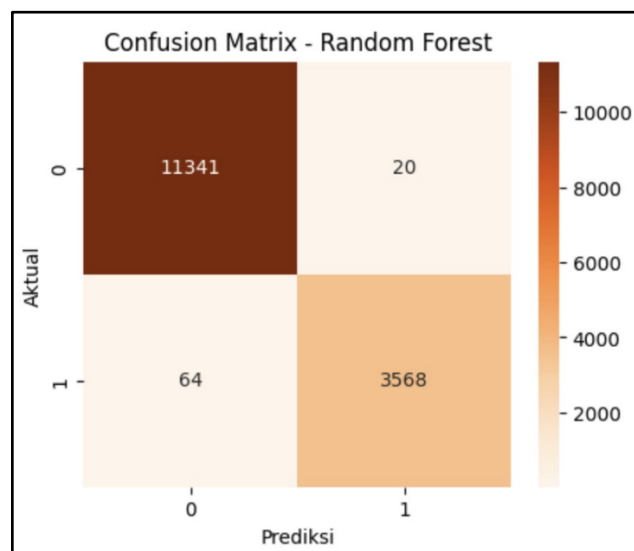
Sebagai pembanding, digunakan Random Forest dan Support Vector Machine (SVM). Hasil eksperimen menunjukkan bahwa CNN–LSTM unggul dalam mendeteksi anomali dibandingkan kedua model baseline, terutama dalam hal recall dan F1-score.



Gambar 3.7 Perbandingan Akurasi CNN–LSTM dan Model Baseline

Penjelasan:

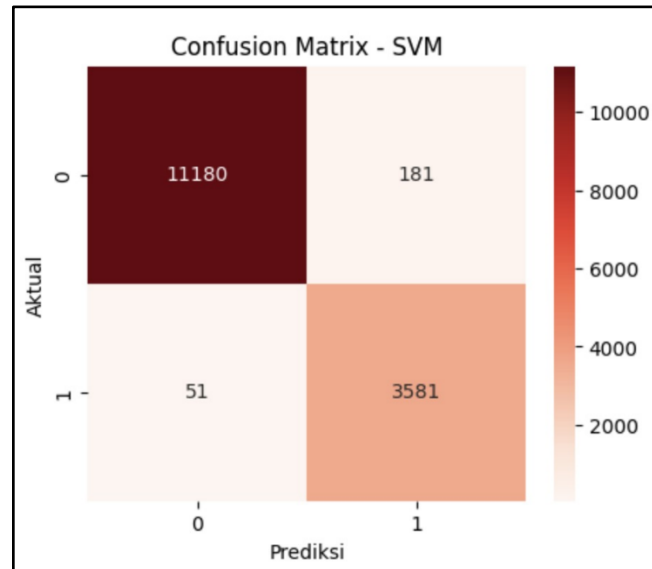
Gambar ini menunjukkan perbandingan akurasi antara model CNN–LSTM dan model baseline Random Forest serta SVM. Hasilnya memperlihatkan bahwa CNN–LSTM memiliki akurasi yang lebih tinggi, karena mampu menangkap pola spasial dan temporal pada data sensor deret waktu secara lebih efektif.



Gambar 3.8 Confusion Matrix Random Forest

Penjelasan:

Gambar ini menampilkan confusion matrix model Random Forest yang menunjukkan bahwa meskipun model mampu mengklasifikasikan data normal dengan baik, masih terdapat kesalahan dalam mendeteksi data anomali, terutama false negative, yang menunjukkan keterbatasan model dalam menangani data time series.



Gambar 3.9 Confusion Matrix SVM

Penjelasan:

Gambar ini menunjukkan confusion matrix model SVM yang memperlihatkan performa lebih rendah dibandingkan CNN-LSTM, khususnya dalam mendeteksi kelas anomali. Hal ini disebabkan oleh keterbatasan SVM dalam memodelkan hubungan temporal jangka panjang pada data sensor IoT.

3.5 Export Model

Model terbaik berhasil dikonversi ke format TensorFlow Lite dengan ukuran file yang relatif kecil, sehingga siap digunakan untuk implementasi pada perangkat IoT berbasis edge computing.

BAB IV

KESIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa model hybrid CNN–LSTM mampu mendeteksi anomali pada data sensor IoT industri dengan performa yang sangat baik. Pendekatan ini lebih unggul dibandingkan model machine learning konvensional seperti Random Forest dan SVM, terutama dalam menangani data time series yang kompleks dan tidak seimbang. Konversi model ke format TensorFlow Lite memungkinkan penerapan sistem deteksi anomali pada perangkat edge IoT secara efisien. Penelitian selanjutnya dapat mengembangkan metode pelabelan yang lebih adaptif serta menguji model pada data real-time.

DAFTAR PUSTAKA

- Chatterjee, A., & Ahmed, B. S. (2022). Internet of Things IoT anomaly detection methods and applications : A survey. *Internet of Things*, 19(June), 100568. <https://doi.org/10.1016/j.iot.2022.100568>
- Duraj, A., Szczepaniak, P. S., & Sadok, A. (2025). *Detection of Anomalies in Data Streams Using the LSTM-CNN Model*. 1–27.
- Gopali, S. (2022). *Deep Learning-Based Time-Series Analysis for Detecting Anomalies in Internet of Things*.
- Krzyszto, E. (2024). *A Comparative Analysis of Anomaly Detection Methods in IoT Networks : An Experimental Study*.
- Li, L., Zhang, Y., Wang, J., & Xiong, K. (2024). *Deep Learning-Based Network Traffic Anomaly Detection : A Study in IoT Environments*. 7(6), 13–26. [https://doi.org/10.53469/wjimt.2024.07\(06\).03](https://doi.org/10.53469/wjimt.2024.07(06).03)
- Singh, A., Singh, S., Alam, M. N., & Singh, G. (2024). *Deep Learning for Anomaly Detection in IoT Systems : Techniques , Applications , and Future Directions*. 6(4), 1–9.
- Wang, F., Jiang, Y., Zhang, R., Wei, A., Xie, J., & Pang, X. (2025). *A Survey of Deep Anomaly Detection in Multivariate Time Series : Taxonomy , Applications , and Directions*. I, 1–27.
- Wu, W. (2024). *Time - series Anomaly Detection Study Based on Deep Learning*. 13(2).