

Machine Learning untuk Predictive Maintenance Motor: Studi Komparasi Isolation Forest, Autoencoder, dan Z-score

Dzaky Rinaldi^{1*)}, Indri Purwita Sary²

^{1,2}Program Studi Teknik Elektro, Fakultas Teknik, Universitas Singaperbangsa Karawang, Jalan H.S. Ronggo Waluyo, Karawang, Indonesia.

*Penulis Korespondensi, e-mail: 2110631160041@student.unsika.ac.id

Received: 17/06/2026

Revised: 18/07/2026

Accepted: 18/07/2026

ABSTRAK

Kerusakan bearing menjadi salah satu penyebab dominan kegagalan pada motor elektrik di lingkungan industri, sehingga deteksi dini kondisi anomali menjadi kebutuhan penting dalam strategi predictive maintenance. Penelitian ini bertujuan membandingkan kemampuan tiga metode deteksi anomali, yaitu Isolation Forest, Autoencoder, dan Z-score, dalam mengidentifikasi kondisi anomali pada sinyal vibrasi bearing motor. Penelitian menggunakan dataset Case Western Reserve University (CWRU) yang berisi 2300 sampel dengan sembilan fitur statistik domain waktu (max, min, mean, standard deviation, RMS, skewness, kurtosis, crest factor, dan form factor) dari sepuluh kondisi bearing, terdiri dari satu kondisi normal dan sembilan jenis kerusakan dengan variasi lokasi dan tingkat keparahan. Ketiga model dilatih hanya menggunakan data kondisi normal, kemudian diuji pada keseluruhan data untuk mengevaluasi kemampuan deteksi terhadap kondisi anomali. Hasil pengujian menunjukkan Z-score mencapai accuracy 0,9961 dan F1-score 0,9978, Isolation Forest mencapai accuracy 0,9900 dan F1-score 0,9945, sedangkan Autoencoder mencapai accuracy 0,8600 dan F1-score 0,9167 dengan recall terendah di antara ketiga metode (0,8556). Temuan ini menunjukkan bahwa metode statistik sederhana dapat memberikan performa setara atau lebih baik dibandingkan metode berbasis deep learning pada kasus fitur domain waktu yang memiliki separabilitas tinggi antarkondisi. Hasil penelitian ini dapat menjadi pertimbangan dalam pemilihan metode deteksi anomali untuk sistem pemantauan kondisi motor elektrik berbasis machine learning.

Kata Kunci: anomaly detection, autoencoder, bearing, isolation forest, predictive maintenance

ABSTRACT

Bearing failure is one of the dominant causes of electric motor breakdown in industrial environments, making early detection of anomalous conditions essential for predictive maintenance strategies. This study aims to compare the capability of three anomaly detection methods, namely Isolation Forest, Autoencoder, and Z-score, in identifying anomalous conditions in bearing vibration signals. The study uses the Case Western Reserve University (CWRU) dataset, consisting of 2300 samples with nine time-domain statistical features (max, min, mean, standard deviation, RMS, skewness, kurtosis, crest factor, and form factor) derived from ten bearing conditions, comprising one normal condition and nine fault types with varying locations and severity levels. All three models were trained exclusively on normal condition data and then evaluated on the entire dataset to assess their anomaly detection capability. The test results show that Z-score achieved an accuracy of 0.9961 and an F1-score of 0.9978, Isolation Forest achieved an accuracy of 0.9900 and an F1-score of 0.9945, while Autoencoder achieved an accuracy of 0.8600 and an F1-score of 0.9167, with the lowest recall among the three methods (0.8556). These findings indicate that simple statistical methods can deliver performance equal to or better than deep learning-based methods when time-domain features exhibit high separability between conditions. The results of this study can serve as a consideration in selecting anomaly detection methods for machine learning-based electric motor condition monitoring systems.

Keywords: anomaly detection, autoencoder, bearing, isolation forest, predictive maintenance



1. PENDAHULUAN

Motor elektrik merupakan komponen penggerak utama pada hampir seluruh proses produksi di sektor industri, sehingga kondisi operasionalnya berpengaruh langsung terhadap kelangsungan produksi secara keseluruhan [2]. Di antara berbagai komponen motor, bearing tercatat sebagai salah satu titik kegagalan yang paling sering terjadi. Survei yang dilakukan oleh Electric Power Research Institute (EPRI) melaporkan bahwa 41% dari total kasus kerusakan motor induksi terjadi pada bagian bearing [1]. Kegagalan bearing yang tidak terdeteksi sejak awal dapat berkembang menjadi kerusakan motor secara menyeluruh, menyebabkan downtime produksi, serta menimbulkan biaya perbaikan yang lebih besar dibandingkan jika ditangani pada fase awal.

Pendekatan pemeliharaan reaktif, yaitu perbaikan setelah kerusakan terjadi, maupun pemeliharaan preventif berbasis jadwal tetap, mulai dianggap kurang efisien dalam konteks Industri 4.0 yang menuntut efisiensi operasional dan minimnya gangguan produksi [7]. Predictive maintenance menjadi alternatif yang lebih sesuai, karena kondisi peralatan dipantau secara berkelanjutan menggunakan data sensor, dan tindakan perawatan dilakukan berdasarkan indikasi kerusakan yang terdeteksi, bukan berdasarkan jadwal yang telah ditentukan sebelumnya. Wisna menunjukkan penerapan predictive maintenance berbasis pemantauan vibrasi pada motor high pressure oil pump di sebuah pembangkit listrik tenaga uap, di mana analisis spektrum vibrasi digunakan untuk menentukan kelayakan operasional motor sebelum kerusakan besar terjadi [2]. Riyanto, Asri, dan Nugraha juga menunjukkan bahwa pemantauan akselerasi getaran dan suhu secara langsung pada motor induksi dapat digunakan sebagai data pendukung untuk memprediksi umur operasional motor, sehingga tindakan perawatan dapat direncanakan lebih awal [9]. Survei mutakhir terhadap penerapan predictive maintenance berbasis Internet of Things juga menegaskan bahwa kombinasi sensor vibrasi, arus, dan suhu dengan algoritma machine learning semakin banyak digunakan untuk mendeteksi pola tidak normal pada motor sebelum kerusakan terjadi, sekaligus memungkinkan tindakan perawatan dilakukan secara proaktif [11].

Penelitian deteksi kerusakan bearing dengan pendekatan machine learning telah banyak dilakukan menggunakan berbagai jenis sinyal dan metode ekstraksi fitur. Naufal dkk. menggunakan ekstraksi fitur Discrete Wavelet Transform pada sinyal arus motor induksi 3 fasa yang dipadukan dengan metode klasterisasi K-Medoids, dan memperoleh akurasi 86,8% pada tiga kondisi bearing [1]. Pada penelitian tersebut juga disebutkan bahwa metode ekstraksi fitur berbasis Empirical Mode Decomposition yang dipadukan dengan Artificial Neural Network mampu mencapai akurasi 91%, sedangkan ekstraksi fitur menggunakan Fast Fourier Transform dengan analisis statistika manual hanya mencapai akurasi 37,04% [1]. Variasi hasil tersebut menunjukkan bahwa performa deteksi kerusakan bearing bergantung pada kombinasi metode ekstraksi fitur dan algoritma klasifikasi atau deteksi anomali yang digunakan.

Pada sisi algoritma deteksi anomali, tiga pendekatan yang umum digunakan adalah metode statistik, metode ensemble berbasis pohon, dan metode deep learning. Z-score merupakan metode statistik klasik yang mengidentifikasi anomali berdasarkan jarak suatu nilai terhadap rata-rata data dalam satuan standar deviasi. Sihombing dkk. menggunakan pendekatan ini untuk mengidentifikasi data pencilan pada data univariat, dengan ambang umum tiga standar deviasi untuk menandai suatu nilai sebagai outlier [5]. Isolation Forest adalah algoritma ensemble unsupervised yang mengidentifikasi anomali berdasarkan kemudahan suatu titik data untuk diisolasi dari titik data lainnya. Amanda dkk. menerapkan Isolation Forest tanpa label pada data kualitas udara dan menunjukkan bahwa metode ini efektif mengidentifikasi kondisi tidak normal secara cepat tanpa memerlukan data berlabel sebelumnya [4]. Autoencoder adalah arsitektur jaringan saraf tiruan yang dilatih untuk merekonstruksi data input, dengan asumsi bahwa data normal dapat direkonstruksi dengan error yang kecil, sedangkan data anomali akan menghasilkan reconstruction error yang lebih besar. Umari dan Supardi menerapkan autoencoder pada data vibrasi tiga unit blower di sebuah industri pupuk, dan melaporkan bahwa model tersebut efektif mendeteksi anomali dengan akurasi tinggi dan keseimbangan yang baik antara deteksi kondisi normal dan anomali [3]. Perbandingan langsung antara Autoencoder dan Isolation Forest pada domain anomaly detection lain, seperti pada lalu lintas jaringan komputer, juga melaporkan bahwa kedua metode tersebut dapat menghasilkan skor anomali yang berbeda secara signifikan



terhadap sampel yang sama, yang mengindikasikan bahwa karakteristik data sangat menentukan metode mana yang lebih unggul untuk suatu kasus tertentu [11], [12].

Pemilihan Isolation Forest, Autoencoder, dan Z-score pada penelitian ini didasarkan pada pertimbangan bahwa ketiganya merepresentasikan tiga paradigma utama deteksi anomali yang paling umum dibandingkan pada literatur, yaitu paradigma statistik, ensemble berbasis pohon, dan deep learning berbasis rekonstruksi [14]. Metode deteksi anomali unsupervised lain yang juga banyak digunakan, seperti One-Class Support Vector Machine (SVM) dan Local Outlier Factor (LOF), tidak dipilih pada penelitian ini karena beberapa pertimbangan praktis. One-Class SVM memiliki dasar teoritis yang kuat, tetapi kebutuhan komputasinya tinggi dan performanya sensitif terhadap pemilihan kernel serta parameter ν , sehingga proses tuning menjadi lebih rumit dibandingkan ketiga metode yang dipilih [14]. LOF juga menunjukkan beban komputasi yang besar dan sensitivitas tinggi terhadap jumlah tetangga (k) yang digunakan, yang dapat menurunkan presisi deteksi apabila parameter tersebut tidak disesuaikan secara cermat dengan karakteristik data [14]. Metode berbasis kepadatan seperti DBSCAN maupun metode berbasis proyeksi linear seperti PCA juga umumnya mengasumsikan struktur data tertentu, seperti kepadatan kluster yang seragam atau hubungan antarfitur yang linear, yang belum tentu sesuai dengan karakteristik fitur statistik domain waktu pada dataset CWRU. Berdasarkan pertimbangan tersebut, Isolation Forest dipilih sebagai representasi metode ensemble yang tidak memerlukan asumsi distribusi tertentu dan relatif efisien secara komputasi [4], Autoencoder dipilih sebagai representasi metode deep learning yang mampu mempelajari pola nonlinear dari data [3], [6], sedangkan Z-score dipilih sebagai representasi metode statistik klasik yang sederhana, murah secara komputasi, dan mudah diinterpretasikan [5]. Ketiga metode ini juga telah banyak digunakan sebagai baseline pembanding pada penelitian deteksi anomali di berbagai domain [7], [12], sehingga perbandingan ketiganya diharapkan memberikan gambaran yang representatif mengenai trade-off antara kompleksitas metode, kebutuhan komputasi, dan performa deteksi anomali pada kasus bearing motor elektrik.

Ketiga metode tersebut juga telah diterapkan pada dataset internasional Case Western Reserve University (CWRU) dalam beberapa penelitian. Zhang dkk. menerapkan variational autoencoder pada dataset CWRU dan IMS untuk klasifikasi semi-supervised kondisi bearing, dan membandingkan performanya dengan autoencoder konvensional serta CNN [6]. Jorry dkk. mengusulkan kerangka deteksi kerusakan bearing berbasis statistical process control yang juga hanya menggunakan data kondisi normal pada tahap pelatihan model, dan menunjukkan bahwa pendekatan statistik mampu mendeteksi seluruh jenis kerusakan pada dataset CWRU tanpa false alarm yang signifikan [10]. Kajian komprehensif oleh Das dkk. mengidentifikasi bahwa metode-metode unsupervised seperti Isolation Forest banyak digunakan sebagai baseline pembanding pada penelitian deteksi anomali berbasis sensor di sektor manufaktur, dan menekankan bahwa pemilihan metode machine learning untuk diagnosis kerusakan mesin rotasi perlu disesuaikan dengan karakteristik data yang tersedia [7]. Meskipun demikian, perbandingan langsung antara metode statistik sederhana seperti Z-score dengan metode yang lebih kompleks seperti Isolation Forest dan Autoencoder pada fitur statistik domain waktu dataset CWRU masih belum banyak dilaporkan secara eksplisit, sehingga belum diketahui secara pasti apakah peningkatan kompleksitas algoritma sebanding dengan peningkatan akurasi deteksi pada jenis data tersebut.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk membandingkan performa tiga metode deteksi anomali, yaitu Isolation Forest, Autoencoder, dan Z-score, dalam mengidentifikasi kondisi anomali pada bearing motor elektrik menggunakan fitur statistik domain waktu dari dataset CWRU. Perbandingan dilakukan berdasarkan metrik accuracy, precision, recall, F1-score, dan ROC-AUC, sehingga dapat memberikan gambaran mengenai trade-off antara kompleksitas metode dan performa deteksi yang dihasilkan.

2. METODE PENELITIAN



Penelitian ini dilakukan melalui beberapa tahapan, yaitu pengumpulan data, ekstraksi dan eksplorasi fitur, praproses data, pelatihan model, dan evaluasi performa ketiga metode deteksi anomali. Tahapan penelitian secara umum mengikuti pendekatan unsupervised anomaly detection, di mana model dilatih hanya menggunakan data kondisi normal, kemudian diuji pada seluruh data untuk menilai kemampuan model dalam membedakan kondisi normal dan kondisi anomali.

2.1 Dataset

Dataset yang digunakan dalam penelitian ini bersumber dari Case Western Reserve University (CWRU) Bearing Data Center, yang merupakan dataset acuan yang banyak digunakan dalam penelitian diagnosis kerusakan bearing [6]. Data diakses melalui versi yang telah diunggah ke platform Kaggle [13], yang menyediakan data sinyal vibrasi beserta hasil ekstraksi fitur statistik domain waktu yang sudah dihitung sebelumnya. Dataset terdiri dari sinyal vibrasi yang direkam menggunakan akselerometer pada drive-end motor dengan frekuensi sampling 48 kHz pada kondisi pembebanan 1 HP. Dari sinyal vibrasi tersebut, telah diekstraksi sembilan fitur statistik domain waktu untuk setiap segmen sinyal, yaitu nilai maksimum (max), nilai minimum (min), rata-rata (mean), standar deviasi (sd), root mean square (rms), skewness, kurtosis, crest factor (crest), dan form factor (form).

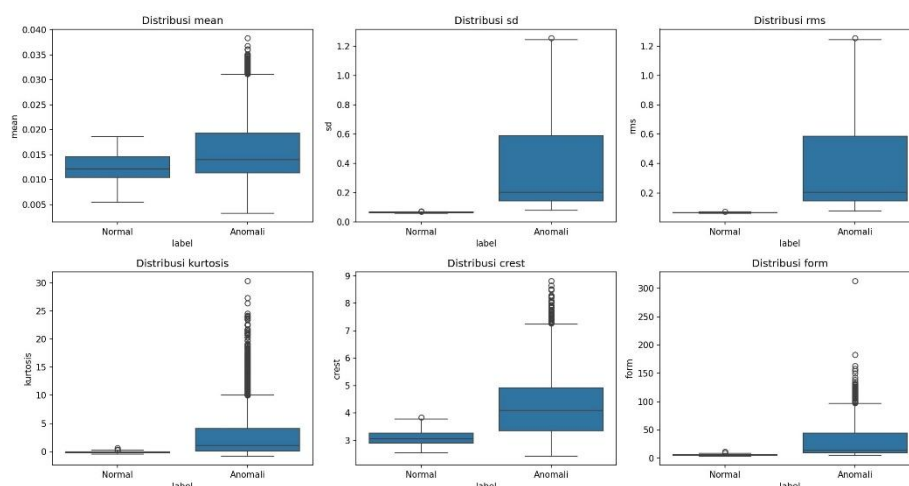
Dataset terdiri dari 2300 sampel yang terbagi rata ke dalam 10 kategori kondisi bearing, masing-masing berjumlah 230 sampel, sebagaimana ditunjukkan pada Tabel I. Kategori Normal merepresentasikan kondisi bearing tanpa cacat, sedangkan sembilan kategori lainnya merepresentasikan kondisi cacat pada tiga lokasi (Ball, Inner Race/IR, dan Outer Race/OR) dengan tiga tingkat diameter cacat (0,007; 0,014; dan 0,021 inci) [6].

TABEL I: DISTRIBUSI KATEGORI KONDISI BEARING

Kategori	Jumlah Sampel	Label Biner
Normal	230	Normal
Ball 0,007"	230	Anomali
Ball 0,014"	230	Anomali
Ball 0,021"	230	Anomali
IR 0,007"	230	Anomali
IR 0,014"	230	Anomali
IR 0,021"	230	Anomali
OR 0,007"	230	Anomali
OR 0,014"	230	Anomali
OR 0,021"	230	Anomali
Total	2300	230 Normal, 2070 Anomali

Untuk keperluan deteksi anomali, label asli yang terdiri dari 10 kategori dikonversi menjadi label biner, yaitu kelas Normal (label 0) dan kelas Anomali (label 1) yang mencakup seluruh kategori kerusakan. Konversi ini menghasilkan distribusi kelas yang tidak seimbang, dengan proporsi 230 sampel Normal berbanding 2070 sampel Anomali. Ketidak seimbangan ini diantisipasi dengan penggunaan metrik evaluasi precision, recall, F1-score, dan ROC-AUC, selain accuracy, karena metrik-metrik tersebut lebih representatif pada kondisi distribusi kelas yang tidak seimbang [7].





Gambar 1: Distribusi nilai enam fitur statistik (mean, sd, rms, kurtosis, crest, dan form) untuk kondisi Normal dan Anomali, ditampilkan dalam bentuk boxplot.

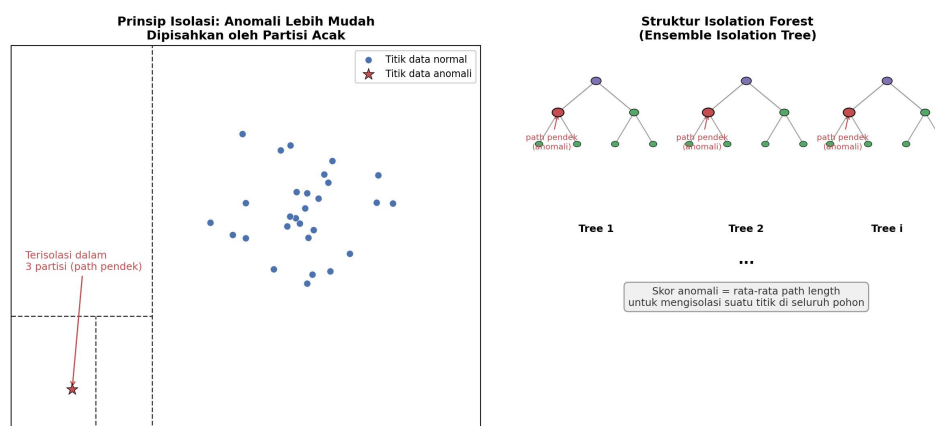
Sebagaimana ditunjukkan pada Gambar 1, fitur-fitur seperti sd, rms, kurtosis, crest, dan form menunjukkan rentang nilai yang relatif sempit dan stabil pada kondisi Normal, sedangkan pada kondisi Anomali rentang nilainya jauh lebih lebar dengan median yang lebih tinggi. Pola ini menunjukkan bahwa fitur statistik domain waktu memiliki kemampuan diskriminatif yang baik terhadap kondisi cacat pada bearing, konsisten dengan karakteristik dataset CWRU yang telah dilaporkan pada penelitian sebelumnya [6].

2.2 Praproses Data

Sebelum pelatihan model, seluruh fitur dinormalisasi menggunakan StandardScaler, yang mentransformasi setiap fitur sehingga memiliki rata-rata 0 dan standar deviasi 1. Ambarwari, Adrian, dan Herdiyeni menunjukkan bahwa teknik standarisasi (zero-mean normalization) memengaruhi performa algoritma machine learning secara berbeda-beda, dan pemilihan teknik scaling yang tepat perlu disesuaikan dengan karakteristik algoritma yang digunakan [8]. Normalisasi pada penelitian ini diperlukan terutama untuk Isolation Forest dan Autoencoder, karena kedua model tersebut sensitif terhadap perbedaan skala antarfitur. Data yang telah dinormalisasi kemudian dipisahkan menjadi subset data Normal, yang digunakan untuk melatih ketiga model, dan keseluruhan data (Normal dan Anomali) yang digunakan untuk evaluasi.

2.3 Isolation Forest

Isolation Forest bekerja dengan membangun sejumlah pohon isolasi (isolation tree), di mana setiap pohon dibentuk dengan melakukan pemisahan data secara acak berdasarkan fitur dan nilai ambang yang dipilih secara acak [4]. Titik data anomali, yang secara karakteristik berada jauh dari titik data normal, cenderung dapat diisolasi dengan jumlah pemisahan yang lebih sedikit dibandingkan titik data normal. Skor anomali dihitung berdasarkan rata-rata kedalaman path yang dibutuhkan untuk mengisolasi suatu titik data di seluruh pohon dalam forest.



Gambar 2: Ilustrasi prinsip kerja Isolation Forest, yaitu partisi acak yang mengisolasi titik anomali dengan path lebih pendek (kiri), dan struktur ensemble beberapa isolation tree yang digunakan untuk menghitung skor anomali (kanan).

Pada penelitian ini, Isolation Forest dilatih menggunakan 230 sampel data Normal dengan parameter contamination sebesar 0,1 dan jumlah estimator sebanyak 100 pohon. Model kemudian digunakan untuk memprediksi label seluruh 2300 sampel data.

Pemilihan nilai contamination sebesar 0,1 mengikuti nilai yang umum digunakan sebagai titik awal pada implementasi Isolation Forest ketika proporsi anomali yang sesungguhnya tidak diketahui secara pasti pada tahap pelatihan, mengingat model pada penelitian ini hanya dilatih menggunakan data kondisi Normal tanpa contoh data Anomali [14]. Nilai tersebut merepresentasikan asumsi konservatif bahwa sebagian kecil data pelatihan berpotensi mengandung variasi alami yang menyerupai pencilan, tanpa mengasumsikan proporsi anomali yang berlebihan. Studi Nicolae dan Korodi menunjukkan bahwa performa Isolation Forest cukup sensitif terhadap pemilihan contamination, dengan pengujian pada nilai yang lebih rendah maupun lebih tinggi dari titik optimalnya sama-sama dapat menurunkan performa deteksi apabila tidak disesuaikan dengan karakteristik data yang dianalisis [14]. Jumlah estimator sebanyak 100 pohon dipilih mengikuti praktik umum pada implementasi Isolation Forest, di mana penambahan jumlah pohon melampaui 100 pada umumnya tidak lagi memberikan peningkatan performa yang signifikan, sementara jumlah pohon yang terlalu sedikit berisiko menghasilkan skor anomali yang kurang stabil antar-pengulangan pelatihan [4].

2.4 Z-score

Z-score menghitung seberapa jauh suatu nilai data menyimpang dari rata-rata data referensi dalam satuan standar deviasi, sebagaimana dirumuskan pada (1).

$$Z = \frac{x - \mu}{\sigma} \quad (1)$$

dengan x merupakan nilai data yang dievaluasi, μ adalah rata-rata data referensi, dan σ adalah standar deviasi data referensi.

Pada penelitian ini, μ dan σ dihitung dari 230 sampel data Normal untuk setiap fitur. Suatu sampel dikategorikan sebagai Anomali apabila terdapat minimal satu fitur dengan nilai $|Z|$ lebih besar dari 3, yang merupakan ambang umum dalam deteksi outlier berbasis Z-score [5]. Skor anomali untuk setiap sampel didefinisikan sebagai nilai $|Z|$ maksimum di antara seluruh fitur.

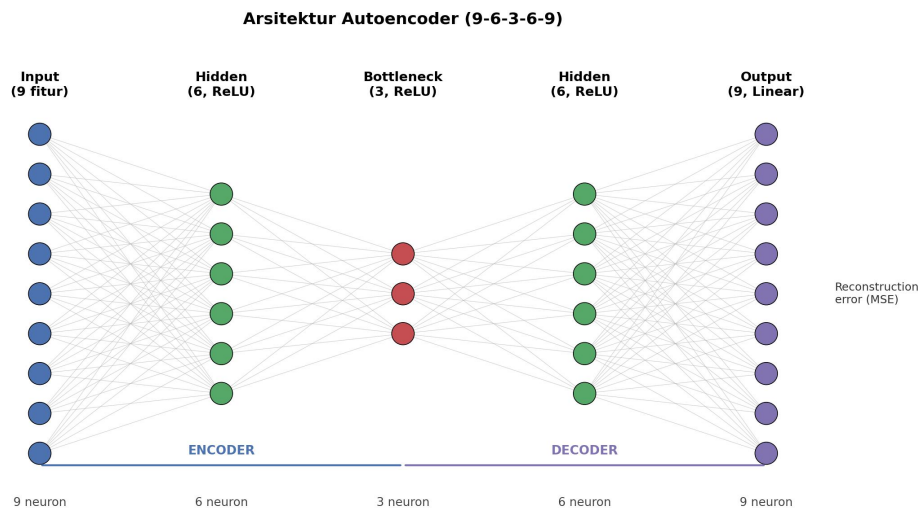
Pemilihan ambang $|Z| > 3$ mengikuti aturan empiris (empirical rule) dalam statistik, di mana pada distribusi yang mendekati normal sekitar 99,7% data berada dalam rentang tiga standar deviasi dari rata-rata, sehingga nilai di luar rentang tersebut dapat dianggap sebagai penyimpangan yang jarang terjadi pada kondisi normal [5]. Ambang ini dipilih karena kesederhanaan dan konsistensinya dengan praktik umum deteksi outlier berbasis Z-score pada berbagai domain [5]. Pengujian sensitivitas terhadap ambang Z-score pada studi Nicolae dan Korodi menunjukkan bahwa ambang yang lebih longgar, misalnya dua standar



deviasi, cenderung meningkatkan recall namun disertai peningkatan false positive, sehingga ambang tiga standar deviasi dipilih pada penelitian ini sebagai titik keseimbangan yang lebih konservatif antara sensitivitas deteksi dan tingkat false alarm [14].

2.5 Autoencoder

Autoencoder yang digunakan dalam penelitian ini memiliki arsitektur encoder-decoder simetris dengan lapisan bottleneck di tengah. Struktur jaringan terdiri dari lapisan input dengan 9 neuron (sesuai jumlah fitur), diikuti lapisan tersembunyi dengan 6 neuron, lapisan bottleneck dengan 3 neuron, lapisan tersembunyi dengan 6 neuron, dan lapisan output dengan 9 neuron. Seluruh lapisan tersembunyi menggunakan fungsi aktivasi ReLU, sedangkan lapisan output menggunakan fungsi aktivasi linear.



Gambar 3: Arsitektur Autoencoder

Arsitektur Autoencoder yang digunakan pada penelitian ini, dengan konfigurasi lapisan 9-6-3-6-9. Encoder memampatkan 9 fitur input menjadi representasi bottleneck berdimensi 3, kemudian decoder merekonstruksinya kembali menjadi 9 fitur output. Selisih antara data input dan hasil rekonstruksi digunakan sebagai reconstruction error. Model dilatih menggunakan 230 sampel data Normal selama 50 epoch dengan batch size 32, optimizer Adam, dan fungsi loss mean squared error (MSE), sebagaimana umum digunakan pada pendekatan reconstruction-based anomaly detection [3]. Ambang batas (threshold) reconstruction error ditentukan berdasarkan persentil ke-90 dari reconstruction error pada data latih (data Normal). Suatu sampel dikategorikan sebagai Anomali apabila reconstruction error-nya melebihi threshold tersebut.

Pemilihan arsitektur bottleneck dengan 3 neuron mengikuti prinsip pengurangan dimensi secara bertahap (progressive dimensionality reduction), di mana lapisan bottleneck perlu cukup kecil untuk memaksa model mempelajari representasi terkompresi dari data normal, namun tidak terlalu kecil sehingga kehilangan informasi penting yang dibutuhkan untuk rekonstruksi [14]. Dengan sembilan fitur input, bottleneck berdimensi 3 dipilih sebagai kompromi antara tingkat kompresi dan kapasitas representasi, mengikuti praktik umum arsitektur autoencoder simetris yang mengurangi dimensi secara bertahap dari lapisan input menuju bottleneck [14]. Pemilihan jumlah epoch sebanyak 50 didasarkan pada pengamatan kurva loss pelatihan, di mana nilai loss menunjukkan kecenderungan konvergen mendekati epoch tersebut tanpa indikasi peningkatan loss validasi yang mengarah pada overfitting, sejalan dengan praktik umum penentuan jumlah epoch berdasarkan pemantauan konvergensi loss pada pendekatan reconstruction-based anomaly detection [14]. Ukuran batch sebesar 32 dipilih sebagai kompromi antara stabilitas estimasi gradien

selama pelatihan dan efisiensi komputasi, mengikuti nilai batch size yang umum digunakan pada pelatihan autoencoder berskala kecil hingga menengah [14].

2.6 Evaluasi Performa

Performa ketiga metode dievaluasi menggunakan lima metrik, yaitu accuracy, precision, recall, F1-score, dan ROC-AUC, sebagaimana umum digunakan dalam evaluasi model klasifikasi berbasis confusion matrix [7]. Kelima metrik tersebut dihitung berdasarkan empat komponen confusion matrix, yaitu True Positive (TP) atau sampel Anomali yang diprediksi benar sebagai Anomali, True Negative (TN) atau sampel Normal yang diprediksi benar sebagai Normal, False Positive (FP) atau sampel Normal yang diprediksi salah sebagai Anomali, dan False Negative (FN) atau sampel Anomali yang diprediksi salah sebagai Normal.

Accuracy mengukur proporsi prediksi yang benar terhadap seluruh data, sebagaimana dirumuskan pada (2).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

Precision mengukur proporsi prediksi Anomali yang benar-benar Anomali, sebagaimana dirumuskan pada (3).

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

Recall mengukur proporsi data Anomali aktual yang berhasil terdeteksi, sebagaimana dirumuskan pada (4).

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

F1-score merupakan rata-rata harmonik dari precision dan recall, sebagaimana dirumuskan pada (5), sehingga memberikan satu nilai tunggal yang menyeimbangkan kedua metrik tersebut.

$$F1 - Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (5)$$

ROC-AUC mengukur kemampuan model dalam membedakan kelas Normal dan Anomali pada berbagai nilai threshold. Nilai ini dihitung sebagai luas area di bawah kurva Receiver Operating Characteristic (ROC), yaitu kurva yang memetakan True Positive Rate (TPR) pada sumbu vertikal terhadap False Positive Rate (FPR) pada sumbu horizontal di setiap kemungkinan nilai threshold, sebagaimana dirumuskan pada (6), (7), dan (8).

$$TPR = \frac{TP}{TP + FN} \quad (6)$$

$$FPR = \frac{FP}{FP + TN} \quad (7)$$

$$ROC - AUC = \int_0^1 TPR d(FPR) \quad (8)$$

Nilai ROC-AUC berkisar antara 0 hingga 1, dengan nilai 1,0 menunjukkan pemisahan sempurna antara kelas Normal dan Anomali pada seluruh nilai threshold, dan nilai 0,5 menunjukkan performa yang setara dengan tebakan acak.

3. HASIL DAN PEMBAHASAN

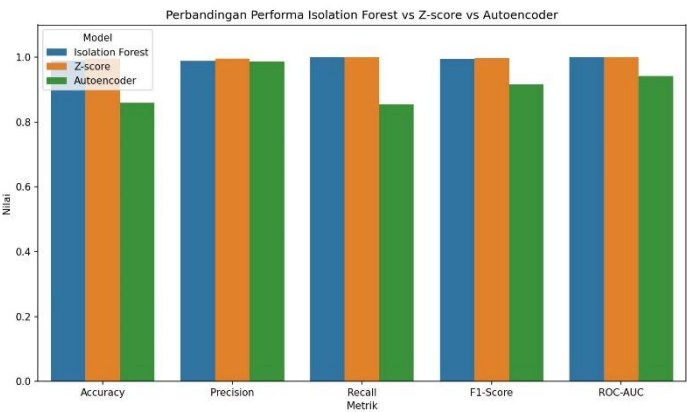
3.1 Hasil Perbandingan Performa



Tabel II menunjukkan hasil evaluasi ketiga metode deteksi anomali berdasarkan lima metrik yang telah ditentukan.

TABEL II: PERBANDINGAN PERFORMA ISOLATION FOREST, Z-SCORE, DAN AUTOENCODER

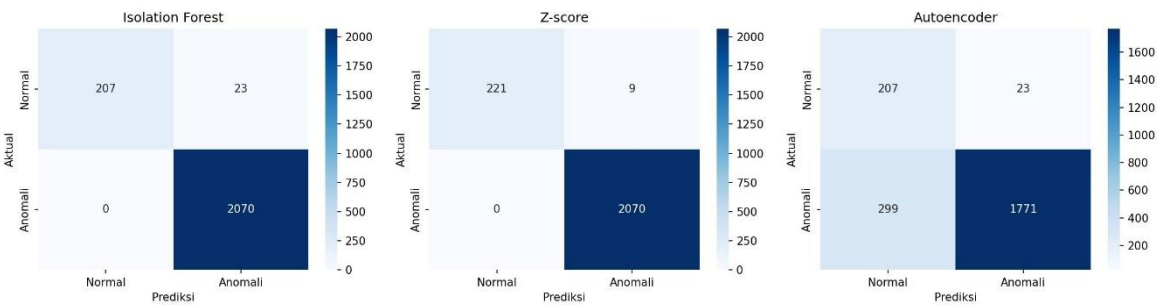
Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Isolation Forest	0,9900	0,9890	1,0000	0,9945	1,0000
Z-score	0,9961	0,9957	1,0000	0,9978	1,0000
Autoencoder	0,8600	0,9872	0,8556	0,9167	0,9421



Gambar 4: Diagram Batang Perbandingan Nilai Accuracy, Precision, Recall, F1-Score, Dan ROC-AUC Untuk Ketiga Metode.

Berdasarkan Tabel II, Z-score memberikan performa terbaik secara keseluruhan, dengan accuracy 0,9961 dan F1-score 0,9978. Isolation Forest berada pada posisi kedua dengan accuracy 0,9900 dan F1-score 0,9945. Autoencoder menunjukkan performa terendah, dengan accuracy 0,8600 dan F1-score 0,9167. Pada metrik recall, Isolation Forest dan Z-score sama-sama mencapai nilai 1,0000, yang berarti seluruh 2070 sampel Anomali berhasil terdeteksi oleh kedua metode tersebut. Sebaliknya, Autoencoder hanya mencapai recall 0,8556.

3.2 Confusion Matrix



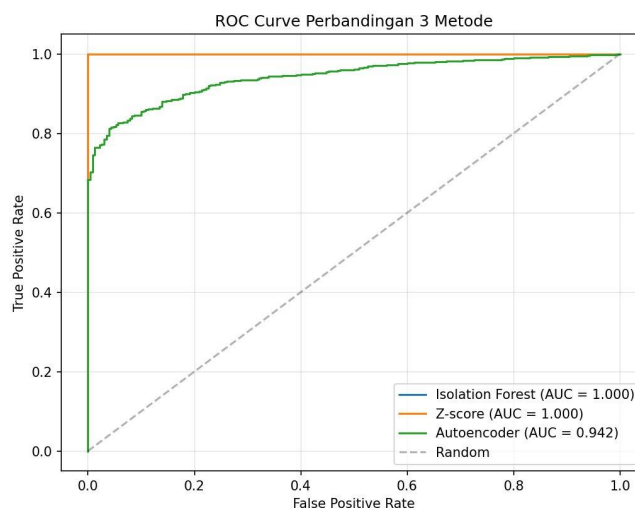
Gambar 5: Confusion Matrix untuk Isolation Forest, Z-Score, Dan Autoencoder.

Confusion matrix pada Gambar 5 memperlihatkan distribusi prediksi setiap model secara lebih rinci. Isolation Forest mengklasifikasikan 207 dari 230 sampel Normal dengan benar dan 23 sampel Normal terklasifikasi sebagai Anomali (false positive), sementara seluruh 2070 sampel Anomali terklasifikasi dengan benar. Z-score menunjukkan hasil yang lebih baik pada sampel Normal, dengan 221 dari 230 sampel terklasifikasi benar dan hanya 9 sampel yang menjadi false positive, sementara seluruh sampel Anomali juga terklasifikasi dengan benar.



Pada Autoencoder, 207 dari 230 sampel Normal terklasifikasi benar dengan 23 false positive, sebanding dengan Isolation Forest. Namun pada sisi Anomali, sebanyak 299 dari 2070 sampel Anomali terklasifikasi sebagai Normal (false negative), sedangkan 1771 sampel terklasifikasi dengan benar. Jumlah false negative inilah yang menjadi penyebab utama rendahnya nilai recall pada Autoencoder.

3.3 ROC Curve



Gambar 6: ROC Curve untuk Isolation Forest, Z-Score, dan Autoencoder.

Gambar 6 menunjukkan kurva ROC ketiga metode. Isolation Forest dan Z-score menghasilkan AUC sebesar 1,000, sedangkan Autoencoder menghasilkan AUC sebesar 0,942. Nilai AUC mendekati atau mencapai 1,000 pada Isolation Forest dan Z-score sejalan dengan pola yang ditunjukkan pada Gambar 1, di mana fitur-fitur seperti sd, rms, kurtosis, crest, dan form memiliki rentang nilai yang sangat berbeda antara kondisi Normal dan Anomali. Separabilitas fitur yang tinggi ini menjadikan kedua metode mampu memisahkan kedua kelas hampir tanpa kesalahan pada berbagai nilai threshold. Karakteristik separabilitas tinggi pada fitur statistik domain waktu CWRU ini konsisten dengan temuan pada penelitian sebelumnya yang menggunakan dataset serupa [6].

3.4 Pembahasan

Temuan utama pada penelitian ini adalah bahwa Z-score, sebagai metode statistik yang paling sederhana di antara ketiga metode yang dibandingkan, memberikan performa tertinggi pada seluruh metrik yang diuji. Hasil ini menunjukkan bahwa pada kasus fitur statistik domain waktu dengan separabilitas tinggi seperti pada dataset CWRU, peningkatan kompleksitas algoritma tidak selalu diikuti dengan peningkatan akurasi deteksi. Hal ini sejalan dengan kajian Das dkk. yang menyatakan bahwa pemilihan metode machine learning untuk diagnosis kerusakan mesin rotasi perlu disesuaikan dengan karakteristik data yang tersedia, bukan semata-mata berdasarkan tingkat kompleksitas model [7]. Temuan serupa juga dilaporkan pada domain anomaly detection lain di luar bearing motor, di mana Smolen dan Benova menunjukkan bahwa Autoencoder dan Isolation Forest dapat menghasilkan keputusan deteksi yang berbeda secara signifikan terhadap sampel yang sama meskipun dilatih pada data yang sama, sehingga keunggulan salah satu metode tidak dapat digeneralisasi tanpa mempertimbangkan karakteristik data yang dianalisis [12].

Performa Autoencoder yang lebih rendah, khususnya pada metrik recall, dapat dijelaskan dari dua sisi. Pertama, arsitektur bottleneck dengan 3 neuron yang digunakan pada penelitian ini relatif sederhana, sehingga kemampuan model dalam menangkap variasi pola dari sembilan jenis kondisi anomali yang berbeda (tiga lokasi cacat dikombinasikan dengan tiga tingkat keparahan) menjadi terbatas. Kondisi anomali dengan tingkat keparahan rendah, misalnya cacat dengan diameter 0,007 inci, kemungkinan menghasilkan



reconstruction error yang masih berada di bawah threshold persentil ke-90, sehingga terklasifikasi sebagai Normal. Kedua, penentuan threshold berdasarkan persentil ke-90 dari reconstruction error data latih bersifat konservatif, sehingga sejumlah sampel anomali dengan reconstruction error yang mendekati distribusi data normal tidak terdeteksi. Pengamatan ini sejalan dengan studi Umari dan Supardi yang juga menekankan pentingnya penentuan ambang reconstruction error yang tepat agar sensitivitas deteksi anomali pada Autoencoder tetap terjaga [3]. Untuk menguji dugaan tersebut secara empiris, dilakukan eksperimen tambahan yang secara sistematis memvariasikan arsitektur bottleneck, jumlah epoch, dan persentil threshold Autoencoder, sebagaimana diuraikan pada Subbagian 3.5.

Nilai AUC sebesar 1,000 pada Isolation Forest dan Z-score perlu diinterpretasikan dengan mempertimbangkan karakteristik dataset yang digunakan. Nilai tersebut bukan menunjukkan bahwa kedua metode tidak memiliki kelemahan dalam konteks umum, melainkan mencerminkan bahwa fitur statistik domain waktu pada dataset CWRU, yang merupakan dataset dengan kondisi cacat yang dibuat secara terkontrol melalui electro-discharge machining [6], memiliki perbedaan karakteristik yang sangat jelas antara kondisi Normal dan kondisi Anomali. Pada kondisi lapangan, di mana kerusakan bearing berkembang secara alami dan bertahap, separabilitas fitur kemungkinan tidak akan sejelas pada dataset terkontrol seperti CWRU, sehingga performa ketiga metode pada kondisi tersebut perlu divalidasi lebih lanjut menggunakan dataset yang merepresentasikan degradasi alami, seperti dataset IMS yang juga telah digunakan pada penelitian sebelumnya [6]. Pendekatan berbasis statistical process control yang diuji pada berbagai kondisi pembebanan motor oleh Jorry dkk. menunjukkan bahwa metode statistik tetap dapat mendeteksi seluruh jenis kerusakan dengan tingkat false alarm yang sangat rendah meskipun kondisi pembebanan bervariasi, yang memperkuat indikasi bahwa metode statistik sederhana cukup robust pada dataset CWRU secara umum [10].

Ketidakseimbangan distribusi kelas, dengan proporsi 230 sampel Normal berbanding 2070 sampel Anomali, juga perlu menjadi perhatian dalam interpretasi hasil accuracy. Meskipun demikian, karena recall pada Isolation Forest dan Z-score mencapai nilai maksimum (1,0000) dan precision keduanya juga berada di atas 0,98, ketidakseimbangan kelas pada penelitian ini tidak menyebabkan bias evaluasi yang signifikan terhadap kedua metode tersebut.

Dari sisi penerapan praktis, hasil penelitian ini mengindikasikan bahwa untuk sistem pemantauan kondisi bearing motor berbasis fitur statistik domain waktu, metode Z-score dapat menjadi pilihan awal yang efisien karena kesederhanaan implementasi, kebutuhan komputasi yang rendah, dan performa yang setara atau lebih baik dibandingkan metode yang lebih kompleks pada penelitian ini. Metode Isolation Forest dapat menjadi alternatif ketika diperlukan pendekatan yang lebih robust terhadap fitur dengan distribusi non-normal, karena Z-score secara teoritis mengasumsikan distribusi data mendekati normal [5]. Autoencoder, meskipun pada penelitian ini menunjukkan performa terendah, tetap memiliki potensi pada skenario dengan jumlah fitur yang lebih besar atau data time-series mentah, di mana kemampuan Autoencoder dalam mempelajari representasi nonlinear dari data berdimensi tinggi dapat memberikan keunggulan dibandingkan metode statistik [3], [6].

3.5 Eksperimen Tambahan: Sensitivitas Konfigurasi Autoencoder

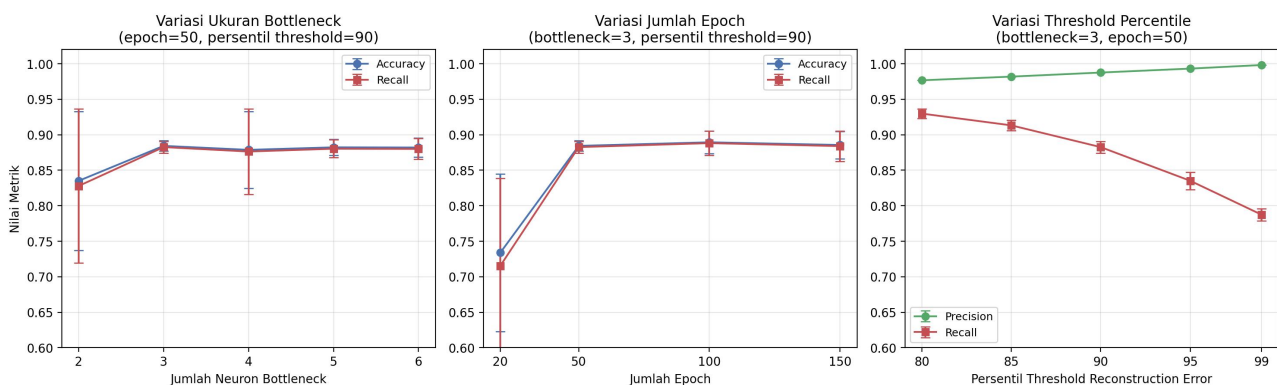
Untuk memperkuat analisis pada Subbagian 3.4 mengenai penyebab performa Autoencoder yang lebih rendah, dilakukan eksperimen tambahan yang secara sistematis memvariasikan tiga aspek konfigurasi, yaitu jumlah neuron pada lapisan bottleneck, jumlah epoch pelatihan, dan persentil threshold reconstruction error. Setiap konfigurasi dilatih ulang menggunakan 10 inisialisasi bobot (random seed) yang berbeda, dengan seluruh pengaturan lain mengikuti konfigurasi dasar pada Subbagian 2.5 (bottleneck 3 neuron, 50 epoch, batch size 32, persentil threshold ke-90), kecuali pada aspek yang sedang divariasikan. Performa dilaporkan sebagai rata-rata dan standar deviasi accuracy serta recall di antara 10 pengulangan tersebut, guna memperhitungkan variasi hasil akibat inisialisasi bobot Autoencoder yang bersifat acak. Perlu dicatat bahwa rata-rata accuracy konfigurasi dasar pada eksperimen ini ($0,8843 \pm 0,0074$) sedikit berbeda dari nilai



tunggal yang dilaporkan pada Tabel II (0,8600), sebagai konsekuensi dari sifat stokastik pelatihan jaringan saraf tiruan; kedua nilai tersebut tetap berada pada rentang performa yang konsisten dan tidak mengubah kesimpulan umum penelitian.

TABEL III: HASIL EKSPERIMEN SENSITIVITAS KONFIGURASI AUTOENCODER (RATA-RATA ± STANDAR DEVIASI, 10 PENGULANGAN)

Variabel yang Divariasikan	Nilai	Accuracy (mean ± std)	Recall (mean ± std)
Jumlah neuron bottleneck (epoch=50, persentil=90)	2	0,8351 ± 0,0976	0,8279 ± 0,1085
	3 (baseline)	0,8843 ± 0,0074	0,8826 ± 0,0082
	4	0,8787 ± 0,0542	0,8763 ± 0,0602
	5	0,8823 ± 0,0113	0,8803 ± 0,0126
	6	0,8820 ± 0,0133	0,8800 ± 0,0148
	20	0,7338 ± 0,1110	0,7154 ± 0,1234
Jumlah epoch (bottleneck=3, persentil=90)	50 (baseline)	0,8843 ± 0,0074	0,8826 ± 0,0082
	100	0,8893 ± 0,0156	0,8882 ± 0,0173
	150	0,8856 ± 0,0193	0,8840 ± 0,0215
Persentil threshold reconstruction error (bottleneck=3, epoch=50)	80	0,9168 ± 0,0058	0,9298 ± 0,0065
	85	0,9067 ± 0,0064	0,9132 ± 0,0071
	90 (baseline)	0,8843 ± 0,0074	0,8826 ± 0,0082
	95	0,8465 ± 0,0110	0,8352 ± 0,0122
	99	0,8073 ± 0,0076	0,7873 ± 0,0085



Gambar 7: Sensitivitas accuracy dan recall Autoencoder terhadap variasi jumlah neuron bottleneck (kiri), jumlah epoch (tengah), dan persentil threshold reconstruction error (kanan). Garis vertikal menunjukkan standar deviasi dari 10 pengulangan dengan inisialisasi bobot berbeda.

Tabel III dan Gambar 7 menunjukkan bahwa jumlah neuron bottleneck sebanyak 2 menghasilkan accuracy rata-rata terendah (0,8351) sekaligus standar deviasi tertinggi (0,0976) di antara seluruh konfigurasi bottleneck yang diuji, yang mengindikasikan bahwa kapasitas representasi yang terlalu kecil tidak hanya menurunkan performa rata-rata, tetapi juga membuat hasil pelatihan menjadi kurang stabil antarpengulangan. Bottleneck sebanyak 3 neuron, sebagaimana digunakan pada konfigurasi awal penelitian ini, menghasilkan accuracy dan recall rata-rata yang kompetitif dengan standar deviasi paling rendah di antara seluruh konfigurasi bottleneck yang diuji, menunjukkan bahwa konfigurasi ini memberikan



keseimbangan yang baik antara kapasitas kompresi dan stabilitas pelatihan. Penambahan neuron bottleneck menjadi 4 hingga 6 tidak memberikan peningkatan performa rata-rata yang konsisten, bahkan bottleneck 4 neuron menunjukkan variabilitas yang lebih tinggi (standar deviasi 0,0542), sehingga penambahan kapasitas bottleneck di luar titik yang digunakan pada penelitian ini tidak terbukti memberikan keuntungan yang berarti pada dataset ini.

Pada variasi jumlah epoch, pelatihan selama 20 epoch menghasilkan accuracy dan recall rata-rata yang jauh lebih rendah (masing-masing 0,7338 dan 0,7154) disertai standar deviasi yang besar, yang mengindikasikan model belum mencapai konvergensi yang memadai pada jumlah epoch tersebut. Peningkatan jumlah epoch menjadi 50, 100, dan 150 menghasilkan performa rata-rata yang relatif serupa (accuracy berkisar 0,884–0,889), dengan epoch sebanyak 50 menunjukkan standar deviasi paling rendah di antara ketiganya. Temuan ini mendukung pemilihan 50 epoch pada konfigurasi awal penelitian sebagai jumlah yang memadai untuk mencapai konvergensi, sekaligus menunjukkan bahwa penambahan epoch lebih lanjut tidak memberikan peningkatan performa yang signifikan pada dataset ini.

Variasi persentil threshold reconstruction error menunjukkan pola trade-off yang jelas antara precision dan recall. Menurunkan persentil threshold dari 90 menjadi 80 meningkatkan recall rata-rata dari 0,8826 menjadi 0,9298 dan accuracy dari 0,8843 menjadi 0,9168, meskipun disertai sedikit penurunan precision (dari 0,9876 menjadi 0,9767). Sebaliknya, menaikkan persentil threshold menjadi 95 dan 99 secara konsisten meningkatkan precision hingga mendekati 1,0 (0,9982 pada persentil ke-99), tetapi menyebabkan penurunan recall yang cukup besar hingga 0,7873. Hasil ini secara empiris mengonfirmasi dugaan pada Subbagian 3.4 bahwa penentuan threshold berdasarkan persentil ke-90 bersifat konservatif dan menjadi salah satu penyebab utama rendahnya recall Autoencoder pada penelitian ini; penggunaan persentil yang lebih rendah, seperti persentil ke-80, berpotensi meningkatkan sensitivitas deteksi terhadap kondisi anomali dengan tingkat keparahan rendah, dengan konsekuensi peningkatan jumlah false positive yang relatif kecil.

Secara keseluruhan, ketiga eksperimen tambahan ini menunjukkan bahwa performa Autoencoder pada penelitian ini lebih dipengaruhi oleh pemilihan persentil threshold reconstruction error dibandingkan oleh jumlah neuron bottleneck atau jumlah epoch pelatihan, selama kedua parameter terakhir berada pada rentang yang wajar (bottleneck lebih besar atau sama dengan 3 neuron dan epoch lebih besar atau sama dengan 50). Temuan ini memperkuat analisis pada Subbagian 3.4 dan menunjukkan bahwa strategi penentuan threshold yang lebih adaptif, alih-alih menggunakan persentil tetap, berpotensi menjadi arah perbaikan yang paling relevan untuk meningkatkan recall Autoencoder pada kasus deteksi anomali bearing motor elektrik dengan karakteristik data serupa.

4. KESIMPULAN

Penelitian ini membandingkan performa tiga metode deteksi anomali, yaitu Isolation Forest, Z-score, dan Autoencoder, dalam mengidentifikasi kondisi anomali pada bearing motor elektrik menggunakan fitur statistik domain waktu dari dataset CWRU. Hasil pengujian pada 2300 sampel menunjukkan Z-score memberikan performa terbaik dengan accuracy 0,9961, F1-score 0,9978, dan ROC-AUC 1,0000, diikuti oleh Isolation Forest dengan accuracy 0,9900, F1-score 0,9945, dan ROC-AUC 1,0000. Autoencoder menunjukkan performa terendah dengan accuracy 0,8600, F1-score 0,9167, dan ROC-AUC 0,9421, dengan recall 0,8556 sebagai metrik terendah di antara ketiga metode.

Temuan ini menunjukkan bahwa metode statistik sederhana seperti Z-score dapat memberikan performa yang setara atau lebih baik dibandingkan metode berbasis deep learning pada kasus fitur statistik domain waktu dengan separabilitas tinggi antara kondisi normal dan anomali. Hasil ini menjadi pertimbangan bahwa pemilihan metode deteksi anomali untuk sistem pemantauan kondisi motor sebaiknya disesuaikan dengan karakteristik fitur yang tersedia, bukan semata-mata berdasarkan kompleksitas algoritma.

Penelitian ini memiliki keterbatasan pada penggunaan dataset dengan kondisi cacat yang dibuat secara terkontrol dan pada satu kondisi pembebanan (1 HP). Eksperimen tambahan yang memvariasikan jumlah neuron bottleneck, jumlah epoch, dan persentil threshold reconstruction error pada Autoencoder



(Subbagian 3.5) menunjukkan bahwa pemilihan persentil threshold memberikan pengaruh yang lebih besar terhadap recall dibandingkan kedua parameter lainnya, sehingga strategi penentuan threshold yang lebih adaptif berpotensi menjadi arah perbaikan yang paling relevan pada penelitian selanjutnya. Penelitian selanjutnya juga dapat diarahkan pada validasi ketiga metode menggunakan dataset dengan degradasi bearing yang berkembang secara alami serta variasi kondisi pembebanan dan kecepatan motor, guna menguji konsistensi temuan ini pada kondisi data yang lebih merepresentasikan kerusakan bearing di lapangan.

DAFTAR PUSTAKA

- [1] E. R. Naufal, G. Priyandoko, and F. Hunaini, "Klasterisasi Kerusakan Bearing Motor Induksi 3 Fasa Menggunakan Metode Transformasi Wavelet Diskrit dan K-Medoids," *ELKHA: Jurnal Teknik Elektro*, vol. 12, no. 2, pp. 54–61, 2020, doi: 10.26418/elkha.v12i2.41511.
- [2] W. S. Alfira WS, "Analisis Predictive Maintenance pada Motor High Pressure Oil Pump (HP Oil Pump) dengan Monitoring Vibrasi di PLTU Barru," *Jurnal Media Elektrik*, vol. 20, no. 2, pp. 14–19, 2023. [Online]. Available: <https://ojs.unm.ac.id/mediaelektrik/article/view/45235>
- [3] Z. Umari and J. Supardi, "Deteksi Anomali Sinyal Vibrasi pada Mesin Industri Menggunakan Autoencoder di PT. Pusri Palembang," *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 4, no. 12, pp. 737–746, 2024, doi: 10.52436/1.jpti.553.
- [4] R. Amanda, E. Helmud, and C. Kirana, "Deteksi Anomali Polusi Udara Menggunakan Algoritma Isolation Forest tanpa Label pada Dataset Kualitas Udara Torino," *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 5, no. 9, pp. 2719–2729, 2025, doi: 10.52436/1.jpti.950.
- [5] P. R. Sihombing, S. Suryadiningrat, D. A. Sunarjo, and Y. P. A. C. Yuda, "Identifikasi Data Outlier (Pencilan) dan Kenormalan Data Pada Data Univariat serta Alternatif Penyelesaiannya," *Jurnal Ekonomi dan Statistik Indonesia*, vol. 2, no. 3, pp. 307–316, 2022. [Online]. Available: <https://jurnaljesi.com/index.php/jurnaljesi/article/view/112>
- [6] S. Zhang, F. Ye, B. Wang, and T. G. Habetler, "Semi-Supervised Learning of Bearing Anomaly Detection via Deep Variational Autoencoders," *arXiv preprint arXiv:1912.01096*, 2019.
- [7] O. Das, D. Bagci Das, and D. Birant, "Machine Learning for Fault Analysis in Rotating Machinery: A Comprehensive Review," *Heliyon*, vol. 9, no. 6, p. e17584, 2023, doi: 10.1016/j.heliyon.2023.e17584.
- [8] A. Ambarwari, Q. J. Adrian, and Y. Herdiyeni, "Analisis Pengaruh Data Scaling Terhadap Performa Algoritma Machine Learning untuk Identifikasi Tanaman," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 4, no. 1, pp. 117–122, 2020. [Online]. Available: <https://jurnal.iaii.or.id/index.php/RESTI/article/view/1517>
- [9] D. K. Riyanto, P. Asri, and A. T. Nugraha, "Monitoring Akselerasi Getaran dan Suhu Motor Induksi," *Elektriese: Jurnal Sains dan Teknologi Elektro*, vol. 11, no. 1, pp. 33–39, 2022. [Online]. Available: <https://jurnal.itscience.org/index.php/elektriese/article/view/1626>
- [10] V. Jorry, Z.-S. Duma, T. Sihvonen, S.-P. Reinikainen, and L. Roininen, "Statistical batch-based bearing fault detection," *Journal of Mathematics in Industry*, vol. 15, art. 4, 2025, doi: 10.1186/s13362-025-00169-w.
- [11] N. A. Mohammed, O. F. Abdulateef, and A. H. Hamad, "An IoT and Machine Learning-Based Predictive Maintenance System for Electrical Motors," *Journal Européen des Systèmes Automatisés*, vol. 56, no. 4, pp. 651–656, 2023, doi: 10.18280/jesa.560414.
- [12] T. Smolen and L. Benova, "Comparing Autoencoder and Isolation Forest in Network Anomaly Detection," in 2023 33rd Conference of Open Innovations Association (FRUCT), 2023.
- [13] brjapon, "CWRU Bearing Datasets," Kaggle. [Online]. Available: <https://www.kaggle.com/datasets/brjapon/cwru-bearing-datasets>
- [14] A. Nicolae and A. Korodi, "Global Anomaly Detection Using Feedforward Symmetrical Autoencoder Neuronal Network: Comparison with Other Methods in a Case Study Using Real Industrial Data," *Applied Sciences*, vol. 16, no. 5, p. 2457, 2026, doi: 10.3390/app16052457

