

Industrial Equipment Monitoring Dataset for Predictive Maintenance Analysis

Kreshna Lucky Pradana, Adlian Jefriza
Politeknik Negeri Batam, Batam, Indonesia

INFORMASI ARTIKEL	ABSTRAK
<p>Sejarah Artikel: Diterima: Juni 2025 Revisi: Juni 2025 Diterima: Juli 2025 Dipublikasi: Juli 2025</p> <p>Kata Kunci: Support Vector Machine (SVM), Klasifikasi, Deteksi Faulty, Data Sensor, Ketidakseimbangan Kelas</p> <p>*Penulis Korespondensi: Kreshnakreshna123@gmail.com</p>	<p>Penelitian ini mengembangkan dan mengevaluasi model <i>Support Vector Machine</i> (SVM) berbasis kernel <i>Radial Basis Function</i> (RBF) untuk mendeteksi kondisi faulty pada sistem menggunakan data sensor (<i>temperature, pressure, vibration, humidity</i>). Data diproses melalui normalisasi dan dibagi untuk pelatihan serta pengujian. Hasil evaluasi menunjukkan akurasi model keseluruhan 0.93. Model sangat efektif dalam mengidentifikasi kondisi normal (presisi 0.93, recall 1.00), namun kurang optimal dalam mendeteksi kondisi faulty (presisi 0.96, recall 0.30), yang mengindikasikan banyak false negatives dan F1-score rendah (0.45) untuk kelas ini. ROC AUC sebesar 0.892 menunjukkan kemampuan diskriminatif yang baik secara umum. Kesenjangan kinerja ini kemungkinan besar disebabkan oleh ketidakseimbangan kelas. Peningkatan deteksi faulty melalui penanganan ketidakseimbangan data atau optimasi model lebih lanjut direkomendasikan untuk aplikasi kritis.</p> <p>ABSTRACT <i>This study develops and evaluates a Support Vector Machine (SVM) model using a Radial Basis Function (RBF) kernel to detect faulty conditions in systems based on sensor data (temperature, pressure, vibration, humidity). The data is processed through normalization and split into training and testing sets. The evaluation results show an overall model accuracy of 0.93. The model is highly effective in identifying normal conditions (precision 0.93, recall 1.00), but less optimal in detecting faulty conditions (precision 0.96, recall 0.30), indicating a high number of false negatives and a low F1-score (0.45) for this class. The ROC AUC score of 0.892 indicates good overall discriminative ability. This performance gap is likely due to class imbalance. Enhancing faulty detection through class imbalance handling or further model optimization is recommended for critical applications.</i></p>

PENDAHULUAN

Di era Revolusi Industri 4.0, penggunaan teknologi digital dan *Internet of Things* (IoT) dalam bidang industri telah membawa perubahan besar, khususnya dalam cara perusahaan mengelola dan merawat peralatan mereka. Salah satu metode perawatan yang kini banyak digunakan adalah pemeliharaan prediktif. Metode ini bekerja dengan memanfaatkan data historis dan data sensor yang dikumpulkan secara langsung dari mesin, lalu dianalisis menggunakan algoritma kecerdasan buatan. Tujuannya adalah untuk memprediksi kapan mesin akan rusak, sehingga perbaikan bisa dilakukan sebelum kerusakan benar-benar terjadi[1].

pendekatan lainya, yaitu:
Pemeliharaan reaktif, yang hanya dilakukan setelah mesin rusak dan Pemeliharaan preventif,

yang dilakukan berdasarkan jadwal tertentu. Dengan prediksi yang akurat, perusahaan bisa mengurangi waktu henti mesin yang tidak direncanakan dan juga menghemat biaya perawatan. Hal ini sangat penting, terutama pada sektor industri seperti otomotif, manufaktur, energi, dan makanan yang sangat bergantung pada kinerja mesin yang stabil[2]. Penelitian ini menggunakan data dari Industrial Equipment Monitoring Dataset yang tersedia di platform Kaggle. Dataset tersebut mencakup berbagai data sensor seperti suhu, tekanan, dan getaran dari peralatan industri. Dengan data ini, peneliti mencoba menerapkan algoritma Support Vector Machine (SVM) untuk mengklasifikasi kondisi mesin ke dalam kategori seperti *normal*, *waspada*, atau *kritis*.

Pendekatan ini diharapkan mampu meningkatkan keandalan sistem monitoring pada industri serta membantu pengambilan keputusan dalam pemeliharaan berbasis data. Selain itu, penerapan metode ini juga bisa menjadi media pembelajaran yang baik bagi siswa dan mahasiswa di bidang teknik dan mekatronika, karena menggabungkan teknologi modern dengan analisis data secara nyata[3].

METODE PENELITIAN

Penelitian ini bertujuan untuk mengembangkan dan mengevaluasi model klasifikasi *Support Vector Machine* (SVM) dengan kernel RBF untuk mendeteksi kondisi '*faulty*' pada peralatan berdasarkan data sensor. Proses dimulai dengan Pengumpulan Data dari file CSV yang memuat empat fitur operasional (suhu, tekanan, getaran, kelembaban) dan variabel target *faulty*. Selanjutnya, tahap Pra-pemrosesan Data dilakukan untuk membersihkan data mentah dengan menghilangkan karakter non-numerik, mengkonversi semua kolom menjadi tipe data numerik, dan menghapus baris-baris dengan nilai yang hilang (NaN). Setelah data bersih, dilakukan Pembagian Data menjadi set pelatihan (80%) dan set pengujian (20%). Langkah krusial berikutnya adalah Normalisasi Data menggunakan standardisasi (rata-rata nol, deviasi satu) pada set fitur, dengan parameter standardisasi dihitung hanya dari set pelatihan untuk mencegah data leakage. Data yang sudah dinormalisasi kemudian digunakan pada tahap Pelatihan Model SVM, yang dikonfigurasi dengan parameter regularisasi $C=1$ dan kemampuan estimasi probabilitas. Terakhir, Evaluasi Model dilakukan pada set pengujian menggunakan berbagai metrik seperti *Confusion Matrix*, *Classification Report* (presisi, recall, f1-score, akurasi), dan ROC AUC Score/Curve. Hasil evaluasi menunjukkan model mencapai akurasi keseluruhan 0.93 dan ROC AUC Score 0.892, mengindikasikan kemampuan prediksi yang baik

HASIL DAN PEMBAHASAN

Berdasarkan analisis data yang dilakukan, model *Support Vector Machine* (SVM) dengan kernel *Radial Basis Function* (RBF) dan parameter $C=1$ telah dilatih dan dievaluasi. Setelah proses penskalaan data menggunakan *StandardScaler* dan pembagian data menjadi set pelatihan dan pengujian (80% pelatihan, 20% pengujian), model menunjukkan kinerja sebagai berikut:

1. Matriks Kebingungan (*Confusion Matrix*):
 - Model berhasil mengklasifikasikan 1377 sampel sebagai kelas '0' (non-faulty) dengan benar dari total 1377 sampel aktual kelas '0'. Ini menunjukkan tingkat recall 100% untuk kelas '0'.
 - Untuk kelas '10' (faulty), model mengklasifikasikan 45% dari 158 sampel aktual dengan benar.
2. Laporan Klasifikasi (*Classification Report*):
 - a) Presisi:
 - Kelas '0': 0.93

- Kelas '10': 0.96
- b) Recall:
- Kelas '0': 1.00
 - Kelas '10': 0.30
- c) F1-Score:
- Kelas '0': 0.96
 - Kelas '10': 0.45
- d) Akurasi Keseluruhan: 0.93
- e) Macro Average: Presisi 0.94, Recall 0.65, F1-Score 0.71
- f) Weighted Average: Presisi 0.93, Recall 0.93, F1-Score 0.91
3. ROC AUC Score: 0.8922579814860777
Kurva ROC menunjukkan kemampuan diskriminasi model yang baik
4. MSE (Mean Squared Error)
MSE untuk prediksi kelas: 0.072
RMSE: 0.269

Hasil yang diperoleh menunjukkan bahwa model SVM memiliki kemampuan yang sangat baik dalam mengidentifikasi sampel *non-faulty* (kelas '0'), dengan recall 100% dan presisi 93%. Ini berarti hampir semua sampel *non-faulty* teridentifikasi dengan benar, dan sebagian besar prediksi *non-faulty* oleh model memang akurat. Namun, kinerja model untuk mendeteksi sampel *faulty* (kelas '10') masih perlu ditingkatkan. Meskipun presisi untuk kelas '10' cukup tinggi (0.96), yang berarti ketika model memprediksi *faulty*, prediksinya cenderung benar, nilai recall yang rendah (0.30) menunjukkan bahwa model hanya mampu mendeteksi 30% dari total sampel *faulty* yang sebenarnya. Ini mengindikasikan adanya sejumlah besar false negatives, di mana sampel *faulty* tidak terdeteksi oleh model. F1-score untuk kelas '10' yang hanya 0.45 juga mengkonfirmasi bahwa ada ketidakseimbangan antara presisi dan recall untuk kelas ini, dan secara keseluruhan kinerja deteksi *faulty* masih kurang optimal.

Akurasi keseluruhan model sebesar 0.93 terlihat tinggi, tetapi ini mungkin menyesatkan karena adanya ketidakseimbangan kelas yang signifikan (jumlah sampel *non-faulty* jauh lebih banyak daripada sampel *faulty*). Model cenderung bias terhadap kelas mayoritas ('0'). Nilai ROC AUC sebesar 0.892 menunjukkan bahwa model memiliki kemampuan yang baik dalam membedakan antara kelas positif (*faulty*) dan negatif (*non-faulty*) secara umum. Kurva ROC yang ditampilkan juga akan memberikan visualisasi lebih

Daftar Pustaka

- [1] Gupta, R., Jain, K., & Jain, R. (2020). Predictive Maintenance using Machine Learning: A Case Study of Industrial Equipment. *Procedia Computer Science*, 167, 2621–2630.
- [2] Javed, K., Gouriveau, R., & Zerhouni, N. (2021). A comprehensive review on data-driven predictive maintenance approaches using deep learning. *IEEE Transactions on Industrial Informatics*, 17(3), 2206-2225.

- [3] Rahman, A., Ismail, M., & Zaini, N. (2022). Integrating Predictive Maintenance in Technical Education Using Smart Datasets. *International Journal of Evaluation and Research in Education (IJERE)*, 11(2), 507-514
- [4] Assagaf, A. Sukandi, A. A. Abdillah, S. Arifin, and J. L. Ga, "Machine Predictive Maintenance by Using Support Vector Machines", *RiESTech*, vol. 1, no. 01, pp. 31–35, Jan. 2023