Terbit *online* pada laman web jurnal : www.journal.unbara.ac.id/index.php/INTECH
INFORMATIKA DAN TEKNOLOGI (INTECH)
ISSN (Online) : 2722-7367



Development of Predictive Maintenance Copilot Based on Physics-Informed Machine Learning for Prediction of Machine Failure Risk and Estimation of Remaining Useful Life

Pengembangan Predictive Maintenance Copilot Berbasis Physics-Informed Machine Learning untuk Prediksi Risiko Kegagalan Mesin dan Estimasi Remaining Useful Life

Rasyid Naufal Saputra[1]*, Heni Ramawati[2], Purwanto[3]
[1*, 2, 3] STIMIK Tunas Bangsa, Banjarnegara, Indonesia
*E-mail : rasyidnaufalsaputra03@gmail.com

ABSTRACT

The oil and gas industry requires maintenance strategies capable of identifying failure risks and supporting maintenance prioritization. This study developed a Predictive Maintenance Copilot as a prototype Decision Support System integrating failure-risk prediction, Physics-Informed Machine Learning, Remaining Useful Life estimation, and an AI Copilot. A mixed-method approach dominated by quantitative-computational analysis was used. Modeling employed the AI4I 2020 Predictive Maintenance Dataset to compare Logistic Regression, Decision Tree, Random Forest, and XGBoost. The physics-informed approach was implemented through derived features and physics guardrails, while Remaining Useful Life was calculated separately based on tool wear, duty cycle, and wear increment. XGBoost achieved the highest F2-score of 0.7373, an accuracy of 0.9489, and a failure-class recall of 0.9412. Evaluation involving 10 respondents resulted in an acceptance score of 90.8%, while all software testing was successfully completed. The results indicate that the system can support more structured maintenance information, although validation using actual industrial machine data is still required.

Keywords: AI Copilot, Decision Support System, Physics-Informed Machine Learning, Predictive Maintenance, Remaining Useful Life

ABSTRAK

Industri minyak dan gas membutuhkan strategi pemeliharaan yang mampu mendeteksi risiko kegagalan dan membantu menentukan prioritas tindakan. Penelitian ini mengembangkan *Predictive Maintenance Copilot* sebagai prototipe Sistem Pendukung Keputusan yang mengintegrasikan prediksi risiko, *Physics-Informed Machine Learning, estimasi Remaining Useful Life,* dan *AI Copilot.* Penelitian menggunakan pendekatan *mixed method* dengan dominasi kuantitatif-komputasional. Pemodelan dilakukan menggunakan AI4I 2020 *Predictive Maintenance Dataset* dengan membandingkan *Logistic Regression, Decision Tree, Random Forest,* dan *XGBoost.* Pendekatan *physics-informed* diterapkan melalui fitur turunan dan *physics guardrails*, sedangkan *Remaining Useful Life* dihitung secara terpisah berdasarkan *tool wear, duty cycle*, dan *wear increment. XGBoost* menghasilkan F2-score tertinggi sebesar 0,7373, *accuracy* 0,9489, dan *recall* kelas gagal 0,9412. Evaluasi terhadap 10 responden memperoleh tingkat penerimaan 90,8%, sedangkan seluruh pengujian perangkat lunak berhasil dijalankan. Hasil penelitian menunjukkan bahwa sistem berpotensi mendukung penyajian informasi pemeliharaan secara lebih terstruktur, tetapi masih memerlukan validasi menggunakan data mesin industri aktual.

Kata Kunci : *AI Copilot, Physics-Informed Machine Learning, Predictive Maintenance, Remaining Useful Life,* Sistem Pendukung Keputusan

Informasi Artikel
Submit 27 Oktober 2026 | Diterima 2 November 2026 | Publish Online 30 November 2026.
This is an open access article under the CC BY-NC 4.0 license.
Published by Universitas Baturaja.

PENDAHULUAN

Industri minyak dan gas memiliki karakteristik operasi yang kompleks dan tingkat risiko yang tinggi. Keberlangsungan produksi sangat bergantung pada keandalan berbagai aset dan peralatan mekanis yang bekerja dalam kondisi operasional berat. Ketika terjadi gangguan, dampaknya tidak hanya berupa penurunan kontinuitas operasi, tetapi juga dapat menimbulkan kerugian ekonomi dan meningkatkan risiko keselamatan. Laporan Kinerja Direktorat Jenderal Minyak dan Gas Bumi Tahun 2025 mencatat 11 kejadian *unplanned downtime* pada kegiatan usaha hulu migas. Pada tahun yang sama juga tercatat tujuh kecelakaan kerja yang menyebabkan *fatality* pada kegiatan usaha hilir [1]. Kondisi tersebut menunjukkan bahwa pemantauan kondisi aset dan deteksi dini terhadap potensi kegagalan masih menjadi bagian penting dalam proses pemeliharaan.

Salah satu pendekatan yang digunakan untuk mendukung kebutuhan tersebut adalah *predictive maintenance*. Berbeda dengan pemeliharaan reaktif yang dilakukan setelah kerusakan terjadi, pendekatan ini memanfaatkan data kondisi mesin untuk mengenali perubahan operasi dan memperkirakan potensi kegagalan sebelum berkembang menjadi gangguan yang lebih serius. Machine learning banyak digunakan dalam *predictive maintenance* karena mampu mempelajari hubungan antarparameter operasional dari data historis. Meskipun demikian, pendekatan yang sepenuhnya *data-driven* memiliki keterbatasan. Hubungan fisik yang sebenarnya terjadi pada mesin tidak selalu menjadi bagian dari proses pembelajaran, sementara keluaran dari beberapa algoritma juga sulit dipahami oleh pengguna teknis karena karakteristiknya yang bersifat *black-box* [2].

Physics-Informed Machine Learning (PIML) menawarkan pendekatan yang menggabungkan pembelajaran berbasis data dengan pengetahuan domain atau hubungan fisik. Informasi mengenai temperatur, kecepatan rotasi, torsi, daya mekanik, dan tingkat keausan tidak hanya dipandang sebagai variabel statistik, tetapi dapat digunakan untuk membentuk representasi kondisi mesin yang lebih sesuai dengan karakteristik fisiknya [3]. Prediksi risiko saja belum cukup untuk menentukan prioritas pemeliharaan. Dua mesin dapat memiliki tingkat risiko yang serupa, tetapi berada pada kondisi degradasi yang berbeda. Karena itu, informasi *Remaining Useful Life* (RUL) diperlukan untuk memberikan gambaran mengenai perkiraan sisa waktu operasi sebelum mesin mencapai batas kondisi tertentu. Informasi risiko dan RUL dapat digunakan secara berdampingan untuk membantu pengguna menentukan mesin yang perlu diperiksa lebih dahulu [4].

Penelitian sebelumnya telah menerapkan berbagai metode pada *predictive maintenance*. Hafidhoh et al. menggunakan Logistic Regression untuk klasifikasi kegagalan mesin dan memperoleh akurasi 96,87% [5], tetapi penelitian tersebut masih berfokus pada klasifikasi jenis kegagalan. Sugito et al. mengintegrasikan FMEA, Statistical Process Control, dan Support Vector Machine untuk prediksi perawatan pompa minyak dengan akurasi 95,48% [6]. Pendekatan tersebut masih melibatkan proses FMEA secara manual dalam identifikasi mode kegagalan. Nwulu et al. menyoroti persoalan interpretabilitas sebagai salah satu tantangan implementasi machine learning di bidang *predictive maintenance[2]*, sedangkan Wang et al. menunjukkan bahwa penggunaan *ensemble learning* mampu menghasilkan performa prediksi yang tinggi dengan konsekuensi peningkatan kompleksitas model dan implementasi [7].

Temuan-temuan tersebut menunjukkan bahwa pengembangan *predictive maintenance* tidak hanya berkaitan dengan peningkatan performa model. Sistem juga perlu mampu menghubungkan hasil prediksi dengan kondisi fisik mesin, tingkat degradasi, interpretasi hasil, serta tindak lanjut yang dapat digunakan dalam proses pemeliharaan.

Kebutuhan serupa diperoleh dari studi pendahuluan dalam penelitian ini. Pengguna pada konteks Operation & Maintenance membutuhkan informasi mengenai status kondisi mesin, tingkat risiko, estimasi RUL, parameter yang berkontribusi terhadap perubahan kondisi, dan rekomendasi tindakan. Ketika kondisi tertentu memerlukan penanganan, informasi tersebut juga perlu diteruskan ke mekanisme tindak lanjut yang lebih terstruktur. Berdasarkan kebutuhan tersebut, penelitian ini mengembangkan Predictive Maintenance Copilot sebagai prototipe Sistem Pendukung Keputusan berbasis web. Prediksi risiko kegagalan dilakukan menggunakan machine learning yang dilengkapi *physics-informed feature engineering* dan *physics guardrails*. Estimasi RUL dihitung secara terpisah berdasarkan tingkat *tool wear*, batas keausan kritis, *duty cycle*, dan *wear increment*. Hasil analitik kemudian digunakan sebagai konteks AI Copilot untuk membantu pengguna memahami kondisi mesin dan rekomendasi pemeliharaan melalui bahasa natural. Sistem juga menyediakan simulasi *maintenance ticket* sebagai mekanisme tindak lanjut.

Penelitian ini bertujuan mengembangkan dan mengevaluasi Predictive Maintenance Copilot dalam mendukung identifikasi risiko kegagalan, estimasi sisa umur pakai, dan penentuan prioritas pemeliharaan. Evaluasi model dilakukan menggunakan AI4I 2020 Predictive Maintenance Dataset dengan membandingkan Logistic Regression, Decision Tree, Random Forest, dan XGBoost. Selain evaluasi komputasional, penelitian juga menilai penerimaan pengguna dan menguji fungsi perangkat lunak yang dikembangkan.

TINJAUAN PUSTAKA

1. Predictive Maintenance dan Remaining Useful Life

Predictive maintenance memanfaatkan data kondisi mesin untuk memperkirakan kemungkinan terjadinya gangguan sebelum kegagalan terjadi. Data seperti temperatur, kecepatan rotasi, torsi, dan tingkat keausan dapat dianalisis menggunakan machine learning untuk mengenali pola yang berkaitan dengan perubahan kondisi mesin [2]. Informasi yang dihasilkan tidak hanya berupa status risiko. Dalam proses pemeliharaan juga diperlukan informasi mengenai tingkat degradasi aset. *Remaining Useful Life* (RUL) digunakan untuk memperkirakan sisa waktu operasi sebelum mesin mencapai batas kondisi tertentu [4]. Pada penelitian ini, prediksi risiko dan RUL digunakan sebagai dua informasi yang berbeda tetapi saling melengkapi. Risiko menunjukkan kemungkinan munculnya kondisi bermasalah, sedangkan RUL digunakan untuk mendukung penentuan urgensi pemeriksaan.

2. Physics-Informed Machine Learning

Physics-Informed Machine Learning mengintegrasikan pendekatan *data-driven* dengan pengetahuan mengenai karakteristik fisik sistem. Integrasi tersebut tidak selalu harus dilakukan melalui persamaan diferensial atau arsitektur *Physics-Informed Neural Network*. Pengetahuan domain juga dapat dimasukkan melalui pembentukan fitur atau aturan yang merepresentasikan hubungan fisik antarparameter [3]. Pada penelitian ini, pendekatan tersebut diterapkan melalui *temperature difference*, *mechanical power*, dan *strain load*. Ketiganya dibentuk dari parameter operasional yang tersedia pada dataset. Selain itu, *physics guardrails* digunakan untuk memeriksa kondisi mesin berdasarkan indikator fisik yang berkaitan dengan beberapa mode kegagalan. Dengan mekanisme tersebut, keluaran analitik tidak hanya bergantung pada probabilitas model machine learning. Informasi fisik tetap digunakan dalam proses penentuan status kondisi mesin.

3. Sistem Pendukung Keputusan dan AI Copilot

Sistem Pendukung Keputusan digunakan untuk menyediakan informasi yang membantu pengguna mempertimbangkan suatu tindakan. Sistem tidak menggantikan keputusan pengguna, tetapi menyajikan informasi yang dapat digunakan sebagai dasar dalam menentukan prioritas [8]. Dalam Predictive Maintenance Copilot, informasi tersebut berupa tingkat risiko, estimasi RUL, indikator kondisi, dan rekomendasi pemeliharaan. AI Copilot kemudian digunakan sebagai antarmuka berbasis bahasa natural untuk menyampaikan informasi tersebut dengan bentuk yang lebih mudah dipahami. AI Copilot pada penelitian ini tidak diposisikan sebagai metode *Explainable Artificial Intelligence* formal yang menghasilkan penjelasan lokal terhadap setiap keputusan model [9]. Fungsinya lebih diarahkan sebagai lapisan interpretasi yang menerima konteks dari hasil analitik sistem, seperti status kondisi, probabilitas risiko, indikator fisik, RUL, dan rekomendasi.

Pendekatan berbasis Large Language Model sebelumnya telah digunakan sebagai antarmuka interaktif pada berbagai sistem informasi, termasuk chatbot berbasis *Retrieval-Augmented Generation* [10]. Perbedaannya, konteks AI Copilot pada penelitian ini tidak berasal dari proses pencarian dokumen, tetapi dari keluaran analitik kondisi mesin. Dengan demikian, chatbot ditempatkan sebagai bagian dari alur Sistem Pendukung Keputusan *predictive maintenance*, bukan sebagai sumber prediksi utama.

METODE PENELITIAN

1. Desain Penelitian

Penelitian menggunakan pendekatan *mixed method* dengan dominasi kuantitatif-komputasional. Komponen kualitatif digunakan pada tahap studi pendahuluan melalui wawancara dengan praktisi Strategy & Consulting dan Operation & Maintenance (O&M) untuk memperoleh kebutuhan pengguna serta *prior knowledge* yang berkaitan dengan kondisi mesin dan proses pemeliharaan.

Tahap kuantitatif mencakup pengolahan dataset, pembentukan fitur berbasis fisik, pelatihan dan evaluasi model machine learning, estimasi *Remaining Useful Life* (RUL), serta evaluasi sistem. Pengembangan Predictive Maintenance Copilot dilakukan secara iteratif menggunakan pendekatan Agile yang mencakup persiapan data dan model, pengembangan layanan sistem, integrasi AI Copilot, serta pengujian [11].

2. Dataset dan Preprocessing

Data komputasional menggunakan AI4I 2020 Predictive Maintenance Dataset dari UCI Machine Learning Repository. Dataset bersifat sintetis dan terdiri atas 10.000 observasi kondisi mesin [12]. Penggunaannya ditujukan sebagai data simulasi dalam pengembangan prototipe, bukan sebagai representasi langsung data operasional industri minyak dan gas. Fitur utama yang digunakan meliputi *Type*, *Air temperature*, *Process temperature*, *Rotational speed*, *Torque*, dan *Tool wear*, sedangkan *Machine failure* digunakan sebagai target klasifikasi. Kolom UDI dan Product ID dikeluarkan karena hanya berfungsi sebagai identitas. Indikator TWF, HDF, PWF, dan OSF tidak digunakan sebagai input model untuk menghindari *data leakage*. Data RNF juga tidak digunakan dalam pelatihan model utama karena merepresentasikan kegagalan acak yang tidak memiliki hubungan fisik yang jelas dengan parameter operasiona.

Distribusi target terdiri atas 9.661 kondisi normal dan 339 kondisi gagal sehingga dataset memiliki ketidakseimbangan kelas. Penanganannya dilakukan melalui pembobotan

kelas menggunakan class_weight atau scale_pos_weight sesuai algoritma yang digunakan. Atribut *Type* dikonversi menjadi representasi numerik sebelum proses pemodelan.

3. Physics-Informed Feature Engineering

Pendekatan *Physics-Informed Machine Learning* diterapkan melalui pembentukan fitur turunan dan *physics guardrails* [3]. Tiga fitur turunan yang digunakan adalah *temperature difference*, *mechanical power*, dan *strain load*.
Selisih temperatur dihitung menggunakan:

$$\Delta T = T_{\text{process}} - T_{\text{air}}$$

Daya mekanik dihitung dengan:

$$P = \frac{2\pi \times \text{RPM}}{60} \times \tau$$

sedangkan *strain load* dihitung menggunakan:

$$S = \tau \times W$$

dengan (P)sebagai daya mekanik, (τ) sebagai torsi, dan (W) sebagai *tool wear*.
Selain digunakan sebagai fitur input, pengetahuan fisik diterapkan melalui *physics guardrails* untuk memeriksa indikasi *Power Failure*, *Heat Dissipation Failure*, *Overstrain Failure*, dan *Tool Wear Failure*. Informasi tersebut digunakan bersama keluaran model untuk mendukung penentuan status mesin menjadi Normal, Warning, atau Critical.

4. Pelatihan dan Evaluasi Model

Dataset dibagi secara *stratified* menjadi 60% data latih, 20% data validasi, dan 20% data uji agar proporsi kelas tetap terjaga pada setiap subset. Nilai random_state=42 digunakan untuk menjaga reproduktibilitas eksperimen. Data validasi digunakan dalam proses pemilihan konfigurasi dan penyesuaian *decision threshold*, sedangkan data uji digunakan untuk evaluasi akhir. Empat algoritma dibandingkan, yaitu Logistic Regression, Decision Tree, Random Forest, dan XGBoost. Konfigurasi utama ditunjukkan pada Tabel 1.

Tabel 1. Konfigurasi Utama Model

Model	Konfigurasi Utama
Logistic Regression	StandardScaler, max_iter=1000, class_weight=balanced
Decision Tree	max_depth=4, min_samples_leaf=10, class_weight=balanced
Random Forest	n_estimators=100, max_depth=5, class_weight=balanced_subsample
XGBoost	n_estimators=120, max_depth=3, learning_rate=0.05, scale_pos_weight

Evaluasi dilakukan menggunakan *accuracy*, *balanced accuracy*, *precision* dan *recall* kelas gagal, F2-score, serta jumlah *false negative*. F2-score digunakan sebagai salah satu pertimbangan utama karena memberi bobot lebih besar terhadap *recall*. Dalam konteks *predictive maintenance*, kegagalan mengenali mesin yang sebenarnya berisiko dianggap lebih kritis dibandingkan pemeriksaan tambahan akibat *false positive*.

5. Estimasi Remaining Useful Life

Estimasi RUL dilakukan terpisah dari model klasifikasi risiko. Perhitungan menggunakan *tool wear*, batas keausan kritis, jam operasi harian, *duty cycle*, dan *wear increment*. Batas keausan kritis pada simulasi ditetapkan sebesar 240 dengan asumsi waktu operasi delapan jam per hari. Konstanta operasi dibedakan berdasarkan tipe produk.

Tabel 2. Konstanta Simulasi RUL

Tipe	Duty Cycle	Wear Increment
L	0,10	2
M	0,15	3
H	0,20	5

Estimasi RUL dihitung menggunakan:

$$\text{RUL} = \frac{W_{\text{critical}} - W_{\text{current}}}{H_{\text{daily}} \times W_{\text{increment}} \times D}$$

dengan (W_{critical})sebagai batas keausan kritis, (W_{current})sebagai *tool wear* saat ini, (H_{daily})sebagai jam operasi harian, $(W_{\text{increment}})$sebagai laju pertambahan keausan, dan (D)sebagai *duty cycle*.

Jika mesin telah berada pada kondisi Critical atau nilai *tool wear* mencapai batas kritis, RUL ditetapkan menjadi nol. Nilai yang dihasilkan merupakan estimasi berbasis simulasi degradasi dan belum dikalibrasi menggunakan histori degradasi mesin industri aktual.

6. Implementasi Sistem

Predictive Maintenance Copilot dikembangkan menggunakan arsitektur modular. Antarmuka pengguna dibangun menggunakan React dan Vite, sedangkan layanan utama menggunakan Node.js dan Express.js. Proses analitik dikembangkan melalui Python dan FastAPI untuk menangani prediksi risiko, estimasi RUL, deteksi anomali, dan penyediaan konteks bagi AI Copilot. AI Copilot menerima konteks dari hasil analitik sistem, antara lain status kondisi, tingkat risiko, RUL, indikator fisik, dan rekomendasi pemeliharaan. Konteks tersebut diteruskan ke Large Language Model melalui Groq API untuk menghasilkan penjelasan berbasis bahasa natural. Sistem juga menyediakan simulasi *maintenance ticket* sebagai mekanisme tindak lanjut terhadap kondisi yang membutuhkan pemeriksaan.

7. Evaluasi Sistem

Evaluasi pengguna melibatkan 10 responden yang berkaitan dengan bidang Operation & Maintenance. Instrumen terdiri atas 10 pernyataan yang mengevaluasi kemudahan penggunaan, dashboard, visualisasi data, estimasi RUL, deteksi anomali, AI Copilot, rekomendasi mitigasi, *maintenance ticket*, dan kelayakan sistem sebagai alat bantu pengambilan keputusan. Penilaian menggunakan skala Likert lima tingkat dengan bobot 1–5. Selain itu, fungsi perangkat lunak diuji melalui 13 skenario *black-box testing* dan 87 *white-box test case* [13]. Pengujian internal dilakukan pada Main System API, Machine Learning API, Model Integration API, Landing Page Chatbot API, dan frontend menggunakan Jest, Pytest, serta Vitest.

HASIL DAN PEMBAHASAN

1. Perbandingan Performa Model Machine Learning

Evaluasi dilakukan terhadap Logistic Regression, Decision Tree, Random Forest, dan XGBoost menggunakan data uji yang tidak digunakan pada proses pelatihan maupun pemilihan konfigurasi model. Karena distribusi target bersifat tidak seimbang, evaluasi tidak hanya menggunakan *accuracy*, tetapi juga *balanced accuracy*, *precision* kelas gagal, *recall* kelas gagal, F2-score, dan jumlah *false negative*. Hasil perbandingan ditunjukkan pada Tabel 3.

Tabel 3. Perbandingan Performa Model

Model	Accuracy	Balanced Accuracy	Precision Failure	Recall Failure	F2-score	FN
Logistic Regression	0,7626	0,8417	0,1184	0,9265	0,3918	5
Decision Tree	0,7882	0,8691	0,1340	0,9559	0,4293	3
Random Forest	0,9384	0,9468	0,3514	0,9559	0,7112	3
XGBoost	0,9489	0,9452	0,3951	0,9412	0,7373	4

Logistic Regression dan Decision Tree menghasilkan *recall* kelas gagal yang tinggi, tetapi memiliki *precision* yang relatif rendah. Kondisi tersebut menunjukkan bahwa kedua model mampu mendeteksi sebagian besar observasi kegagalan, tetapi juga menghasilkan jumlah prediksi positif palsu yang cukup besar. Random Forest menunjukkan peningkatan yang signifikan dengan *balanced accuracy* sebesar 0,9468 dan *recall* sebesar 0,9559, sedangkan XGBoost menghasilkan *accuracy* tertinggi sebesar 0,9489 dan F2-score tertinggi sebesar 0,7373.

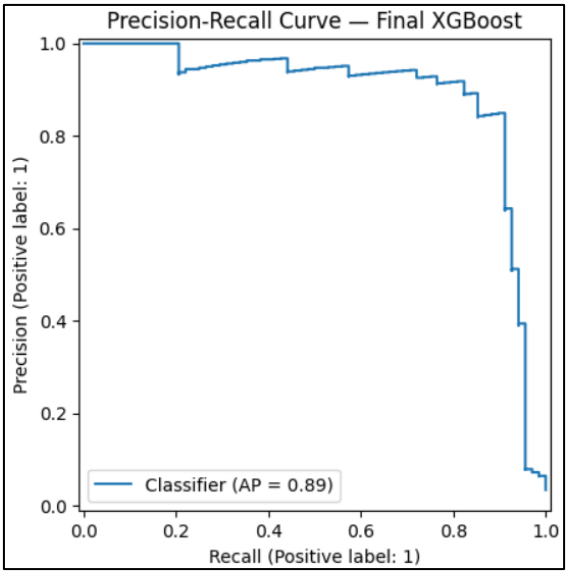
XGBoost dipilih sebagai model akhir karena memberikan keseimbangan performa yang paling sesuai dengan tujuan sistem. Meskipun *recall* Random Forest sedikit lebih tinggi, XGBoost memiliki *precision* kelas gagal dan F2-score yang lebih tinggi. Penggunaan F2-score menempatkan bobot lebih besar pada *recall* sehingga sesuai dengan karakteristik *predictive maintenance* yang lebih memprioritaskan kemampuan menemukan mesin berisiko daripada meminimalkan seluruh *false positive*.

Nilai *recall* kelas gagal sebesar 0,9412 menunjukkan bahwa sebagian besar data kegagalan dapat dikenali oleh XGBoost. Namun, *precision* sebesar 0,3951 menunjukkan bahwa hanya sebagian prediksi kelas gagal yang benar-benar merupakan kegagalan. Kondisi tersebut menandakan adanya *trade-off* antara sensitivitas terhadap kegagalan dan jumlah *false alarm*. Dalam konteks prototipe ini, kesalahan berupa mesin normal yang ditandai berisiko masih dapat ditindaklanjuti melalui pemeriksaan tambahan, sedangkan kegagalan yang diprediksi sebagai kondisi normal berpotensi menyebabkan risiko operasional yang lebih besar.

Hasil ini juga memperlihatkan bahwa *accuracy* tidak cukup digunakan sebagai satu-satunya indikator keberhasilan. Penelitian Hafidhoh et al [5]. dan Sugito et al [6]. melaporkan akurasi model di atas 95%, tetapi tugas klasifikasi, konfigurasi data, dan pendekatan pemodelannya berbeda sehingga nilai tersebut tidak dapat dibandingkan secara langsung dengan penelitian ini. Pada dataset dengan ketidakseimbangan kelas, penggunaan *recall*, F2-score, dan *false negative* memberikan gambaran yang lebih relevan terhadap kemampuan model mengenali kondisi kegagalan.

2. Analisis Precision-Recall dan Karakteristik Kelas Minoritas

Kurva *Precision-Recall* digunakan untuk mengevaluasi kemampuan model XGBoost terhadap kelas kegagalan sebagai kelas minoritas. Hasil evaluasi menghasilkan nilai *Average Precision* sekitar 0,89. Nilai tersebut menunjukkan bahwa model masih mampu mempertahankan pemisahan yang baik antara kelas normal dan gagal pada berbagai titik *decision threshold*, meskipun distribusi kelas sangat tidak seimbang.



Gambar 1. Kurva Precision-Recall Model XGBoost

Kurva tersebut memperlihatkan adanya *trade-off* antara *precision* dan *recall*. Peningkatan sensitivitas terhadap kegagalan dapat diikuti oleh bertambahnya *false positive*. Oleh karena itu, penentuan *threshold* perlu mempertimbangkan konsekuensi operasional dari kedua jenis kesalahan, bukan hanya mencari satu nilai metrik tertinggi.

3. Kontribusi Fitur Operasional dan Physics-Informed Features

Analisis *feature importance* dilakukan untuk mengetahui kontribusi relatif setiap fitur terhadap prediksi XGBoost. Hasilnya ditunjukkan pada Tabel 4.

Tabel 4. Feature Importance Model XGBoost

Peringkat	Fitur	Importance
1	Rotational speed	0,2421
2	Torque	0,1851
3	Tool wear	0,1592
4	Mechanical power	0,1351
5	Temperature difference	0,1241
6	Strain load	0,0929
7	Air temperature	0,0441
8	Process temperature	0,0133
9	Product type	0,0040

Fitur *rotational speed*, *torque*, dan *tool wear* memberikan kontribusi terbesar terhadap model. Ketiganya berkaitan langsung dengan kondisi mekanik, beban operasi, dan tingkat keausan mesin. Menariknya, tiga fitur hasil *physics-informed feature engineering* juga memiliki kontribusi yang cukup terlihat, yaitu *mechanical power* sebesar 0,1351, *temperature difference* sebesar 0,1241, dan *strain load* sebesar 0,0929.

Hal ini menunjukkan bahwa representasi berbasis hubungan fisik digunakan oleh model akhir dalam proses prediksi. Selain pada fitur input, informasi fisik diterapkan melalui *physics guardrails* untuk memeriksa indikasi *Power Failure*, *Heat Dissipation Failure*, *Overstrain Failure*, dan *Tool Wear Failure*. Mekanisme ini membuat penentuan status mesin

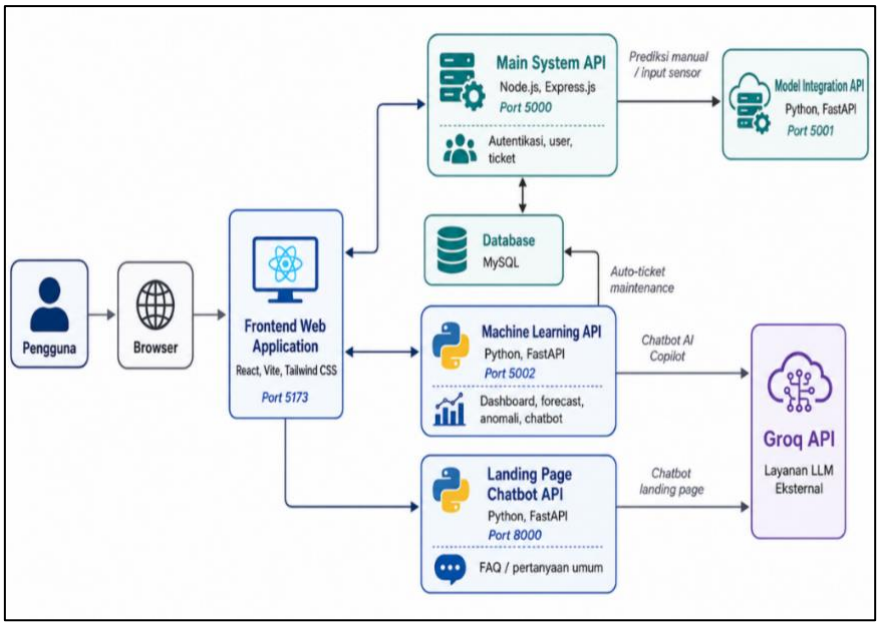
tidak hanya bergantung pada probabilitas model, tetapi juga pada indikator kondisi fisik yang digunakan dalam sistem.

4. Integrasi Prediksi Risiko dan Estimasi Remaining Useful Life

XGBoost digunakan untuk menghasilkan probabilitas risiko kegagalan, sedangkan RUL dihitung melalui simulasi degradasi berbasis *tool wear*. Kedua proses tersebut sengaja dipisahkan karena memiliki fungsi analitik yang berbeda. Nilai RUL menurun ketika *tool wear* mendekati batas kritis. Perbedaan *duty cycle* dan *wear increment* menyebabkan setiap tipe mesin memiliki laju degradasi simulasi yang berbeda. Informasi tersebut kemudian digunakan bersama probabilitas risiko untuk membantu menentukan prioritas pemeriksaan. Mesin dengan risiko tinggi dan RUL rendah dapat diprioritaskan lebih awal dibandingkan mesin dengan risiko rendah dan estimasi sisa umur yang masih panjang. Namun, RUL pada penelitian ini masih merupakan hasil simulasi dan belum dikalibrasi menggunakan histori degradasi mesin industri aktual. Karena itu, hasil RUL belum dapat diposisikan sebagai prognostik yang siap digunakan langsung pada aset nyata.

5. Implementasi Predictive Maintenance Copilot sebagai Sistem Pendukung Keputusan

Hasil analitik diintegrasikan ke dalam Predictive Maintenance Copilot berbasis web. Sistem menghubungkan antarmuka pengguna, backend utama, layanan machine learning, layanan integrasi model, basis data, dan Large Language Model dalam arsitektur modular.



Gambar 2. Arsitektur Predictive Maintenance Copilot

AI Copilot digunakan untuk menyampaikan keluaran analitik dalam bahasa natural. Konteks yang diteruskan kepada LLM berasal dari hasil analisis sistem, seperti kondisi mesin, tingkat risiko, estimasi RUL, parameter bermasalah, dan rekomendasi pemeliharaan. Dengan demikian, LLM tidak digunakan sebagai pengganti model prediksi, tetapi sebagai lapisan interaksi untuk membantu pengguna memahami informasi yang telah dihasilkan komponen analitik.

Pendekatan tersebut berbeda dengan chatbot informasional berbasis RAG yang dikembangkan Tutu et al. [10], karena konteks percakapan Predictive Maintenance Copilot berasal dari kondisi dan keluaran analitik mesin. Walaupun AI Copilot membantu meningkatkan keterbacaan informasi, fitur ini juga tidak diposisikan sebagai metode XAI

formal. Penelitian Dereci dan Tuzkaya [9] menunjukkan bahwa XAI dapat digunakan untuk meningkatkan transparansi model *predictive maintenance*, sedangkan penelitian ini menggunakan *feature importance*, indikator fisik, status risiko, dan informasi analitik sebagai konteks yang kemudian dijelaskan melalui bahasa natural.

Sistem juga menyediakan simulasi *maintenance ticket*. Ketika mesin berada pada kondisi yang membutuhkan perhatian, informasi risiko dapat diteruskan menjadi tiket pemeliharaan yang memuat identitas mesin, tingkat risiko, deskripsi masalah, rekomendasi tindakan, prioritas, dan status penanganan. Integrasi tersebut membuat alur sistem tidak berhenti pada prediksi, tetapi dilanjutkan sampai tahap dukungan tindak lanjut.

6. Evaluasi Pengguna

Evaluasi dilakukan terhadap 10 responden yang memiliki keterkaitan dengan bidang Operation & Maintenance.

Tabel 5. Hasil Evaluasi Pengguna

Aspek	Persentase
Kemudahan penggunaan sistem	92%
Kejelasan dashboard	88%
Visualisasi data sensor	90%
Pemahaman estimasi RUL	88%
Manfaat deteksi anomali	92%
AI Copilot membantu memahami prediksi	94%
Penjelasan AI Copilot meningkatkan pemahaman	90%
Kesesuaian rekomendasi mitigasi	88%
Manfaat maintenance ticket	92%
Kelayakan sebagai alat bantu keputusan	94%

Total skor evaluasi pengguna adalah 90,8% dan termasuk kategori Sangat Baik. Nilai tertinggi diperoleh pada pernyataan mengenai manfaat AI Copilot dalam membantu memahami hasil prediksi dan kelayakan sistem sebagai alat bantu pengambilan keputusan, masing-masing sebesar 94%. Hasil tersebut menunjukkan bahwa fungsi interpretasi dan dukungan keputusan memperoleh respons positif dari responden. Penilaian 90,8% tidak dapat diinterpretasikan sebagai bukti bahwa sistem telah efektif secara operasional pada industri minyak dan gas. Evaluasi dilakukan pada jumlah responden yang terbatas dan berfokus pada penerimaan terhadap prototipe. Penelitian lanjutan masih diperlukan dengan jumlah pengguna yang lebih besar, serta pengukuran dampak terhadap waktu pengambilan keputusan, *downtime*, atau efektivitas tindakan pemeliharaan.

7. Pengujian Fungsional dan Internal Sistem

Pengujian black-box dilakukan terhadap 13 skenario fungsi utama dan seluruh skenario menghasilkan status berhasil. Pengujian white-box dilakukan menggunakan pendekatan *unit testing* terhadap lima komponen utama. Rekapitulasi pengujian ditunjukkan pada Tabel 6.

Tabel 6. Rekapitulasi White-Box Testing

Komponen	Framework	Test Case	Hasil
Main System API	Jest	63	PASS
Machine Learning API	Pytest	7	PASS
Model Integration API	Pytest	6	PASS
Landing Page Chatbot API	Pytest	3	PASS
Frontend	Vitest	8	PASS

Seluruh 87 *test case* berhasil dijalankan. Hasil tersebut menunjukkan bahwa fungsi internal, integrasi layanan, serta mekanisme validasi yang diuji dapat berjalan sesuai skenario pengujian. Black-box dan white-box testing menunjukkan bahwa fungsi perangkat lunak bekerja sesuai implementasi dan skenario pengujian yang ditetapkan. Secara keseluruhan, hasil penelitian menunjukkan bahwa Predictive Maintenance Copilot dapat mengintegrasikan prediksi risiko berbasis machine learning, informasi fisik, estimasi degradasi, interaksi AI, dan tindak lanjut pemeliharaan ke dalam satu prototipe Sistem Pendukung Keputusan. Kontribusi utama penelitian tidak terletak pada algoritma XGBoost secara tunggal, tetapi pada integrasi komponen analitik dan pengetahuan domain menjadi alur pengambilan keputusan yang lebih terstruktur. Meskipun demikian, validasi menggunakan data sensor dan histori kegagalan aktual tetap diperlukan sebelum pendekatan ini dapat diterapkan pada lingkungan industri.

KESIMPULAN DAN SARAN

Penelitian ini berhasil mengembangkan Predictive Maintenance Copilot sebagai prototipe Sistem Pendukung Keputusan yang mengintegrasikan prediksi risiko kegagalan, *physics-informed feature engineering*, *physics guardrails*, estimasi *Remaining Useful Life* (RUL), AI Copilot, dan simulasi *maintenance ticket*. Dari empat algoritma yang dibandingkan, XGBoost dipilih sebagai model akhir dengan F2-score tertinggi sebesar 0,7373 dan *recall* kelas gagal sebesar 0,9412. Pendekatan Physics-Informed Machine Learning diterapkan melalui fitur *mechanical power*, *temperature difference*, dan *strain load*, serta melalui *physics guardrails* yang melengkapi hasil prediksi berbasis data. Estimasi RUL dihitung secara terpisah melalui simulasi degradasi berbasis *tool wear*, *duty cycle*, dan *wear increment*. Integrasi kedua informasi tersebut memungkinkan sistem memberikan gambaran mengenai tingkat risiko dan urgensi pemeliharaan secara bersamaan. Evaluasi terhadap 10 responden yang berkaitan dengan Operation & Maintenance menghasilkan tingkat penerimaan sebesar 90,8% dengan kategori Sangat Baik. Seluruh 13 skenario *black-box testing* dan 87 *white-box test case* juga berhasil dijalankan. Meskipun demikian, penggunaan dataset sintetis, estimasi RUL berbasis simulasi, serta belum dilakukannya *ablation study* menjadi keterbatasan utama penelitian.

Penelitian selanjutnya disarankan menggunakan data sensor dan histori kegagalan aktual untuk memvalidasi performa model serta mengkalibrasi estimasi RUL. Eksperimen *ablation* juga diperlukan untuk mengukur kontribusi fitur berbasis fisik secara lebih objektif. Selain itu, metode interpretabilitas lokal seperti SHAP dan pendekatan prognostik berbasis *time series* dapat dieksplorasi untuk meningkatkan transparansi dan kemampuan prediksi sistem.

DECLARATION

1. *Statement on the Use of Artificial Intelligence*

Dalam penyusunan artikel ini, penulis menggunakan *OpenAI ChatGPT* untuk membantu penyempurnaan bahasa dan pemeriksaan konsistensi. Seluruh konten yang dibantu oleh AI telah ditinjau, diverifikasi, dan direvisi oleh penulis. Penulis bertanggung jawab penuh atas keakuratan, originalitas, integritas, dan isi akhir artikel.

2. *Author Contribution Statement*

Rasyid Naufal Saputra berkontribusi dalam konseptualisasi penelitian, penyusunan metodologi, pengembangan sistem, pengolahan dan analisis data, evaluasi hasil, serta penyusunan artikel. Heni Ramawati dan Purwanto berkontribusi dalam supervisi penelitian, validasi metodologi dan hasil, serta peninjauan dan penyempurnaan artikel. Seluruh penulis telah membaca dan menyetujui versi akhir artikel.

3. *Funding Statement*

Penelitian ini tidak menerima pendanaan atau hibah khusus dari lembaga pemerintah, komersial, maupun organisasi pendanaan lainnya.

4. *Conflict of Interest*

Penulis menyatakan bahwa tidak terdapat konflik kepentingan yang berkaitan dengan penelitian dan penerbitan artikel ini.

DAFTAR PUSTAKA

[1] Direktorat Jenderal Minyak dan Gas Bumi, "LAPORAN KINERJA Tahun 2025 Direktorat Jenderal Minyak dan Gas Bumi," Indonesia, 2025. Accessed: Mar. 12, 2026. [Online]. Available: migas.esdm.go.id

[2] E. O. Nwulu, T. Y. Elete, O. V. Erhueh, O. A. Akano, and K. O. Omomo, "Machine learning applications in predictive maintenance: Enhancing efficiency across the oil and gas industry," *International Journal of Engineering Research Updates*, vol. 5, no. 1, pp. 013–027, Sep. 2023, doi: 10.53430/ijeru.2023.5.1.0017.

[3] Y. Wu, B. S. Sicard, and S. A. Gadsden, "Physics-informed machine learning: A comprehensive review on applications in anomaly detection and condition monitoring," *Expert Syst. Appl.*, vol. 255, p. 124678, 2024, doi: 10.1016/j.eswa.2024.124678.

[4] M. H. de Beaulieu, M. S. Jha, H. Garnier, and F. Cerbah, "Remaining Useful Life prediction based on physics-informed data augmentation," *Reliab. Eng. Syst. Saf.*, vol. 252, p. 110451, 2024, doi: 10.1016/j.ress.2024.110451.

[5] N. Hafidhoh, A. P. Atmaja, G. N. Syaifuddiin, I. B. Sumafta, S. M. Pratama, and H. N. Khasanah, "Machine Learning untuk Prediksi Kegagalan Mesin dalam Predictive Maintenance System," *Jurnal Masyarakat Informatika*, vol. 15, no. 1, pp. 56–66, May 2024, doi: 10.14710/jmasif.15.1.63641.

[6] B. Sugito, M. Iqbal, and Z. Sitorus, "Analisis Metode FMEA Dan SPC Pada Proyeksi Losses Produksi Dan Prediksi Perawatan Pompa Minyak Di PT.Pertamina Ep Zona 1 Rantau Field Kuala Simpang," *Jurnal Riset Sistem Informasi Dan Teknik Informatika (JURASIK)*, vol. 10, pp. 468–479, Feb. 2025, doi: 10.30645/jurasik.v10i1.889.

[7] M. Wang, X. Su, H. Song, Y. Wang, and X. Yang, "Enhancing predictive maintenance strategies for oil and gas equipment through ensemble learning modeling," *J. Pet. Explor. Prod. Technol.*, vol. 15, no. 3, Mar. 2025, doi: 10.1007/s13202-025-01931-x.

[8] Y. Meinard and A. Tsoukiàs, "What Is Legitimate Decision Support?," in *Intelligent Decision Support Systems: Combining Operations Research and Artificial Intelligence – Essays in Honor of Roman Słowiński*, S. Greco, V. Mousseau, J. Stefanowski, and C. Zopounidis, Eds., Cham: Springer, 2022, pp. 207–224. doi: 10.1007/978-3-030-96318-7_11.

[9] U. Dereci and G. Tuzkaya, "An explainable artificial intelligence model for predictive maintenance and spare parts optimization," *Supply Chain Analytics*, vol. 8, p. 100078, 2024, doi: 10.1016/j.sca.2024.100078.

[10] E. M. I. Tutu, B. Pinontoan, E. Tenda, C. A. J. Soewoeh, M. I. Takaendengan, and S. C. W. Ngangi, "Pengembangan Chatbot Perpustakaan Berbasis Retrieval-Augmented Generation (RAG) pada Telegram," *INTECH*, vol. 7, no. 1, pp. 55–65, May 2026, doi: 10.54895/intech.v7i1.3435.

[11] F. A. Dzaky and D. Kurniawan, "Implementation of Scrum Framework Agile Method to Develop Integrated Asset Management Information System at Universitas Diponegoro Inventory Module," *Jurnal Masyarakat Informatika*, vol. 14, no. 1, pp. 53–69, 2023, doi: 10.14710/jmasif.14.1.52605.

[12] S. Matzka, "AI4I 2020 Predictive Maintenance Dataset," 2020, *UCI Machine Learning Repository*. doi: 10.24432/C5HS5C.

[13] A. C. Praniffa, A. Syahri, F. Sandes, U. Fariha, and Q. A. Giansyah, "PENGUJIAN BLACK BOX DAN WHITE BOX SISTEM INFORMASI PARKIR BERBASIS WEB," *Jurnal Testing dan Implementasi Sistem Informasi*, vol. 1, no. 1, pp. 1–16, 2023, doi: 10.55583/jtisi.v1i1.321.