
Prediksi Remaining Useful Life (RUL) Mesin Produksi UMKM Menggunakan Model Machine Learning dengan XGBoost

Muhammad Tiflanil Azhar Rudin

Department of Computer Science
Bina Nusantara University
Palmerah, Jakarta Barat 11530
muhammad.rudin@binus.ac.id

Jonathan Dharma Ratana

Department Of Computer Science
Bina Nusantara University
Palmerah, Jakarta Barat 11530
jonathan.ratana@binus.ac.id

Abstract

Mesin produksi adalah hal yang sangat krusial dalam bisnis UMKM untuk menjaga kontinuitas bisnis. Namun seringkali keandalan mesin menjadi penyebab utama terhambatnya operasional pada sektor bisnis usaha mikro, kecil dan menengah, hal ini disebabkan oleh keterbatasan sumber daya untuk perawatan mesin sehingga menimbulkan downtime dan kerusakan mendadak, hal ini tentu sangat merugikan bagi para pengusaha skala menengah kebawah. mengusulkan model prediksi Remaining Useful Life (RUL) berbasis algoritma Extreme Gradient Boosting (XGBoost) untuk memperkirakan umur sisa pakai mesin secara lebih akurat. Data historis operasional mesin, termasuk jumlah produk, jam kerja, dan catatan kerusakan, implementasi sistem ini berpotensi membantu pelaku UMKM dalam merencanakan perawatan preventif secara lebih efisien dan mengurangi risiko gangguan produksi.

1 Pendahuluan

Usaha Mikro, Kecil, dan Menengah (UMKM) memegang peranan penting dalam perekonomian nasional, terutama sebagai penyedia lapangan kerja dan penggerak ekonomi lokal. Namun, salah satu tantangan utama yang dihadapi oleh sektor UMKM adalah keterbatasan dalam hal pemeliharaan dan manajemen aset produksi, khususnya mesin. Kegagalan mesin yang tidak terprediksi dapat menyebabkan downtime yang signifikan, menurunkan produktivitas, dan meningkatkan biaya operasional.

Remaining Useful Life (RUL) atau umur sisa pakai adalah estimasi waktu sebelum suatu komponen atau sistem mengalami kegagalan. Prediksi RUL menjadi penting untuk mendukung strategi pemeliharaan preventif dan meminimalkan risiko kerusakan mendadak. Di sektor industri, prediksi RUL telah banyak diterapkan menggunakan metode data-driven, seperti Machine Learning. Namun, penerapannya dalam konteks UMKM masih sangat terbatas, baik karena keterbatasan data maupun akses terhadap teknologi sehingga menggunakan data manual (*logging*) akan lebih diutamakan.

Dalam penelitian ini, kami mengusulkan penggunaan model *Extreme Gradient Boosting (XGBoost)* untuk memprediksi RUL mesin produksi UMKM berbasis data historis operasional mesin, seperti jam kerja, jumlah produk yang dihasilkan, dan frekuensi kerusakan, metode XGBoost dipilih karena *versatile* dan dapat menangani data tabular kompleks

46 **2 Kajian Teori**

47 **2.1 Remaining Useful Life (RUL)**

48 Remaining Useful Life (RUL) adalah estimasi waktu atau siklus operasi yang tersisa
49 sebelum suatu komponen mengalami kegagalan. Prediksi RUL menjadi komponen penting
50 dalam sistem *prognostics and health management* (PHM) karena memungkinkan perawatan
51 dilakukan secara proaktif sebelum kerusakan terjadi. Terdapat dua pendekatan utama dalam
52 prediksi RUL: pendekatan berbasis model fisik (physics-based) dan pendekatan berbasis data
53 (data-driven). Untuk mesin UMKM, pendekatan data-driven lebih relevan karena
54 keterbatasan informasi desain dan biaya sensorisasi.

56 **2.2 XGBoost (Extreme Gradient Boosting)**

57 XGBoost merupakan algoritma *ensemble* berbasis gradient boosting yang terkenal karena
58 efisiensi dan akurasi dalam berbagai kompetisi dan aplikasi nyata. XGBoost
59 menggunakan teknik boosting dengan pohon keputusan sebagai *base learner*, dan
60 dioptimalkan menggunakan metode regularisasi untuk menghindari overfitting. XGBoost
61 mendukung penanganan missing value, fitur penting dalam konteks industri kecil yang
62 sering kali memiliki data tidak lengkap atau berkualitas rendah.
63 XGBoost telah terbukti efektif dalam prediksi kegagalan peralatan, termasuk pada data
64 sensor dan kondisi mesin, serta mampu memberikan estimasi RUL dengan performa yang
65 kompetitif dibanding metode lain.

68 **3 Solusi Usulan**

69 Sebuah sistem yang memprediksi Remaining Useful Life (RUL) mesin produksi yang
70 dirancang khusus untuk kebutuhan UMKM dengan pendekatan berbasis data dan algoritma
71 Machine Learning. Sistem ini bertujuan untuk membantu pelaku UMKM dalam melakukan
72 perawatan mesin secara efisien, dengan memanfaatkan data historis operasional melalui log
73 manual.

75 **3.1 Pengumpulan Data**

76 Data dikumpulkan dari mesin produksi UMKM yang mencakup:

- 77 • Tanggal
- 78 • Id_mesin
- 79 • Jenis_mesin
- 80 • Jam_operasi
- 81 • Jumlah_produksi
- 82 • usia_mesin (dalam bulan)
- 83 • terakhir_servis
- 84 • hari_sejak_servis
- 85 • kerusakan (berupa Boolean apakah terjadi rusak atau tidak)

86

dan data saat ini yang kami gunakan adalah hasil dari *generated data* untuk melakukan simulasi.

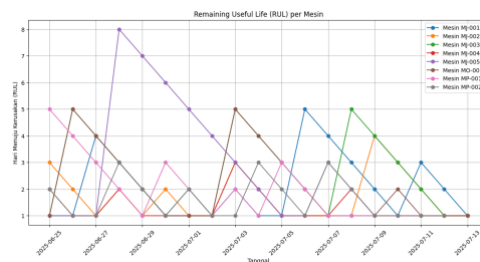
3.1 Membuat model dengan XGBoost

Data historis operasional yang dikumpulkan dari mesin produksi UMKM akan dibangun menjadi sebuah model dengan XGBRegressor, walaupun data kerusakan berupa biner tapi kita membangun sebuah kolom *days_to_failure* secara otomatis dalam program yang dapat digunakan sebagai target regresi

4 Hasil Eksperimen dan Pengujian

Diberikan dataset berupa data uji coba berformat CSV berisikan data yang dibutuhkan. berisikan mesin produksi UMKM produksi baju yaitu: mesin jahit, obras dan pemotong. Setelah itu akan di bangun sebuah model dengan XGBRegressor dengan target. Dengan data uji coba yang kami miliki, ini menghasilkan beberapa hasil analisis dan prediksi.

4.1 RUL (remaining useful life)



Plot diatas merupakan visualisasi dari data Remaining Useful Life (RUL) per mesin yang menggambarkan sisa usia mesin sebelum diperbaiki

4.2 RUL (remaining useful life)



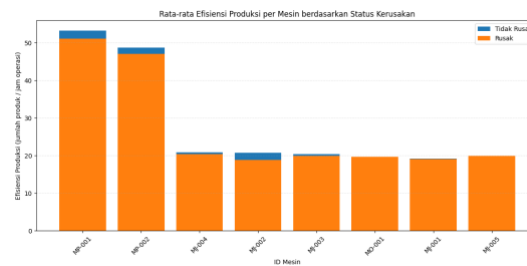
Plot diatas menunjukan prediksi hari menuju kerusakan dengan versi data asli yang menggambarkan keakuratan prediksi, data yang buruk (seperti *data loss* atau ketidaktepatan) kemungkinan akan menghasilkan hasil yang kurang baik

4.3 Prediksi tanggal servis direkomendasikan

	id_mesin	ru_prediksi	tanggal	tanggal_terakhir_kerusakan	tanggal_sarat_direkomendasikan	rekomendasi_sarat
8	MJ-001	2.108041	2025-07-12	2025-07-12 02:35:34.767700878	2025-07-12 02:35:34.767700878	red
10	MJ-002	2.676045	2025-07-12	2025-07-12 06:36:20.077001441	2025-07-12 06:36:20.077001441	red
7	MJ-001	1.666423	2025-07-12	2025-07-12 06:02:01.70889721	2025-07-09 06:02:01.70889721	red
11	MJ-003	2.235743	2025-07-12	2025-07-12 08:38:28.170058249	2025-07-09 08:38:28.170058249	red
4	MJ-004	1.776951	2025-07-12	2025-07-12 08:03:48.84447228	2025-07-07 08:03:48.84447228	red
6	MJ-001	3.096708	2025-07-12	2025-07-12 08:50:16.345891688	2025-07-08 08:50:16.345891688	red
10	MJ-003	2.218777	2025-07-08	2025-07-12 06:09:51.27731232	2025-07-08 06:09:51.27731232	red
1	MP-002	1.020086	2025-07-08	2025-07-08 00:04:36.48871688	2025-07-04 00:04:36.48871688	red
10	MJ-001	1.686309	2025-07-07	2025-07-04 06:28:17.086670071	2025-07-03 06:28:17.086670071	red
14	MJ-002	1.770276	2025-07-07	2025-07-08 08:29:11.738630797	2025-07-03 08:29:11.738630797	red
0	MJ-004	1.020086	2025-07-07	2025-07-08 06:12:28.800600622	2025-07-03 06:12:28.800600622	red
12	MJ-001	1.646742	2025-07-06	2025-07-07 00:43:10.968976977	2025-07-02 00:43:10.968976977	red
10	MJ-001	1.302087	2025-07-05	2025-07-06 07:05:43.500491224	2025-07-01 07:05:43.500491224	red
8	MJ-001	1.818181	2025-07-04	2025-07-05 02:12:32.324891205	2025-06-30 02:12:32.324891205	red
23	MJ-004	4.109881	2025-07-04	2025-07-08 04:04:48.371643851	2025-07-03 04:04:48.371643851	yellow
21	MP-002	1.488813	2025-07-04	2025-07-05 01:12:21.281888987	2025-06-30 01:12:21.281888987	red
8	MJ-008	3.856657	2025-07-03	2025-07-08 20:32:26.187830891	2025-07-07 20:32:26.187830891	yellow
2	MJ-002	1.806477	2025-07-03	2025-07-04 10:12:41.248108005	2025-06-29 10:12:41.248108005	red
5	MP-001	1.888874	2025-07-02	2025-07-03 04:47:57.748888379	2025-06-28 04:47:57.748888379	red
25	MJ-003	2.008887	2025-07-01	2025-07-04 07:04:10.648716712	2025-06-29 07:04:10.648716712	red
17	MP-003	1.130381	2025-07-01	2025-07-01 03:53:13.83789331	2025-06-30 03:53:13.83789331	red

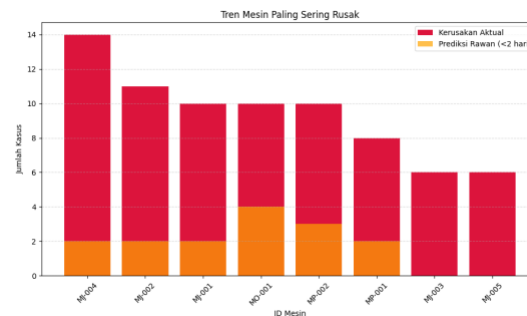
Akan mencetak data tabular berisikan rekomendasi kapan mesin harus diperbaiki dan level kedaruratan mesin tersebut diperbaiki (tingkatan berupa warna: red, yellow dan green) prioritas sesuai urutan warna tersebut

4.4 Rata-rata efisiensi produksi mesin berdasarkan status kerusakan



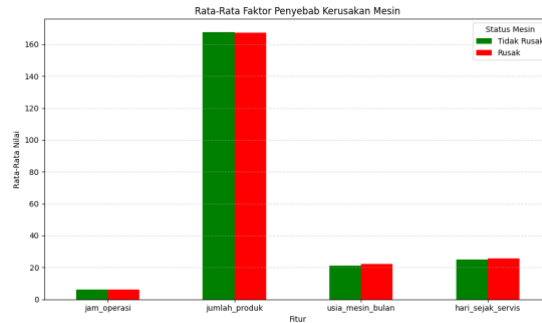
Disini rusak bukan berarti tidak dapat menghasilkan barang sama sekali bisa saja memproduksi barang walaupun jumlahnya berkurang, di data uji coba ini barang tetap dapat diproduksi dalam skala besar jadi tidak ada pengurangan signifikan.

4.5 Tren mesin paling sering rusak



Menampilkan data tren mesin yang sering rusak, hal ini bertujuan untuk melihat mesin apa yang harus diperhatikan lebih sehingga operasional akan lebih efisien.

4.6 Rata-rata faktor penyebab kerusakan



Menampilkan data berbentuk bar untuk melihat faktor apa saja (yang berada di dalam kolom data) yang mempengaruhi kerusakan pada mesin, saat ini hanya terdapat 4 kolom saja tapi ini dapat ditambah sesuai kebutuhan nanti

5 Analisis data

Model ini menggunakan data historis operasional untuk memprediksi Remaining Useful Life (RUL) atau sisa umur pakai mesin hingga kerusakan berikutnya terjadi. Model regresi XGBoost Regressor dipilih untuk memodelkan hubungan antara fitur operasional mesin dengan waktu menuju kerusakan. Kolom (atau bisa disebut fitur) yang digunakan dalam *training model* meliputi **jam operasi harian, jumlah produk yang dihasilkan, usia mesin dalam bulan, serta jumlah hari sejak servis terakhir dilakukan**. Fitur-fitur ini dinilai relevan karena secara teoritis memiliki keterkaitan dengan keausan dan penurunan performa mesin.

Target prediksi model, yaitu *days_to_failure*, dihitung berdasarkan selisih antara tanggal pengamatan saat ini dengan tanggal kerusakan terdekat berikutnya untuk setiap mesin. memungkinkan model untuk belajar dari pola historis kerusakan yang pernah terjadi. Meski demikian, hasil analisis menunjukkan bahwa distribusi nilai target cenderung tidak merata—beberapa mesin mengalami kerusakan secara berkala, sementara mesin lain relatif stabil. Hal ini bisa saja menyebabkan ketidakseimbangan data, yang dapat berdampak pada keakuratan prediksi.

Lalu untuk bagian prediksi kerusakan adalah hanya penjumlahan tanggal data terakhir dengan *prediksi_rul* dari XGBoost Regressor. Setelah itu untuk bagian *tanggal_servis_direkomendasikan* memiliki toleransi < 2 hari sebelum terjadi kerusakan. Hanya saja terkadang toleransi yang diberikan bisa terlalu dekat atau terlalu jauh karena diberikan batas toleransi secara konstan.

Bagian kerusakan mesin menunjukkan jumlah kerusakan aktual yang tercatat untuk tiap mesin, serta berapa kali model memprediksi bahwa mesin tersebut berada dalam kondisi "rawan rusak" yaitu ketika nilai *Remaining Useful Life (RUL)* diperkirakan tinggal 2 hari atau kurang, secara umum, beberapa mesin terlihat memiliki angka kerusakan yang memang cukup tinggi, dan ini sejalan dengan prediksi model yang juga sering mengindikasikan risiko, tapi disini juga dapat terjadi false alarm karena model terlalu "waspada"

6 Kesimpulan atau saran

Penggunaan XGBoost dalam menganalisa dan memprediksi data operasional mesin sebenarnya cukup akurat namun sangat bergantung pada data yang diberikan dan Visualisasi tren kerusakan menunjukkan bahwa beberapa mesin memiliki tingkat kerusakan yang konsisten tinggi, sementara beberapa lainnya meskipun tidak terlalu sering rusak, tetap masuk kategori rawan karena sering diprediksi akan segera rusak. Ini menunjukkan bahwa model mampu menangkap risiko lebih awal, meskipun perlu evaluasi lebih lanjut terhadap akurasi peringatan tersebut.

174 Dan juga fitur (kolom) mungkin akan terasa terlalu sedikit, tapi mungkin jika memang
175 dibutuhkan akan ditambahkan kolom yang berisikan data sensorik seperti temperatur, rpm,
176 dll tapi untuk saat ini fokus utama adalah menyederhanakan hal tersebut sehingga model
177 dapat menganalisa dan memprediksi berdasarkan pencatatan manual bukan sensorik dan
178 elektronik.

179 Untuk bagian teknis seperti *rul_prediksi* mungkin perlu dikalibrasi ulang untuk menghindari
180 false alarm karena memang itu adalah data yang cukup “rawan”

181

182 **References**

183

184 Himansuda. Predictive Maintenance Dataset. Kaggle,
185 <https://www.kaggle.com/datasets/hiimanshuagarwal/predictive-maintenance-dataset>

186 Diakses 10 juli 21.53.

187

188 Nvidia. XGBoost.

189 <https://www.nvidia.com/en-us/glossary/xgboost/>.

190 Diakses 10 juli 21.53.

191

192 Partneresi. Partneresi.

193 <https://www.partneresi.com/resources/glossary/remaining-useful-life-rul/>

194 Diakses 10 juli 21.55

195

196 Katser, lurii. Medium.

197 [https://medium.com/@katser/all-you-want-to-know-about-determining-remaining-useful-](https://medium.com/@katser/all-you-want-to-know-about-determining-remaining-useful-life-rul-of-industrial-equipment-6a88dc71a0ac)
198 [life-rul-of-industrial-equipment-6a88dc71a0ac](https://medium.com/@katser/all-you-want-to-know-about-determining-remaining-useful-life-rul-of-industrial-equipment-6a88dc71a0ac)

199 Diakses 10 july 22.00