

TUGAS AKHIR - TF234801

**ESTIMASI *REMAINING USEFUL LIFE* (RUL) SISTEM
TRAKSI LRT JABODEBEK MENGGUNAKAN *SIMILARITY
BASED POLYNOMIAL***

MUHAMMAD JAMIL UWAIS

NRP 5009211123

Dosen Pembimbing

Prof. Dr. Ir. Katherin Indriawati, S.T., M.T., IPU

NIP 19760523 200012 2 001

Natasya Ariola Santoso Puteri, S.T., M.T.

Program Studi S-1 Teknik Fisika

Departemen Teknik Fisika

Fakultas Teknologi Industri dan Rekayasa Sistem

Institut Teknologi Sepuluh Nopember

Surabaya

2026



TUGAS AKHIR - TF 234801

**ESTIMASI *REMAINING USEFUL LIFE* (RUL) SISTEM
TRAKSI LRT JABODEBEK MENGGUNAKAN *SIMILARITY
BASED POLYNOMIAL***

MUHAMMAD JAMIL UWAIS

NRP 5009211123

Dosen Pembimbing

Prof Dr. Ir. Katherin Indriawati, S.T., M.T. IPU

NIP 19760523 200012 2 001

Natasya Ariola Santoso Puteri, S.T., M.T.

Program Studi S1 Teknik Fisika

Departemen Teknik Fisika

Fakultas Teknologi Industri dan Rekayasa Sistem

Institut Teknologi Sepuluh Nopember

Surabaya

2026



FINAL PROJECT - TF 234801

REMAINING USEFUL LIFE (RUL) ESTIMATION FOR THE LRT JABODEBEK TRACTION SYSTEM USING A SIMILARITY BASED POLYNOMIAL

MUHAMMAD JAMIL UWAIS
5009211123

Advisor

Prof Dr. Ir. Katherin Indriawati, S.T., M.T. IPU

NIP 19760523 200012 2 001

Natasya Ariola Santoso Puteri, S.T., M.T.

Study Program Bachelor of Engineering Physics

Department of Engineering Physics

Faculty of Industrial Technology and System Engineering

Institut Teknologi Sepuluh Nopember

Surabaya

2026

LEMBAR PENGESAHAN

ESTIMASI *REMAINING USEFUL LIFE* (RUL) SISTEM TRAKSI LRT JABODEBEK MENGGUNAKAN *SIMILARITY BASED POLYNOMIAL*

TUGAS AKHIR

Diajukan untuk memenuhi salah satu syarat memperoleh gelar Sarjana Teknik pada
Program Studi Sarjana Teknik Fisika Departemen Teknik Fisika
Fakultas Teknologi Industri dan Rekayasa Sistem
Institut Teknologi Sepuluh Nopember

Oleh: **Muhammad Jamil Uwais**

NRP. 5009211123

Disetujui oleh Tim Penguji Tugas Akhir:

- | | |
|--|--------------|
| 1. Prof. Dr. Ir. Katherin Indriawati, S.T., M.T., IPU Pembimbing | |
| 2. Dr. Bambang Lelono Widjiantoro, S.T., M.T. | Ketua Sidang |
| 3. Prof. Totok Ruki Biyanto, S.T., M.T., Ph.D. | Penguji I |
| 4. Iwan Cony Setiadi, S.T., M.T. | Penguji II |
| 5. Tasya Yusiasfa Christnantasari, S.T., M.Sc. | Sekretaris |

(.....*Katherin*.....)
(.....*Bambang*.....)
(.....*Totok*.....)
(.....*Iwan*.....)
(.....*Tasya*.....)

Mengetahui,

Kepala Departemen Teknik Fisika FT-IRS ITS



Prof. Gunawan Nugroho, S.T., M.T., Ph.D.
197711272002121002

Halaman ini sengaja dikosongkan

APPROVAL SHEET

REMAINING USEFUL LIFE (RUL) ESTIMATION FOR THE LRT JABODEBEK TRACTION SYSTEM USING SIMILARITY-BASED POLYNOMIAL

FINAL PROJECT

Submitted to fulfill one of the requirements for
Bachelor of Engineering in Physics Program, Department of Physics Engineering
Faculty of Industrial Technology and Systems Engineering
Institut Teknologi Sepuluh Nopember

By: **Muhammad Jamil Uwais**

ID. 5009211123

Approved by the Final Project Examination Committee:

- | | |
|---|---------|
| 1. Prof. Dr. Ir. Katherin Indriawati, S.T., M.T., IPU Advisor | (.....) |
| 2. Dr. Bambang Lelono Widjiantoro, S.T., M.T. Chief Examiner | (.....) |
| 3. Prof. Totok Ruki Biyanto, S.T., M.T., Ph.D. Examiner I | (.....) |
| 4. Iwan Cony Setiadi, S.T., M.T. Examiner II | (.....) |
| 5. Tasya Yusiasfa Christnantasari, S.T., M.Sc. Secretary | (.....) |

Acknowledged by,
Head of the Department of Engineering Physics FT-IRS ITS


Prof. Sudhawan Nugroho, S.T., M.T., Ph.D.
197711272002121002

Halaman ini sengaja dikosongkan

PERNYATAAN ORISINALITAS
STATEMENT OF ORIGINALITY

Yang bertanda tangan di bawah ini
The undersigned below

Nama Mahasiswa / NRP : Muhammad Jamil Uwais / 5009211123
Student Full Name / Student ID

Program Studi : S1 Teknik Fisika
Study Programme

dengan ini menyatakan bahwa Tugas Akhir dengan judul “Estimasi *Remaining Useful Life* (RUL) Sistem Traksi LRT Jabodebek Menggunakan *Similarity Based Polynomial*” adalah hasil karya saya sendiri, bersifat orisinal, dan ditulis dengan mengikuti kaidah penulisan ilmiah. Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia menerima sanksi sesuai dengan ketentuan yang berlaku di Institut Teknologi Sepuluh Nopember.

hereby declare that the Final Project with the title of “Remaining useful life (RUL) estimation for the LRT Jabodebek traction system using a similarity based polynomial” is the result of my own work, is original, and is written by following the rules of scientific writing.

If in the future there is a discrepancy with this statement, then I am willing to accept sanctions in accordance with the provisions that apply at Institut Teknologi Sepuluh Nopember.

Surabaya, 29 Juli 2026

Mahasiswa,



Muhammad Jamil Uwais

NRP. 5009211123

Halaman ini sengaja dikosongkan

PERNYATAAN KODE ETIK PENGGUNAAN AI GENERATIF
CODE OF CONDUCT STATEMENT: GENERATIVE AI OR AI-ASSISTED USAGE

Saya yang bertanda tangan di bawah ini

I, the undersigned

Nama Mahasiswa / NRP : Muhammad Jamil Uwais / 5009211123

Full Name / Student ID

Program Studi : S-1 Teknik Fisika

Study Programme

Judul Tugas Akhir : Estimasi *Remaining Useful Life* (RUL) Sistem Traksi LRT

Final Project Title

Jabodebek Menggunakan *Similarity-Based Polynomial*

dengan ini menyatakan bahwa pada Tugas Akhir dengan judul di atas tersebut:

hereby declare that in the Final Project with the above title:

No.	Pernyataan <i>Statement</i>	(√)
1	Saya tidak menggunakan AI generatif sama sekali <i>I did not use generative AI at all</i>	
2	Saya hanya menggunakan AI generatif sebagai alat bantu untuk memperbaiki tata bahasa. AI generatif tidak digunakan untuk membuat isi Tugas Akhir. <i>I only used generative AI as a tool to improve the readability or language of the text in my Final Project. It was not used to generate a complete text of my work.</i>	√
3	Saya telah memeriksa dan/atau memperbaiki seluruh bagian dari Tugas Akhir saya yang dibantu oleh AI generatif agar sesuai dengan baku mutu penulisan karya ilmiah. <i>I have reviewed and refined all aspects of my work that generative AI assists with, ensuring it adheres to the standards of academic writing.</i>	√
4	Saya tidak menggunakan AI generatif untuk pembuatan data primer, grafik dan/atau tabel pada Tugas Akhir saya. <i>I did not use generative AI to generate primary data, figures, and/or tables in my work.</i>	√
5	Saya telah memberikan atribusi/pengakuan terhadap alat AI yang digunakan, secara rinci pada suatu bagian pada lampiran. <i>I have acknowledged the use of generative AI in any part of the work in the specific appendix page.</i>	
6	Saya memastikan tidak ada plagiarisme, termasuk hal yang berasal dari penggunaan AI generatif. <i>I have ensured that there is no plagiarism issue in the work, including any parts generated by AI.</i>	√

Surabaya, 29 Juli 2026

Mahasiswa,



Muhammad Jamil Uwais
NRP. 5009211123

Halaman ini sengaja dikosongkan

ESTIMASI *REMAINING USEFUL LIFE (RUL)* SISTEM TRAKSI LRT JABODEBEK MENGGUNAKAN *SIMILARITY BASED POLYNOMIAL*

Nama Mahasiswa / NRP : Muhammad Jamil Uwais / 5009211123
Departemen : Teknik Fisika FTIRS – ITS
Dosen Pembimbing : Prof. Dr. Ir. Katherin Indriawati, S.T., M.T., IPU
Natasya Ariola Santoso Puteri, S.T., M.T.

ABSTRAK

Sistem traksi merupakan salah satu subsistem utama pada LRT Jabodebek karena berperan langsung dalam mengubah energi listrik menjadi energi mekanik untuk menggerakkan kereta. Penurunan kondisi pada sistem traksi dapat memengaruhi keandalan operasi, sehingga diperlukan pendekatan *predictive maintenance* melalui estimasi *Remaining Useful Life (RUL)*. Penelitian ini bertujuan untuk menentukan formulasi *Health Indicator (HI)* dan merancang estimator RUL pada sistem traksi LRT Jabodebek menggunakan pendekatan *similarity-based polynomial*. Data yang digunakan merupakan data historis *run-to-failure* sistem traksi dengan ukuran 32.845 baris data, 17 kolom variabel, 133 ID pengamatan, tiga variabel operasi, dan dua belas sensor utama. Tahapan penelitian meliputi pembagian data *train* dan *test*, pengujian variabel operasi menggunakan *mean silhouette*, *clustering* kondisi operasional dengan K-Means, normalisasi sensor per *cluster*, pemilihan sensor *trendable*, pembentukan HI, pemodelan degradasi polinomial, serta estimasi RUL menggunakan *residual-based weighted KNN*. Hasil penelitian menunjukkan bahwa Op3 atau torsi motor menjadi variabel operasi terbaik untuk proses *clustering*. Sensor yang digunakan dalam pembentukan HI adalah Sensor3, Sensor2, Sensor6, Sensor5, dan Sensor8, dengan target *health condition* ideal berbentuk polinomial $HC(\tau) = 1 - \tau^3$. Evaluasi model menunjukkan bahwa nilai MAPE pada *breakpoint* 50%, 70%, dan 90% masing-masing sebesar 27,1163%, 24,7927%, dan 20,6347%. Nilai MAEfrac pada ketiga *breakpoint* tersebut sebesar 13,5278%, 7,4326%, dan 2,0686%, sedangkan RMSEfrac sebesar 16,8350%, 9,7830%, dan 2,6400%. Seluruh nilai MAPE berada di bawah target performansi 30%, sehingga estimator RUL yang dirancang memenuhi kriteria performansi penelitian. Hasil MAE, RMSE, MAEfrac, dan RMSEfrac juga menunjukkan bahwa model menghasilkan peningkatan performa dibandingkan penelitian sebelumnya, terutama ketika *trajectory* degradasi yang tersedia semakin lengkap. Selain itu, *dashboard* prediksi RUL berhasil dikembangkan sebagai media visualisasi hasil model untuk mendukung pemantauan kondisi sistem traksi.

Kata kunci: *Health Indicator*, LRT Jabodebek, *Remaining Useful Life*, *Residual-Based Weighted KNN*, *Similarity-Based Polynomial*.

Halaman ini sengaja dikosongkan

REMAINING USEFUL LIFE (RUL) ESTIMATION FOR THE LRT JABODEBEK TRACTION SYSTEM USING A SIMILARITY-BASED POLYNOMIAL

Student Name /NRP : Muhammad Jamil Uwais / 5009211123
Department : Engineering Physics FTIRS – ITS
Advisor : Prof. Dr. Ir. Katherin Indriawati, S.T., M.T., IPU
Natasya Ariola Santoso Puteri, S.T., M.T.

ABSTRACT

The traction system is one of the main subsystems of the LRT Jabodebek because it directly converts electrical energy into mechanical energy to drive the train. Degradation in the traction system may affect operational reliability; therefore, a *predictive maintenance* approach through *Remaining Useful Life* (RUL) estimation is required. This study aims to formulate a *Health Indicator* (HI) and design an RUL estimator for the LRT Jabodebek traction system using a *similarity-based polynomial* approach. The dataset used in this study is historical *run-to-failure* traction system data consisting of 32,845 rows, 17 variables, 133 observation IDs, three operational variables, and twelve main sensors. The research stages include *train-test* data splitting, operational variable evaluation using *mean silhouette*, operational condition *clustering* using K-Means, sensor normalization per *cluster*, *trendable* sensor selection, HI construction, polynomial degradation modeling, and RUL estimation using *residual-based weighted KNN*. The results show that Op3, representing motor torque, is selected as the best operational variable for *clustering*. The selected sensors for HI construction are Sensor3, Sensor2, Sensor6, Sensor5, and Sensor8, with a polynomial ideal *health condition* target defined as $HC(\tau) = 1 - \tau^3$. The model evaluation shows that the MAPE values at the 50%, 70%, and 90% *breakpoints* are 27.1163%, 24.7927%, and 20.6347%, respectively. The MAEfrac values at the same *breakpoints* are 13.5278%, 7.4326%, and 2.0686%, while the RMSEfrac values are 16.8350%, 9.7830%, and 2.6400%. All MAPE values are below the 30% performance target, indicating that the proposed RUL estimator satisfies the performance criteria of this study. The MAE, RMSE, MAEfrac, and RMSEfrac results also indicate that the proposed model improves RUL prediction performance compared with the previous study, especially when a longer degradation *trajectory* is available. In addition, an RUL prediction *dashboard* is developed as a visualization medium for the model results to support traction system condition monitoring.

Keywords: Health Indicator, LRT Jabodebek, Remaining Useful Life, Residual-Based Weighted KNN, Similarity-Based Polynomial.

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Puji Puji syukur penulis panjatkan kehadirat Allah SWT atas rahmat, karunia, dan pertolongan-Nya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul “**Estimasi Remaining Useful Life (RUL) Sistem Traksi LRT Jabodebek Menggunakan Similarity-Based Polynomial**”. Tugas Akhir ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik pada Program Studi Sarjana Teknik Fisika, Departemen Teknik Fisika, Fakultas Teknologi Industri dan Rekayasa Sistem, Institut Teknologi Sepuluh Nopember.

Penyusunan Tugas Akhir ini tidak terlepas dari bimbingan, dukungan, doa, serta bantuan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Allah SWT atas segala rahmat, kemudahan, kekuatan, dan pertolongan yang diberikan kepada penulis selama proses penyusunan Tugas Akhir ini.
2. Orang tua penulis, Bapak **Jawirman Rama Indra, S.T.** dan Ibu **Afifah Hasyim, S.E.**, atas doa, kasih sayang, dukungan moral maupun materiil, serta kepercayaan yang selalu diberikan kepada penulis. Terima kasih juga kepada adik penulis, **Muhammad Zidane Hasymi Saragih, S.H.** dan keluarga di rumah yang turut memberikan dukungan dan semangat selama proses penyusunan Tugas Akhir ini.
3. Ibu **Prof. Dr. Ir. Katherin Indriawati, S.T., M.T., IPU** selaku dosen pembimbing pertama yang telah memberikan bimbingan, arahan, ilmu, saran, dan masukan yang sangat berarti kepada penulis selama proses penyusunan Tugas Akhir ini.
4. Mba **Natasya Ariola Santoso Puteri, S.T., M.T.** selaku pembimbing dari PT INKA yang telah memberikan arahan, pengetahuan, masukan, dan bantuan kepada penulis, khususnya dalam memahami topik penelitian dan pengolahan data yang digunakan dalam Tugas Akhir ini.
5. **Pirda Hestiani** yang telah memberikan dukungan, semangat, doa, dan motivasi kepada penulis selama proses penyusunan Tugas Akhir ini.
6. **Satria, Arif, Didit, Austin, Nabil, Rey**, serta teman-teman yang telah menemani, membantu, dan memberikan semangat kepada penulis selama proses pengerjaan Tugas Akhir.

7. Ibu **Dr. Erna Septyaningrum, S.T., M.T.** dan Bapak **Ir. Harsono, M.Sc., Ph.D.** selaku dosen wali yang telah memberikan arahan, nasihat, dan bimbingan selama penulis menempuh masa perkuliahan.
8. Seluruh pihak yang pernah hadir dan memberikan warna dalam perjalanan penulis, baik dalam bentuk dukungan, pengalaman, pelajaran, maupun kebersamaan yang turut membentuk penulis hingga dapat menyelesaikan tahap ini.

Penulis menyadari bahwa Tugas Akhir ini masih jauh dari sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun untuk perbaikan di masa mendatang. Semoga Tugas Akhir ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan, khususnya dalam bidang estimasi *Remaining Useful Life* dan *predictive maintenance* pada sistem traksi kereta listrik.

Surabaya, Juli 2026
Muhammad Jamil Uwais

DAFTAR ISI

HALAMAN JUDUL.....	i
LEMBAR PENGESAHAN.....	v
APPROVAL SHEET.....	vii
PERNYATAAN ORISINALITAS	ix
PERNYATAAN KODE ETIK PENGGUNAAN AI GENERATIF.....	xi
ABSTRAK	xiii
ABSTRACT	xv
KATA PENGANTAR.....	xvii
DAFTAR ISI	xix
DAFTAR GAMBAR.....	xxi
DAFTAR TABEL	xxiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan	4
1.4 Batasan Masalah	4
1.5 Manfaat dan Relevansi terhadap SDGs	5
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI	7
2.1 Sistem Traksi Pada LRT	7
2.2 <i>Predictive Maintenance</i>	11
2.3 <i>Remaining Useful Life</i>	13
2.4 <i>Mean Silhouette</i>	15
2.5 <i>K-Means Clustering</i>	17
2.6 <i>Health Indicator</i>	18
2.7 <i>Similarity Based Polynomial</i>	21
2.8 <i>K-Nearest Neighbor (KNN)</i>	22
2.9 Metrik Evaluasi Model	25
BAB III METODE PENELITIAN	29
3.1 Pengumpulan Data Historis Sistem Traksi	30
3.2 Pembagian Data <i>Train</i> dan <i>Test</i>	34
3.3 Pengujian <i>Mean Silhouette</i>	35
3.4 <i>Clustering</i> Kondisi Operasional dengan <i>K-Means</i>	35
3.5 Normalisasi Sensor Per <i>Cluster</i>	36

3.6	Pemilihan Sensor <i>Trendable</i> untuk Pemodelan Degradasi	37
3.7	Pengujian dan Penetapan <i>Health Condition</i> Ideal	39
3.8	Pembentukan <i>Health Indicator</i>	41
3.9	<i>Polynomial Degradation Model</i>	42
3.10	Estimasi RUL dengan <i>Residual-based Weighted KNN</i>	43
3.11	Evaluasi Model Estimator RUL	46
3.12	Dashboard Prediksi RUL	48
BAB IV PEMBAHASAN		49
4.1	Hasil <i>Pre-Processing</i> Data	49
4.1.1	<i>Splitting</i> Data Menjadi Unit Train dan Test	49
4.1.2	Pengujian Variabel Operasi Untuk <i>Clustering</i>	50
4.1.3	<i>Clustering</i> Kondisi Operasional dengan <i>K-Means</i>	54
4.1.4	Normalisasi Sensor Per <i>Cluster</i>	58
4.2	Hasil Pembentukan <i>Health Indicator</i>	61
4.2.1	Pemilihan Sensor <i>Trendable Polynomial</i>	62
4.2.2	Pengujian dan Penetapan <i>Health Condition</i> Ideal	66
4.2.3	Pemodelan <i>Health Indicator</i>	70
4.3	Hasil Estimasi <i>Remaining Useful Life</i> (RUL)	74
4.3.1	<i>Polynomial Degradation Model</i>	74
4.3.2	Estimasi RUL Per <i>Breakpoint</i>	77
4.3.3	Evaluasi Model Estimator RUL	83
4.4	Hasil Visualisasi <i>Dashboard</i>	86
BAB V KESIMPULAN DAN SARAN		89
5.1	Kesimpulan	89
5.2	Saran	90
DAFTAR PUSTAKA		91
LAMPIRAN A		93
LAMPIRAN B		95
LAMPIRAN C		97
BIODATA PENULIS		99

DAFTAR GAMBAR

Gambar 2.1 Konfigurasi Satu Set Rangkaian LRT Jabodebek (Iqbal Arjiansah, 2021)	7
Gambar 2.2 <i>Current Collector Device</i> (Iqbal Arjiansah, 2021)	8
Gambar 2.3 Struktur <i>Third Rail</i> pada Jalur Kereta (Paudel et al., 2021).....	9
Gambar 2.4 Jenis Kontak antara <i>Third Rail</i> dan <i>Collector Shoe</i> (Paudel et al., 2021).....	9
Gambar 3.1 Diagram Alir Penelitian	29
Gambar 3.2 Skematik Sistem Traksi LRT	31
Gambar 4.1 Perbandingan <i>Mean Silhouette</i> Tujuh Skenario Variabel Operasi.....	51
Gambar 4.2 <i>Silhouette Plot</i> Hasil <i>Clustering</i> Berdasarkan <i>Op3</i>	53
Gambar 4.3 Hasil <i>Clustering</i> Kondisi Operasional Menggunakan Variabel <i>Op3</i>	55
Gambar 4.4 Kurva HC-3 $\tau = 1 - 1 - \tau_2$	67
Gambar 4.5 Kurva HC-4 $\tau = 1 - 1 - \tau_3$	68
Gambar 4.6 Kurva HC-1 $\tau = 1 - \tau_2$	69
Gambar 4.7 Kurva HC-2 $\tau = 1 - \tau_3$	69
Gambar 4.8 <i>Full Trajectory Health Indicator</i> pada Beberapa ID Unit Data <i>Train</i>	72
Gambar 4.9 Tahapan Pembentukan <i>Health Indicator</i> pada ID 59 Data <i>Train</i>	73
Gambar 4.10 Perbandingan Kurva HI <i>Smooth</i> dan <i>Polynomial Fit</i> pada ID 84 Data <i>Train</i> ..	75
Gambar 4.11 <i>K Nearest Neighbor Plot</i> pada Breakpoint 50% untuk ID 13 Data <i>Test</i>	79
Gambar 4.12 <i>K Nearest Neighbor Plot</i> pada Breakpoint 70% untuk ID 13 Data <i>Test</i>	79
Gambar 4.13 <i>K Nearest Neighbor Plot</i> pada Breakpoint 90% untuk ID 13 Data <i>Test</i>	80
Gambar 4.14 Tampilan Awal Dashboard Prediksi RUL	86
Gambar 4.15 Pemilihan Folder File Input pada Dashboard	87
Gambar 4.16 Pemilihan File Input Data <i>Test</i> pada Salah Satu Folder <i>Breakpoint</i>	87
Gambar 4.17 Tampilan Hasil Prediksi pada Dashboard	88

Halaman ini sengaja dikosongkan

DAFTAR TABEL

Tabel 3.1 Deskripsi Variabel Operasi dan Sensor pada Data Penelitian	32
Tabel 3. 2 Ringkasan Data <i>Run-to-Failure</i>	33
Tabel 3. 3 Skema Pengujian Health Condition Ideal	40
Tabel 4.1 Pembagian ID Unit ke dalam Data <i>Train</i> dan Data <i>Test</i>	49
Tabel 4.2 Hasil Pengujian Variabel Operasi Berdasarkan Nilai <i>Mean Silhouette</i>	51
Tabel 4.3 Ringkasan Hasil <i>K-Means Clustering</i> Kondisi Operasional.....	56
Tabel 4.4 Nilai Rata-Rata dan Simpangan Baku Sensor pada <i>Cluster 1</i>	59
Tabel 4.5 Nilai Rata-Rata dan Simpangan Baku Sensor pada <i>Cluster 2</i>	60
Tabel 4.6 Peringkat Sensor Berdasarkan Nilai <i>Trend Score</i>	63
Tabel 4.7 Hasil Pengujian Empat Bentuk <i>Health Condition</i> Ideal (Bagian 1)	66
Tabel 4.8 Hasil Pengujian Empat Bentuk <i>Health Condition</i> Ideal (Bagian 2)	67
Tabel 4.9 Ringkasan Model Pembentukan Health Indicator	71
Tabel 4.10 Ringkasan Hasil <i>Polynomial Fit</i>	76
Tabel 4.11 Ringkasan <i>RMSE Polynomial Fit</i> pada Data <i>Train</i>	76
Tabel 4.12 Ringkasan Hasil Estimasi RUL pada ID 13 Data <i>Test</i>	78
Tabel 4.13 <i>Nearest Neighbors</i> pada Breakpoint 50% untuk ID 13 Data <i>Test</i>	81
Tabel 4.14 <i>Nearest Neighbors</i> pada Breakpoint 70% untuk ID 13 Data <i>Test</i>	81
Tabel 4.15 <i>Nearest Neighbors</i> pada Breakpoint 90% untuk ID 13 Data <i>Test</i>	81
Tabel 4.16 Hasil Evaluasi Error Model Estimator RUL	83
Tabel 4.17 Perbandingan Hasil Evaluasi terhadap Penelitian Terdahulu	84

Halaman ini sengaja dikosongkan

DAFTAR SIMBOL

Simbol	Keterangan
t	Waktu atau indeks sampel dalam satu <i>trajectory</i> .
τ	Fraksi umur operasi yang telah dinormalisasi, dengan $\tau \in [0,1]$.
$s(i)$	Nilai <i>silhouette</i> untuk sampel ke- i .
$a(i)$	Rata-rata jarak sampel ke- i terhadap anggota dalam <i>cluster</i> -nya sendiri.
$b(i)$	Rata-rata jarak minimum sampel ke- i terhadap <i>cluster</i> lain terdekat.
$\bar{s}$	Nilai rata-rata <i>silhouette</i> untuk seluruh sampel pada suatu skenario.
J	Fungsi objektif pada proses <i>K-means clustering</i> .
x_i	Vektor data atau sampel ke- i yang digunakan pada proses <i>clustering</i> .
C_j	Himpunan data pada <i>cluster</i> ke- j .
μ_j	Titik pusat (<i>centroid</i>) pada <i>cluster</i> ke- j .
K	Jumlah <i>cluster</i> pada proses <i>K-means</i> .
$s_{i,j}$	Nilai asli sensor ke- j pada sampel ke- i .
$\tilde{s}_{i,j}^{(k)}$	Nilai sensor ke- j pada sampel ke- i yang telah dinormalisasi pada <i>cluster</i> ke- k .
$\mu_{k,j}$	Nilai rata-rata sensor ke- j pada <i>cluster</i> ke- k .
$\sigma_{k,j}$	Nilai simpangan baku sensor ke- j pada <i>cluster</i> ke- k .
$s_j(t)$	<i>Trajectory</i> sensor ke- j terhadap waktu.
a, b, c	Koefisien polinomial orde dua pada pendekatan <i>trend</i> sensor.
T	Waktu akhir <i>trajectory</i> yang digunakan pada perhitungan <i>trend score</i> .
TrendScore	Skor kecenderungan tren degradasi sensor.
$h(t), HI(t)$	<i>Health Indicator</i> sebagai representasi kondisi kesehatan sistem terhadap waktu.
β_0	Intersep atau bias pada model pembentukan <i>Health Indicator</i> .
β_j	Koefisien atau bobot sensor ke- j pada model pembentukan <i>Health Indicator</i> .
$x_j(t)$	Nilai sensor atau fitur ke- j pada waktu t yang digunakan dalam pembentukan <i>Health Indicator</i> .
M	Nilai <i>monotonicity</i> dari <i>Health Indicator</i> .

Simbol	Keterangan
N	Jumlah titik pengamatan pada satu <i>trajectory</i> atau jumlah data yang digunakan pada suatu perhitungan.
HI_k	Nilai <i>Health Indicator</i> pada titik pengamatan ke- k .
$c_0, c_1, \dots, c_p$	Koefisien <i>polynomial degradation model</i> orde p .
p	Orde polinomial yang digunakan pada model degradasi.
$\hat{h}_i(t)$	Hasil aproksimasi polinomial dari <i>Health Indicator</i> unit historis ke- i .
a_i, b_i, c_i	Koefisien polinomial orde dua untuk unit historis ke- i .
$d(x, y)$	Jarak Euclidean antara dua vektor x dan y .
d_m	<i>Residual distance</i> antara unit uji dan <i>neighbor</i> ke- m .
k	Jumlah <i>nearest neighbors</i> yang digunakan dalam estimasi RUL.
w_m	Bobot <i>inverse-distance</i> untuk <i>neighbor</i> ke- m .
RUL_m	Nilai RUL aktual dari <i>neighbor</i> ke- m .
C_m	Kontribusi berbobot dari <i>neighbor</i> ke- m terhadap hasil prediksi RUL.
$\overline{RUL}$	Nilai estimasi <i>Remaining Useful Life</i> untuk unit yang diprediksi.
y_i	Nilai aktual RUL pada data ke- i .
$\hat{y}_i$	Nilai prediksi RUL pada data ke- i .
$N_{life,i}$	Umur total atau panjang <i>trajectory</i> penuh dari unit ke- i .
n	Jumlah data atau jumlah unit yang digunakan dalam perhitungan metrik evaluasi.
ε	Konstanta kecil untuk menghindari pembagian dengan nol pada perhitungan bobot.
MAE	<i>Mean Absolute Error</i> .
$RMSE$	<i>Root Mean Square Error</i> .
$MAPE$	<i>Mean Absolute Percentage Error</i> .
MAEfrac	<i>Mean Absolute Error fraction</i> , yaitu rata-rata <i>error</i> absolut yang dinormalisasi terhadap umur total unit.
RMSEfrac	<i>Root Mean Square Error fraction</i> , yaitu RMSE yang dinormalisasi terhadap umur total unit.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pertumbuhan mobilitas di kawasan perkotaan menuntut tersedianya moda transportasi umum yang tidak hanya mampu mengurangi kemacetan, tetapi juga memiliki tingkat keandalan operasional yang tinggi (Davari et al., 2021). Dalam konteks tersebut, LRT Jabodebek hadir sebagai salah satu sistem transportasi perkotaan berbasis rel yang mengandalkan tenaga listrik sebagai sumber penggerak utama. Sistem elektrifikasinya menggunakan suplai 750 VDC melalui *third rail*, sehingga performa subsistem kelistrikan dan traksi menjadi sangat penting dalam menjaga kelancaran operasi kereta (Iqbal Arjiansah, 2021). Apabila terjadi gangguan pada sistem traksi, dampaknya tidak hanya dirasakan pada penurunan performa kereta, tetapi juga dapat memengaruhi ketepatan jadwal, efisiensi operasi, dan kualitas layanan transportasi secara keseluruhan (Iqbal Arjiansah, 2021; Davari et al., 2021).

Sistem traksi pada kereta listrik pada dasarnya berfungsi mengubah energi listrik menjadi energi mekanik untuk menggerakkan kereta. Pada sistem seperti ini, komponen-komponen seperti penghantar daya, *high-speed circuit breaker*, inverter traksi, dan motor traksi bekerja secara terintegrasi dalam kondisi operasi yang berubah-ubah, misalnya akibat variasi kecepatan, beban, dan pola pengereman (Iqbal Arjiansah, 2021). Oleh karena itu, sistem traksi menjadi salah satu subsistem yang rentan mengalami penurunan performa seiring waktu, sehingga pemantauan kondisi perlu dilakukan agar gejala degradasi dapat dikenali sebelum berkembang menjadi kegagalan yang lebih serius (Davari et al., 2021).

Pada sistem dengan karakter operasi yang dinamis, pendekatan perawatan berbasis jadwal belum tentu selalu mampu mencerminkan kondisi aktual komponen. Oleh karena itu, konsep *predictive maintenance* menjadi semakin relevan karena keputusan perawatan dapat dilakukan berdasarkan kondisi nyata peralatan, bukan semata-mata berdasarkan interval waktu tertentu (Davari et al., 2021; Schwartz et al., 2022). Dalam industri perkeretaapian, pendekatan data-driven *predictive maintenance* telah banyak dikaji karena memungkinkan

pemanfaatan data sensor untuk mendeteksi anomali, memantau degradasi, dan mendukung perencanaan perawatan yang lebih efisien (Davari et al., 2021). Dengan pendekatan ini, risiko *downtime* mendadak dapat ditekan, sementara penggantian komponen yang terlalu dini juga dapat dihindari (Davari et al., 2021; Schwartz et al., 2022).

Salah satu elemen penting dalam predictive maintenance adalah *Remaining Useful Life* (RUL), yaitu estimasi sisa umur pakai suatu komponen atau sistem sebelum mencapai batas kegagalan tertentu. Informasi RUL dibutuhkan karena dapat membantu operator menentukan kapan tindakan inspeksi, perawatan, atau penggantian komponen perlu dilakukan. Literatur juga menunjukkan bahwa estimasi RUL yang baik berkontribusi pada peningkatan keandalan, keselamatan, dan efisiensi biaya pemeliharaan (Schwartz et al., 2022). Namun, prediksi RUL bukan persoalan yang sederhana, karena proses degradasi di lapangan umumnya dipengaruhi oleh banyak faktor dan sering kali tidak mengikuti pola yang benar-benar linear (Zhu et al., 2025; Lin et al., 2023).

Dalam pendekatan prognosis berbasis data, sebelum RUL diestimasi biasanya terlebih dahulu dibentuk suatu *Health Indicator* (HI) sebagai representasi ringkas kondisi kesehatan sistem (Xu et al., 2023). HI berfungsi mengubah data sensor mentah yang kompleks menjadi sebuah variabel yang lebih mudah dibaca sebagai tren degradasi (Xu et al., 2023; Zhu et al., 2025). Oleh karena itu, kualitas HI sangat menentukan kualitas model prediksi RUL yang dibangun setelahnya (Xu et al., 2023). HI yang baik seharusnya mampu menggambarkan proses degradasi secara konsisten, cukup stabil terhadap gangguan lokal, dan memiliki kecenderungan yang monoton terhadap waktu degradasi (Xu et al., 2023; Zhu et al., 2025). Jika HI yang terbentuk tidak mampu merepresentasikan penurunan kondisi dengan baik, maka prediksi RUL yang dihasilkan juga cenderung kurang andal (Xu et al., 2023).

Selain tantangan dalam pembentukan HI, estimasi RUL juga menghadapi fakta bahwa pola degradasi antarunit sering kali tidak identik (Zhu et al., 2025; Lin et al., 2023). Satu unit dapat menunjukkan penurunan performa yang lebih cepat atau lebih lambat dibanding unit lain, meskipun berasal dari sistem yang sama (Lin

et al., 2023). Kondisi tersebut membuat pendekatan berbasis kemiripan atau *similarity-based* menjadi menarik, karena estimasi dilakukan dengan membandingkan pola degradasi unit yang sedang diamati dengan pola historis unit-unit sebelumnya (Lin et al., 2023; Berthold Michael R., 2020). Dalam pendekatan ini, asumsi dasarnya adalah bahwa unit yang memiliki perkembangan degradasi yang serupa hingga suatu titik cenderung memiliki kelanjutan perilaku degradasi yang juga serupa (Berthold Michael R., 2020).

Meskipun demikian, pendekatan berbasis *similarity-based* tetap membutuhkan representasi degradasi yang cukup fleksibel (Lin et al., 2023). Pada banyak kasus, penurunan kondisi komponen tidak berlangsung secara linear, melainkan menunjukkan pola non-linear yang berubah seiring waktu (Zhu et al., 2025; Lin et al., 2023). Karena itu, pendekatan *similarity based polynomial* menjadi relevan karena mampu memberikan bentuk representasi yang lebih lentur terhadap dinamika degradasi tersebut (Lin et al., 2023). Dengan memanfaatkan data historis sistem traksi LRT Jabodebek, pembentukan HI, dan pencarian *similarity-based* terhadap *trajectory* degradasi historis, pendekatan ini diharapkan dapat menghasilkan estimasi RUL yang lebih representatif terhadap karakter nyata sistem traksi (Davari et al., 2021; Lin et al., 2023).

Berdasarkan uraian tersebut, penelitian ini berfokus pada pengembangan model estimasi RUL sistem traksi LRT Jabodebek menggunakan pendekatan *similarity based polynomial*. Penelitian dilakukan dengan memanfaatkan data historis *run-to-failure*, membentuk Health Indicator dari variabel sensor yang relevan, lalu menggunakan pola degradasi historis sebagai dasar estimasi sisa umur pakai. Hasil penelitian ini diharapkan dapat mendukung penerapan *predictive maintenance* pada sistem traksi kereta listrik agar proses perawatan dapat dilakukan secara lebih tepat waktu, lebih efisien, dan lebih mendukung keandalan operasional LRT Jabodebek (Davari et al., 2021; Schwartz et al., 2022; Xu et al., 2023).

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka permasalahan yang ingin diselesaikan dalam penelitian ini antara lain

- Bagaimana formulasi atau rumusan *Health Indicator* yang merepresentasikan degradasi sistem traksi LRT Jabodebek?
- Bagaimana rancangan estimator *Remaining Useful Life* (RUL) pada sistem traksi LRT Jabodebek?

1.3 Tujuan

Berdasarkan masalah yang ingin diselesaikan, tujuan yang ingin dicapai pada penelitian ini adalah

- Menentukan formulasi atau rumusan *Health Indicator* pada sistem traksi LRT Jabodebek.
- Merancang estimator *Remaining Useful Life* (RUL) pada Sistem Traksi LRT Jabodebek.

1.4 Batasan Masalah

Batasan masalah ditetapkan agar penelitian ini lebih berfokus pada permasalahan yang ingin diselesaikan. Oleh karena itu ditetapkan batasan masalah yang terdapat beberapa poin sebagai berikut :

- Penelitian ini hanya berfokus pada sistem traksi kereta listrik pada LRT Jabodebek dan tidak membandingkannya dengan sistem traksi kereta listrik lainnya.
- Data yang digunakan merupakan data historis sistem traksi yang bersifat *run-to-failure* atau *condition monitoring* periode Agustus 2023 sampai Mei 2025.
- Data penelitian diperoleh dari PT INKA Madiun.
- Variabel yang digunakan pada penelitian ini terbatas pada tiga variabel operasi dan dua belas sensor utama yang tersedia pada dataset.
- Metode yang digunakan dalam penelitian ini dibatasi pada pembentukan *Health Indicator*, *clustering* kondisi operasional, *polynomial degradation model*, dan estimasi RUL menggunakan *residual-based weighted K-Nearest Neighbour* dalam kerangka *similarity based polynomial*.

- Dashboard pada penelitian ini berfungsi sebagai media visualisasi hasil prediksi RUL berdasarkan model yang telah dibangun.

1.5 Manfaat dan Relevansi terhadap SDGs

Penelitian ini diharapkan dapat memberikan manfaat bagi pengembangan ilmu pengetahuan, industri perkeretaapian, dan masyarakat. Bagi pengembangan ilmu pengetahuan, penelitian ini memberikan kontribusi terhadap kajian *predictive maintenance*, khususnya pada estimasi *Remaining Useful Life* (RUL) sistem traksi LRT Jabodebek berbasis data historis. Penelitian ini menyusun alur estimasi RUL melalui pembentukan *Health Indicator* (HI), segmentasi kondisi operasional, normalisasi sensor, pemilihan sensor *trendable*, pemodelan degradasi polinomial, serta estimasi RUL menggunakan pendekatan *residual-based weighted KNN*. Dengan demikian, penelitian ini dapat menjadi acuan awal dalam penerapan metode *data-driven prognostics* pada sistem traksi kereta listrik.

Bagi industri perkeretaapian, hasil penelitian ini dapat memberikan gambaran mengenai pemanfaatan data sensor dan variabel operasi untuk mendukung pemantauan kondisi sistem traksi. Estimasi RUL yang dihasilkan dapat digunakan sebagai salah satu dasar dalam merencanakan inspeksi, perawatan, atau penggantian komponen secara lebih tepat waktu. Dengan adanya informasi RUL, proses perawatan tidak hanya bergantung pada jadwal tetap, tetapi juga dapat mempertimbangkan kecenderungan degradasi aktual sistem. Hal ini diharapkan dapat membantu mengurangi risiko *downtime* mendadak, meningkatkan keandalan operasi, serta mendukung efisiensi perawatan pada sistem traksi LRT Jabodebek.

Penelitian ini juga memiliki manfaat praktis melalui pengembangan *dashboard* prediksi RUL. *Dashboard* tersebut berfungsi sebagai media visualisasi hasil model, sehingga informasi mengenai data input, *Health Indicator*, hasil pemodelan degradasi, *nearest neighbors*, dan estimasi RUL dapat ditampilkan secara lebih terstruktur. Bagi penelitian selanjutnya, hasil penelitian ini dapat menjadi acuan untuk pengembangan model estimasi RUL yang lebih kompleks, misalnya dengan menambahkan analisis ketidakpastian, validasi pada data operasi yang lebih beragam, atau integrasi dengan sistem pemantauan kondisi secara *real-time*.

Penelitian ini memiliki keterkaitan dengan beberapa tujuan *Sustainable Development Goals* (SDGs). Kontribusi utama penelitian berkaitan dengan SDG 9: *Industry, Innovation and Infrastructure*, karena penelitian ini mengembangkan pendekatan berbasis data untuk mendukung inovasi pemantauan kondisi dan perawatan pada infrastruktur transportasi. Sistem traksi LRT merupakan bagian penting dari infrastruktur transportasi perkotaan, sehingga pengembangan model estimasi RUL dapat mendukung pengelolaan aset infrastruktur yang lebih tangguh, cerdas, dan berkelanjutan.

Penelitian ini juga mendukung SDG 11: *Sustainable Cities and Communities*, karena LRT Jabodebek merupakan bagian dari sistem transportasi publik perkotaan. Keandalan sistem traksi berperan penting dalam menjaga kelancaran operasi transportasi publik, sehingga hasil penelitian ini secara tidak langsung dapat mendukung mobilitas perkotaan yang lebih berkelanjutan. Selain itu, penelitian ini memiliki relevansi terhadap SDG 12: *Responsible Consumption and Production*, karena estimasi RUL dapat membantu proses perawatan agar komponen tidak diganti terlalu dini maupun terlambat hingga menyebabkan kegagalan.

Pada tingkat lokal, hasil penelitian ini dapat mendukung peningkatan keandalan operasi LRT Jabodebek melalui pemanfaatan data historis sistem traksi sebagai dasar pemantauan kondisi dan estimasi umur sisa. Pada tingkat nasional, penelitian ini dapat menjadi referensi dalam pengembangan sistem *predictive maintenance* pada moda transportasi rel listrik di Indonesia. Pada tingkat global, penelitian ini memberikan tambahan kajian mengenai penerapan metode *data-driven prognostics* dan *similarity-based estimation* untuk mendukung infrastruktur transportasi yang inovatif, tangguh, dan berkelanjutan.

BAB II

TINJAUAN PUSTAKA DAN DASAR TEORI

2.1 Sistem Traksi Pada LRT

Sistem traksi merupakan salah satu subsistem utama pada kereta listrik yang berfungsi mengubah energi listrik menjadi energi mekanik agar kereta dapat bergerak. Pada *Light Rail Transit* (LRT), sistem ini berhubungan langsung dengan kemampuan kereta untuk mulai berjalan, berakselerasi, mempertahankan kecepatan, dan melakukan pengereman secara terkendali. Oleh karena itu, performa operasi LRT sangat dipengaruhi oleh kondisi dan keandalan sistem traksinya (Iqbal Arjiansah, 2021).

Sistem traksi pada LRT tidak hanya terdiri atas satu komponen, tetapi terintegrasi dengan konfigurasi rangkaian kereta secara keseluruhan. Pada LRT Jabodebek, satu set rangkaian terdiri atas enam kereta, yaitu MC1, M1, T1, T2, M2, dan MC2. Setiap jenis kereta memiliki fungsi yang berbeda dalam mendukung pengoperasian rangkaian. Untuk memperjelas posisi kereta penggerak dalam satu set rangkaian LRT Jabodebek, konfigurasi rangkaian ditunjukkan pada Gambar 2.1.



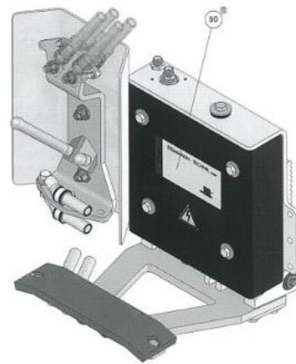
Gambar 2.1 Konfigurasi Satu Set Rangkaian LRT Jabodebek (Iqbal Arjiansah, 2021)

Berdasarkan Gambar 2.1, MC1 dan MC2 merupakan kereta yang dilengkapi ruang kemudi dan sistem kontrol. Motor traksi terpasang pada kereta MC1, M1, M2, dan MC2, sedangkan T1 dan T2 merupakan kereta tanpa motor traksi yang mendukung sistem pembantu atau *auxiliary*. Konfigurasi tersebut menunjukkan bahwa kemampuan gerak rangkaian dihasilkan oleh beberapa motor traksi yang tersebar pada kereta penggerak. Dengan demikian, kondisi sistem traksi pada setiap kereta penggerak dapat memengaruhi performa operasi rangkaian secara keseluruhan (Iqbal Arjiansah, 2021).

Agar motor traksi dapat menghasilkan gerak, sistem membutuhkan suplai energi listrik yang berasal dari sistem elektrifikasi jalur. Pada LRT Jabodebek, suplai daya diberikan melalui *third rail* dengan tegangan 750 VDC. Energi listrik

dari *third rail* kemudian diteruskan ke rangkaian kelistrikan kereta melalui perangkat pengambil arus sebelum masuk ke sistem propulsi dan sistem *auxiliary*. Penggunaan suplai DC melalui *third rail* menunjukkan bahwa kinerja sistem traksi berkaitan langsung dengan kontinuitas penyaluran daya dari infrastruktur jalur menuju kereta (Iqbal Arjiansah, 2021).

Perangkat yang digunakan untuk mengambil arus dari *third rail* disebut *Current Collector Device* (CCD). CCD berfungsi sebagai penghubung antara rel ketiga sebagai sumber energi listrik dan rangkaian kelistrikan pada kereta. Energi yang diterima melalui perangkat tersebut kemudian disalurkan menuju sistem propulsi sebagai sumber tenaga penggerak serta menuju sistem *auxiliary* untuk memenuhi kebutuhan operasional lainnya. Bentuk *Current Collector Device* yang digunakan sebagai perangkat pengambil arus ditunjukkan pada Gambar 2.2.

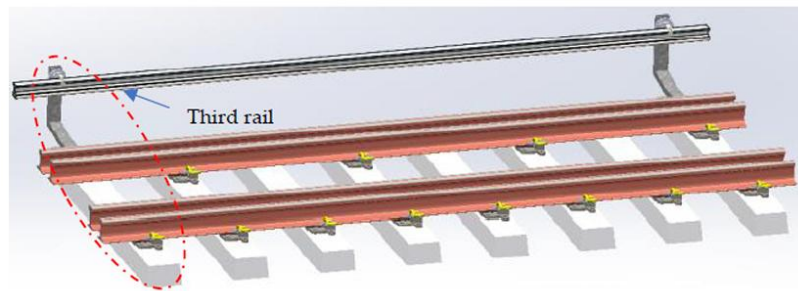


Gambar 2.2 *Current Collector Device* (Iqbal Arjiansah, 2021)

Berdasarkan Gambar 2.2, CCD memiliki bagian kontak yang berinteraksi secara langsung dengan konduktor pada *third rail*. Kontak tersebut memungkinkan energi listrik dari infrastruktur jalur masuk ke sistem kelistrikan kereta selama kereta beroperasi. Oleh karena itu, CCD menjadi titik awal penyaluran energi dari sistem elektrifikasi menuju sistem propulsi. Kualitas kontak pada perangkat ini perlu dijaga karena gangguan pada proses pengambilan arus dapat memengaruhi kontinuitas suplai daya menuju sistem traksi (Iqbal Arjiansah, 2021).

Selain kondisi perangkat pengambil arus, posisi dan struktur *third rail* juga berpengaruh terhadap kelancaran penyaluran daya. *Third rail* umumnya dipasang sejajar dengan rel utama dan berada di sisi jalur kereta. Posisi tersebut dirancang agar *collector shoe* pada kereta dapat mempertahankan kontak dengan *third rail*

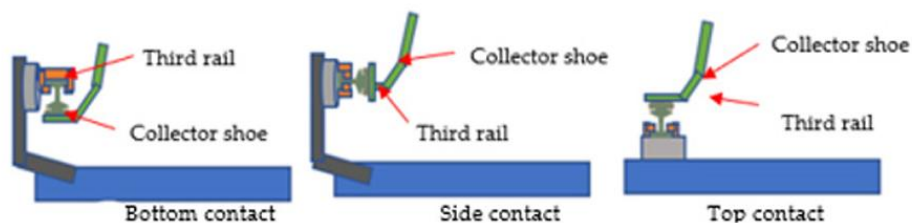
ketika kereta bergerak. Untuk memperjelas posisi *third rail* terhadap rel utama dan kereta, struktur pemasangannya ditunjukkan pada Gambar 2.3.



Gambar 2.3 Struktur *Third Rail* pada Jalur Kereta (Paudel et al., 2021)

Berdasarkan Gambar 2.3, *third rail* ditempatkan sejajar dengan rel utama sebagai konduktor penyedia daya listrik. Sementara itu, *collector shoe* berfungsi sebagai media pengambilan arus dari *third rail* menuju sistem kelistrikan kereta. Susunan tersebut memungkinkan suplai daya menuju sistem traksi berlangsung secara kontinu selama kontak antara kedua komponen tetap terjaga. Oleh karena itu, kestabilan posisi *third rail* dan kualitas kontak dengan *collector shoe* menjadi bagian penting dalam menjaga kelancaran operasi sistem propulsi (Paudel et al., 2021).

Mekanisme kontak antara *third rail* dan *collector shoe* dapat dirancang dalam beberapa konfigurasi. Perbedaan konfigurasi tersebut disesuaikan dengan kebutuhan sistem elektrifikasi, kondisi lingkungan, keselamatan operasi, dan kebutuhan perlindungan terhadap bagian konduktor. Secara umum, posisi kontak dapat dibedakan menjadi *bottom contact*, *side contact*, dan *top contact*. Ketiga jenis konfigurasi tersebut ditunjukkan pada Gambar 2.4.



Gambar 2.4 Jenis Kontak antara *Third Rail* dan *Collector Shoe* (Paudel et al., 2021)

Berdasarkan Gambar 2.4, konfigurasi *top contact* menempatkan bidang kontak pada bagian atas *third rail* sehingga lebih terbuka terhadap air, debu, dan kontaminasi. Sementara itu, konfigurasi *side contact*

dan *bottom contact* memungkinkan bagian konduktor diberi pelindung sehingga lebih terlindungi dari lingkungan dan risiko kontak langsung. Terlepas dari konfigurasi yang digunakan, gangguan pada hubungan antara *collector shoe* dan *third rail* dapat mengurangi kontinuitas suplai daya menuju sistem propulsi sehingga memengaruhi keandalan sistem traksi (Paudel et al., 2021).

Setelah diterima melalui perangkat pengambil arus, daya DC diteruskan menuju komponen proteksi dan pengaturan. Salah satu komponen proteksi tersebut adalah *high-speed circuit breaker* (HSCB), yang memutus aliran listrik secara cepat ketika terjadi gangguan untuk melindungi perangkat traksi dari kerusakan lebih lanjut. Daya kemudian diproses oleh inverter *Variable Voltage Variable Frequency* (VVVF) untuk mengatur tegangan dan frekuensi keluaran sesuai kebutuhan motor traksi. Melalui pengaturan tersebut, inverter dapat mengendalikan kecepatan dan torsi motor pada berbagai kondisi operasi kereta (Iqbal Arjiansah, 2021).

Motor traksi menerima daya yang telah diatur oleh inverter dan mengubahnya menjadi energi mekanik dalam bentuk putaran. Putaran tersebut diteruskan melalui sistem transmisi menuju roda untuk menghasilkan gaya gerak kereta. Dalam operasinya, motor menghadapi perubahan beban, kecepatan, kondisi lintasan, akselerasi, dan pengereman. Karena bekerja secara terus-menerus pada kondisi yang dinamis, motor traksi menjadi salah satu komponen utama yang perlu dipantau kondisinya (Iqbal Arjiansah, 2021; Davari et al., 2021).

Kondisi operasi yang dinamis menyebabkan nilai arus, tegangan, temperatur, kecepatan, dan parameter lainnya berubah selama perjalanan. Perubahan sinyal tersebut dapat berasal dari variasi operasi maupun penurunan kesehatan komponen. Oleh karena itu, analisis sistem traksi perlu membedakan perubahan akibat operasi normal dari perubahan yang benar-benar menunjukkan degradasi. Jika variasi operasi tidak dipertimbangkan, kondisi normal dapat keliru dibaca sebagai penurunan kesehatan atau justru menutupi pola degradasi yang sebenarnya (Davari et al., 2021).

Sistem traksi termasuk subsistem kritis karena gangguan pada inverter, motor traksi, jalur penyaluran daya, maupun kontak antara *third rail* dan *collector shoe* dapat menurunkan performa kereta dan mengganggu layanan. Oleh sebab itu, sistem traksi tidak hanya berperan sebagai penggerak, tetapi juga menjadi objek

penting dalam *condition monitoring* dan *predictive maintenance* (Davari et al., 2021; Paudel et al., 2021).

Data historis sistem traksi dapat dimanfaatkan untuk mengidentifikasi kecenderungan perubahan kondisi, membentuk *Health Indicator*, dan mengestimasi *Remaining Useful Life* (RUL). Dalam penelitian ini, sistem traksi dipilih sebagai objek studi karena menghubungkan sumber energi listrik dengan gerak mekanik melalui *third rail*, *Current Collector Device*, HSCB, inverter VVVF, motor traksi, dan sistem transmisi menuju roda. Pemantauan terhadap subsistem tersebut digunakan sebagai dasar untuk menggambarkan pola degradasi dan mendukung perencanaan perawatan sebelum sistem mencapai kondisi gagal (Schwartz et al., 2022).

2.2 Predictive Maintenance

Maintenance atau perawatan merupakan serangkaian kegiatan untuk mempertahankan fungsi, keandalan, dan ketersediaan suatu sistem. Kegiatan perawatan tidak hanya dilakukan setelah kerusakan terjadi, tetapi juga mencakup tindakan pencegahan dan pemantauan kondisi untuk mengurangi risiko gangguan operasi. Pada sistem transportasi berbasis rel, perawatan memiliki peran penting karena gangguan pada satu subsistem dapat memengaruhi kinerja dan ketersediaan rangkaian secara keseluruhan (Davari et al., 2021).

Secara umum, strategi perawatan dibedakan menjadi *corrective maintenance*, *preventive maintenance*, dan *predictive maintenance*. *Corrective maintenance* dilakukan setelah komponen mengalami kerusakan sehingga berisiko menimbulkan *downtime* tidak terencana dan biaya perbaikan yang lebih besar. Sementara itu, *preventive maintenance* dilakukan berdasarkan interval waktu, jam operasi, atau jadwal tertentu. Pendekatan tersebut dapat mengurangi risiko kegagalan mendadak, tetapi waktu pelaksanaannya belum tentu sesuai dengan kondisi aktual komponen (Davari et al., 2021; Schwartz et al., 2022).

Berbeda dengan kedua pendekatan tersebut, *predictive maintenance* menetapkan tindakan perawatan berdasarkan kondisi aktual peralatan yang diperoleh melalui proses pemantauan. Data sensor, data operasi, dan data historis digunakan untuk mendeteksi gejala awal kerusakan, mengamati perkembangan degradasi, serta memperkirakan waktu ketika komponen mendekati kondisi gagal.

Dengan demikian, perawatan dapat direncanakan berdasarkan kebutuhan sistem, bukan hanya berdasarkan jadwal tetap atau setelah kegagalan terjadi (Davari et al., 2021; Schwartz et al., 2022).

Perkembangan teknologi sensor, komputasi, dan penyimpanan data mendukung penerapan *data-driven predictive maintenance*. Pada pendekatan ini, data selama sistem beroperasi digunakan untuk membangun model yang menggambarkan kondisi peralatan. Model tersebut dapat digunakan untuk mendeteksi anomali, mengenali kecenderungan degradasi, dan mendukung pengambilan keputusan perawatan berdasarkan perubahan kondisi yang terekam pada sistem (Davari et al., 2021).

Dalam industri perkeretaapian, *predictive maintenance* penting karena kendaraan, infrastruktur, serta komponen kelistrikan dan mekanik bekerja secara terintegrasi pada kondisi operasi yang dinamis. Penurunan kondisi dapat dipengaruhi oleh penggunaan berulang, perubahan beban, pola akselerasi dan pengereman, kondisi lintasan, serta faktor lingkungan. Ketersediaan data sensor memungkinkan perubahan tersebut dianalisis untuk meningkatkan efisiensi operasi, menekan biaya perawatan, dan mendukung keselamatan sistem perkeretaapian (Davari et al., 2021).

Penerapan *predictive maintenance* dapat mengurangi risiko *unexpected failure* karena kecenderungan degradasi dapat dikenali sebelum berkembang menjadi kegagalan yang lebih serius. Inspeksi, perbaikan, atau penggantian komponen dapat direncanakan lebih awal sehingga *downtime* tidak terencana dapat ditekan dan umur penggunaan komponen dapat dioptimalkan. Pendekatan ini juga mengurangi risiko *over-maintenance*, yaitu perawatan atau penggantian komponen ketika kondisinya masih layak digunakan (Davari et al., 2021; Schwartz et al., 2022).

Meskipun memberikan berbagai manfaat, keberhasilan *predictive maintenance* sangat bergantung pada kualitas data dan metode analisis yang digunakan. Data yang tidak lengkap, gangguan pengukuran, dan keterbatasan histori kegagalan dapat memengaruhi hasil prediksi. Selain itu, perubahan sinyal sensor tidak selalu menunjukkan degradasi karena dapat pula disebabkan oleh variasi kondisi operasi. Oleh sebab itu, metode yang digunakan perlu mampu

membedakan pengaruh kondisi operasi dari perubahan yang benar-benar menunjukkan penurunan kesehatan komponen (Davari et al., 2021; Schwartz et al., 2022).

Informasi dari data sensor selanjutnya dapat diolah menjadi *Health Indicator* yang menggambarkan perkembangan kondisi sistem dari kondisi awal hingga mendekati kegagalan. Pola perubahan *Health Indicator* kemudian digunakan sebagai dasar untuk memperkirakan *Remaining Useful Life*, yaitu sisa waktu operasi sebelum sistem mencapai kondisi gagal (Schwartz et al., 2022).

Dalam penelitian ini, *predictive maintenance* digunakan sebagai landasan untuk menganalisis degradasi sistem traksi LRT Jabodebek melalui data historis. Data kondisi dan operasi diolah untuk membentuk *Health Indicator*, kemudian pola degradasinya digunakan dalam estimasi *Remaining Useful Life*. Dengan demikian, *predictive maintenance* menjadi kerangka yang menghubungkan pemantauan kondisi, pembentukan indikator kesehatan, dan estimasi umur sisa pakai untuk mendukung penentuan waktu perawatan berdasarkan kondisi aktual sistem.

2.3 *Remaining Useful Life*

Remaining Useful Life (RUL) merupakan estimasi sisa umur pakai suatu komponen atau sistem sebelum mencapai kondisi gagal atau tidak lagi memenuhi kebutuhan operasi. Informasi RUL digunakan untuk menentukan waktu inspeksi, perawatan, maupun penggantian komponen berdasarkan kondisi aktual sistem. Dengan demikian, keputusan perawatan tidak hanya mengacu pada jadwal tetap, tetapi juga mempertimbangkan perkembangan degradasi selama sistem beroperasi (Schwartz et al., 2022).

Dalam kerangka *Prognostics and Health Management* (PHM), estimasi RUL berperan dalam mengurangi risiko kegagalan yang tidak terduga. Apabila sisa umur pakai dapat diperkirakan lebih awal, tindakan perawatan dapat direncanakan sebelum sistem mencapai kondisi gagal sehingga *downtime* yang tidak direncanakan dapat ditekan. Estimasi RUL juga membantu mengoptimalkan umur penggunaan komponen karena penggantian tidak dilakukan terlalu cepat maupun terlambat (Schwartz et al., 2022; Zhu et al., 2025).

Secara umum, estimasi RUL dapat dilakukan menggunakan pendekatan *physics-based*, *data-driven*, dan *hybrid*. Pendekatan *physics-based* membangun

model berdasarkan mekanisme fisik kerusakan atau degradasi. Pendekatan ini memiliki interpretasi yang kuat, tetapi membutuhkan pemahaman rinci mengenai mekanisme kegagalan, karakteristik komponen, dan kondisi operasi. Kebutuhan tersebut membuat pendekatan ini sulit digunakan ketika mekanisme degradasi terlalu kompleks atau tidak diketahui secara lengkap (Zhu et al., 2025).

Pendekatan *data-driven* membangun hubungan antara kondisi sistem dan umur sisa berdasarkan pola pada data sensor serta data historis. Pendekatan ini lebih fleksibel ketika model fisik degradasi tidak tersedia, tetapi sistem memiliki data pemantauan yang memadai. Sementara itu, pendekatan *hybrid* menggabungkan informasi model fisik dan pola yang dipelajari dari data untuk menghasilkan prediksi yang lebih representatif (Lin et al., 2023; Zhu et al., 2025).

Pada pendekatan *data-driven*, estimasi RUL dapat dilakukan secara langsung dari data sensor atau melalui pembentukan *Health Indicator* terlebih dahulu. *Health Indicator* menggabungkan informasi dari beberapa sensor menjadi satu indikator yang menggambarkan perkembangan kondisi kesehatan sistem. Pola perubahan indikator tersebut membentuk *trajectory* degradasi yang kemudian digunakan sebagai dasar untuk mengestimasi umur sisa pakai. Oleh karena itu, kualitas *Health Indicator* sangat memengaruhi kemampuan model dalam menggambarkan degradasi dan menghasilkan estimasi RUL yang konsisten (Xu et al., 2023; Zhu et al., 2025).

Proses estimasi RUL memiliki tantangan karena pola degradasi pada kondisi nyata tidak selalu linear dan dapat berbeda antarunit meskipun berasal dari jenis sistem yang sama. Perbedaan tersebut dapat dipengaruhi oleh riwayat operasi, tingkat beban, kondisi lingkungan, dan karakteristik komponen. Selain itu, waktu awal terjadinya degradasi tidak selalu diketahui secara pasti. Kondisi ini menyebabkan perkembangan kesehatan seluruh unit tidak selalu dapat direpresentasikan menggunakan satu model degradasi global (Lin et al., 2023; Zhu et al., 2025).

Variasi pola degradasi tersebut mendorong penggunaan pendekatan *similarity-based*. Pada pendekatan ini, *trajectory* degradasi unit yang diamati dibandingkan dengan *trajectory* historis dari unit yang telah beroperasi sampai gagal. Unit historis dengan pola paling mirip kemudian digunakan sebagai referensi

untuk memperkirakan kelanjutan degradasi dan sisa umur pakai. Pendekatan ini didasarkan pada asumsi bahwa unit dengan perkembangan kondisi yang serupa hingga suatu titik pengamatan cenderung memiliki pola lanjutan dan umur sisa yang serupa (Schwartz et al., 2022; Lin et al., 2023; Berthold Michael R., 2020).

Pendekatan *similarity-based* dapat memanfaatkan data *run-to-failure* tanpa memerlukan model fisik degradasi secara lengkap. Selain itu, variasi pola antarunit dapat diakomodasi melalui perbandingan dengan beberapa referensi historis. Namun, hasil estimasinya tetap bergantung pada kualitas representasi kondisi, ukuran kemiripan, dan ketersediaan data historis yang relevan (Lin et al., 2023).

Dalam penelitian ini, estimasi RUL pada sistem traksi LRT Jabodebek dilakukan menggunakan pendekatan *data-driven* berbasis *Health Indicator* dan kemiripan pola degradasi historis. Data sensor dan kondisi operasi diolah untuk membentuk *Health Indicator* yang merepresentasikan perubahan kesehatan sistem. *Trajectory* tersebut kemudian dimodelkan dan dibandingkan dengan data historis unit *run-to-failure* untuk memperkirakan sisa umur pakai. Pendekatan ini dipilih karena pola degradasi antarunit tidak selalu identik, sehingga perbandingan pola dinilai lebih sesuai dibandingkan penggunaan satu model global untuk seluruh unit (Schwartz et al., 2022; Xu et al., 2023; Lin et al., 2023).

2.4 Mean Silhouette

Dalam proses *clustering*, kualitas kelompok yang terbentuk perlu dievaluasi untuk memastikan bahwa setiap data memiliki kedekatan yang baik dengan *cluster*-nya dan terpisah dari *cluster* lain. Pada penelitian ini, evaluasi tersebut dilakukan menggunakan *mean silhouette*, yaitu ukuran yang membandingkan tingkat kekompakan data dalam satu *cluster* dengan tingkat keterpisahannya terhadap *cluster* terdekat. Nilai ini dapat digunakan untuk menilai kualitas hasil *clustering* sekaligus memilih skenario pengelompokan yang paling sesuai (Shutaywi & Kachouie, 2021; Gere, 2023).

Nilai *silhouette* setiap data dihitung berdasarkan dua komponen, yaitu *cohesion* dan *separation*. *Cohesion* menyatakan rata-rata jarak suatu data terhadap data lain dalam *cluster* yang sama, sedangkan *separation* menyatakan rata-rata jaraknya terhadap *cluster* lain yang paling dekat. Berdasarkan kedua komponen tersebut, nilai *silhouette* untuk data ke-*i* dirumuskan pada persamaan (2.1).

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (2.1)$$

dengan $a(i)$ merupakan rata-rata jarak data ke- i terhadap data lain dalam *cluster* yang sama, sedangkan $b(i)$ merupakan rata-rata jarak data tersebut terhadap *cluster* terdekat. Nilai $s(i)$ berada pada rentang -1 sampai 1 . Nilai yang mendekati 1 menunjukkan bahwa data lebih dekat dengan *cluster*-nya sendiri dan terpisah dengan baik dari *cluster* lain. Nilai yang mendekati 0 menunjukkan bahwa data berada di sekitar batas antar-*cluster*, sedangkan nilai negatif menunjukkan bahwa data lebih dekat dengan *cluster* lain sehingga penempatannya berpotensi kurang tepat (Shutaywi & Kachouie, 2021).

Setelah nilai *silhouette* seluruh data diperoleh, kualitas hasil *clustering* secara keseluruhan dihitung menggunakan nilai rata-ratanya. Nilai tersebut disebut *mean silhouette* dan dirumuskan pada persamaan (2.2).

$$\bar{s} = \frac{1}{n} \sum_{i=1}^n s(i) \quad (2.2)$$

dengan n merupakan jumlah data yang dievaluasi. Nilai *mean silhouette* yang lebih tinggi menunjukkan bahwa data dalam satu *cluster* cenderung lebih kompak dan memiliki pemisahan yang lebih jelas terhadap *cluster* lain. Sebaliknya, nilai yang rendah menunjukkan bahwa kelompok yang terbentuk masih saling berdekatan atau tumpang tindih sehingga kualitas segmentasinya kurang baik (Shutaywi & Kachouie, 2021; Gere, 2023).

Pada penelitian ini, *mean silhouette* digunakan untuk membandingkan beberapa skenario variabel operasi dalam proses *clustering* kondisi operasional sistem traksi. Tahap ini diperlukan karena perubahan sinyal sensor tidak hanya dipengaruhi oleh degradasi, tetapi juga oleh variasi kondisi operasi. Skenario dengan nilai *mean silhouette* tertinggi dipilih karena menghasilkan kelompok data yang paling kompak dan memiliki pemisahan paling jelas.

Dengan demikian, penggunaan *mean silhouette* membantu menentukan variabel operasi yang mampu membentuk *cluster* kondisi operasional secara lebih representatif. Hasil evaluasi tersebut selanjutnya digunakan sebagai dasar proses *clustering* sebelum normalisasi sensor dan pembentukan *Health Indicator*.

2.5 *K-Means Clustering*

K-Means Clustering merupakan metode *unsupervised learning* yang digunakan untuk mengelompokkan data ke dalam sejumlah *cluster* berdasarkan tingkat kedekatannya. Metode ini banyak digunakan karena konsepnya relatif sederhana, efisien secara komputasi, dan dapat diterapkan pada berbagai jenis data numerik. Dalam penelitian ini, *K-Means* digunakan untuk mengelompokkan kondisi operasional sistem traksi berdasarkan variabel operasi terpilih sebelum normalisasi sensor dan pembentukan *Health Indicator* dilakukan (Ahmed et al., 2020; Wani, 2024).

Tujuan utama *K-Means* adalah membagi data ke dalam K *cluster* sehingga data dalam kelompok yang sama memiliki karakteristik lebih mirip dibandingkan data pada kelompok lain. Kedekatan tersebut umumnya dihitung menggunakan jarak Euclidean. Setiap data kemudian ditempatkan pada *cluster* yang memiliki pusat kelompok atau *centroid* paling dekat (Ahmed et al., 2020; Wani, 2024).

Proses *K-Means* dimulai dengan menentukan jumlah *cluster* sebanyak K dan memilih *centroid* awal untuk setiap kelompok. Jarak setiap data terhadap seluruh *centroid* kemudian dihitung, lalu data ditempatkan pada *cluster* dengan jarak terdekat. Setelah seluruh data memperoleh label *cluster*, posisi *centroid* diperbarui berdasarkan nilai rata-rata data pada masing-masing kelompok. Tahapan penempatan data dan pembaruan *centroid* dilakukan secara berulang hingga perubahan posisi *centroid* tidak lagi signifikan atau kondisi konvergen tercapai (Ahmed et al., 2020; Wani, 2024).

Secara matematis, proses tersebut bertujuan meminimalkan jumlah kuadrat jarak antara setiap data dan *centroid cluster*-nya. Fungsi objektif *K-Means* dituliskan pada persamaan (2.3).

$$J = \sum_{j=1}^K \sum_{x_i \in C_j} \|x_i - \mu_j\|^2 \quad (2.3)$$

Pada persamaan (2.3), K merupakan jumlah *cluster*, C_j merupakan kumpulan data pada *cluster* ke- j , x_i merupakan data ke- i , dan μ_j merupakan *centroid cluster* ke- j . Nilai fungsi objektif yang semakin kecil menunjukkan bahwa data dalam setiap *cluster* semakin dekat terhadap pusat kelompoknya. Dengan demikian, proses iteratif *K-Means* bertujuan menghasilkan kelompok dengan

kedekatan internal yang tinggi dan pemisahan antar-*cluster* yang lebih jelas (Ahmed et al., 2020).

K-Means memiliki kelebihan berupa proses komputasi yang relatif cepat dan hasil pengelompokan yang mudah dipahami karena setiap *cluster* direpresentasikan oleh sebuah *centroid*. Namun, hasilnya dapat dipengaruhi oleh pemilihan *centroid* awal, sehingga inisialisasi yang berbeda dapat menghasilkan pembagian kelompok yang berbeda. Metode ini juga sensitif terhadap *outlier* dan memerlukan jumlah *cluster* ditentukan sebelum proses pengelompokan dilakukan (Ahmed et al., 2020; Wani, 2024).

Pada penelitian ini, *K-Means* digunakan untuk memisahkan data berdasarkan kondisi operasional karena perubahan sinyal sistem traksi tidak selalu berasal dari degradasi komponen, tetapi dapat pula disebabkan oleh perbedaan mode operasi. Pengelompokan tersebut diperlukan agar normalisasi sensor dan analisis tren degradasi dilakukan pada kelompok operasi yang lebih homogen. Variabel operasi yang digunakan dalam proses *clustering* terlebih dahulu dievaluasi menggunakan *mean silhouette* untuk memperoleh skenario dengan kekompakan dan pemisahan *cluster* terbaik (Davari et al., 2021; Ahmed et al., 2020; Shutaywi & Kachouie, 2021).

Dengan demikian, *K-Means* pada penelitian ini tidak digunakan untuk mendeteksi degradasi secara langsung, tetapi sebagai tahap awal untuk memisahkan pengaruh kondisi operasional dari perubahan kondisi kesehatan sistem. Hasil pengelompokan tersebut selanjutnya digunakan sebagai dasar normalisasi sensor per *cluster* agar informasi yang digunakan dalam pembentukan *Health Indicator* lebih merepresentasikan degradasi komponen daripada variasi akibat perubahan mode operasi.

2.6 *Health Indicator*

Dalam pendekatan prognosis berbasis data, salah satu tahap penting sebelum melakukan estimasi *Remaining Useful Life* (RUL) adalah membentuk *Health Indicator* (HI). HI merupakan indikator yang menyederhanakan informasi dari sejumlah data sensor menjadi representasi kondisi kesehatan sistem yang lebih mudah diamati. Melalui HI, perubahan kondisi dari waktu ke waktu dapat

ditampilkan sebagai satu tren degradasi yang menjadi penghubung antara data kondisi dan proses estimasi RUL (Xu et al., 2023; Zhu et al., 2025).

HI dibangun untuk menggambarkan perkembangan kondisi sistem dari keadaan sehat hingga mendekati kegagalan. Pada kondisi awal, nilai HI merepresentasikan kondisi kesehatan awal sistem, kemudian berubah mengikuti penurunan performa yang terjadi. Oleh karena itu, kualitas HI sangat menentukan kemampuan model prognosis dalam membaca pola degradasi. HI yang tidak mampu menggambarkan penurunan kondisi dengan baik dapat menghasilkan estimasi RUL yang kurang andal (Xu et al., 2023).

HI dapat dibentuk menggunakan satu fitur maupun gabungan beberapa fitur. Pada sistem sederhana, satu parameter fisik dapat digunakan secara langsung sebagai indikator kesehatan. Namun, pada sistem yang lebih kompleks, satu sensor sering kali belum cukup untuk menggambarkan kondisi sistem secara menyeluruh. Oleh sebab itu, digunakan *virtual Health Indicator*, yaitu indikator kesehatan yang dibentuk dari kombinasi beberapa variabel sensor atau fitur hasil ekstraksi sehingga informasi degradasi dapat direpresentasikan dalam satu nilai yang lebih ringkas (Zhu et al., 2025).

Secara umum, pembentukan HI dari beberapa variabel sensor dapat dinyatakan pada persamaan (2.4).

$$HI(t) = \beta_0 + \sum_{j=1}^m \beta_j x_j(t) \quad (2.4)$$

Pada persamaan (2.4), $HI(t)$ merupakan nilai indikator kesehatan pada waktu t , β_0 merupakan konstanta, β_j merupakan koefisien kontribusi sensor ke- j , $x_j(t)$ merupakan nilai sensor ke- j , dan m merupakan jumlah sensor yang digunakan. Persamaan tersebut menunjukkan bahwa setiap sensor dapat memberikan kontribusi yang berbeda terhadap HI. Oleh karena itu, sensor perlu dipilih berdasarkan kemampuannya dalam merepresentasikan perubahan kondisi sistem karena tidak seluruh sensor memiliki keterkaitan yang sama terhadap degradasi (Xu et al., 2023; Zhu et al., 2025).

HI yang baik perlu memiliki kecenderungan perubahan yang konsisten, cukup stabil, dan mampu merepresentasikan kondisi kesehatan sistem. Salah satu karakteristik utamanya adalah *monotonicity*, yaitu kemampuan HI untuk bergerak

secara relatif konsisten mengikuti arah degradasi. Selain itu, HI tidak seharusnya terlalu dipengaruhi oleh fluktuasi lokal, perubahan mode operasi, atau gangguan acak yang tidak berkaitan langsung dengan penurunan kondisi (Xu et al., 2023; Zhu et al., 2025).

Pada penelitian ini, HI dirancang memiliki arah penurunan seiring dengan berlangsungnya degradasi. Oleh karena itu, nilai *monotonicity* dihitung berdasarkan proporsi perubahan HI yang bergerak ke arah menurun, sebagaimana ditunjukkan pada persamaan (2.5).

$$M = \frac{1}{N-1} \sum_{k=1}^{N-1} \mathbf{1}(HI_{k+1} - HI_k \leq 0) \quad (2.5)$$

Pada persamaan (2.5), M merupakan nilai *monotonicity*, N merupakan jumlah titik pengamatan, dan $\mathbf{1}(\cdot)$ merupakan fungsi indikator. Nilai *monotonicity* yang mendekati 1 menunjukkan bahwa HI semakin konsisten bergerak mengikuti arah degradasi. Sebaliknya, nilai yang rendah menunjukkan bahwa HI masih banyak mengalami fluktuasi sehingga kurang sesuai digunakan sebagai dasar estimasi RUL.

Dalam penerapannya, HI dapat dibentuk melalui metode statistik, reduksi dimensi, pembobotan fitur, maupun model pembelajaran tertentu. Pada penelitian ini, HI dibentuk dari sensor terpilih dengan mempertimbangkan arah dan kecenderungan degradasinya terhadap waktu operasi. Nilai HI yang diperoleh kemudian melalui proses *smoothing* untuk mengurangi fluktuasi lokal dan menghasilkan *trajectory* degradasi yang lebih stabil (Xu et al., 2023; Zhu et al., 2025).

Dalam penelitian ini, HI digunakan untuk merepresentasikan kondisi sistem traksi LRT Jabodebek sebelum dilakukan pemodelan degradasi dan estimasi RUL. Pembentukan HI diperlukan karena data sensor sistem traksi terdiri atas beberapa variabel dan dipengaruhi oleh kondisi operasi yang dinamis. Informasi tersebut diringkas menjadi satu indikator kesehatan yang selanjutnya digunakan sebagai dasar pemodelan *similarity-based polynomial* dan estimasi *Remaining Useful Life* (Schwartz et al., 2022; Xu et al., 2023; Zhu et al., 2025).

2.7 *Similarity Based Polynomial*

Pada estimasi *Remaining Useful Life* (RUL) berbasis data, pola degradasi antarunit tidak selalu identik. Unit yang berasal dari sistem yang sama dapat memiliki laju penurunan kondisi berbeda akibat variasi riwayat operasi, beban, dan lingkungan. Kondisi tersebut menyebabkan satu model degradasi global belum tentu mampu merepresentasikan seluruh unit secara baik. Pendekatan *similarity-based* digunakan untuk mengatasi permasalahan tersebut dengan membandingkan pola degradasi unit yang diamati terhadap data historis unit lain (Schwartz et al., 2022; Lin et al., 2023; Zhu et al., 2025).

Pendekatan *similarity-based* didasarkan pada asumsi bahwa unit yang memiliki pola degradasi serupa hingga suatu titik pengamatan cenderung mempunyai perkembangan lanjutan dan sisa umur pakai yang serupa. Data historis *run-to-failure* digunakan sebagai referensi untuk memperkirakan RUL unit yang sedang diamati. Pendekatan ini tidak memerlukan model fisik degradasi secara lengkap sehingga dapat diterapkan ketika mekanisme dan bentuk penurunan kondisi sulit dijelaskan secara eksplisit (Schwartz et al., 2022; Lin et al., 2023; Berthold Michael R., 2020).

Pembandingan antarunit memerlukan representasi degradasi yang ringkas, tetapi tetap mempertahankan karakteristik utama perubahan kondisi. Pada penelitian ini, representasi tersebut diperoleh dari *Health Indicator* (HI) yang telah merangkum informasi beberapa sensor menjadi satu *trajectory* kesehatan. Agar bentuk *trajectory* dapat dinyatakan secara matematis, HI masing-masing unit dimodelkan menggunakan fungsi polinomial. Fungsi tersebut dipilih karena mampu merepresentasikan degradasi yang tidak selalu berbentuk linear (Xu et al., 2023; Lin et al., 2023; Zhu et al., 2025). Model polinomial HI terhadap waktu dinyatakan pada persamaan (2.6).

$$HI(t) = c_0 + c_1t + c_2t^2 + \dots + c_pt^p \quad (2.6)$$

Pada persamaan (2.6), $HI(t)$ merupakan nilai indikator kesehatan pada waktu t , $c_0, c_1, \dots, c_p$ merupakan koefisien polinomial, dan p merupakan orde model. Melalui pemodelan tersebut, *trajectory* setiap unit dapat direpresentasikan oleh sekumpulan koefisien yang menggambarkan bentuk dan laju degradasinya. Dengan demikian, pembandingan antarunit dapat dilakukan menggunakan

representasi yang lebih ringkas daripada membandingkan seluruh titik pengamatan secara langsung.

Setelah model polinomial diperoleh, tingkat kemiripan antara unit yang diamati dan unit referensi ditentukan berdasarkan jarak antara representasi degradasinya. Salah satu ukuran yang dapat digunakan adalah *Euclidean distance*, sebagaimana ditunjukkan pada Persamaan (2.7).

$$d(x, y) = \sqrt{\sum_{k=1}^n (x_k - y_k)^2} \quad (2.7)$$

Pada persamaan (2.7), x_k dan y_k merupakan elemen representasi degradasi dari dua unit yang dibandingkan, sedangkan n merupakan jumlah elemen yang digunakan. Nilai jarak yang semakin kecil menunjukkan tingkat kemiripan yang semakin tinggi. Oleh karena itu, unit historis dengan jarak terkecil dianggap memiliki pola degradasi paling relevan untuk digunakan sebagai referensi estimasi RUL (Lin et al., 2023).

Pendekatan *similarity-based polynomial* menggabungkan fleksibilitas model polinomial dalam merepresentasikan degradasi nonlinier dengan kemampuan pendekatan berbasis kemiripan dalam memanfaatkan data historis *run-to-failure*. Estimasi tidak hanya bergantung pada satu hubungan umum antara waktu dan degradasi, tetapi menggunakan pola yang pernah terjadi pada unit-unit referensi untuk memperkirakan kelanjutan kondisi unit yang diamati (Schwartz et al., 2022; Lin et al., 2023).

Dalam penelitian ini, HI setiap unit direpresentasikan menggunakan model polinomial, kemudian pola degradasi unit *test* dibandingkan dengan unit-unit pada data *train*. Konsep tersebut menjadi dasar pemodelan degradasi dan estimasi RUL yang implementasinya dijelaskan lebih lanjut pada Bab III melalui tahap *polynomial degradation model* dan estimasi RUL menggunakan *residual-based weighted KNN*.

2.8 *K-Nearest Neighbor (KNN)*

K-Nearest Neighbor (KNN) merupakan metode *supervised learning* yang digunakan untuk klasifikasi maupun prediksi berdasarkan kedekatan antar data. Prinsip dasarnya adalah bahwa data baru diprediksi menggunakan sejumlah data

referensi yang memiliki karakteristik paling mirip. KNN termasuk metode berbasis jarak yang memanfaatkan data referensi secara langsung pada saat prediksi tanpa membangun model parametrik yang kompleks terlebih dahulu (Halder et al., 2024).

Pada proses KNN, setiap data *test* dibandingkan dengan data referensi menggunakan suatu metrik jarak. Sejumlah k data yang memiliki jarak paling kecil dipilih sebagai *nearest neighbors*. Pada kasus regresi, nilai target data *test* dapat diperkirakan menggunakan rata-rata nilai target dari *neighbors* tersebut, sebagaimana ditunjukkan pada persamaan (2.8).

$$\hat{y} = \frac{1}{k} \sum_{i=1}^k y_i \quad (2.8)$$

Pada persamaan (2.8), $\hat{y}$ merupakan nilai prediksi, k merupakan jumlah *neighbors*, dan y_i merupakan nilai target dari *neighbor* ke- i . Prediksi yang dihasilkan akan semakin representatif apabila data yang dipilih sebagai *neighbors* memiliki karakteristik yang benar-benar mirip dengan data *test* (Seyghaly et al., 2024).

Penggunaan rata-rata biasa menganggap seluruh *neighbors* memiliki kontribusi yang sama. Padahal, setiap *neighbor* dapat memiliki tingkat kedekatan yang berbeda terhadap data *test*. Oleh karena itu, digunakan *weighted KNN* yang memberikan bobot lebih besar kepada *neighbor* dengan jarak lebih kecil dan bobot lebih kecil kepada *neighbor* yang lebih jauh. Salah satu metode pembobotan yang umum digunakan adalah *inverse distance weighting*, sebagaimana ditunjukkan pada persamaan (2.9).

$$w_i = \frac{\frac{1}{d_i}}{\sum_{j=1}^k \frac{1}{d_j}} \quad (2.9)$$

Pada persamaan (2.9), w_i merupakan bobot *neighbor* ke- i , sedangkan d_i merupakan jarak *neighbor* tersebut terhadap data *test*. Nilai bobot dinormalisasi agar jumlah seluruh bobot sama dengan satu. Dengan demikian, *neighbor* yang paling dekat akan memberikan kontribusi lebih besar terhadap hasil prediksi dibandingkan *neighbor* yang jaraknya lebih jauh (Seyghaly et al., 2024). Setelah bobot setiap *neighbor* diperoleh, nilai prediksi pada *weighted KNN* dihitung melalui

penjumlahan berbobot nilai target seluruh *neighbors*, sebagaimana dinyatakan pada persamaan (2.10).

$$\hat{y} = \sum_{i=1}^k w_i y_i \quad (2.10)$$

Persamaan (2.10) menunjukkan bahwa prediksi akhir tidak lagi diperoleh dari rata-rata sederhana, tetapi dari kontribusi setiap *neighbor* berdasarkan tingkat kedekatannya. Pendekatan ini memungkinkan data referensi yang paling mirip memberikan pengaruh lebih besar terhadap hasil prediksi.

KNN memiliki kelebihan berupa konsep yang sederhana, mudah dipahami, dan tidak membutuhkan proses pelatihan model yang rumit karena data referensi digunakan secara langsung saat prediksi. Metode ini juga cukup fleksibel untuk diterapkan pada kasus klasifikasi maupun regresi. Namun, kinerjanya dipengaruhi oleh nilai k , metrik jarak, skala variabel, serta kualitas data referensi. Nilai k yang terlalu kecil dapat membuat prediksi sensitif terhadap *noise*, sedangkan nilai k yang terlalu besar dapat menghasilkan prediksi yang terlalu umum dan kurang mampu merepresentasikan pola lokal data (Halder et al., 2024; Seyghaly et al., 2024).

Dalam estimasi *Remaining Useful Life* (RUL), KNN dapat digunakan pada pendekatan *similarity-based* dengan membandingkan unit yang diamati terhadap unit-unit historis yang memiliki pola degradasi serupa. Setelah representasi degradasi diperoleh, unit referensi diurutkan berdasarkan tingkat kedekatannya terhadap unit *test*. Informasi umur sisa dari beberapa unit terdekat kemudian digunakan untuk menghasilkan estimasi RUL (Schwartz et al., 2022; Lin et al., 2023; Berthold Michael R., 2020).

Pada penelitian ini, KNN diterapkan dalam bentuk *residual-based weighted KNN*. Kedekatan antarunit tidak ditentukan hanya berdasarkan jarak parameter model, tetapi menggunakan *residual distance* antara HI parsial unit *test* dan model degradasi unit *train*. Nilai RUL selanjutnya dihitung menggunakan *inverse distance weighting* agar unit historis dengan pola paling mirip memberikan kontribusi lebih besar terhadap hasil estimasi. Implementasi metode tersebut dijelaskan lebih lanjut pada Bab III.

2.9 Metrik Evaluasi Model

Setelah model prediksi dibangun, langkah berikutnya adalah mengevaluasi seberapa baik model tersebut dalam memperkirakan nilai *Remaining Useful Life* (RUL). Pada penelitian prediksi RUL, evaluasi model umumnya dilakukan dengan membandingkan nilai prediksi terhadap nilai aktual, kemudian menghitung besar kesalahannya menggunakan metrik tertentu. Beberapa metrik yang umum digunakan pada studi RUL adalah *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE) (Berghout & Benbouzid, 2022; De Pater & Mitici, 2022).

Pada penelitian ini, evaluasi model difokuskan pada MAE, RMSE, dan MAPE karena ketiga metrik tersebut dapat memberikan gambaran kesalahan prediksi dari sudut pandang yang saling melengkapi. Secara matematis, *Mean Absolute Error* (MAE) dapat dituliskan sebagai berikut (Berghout & Benbouzid, 2022).

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (2.11)$$

Persamaan tersebut menunjukkan bahwa MAE dihitung sebagai rata-rata selisih absolut antara nilai prediksi dan nilai aktual pada seluruh data. Karena menggunakan nilai absolut, metrik ini hanya memperhatikan besar kesalahan tanpa membedakan arah kesalahan prediksi. Oleh karena itu, semakin kecil nilai MAE, maka secara umum model menghasilkan prediksi yang semakin dekat dengan nilai aktual. Sementara itu, *Root Mean Square Error* (RMSE) dapat dituliskan sebagai berikut (Berghout & Benbouzid, 2022).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (2.12)$$

Persamaan tersebut menunjukkan bahwa RMSE dihitung dari akar rata-rata kuadrat kesalahan prediksi. Karena galat dikuadratkan terlebih dahulu, metrik ini memberikan penalti yang lebih besar pada *error* yang besar. Oleh sebab itu, RMSE sering digunakan ketika peneliti ingin melihat apakah model memiliki kesalahan ekstrem yang cukup signifikan. Semakin kecil nilai RMSE, maka performa model dapat dianggap semakin baik (Berghout & Benbouzid, 2022). Adapun *Mean*

Absolute Percentage Error (MAPE) dapat dituliskan sebagai berikut (Berghout & Benbouzid, 2022; De Pater & Mitici, 2022).

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (2.13)$$

Persamaan tersebut menunjukkan bahwa MAPE merepresentasikan rata-rata persentase kesalahan prediksi terhadap nilai aktual. Dengan menggunakan bentuk persentase, MAPE memudahkan interpretasi tingkat akurasi model karena hasilnya dapat langsung dibaca sebagai besar kesalahan relatif terhadap nilai sebenarnya (Berghout & Benbouzid, 2022; De Pater & Mitici, 2022). Semakin kecil nilai MAPE, maka secara umum akurasi model dapat dianggap semakin baik.

Selain MAE, RMSE, dan MAPE, penelitian ini juga menggunakan MAEfrac dan RMSEfrac sebagai metrik evaluasi tambahan. Kedua metrik ini digunakan untuk menunjukkan besar *error* prediksi yang dinormalisasi terhadap umur total unit, bukan terhadap nilai RUL aktual pada *breakpoint* tertentu. Dengan demikian, MAEfrac dan RMSEfrac dapat digunakan sebagai metrik pembandingan terhadap penelitian terdahulu yang menyajikan *error* dalam bentuk fraksi terhadap umur total unit. Secara matematis, MAEfrac dapat dituliskan sebagai berikut.

$$MAE_{frac} = \frac{100\%}{n} \sum_{i=1}^n \left(\frac{|\hat{y}_i - y_i|}{N_{life,i}} \right) \quad (2.14)$$

Pada persamaan tersebut, $N_{life,i}$ merupakan panjang umur total atau jumlah data penuh dari unit ke- i . Berbeda dengan MAPE yang membagi *error* terhadap RUL aktual, MAEfrac membagi *error* terhadap umur total unit. Oleh karena itu, MAEfrac menunjukkan proporsi rata-rata *error* absolut terhadap total umur unit. Semakin kecil nilai MAEfrac, maka semakin kecil pula *error* prediksi jika dibandingkan terhadap keseluruhan umur operasi unit. Sementara itu, RMSEfrac dapat dituliskan sebagai berikut.

$$RMSE_{frac} = 100\% \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{\hat{y}_i - y_i}{N_{life,i}} \right)^2} \quad (2.15)$$

Persamaan tersebut menunjukkan bahwa RMSEfrac merupakan bentuk RMSE yang dinormalisasi terhadap umur total unit. Karena menggunakan kuadrat *error*, RMSEfrac tetap memberikan penalti lebih besar terhadap kesalahan prediksi

yang besar, tetapi hasilnya dinyatakan dalam bentuk persentase terhadap total umur unit. Dengan demikian, RMSEfrac dapat digunakan untuk melihat besar *error* relatif terhadap umur total sekaligus tetap mempertahankan sensitivitas RMSE terhadap *error* yang besar. Berikut penjelasan pada Tabel 2.1 Untuk memperjelas fungsi setiap metrik evaluasi yang digunakan pada penelitian ini.

Tabel 2.1 Tujuan Penggunaan Metrik Evaluasi Model

Metrik	Dasar Perhitungan	Tujuan Penggunaan	Posisi dalam Penelitian
MAE	Rata-rata error absolut antara RUL prediksi dan RUL aktual	Mengetahui rata-rata besar kesalahan prediksi dalam satuan RUL	Metrik utama
RMSE	Akar rata-rata kuadrat error	Mengetahui sensitivitas model terhadap error besar atau prediksi yang menyimpang jauh	Metrik utama
MAPE	Rata-rata error absolut relatif terhadap RUL aktual	Menilai kesalahan prediksi dalam bentuk persentase dan menjadi dasar target performansi model	Metrik utama
MAEfrac	Rata-rata error absolut relatif terhadap umur total unit	Menilai proporsi error terhadap keseluruhan umur unit dan mendukung perbandingan dengan penelitian terdahulu	Metrik tambahan
RMSEfrac	Akar rata-rata kuadrat error relatif terhadap umur total unit	Menilai error relatif terhadap umur total unit dengan penalti lebih besar pada error besar	Metrik tambahan

Berdasarkan Tabel 2.1, MAE dan RMSE digunakan untuk mengevaluasi besar error dalam satuan RUL. MAE memberikan gambaran rata-rata kesalahan absolut, sedangkan RMSE memberikan penekanan lebih besar terhadap error yang

tinggi. MAPE digunakan untuk menilai error relatif terhadap RUL aktual dan menjadi dasar penetapan target performansi model. Sementara itu, MAEfrac dan RMSEfrac digunakan sebagai metrik tambahan untuk menilai error relatif terhadap umur total unit. Dengan demikian, MAE, RMSE, dan MAPE digunakan sebagai metrik evaluasi utama, sedangkan MAEfrac dan RMSEfrac digunakan sebagai metrik pendukung untuk analisis perbandingan performa terhadap penelitian terdahulu.

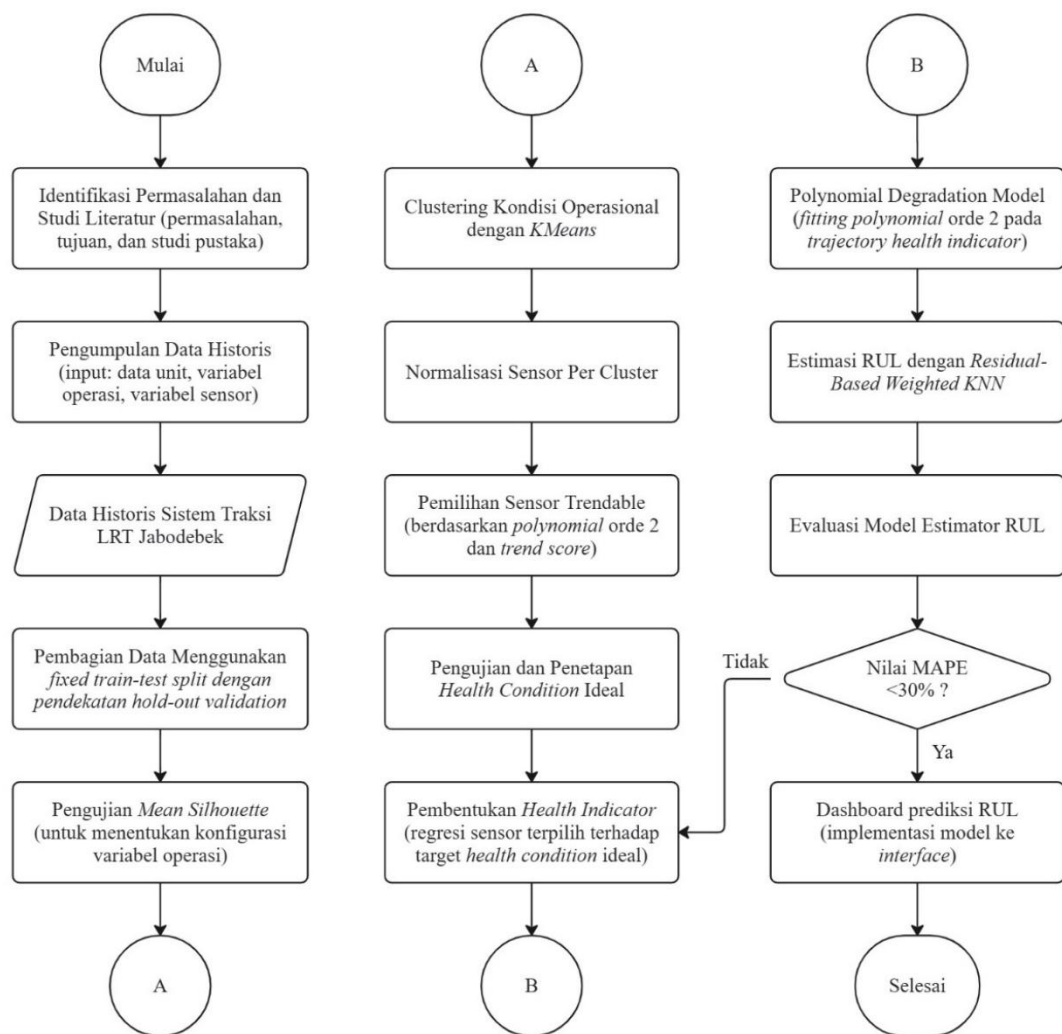
Jika dikaitkan dengan penelitian sebelumnya, penggunaan metrik pada penelitian ini juga disusun agar tetap memiliki keterbandingan dengan studi terdahulu sekaligus memberikan evaluasi yang lebih lengkap. Pada penelitian Natasya digunakan pendekatan *data-driven* berbasis K-Nearest Neighbor untuk estimasi RUL sistem traksi kereta dengan evaluasi pada *breakpoint* 0,5, 0,7, dan 0,9 serta analisis *Probability Density Function* (PDF). Sementara itu, penelitian Husain mengembangkan pendekatan estimasi RUL sistem traksi LRT berbasis *similarity-based exponential* dengan segmentasi k-means dan *Isolation Forest*, serta mengevaluasi performa model menggunakan MAE, RMSE, MAEfrac, dan RMSEfrac. Oleh karena itu, pada penelitian ini MAEfrac dan RMSEfrac tetap dipertahankan agar hasil evaluasi masih dapat dibandingkan dengan penelitian Husain, sedangkan MAPE ditambahkan sebagai metrik utama tambahan untuk memberikan interpretasi kesalahan relatif terhadap RUL aktual dalam bentuk persentase.

Dalam konteks prediksi RUL, penggunaan MAE, RMSE, MAPE, MAEfrac, dan RMSEfrac memberikan gambaran evaluasi yang lebih lengkap. MAE dan RMSE menunjukkan besar *error* dalam satuan RUL, MAPE menunjukkan *error* relatif terhadap RUL aktual, sedangkan MAEfrac dan RMSEfrac menunjukkan *error* relatif terhadap umur total unit. Pada penelitian ini, MAE, RMSE, dan MAPE digunakan sebagai metrik utama untuk menilai performa model, sedangkan MAEfrac dan RMSEfrac digunakan sebagai metrik tambahan untuk mendukung perbandingan performa dengan penelitian terdahulu.

BAB III

METODE PENELITIAN

Penelitian ini dilakukan untuk mengembangkan model estimasi *Remaining Useful Life* (RUL) pada sistem traksi LRT Jabodebek menggunakan pendekatan *similarity-based polynomial* dan *residual-based KNN*. Secara umum, penelitian ini mencakup tahapan identifikasi permasalahan, pengolahan data historis, pembentukan *Health Indicator* (HI), pemodelan degradasi, estimasi RUL, evaluasi model, serta implementasi model dalam bentuk dashboard prediksi. Adapun alur penelitian secara keseluruhan disajikan pada gambar 3.1.



Gambar 3.1 Diagram Alir Penelitian

Berdasarkan gambar 3.1, penelitian diawali dengan identifikasi permasalahan dan studi literatur, kemudian dilanjutkan dengan pengumpulan data historis sistem traksi LRT Jabodebek. Data yang diperoleh selanjutnya dibagi menggunakan *fixed train-test split*. Pada data *train*, dilakukan pengujian *mean silhouette* untuk menentukan konfigurasi variabel operasi yang paling sesuai pada tahap *clustering* kondisi operasional. Setelah itu, data dikelompokkan menggunakan *K-Means* dan dinormalisasi per *cluster* agar pengaruh perbedaan kondisi operasi terhadap sinyal sensor dapat dikurangi.

Tahap berikutnya adalah pemilihan sensor *trendable* untuk memperoleh sensor-sensor yang paling relevan dalam merepresentasikan degradasi sistem. Sensor terpilih tersebut kemudian digunakan pada pengujian dan penetapan *health condition ideal*, lalu dilanjutkan dengan pembentukan *Health Indicator* melalui regresi sensor terhadap target *health condition ideal* yang telah ditetapkan. Setelah HI terbentuk, dilakukan *polynomial degradation model* orde dua untuk memperoleh representasi tren degradasi yang lebih ringkas pada setiap unit. Selanjutnya, estimasi RUL dilakukan menggunakan pendekatan *residual-based KNN* dengan membandingkan HI unit *test* terhadap model degradasi historis unit-unit pada data *train*. Hasil estimasi yang diperoleh kemudian dievaluasi menggunakan MAE, RMSE, dan MAPE untuk menilai kinerja model dari sisi error absolut, sensitivitas terhadap error besar, dan error relatif dalam bentuk persentase. Pada penelitian ini, MAPE digunakan sebagai acuan keputusan performansi pada diagram alir dengan target nilai di bawah 30% karena metrik ini lebih mudah diinterpretasikan sebagai tingkat kesalahan relatif model. Sementara itu, MAEfrac dan RMSEfrac tidak digunakan sebagai dasar keputusan, tetapi digunakan sebagai metrik evaluasi tambahan untuk mendukung perbandingan hasil penelitian ini dengan penelitian terdahulu. Model yang telah dibangun diimplementasikan ke dalam dashboard prediksi RUL agar dapat digunakan dalam bentuk yang lebih aplikatif.

3.1 Pengumpulan Data Historis Sistem Traksi

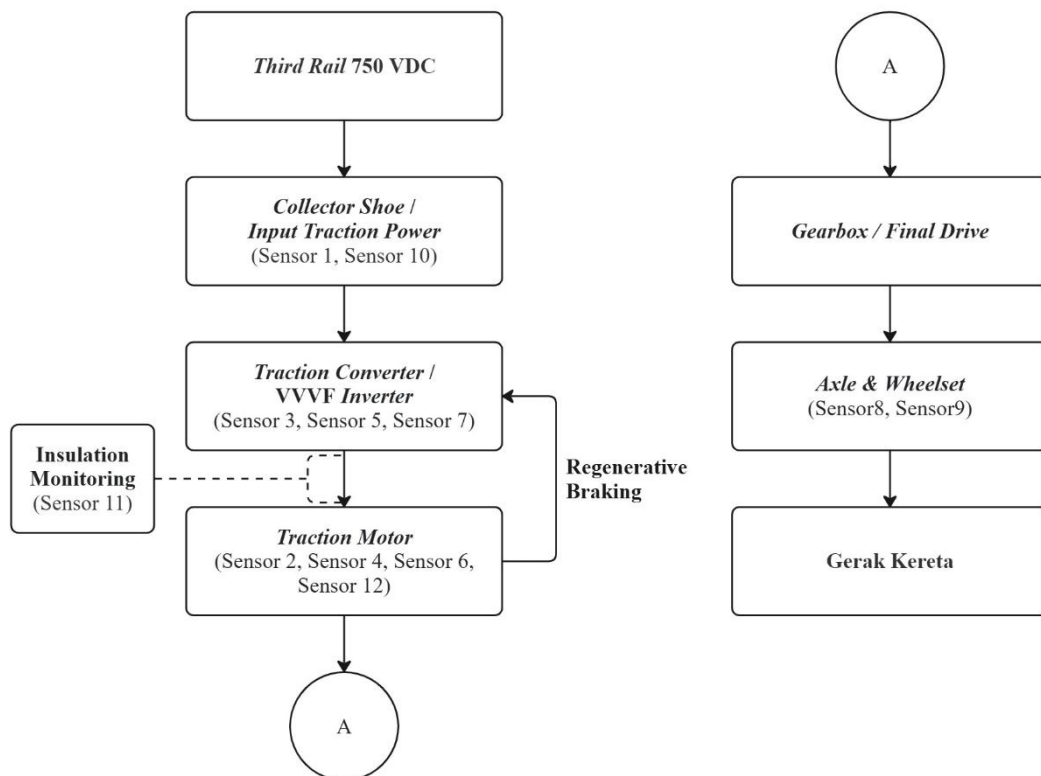
Tahap awal penelitian ini adalah pengumpulan data historis sistem traksi LRT Jabodebek sebagai dasar pembentukan model estimasi *Remaining Useful Life* (RUL). Data historis digunakan dalam pendekatan *data-driven prognostics* untuk

mempelajari perkembangan degradasi dan memanfaatkan riwayat perubahan antarunit sebagai referensi estimasi RUL.

Data memiliki struktur *run-to-failure*, yaitu rangkaian pengamatan berurutan sepanjang waktu operasi hingga mendekati kondisi akhir. Struktur ini membentuk *trajectory* degradasi setiap unit dan mendukung pendekatan *similarity-based*, karena pola unit pada data *train* digunakan sebagai referensi bagi unit pada data *test*.

Dataset terdiri atas 17 variabel, yaitu satu variabel identitas, satu variabel waktu, tiga variabel operasi, dan dua belas variabel sensor. Variabel *id* membedakan unit, sedangkan *TimeStamp* menunjukkan urutan pengamatan. Variabel operasi mencakup kecepatan, *brake force*, dan torsi motor, sementara variabel sensor merepresentasikan parameter elektrik, termal, dan mekanik sistem traksi.

Untuk memperjelas keterkaitan sensor dengan subsistem yang diamati, skematik sistem traksi dan posisi fungsional parameter sensor ditunjukkan pada gambar 3.2. Skematik tersebut merupakan ilustrasi hubungan parameter dataset dengan subsistem traksi, bukan *wiring diagram* resmi pabrikan.



Gambar 3.2 Skematik Sistem Traksi LRT

Berdasarkan gambar 3.2, energi listrik berasal dari *third rail* 750 VDC dan diterima melalui *collector shoe* sebelum diteruskan menuju *traction converter* atau inverter VVVF. Sensor 1 dan Sensor 10 berkaitan dengan sisi *input* daya, sedangkan Sensor 3, Sensor 5, dan Sensor 7 merepresentasikan kondisi *output* dan temperatur inverter VVVF. Daya yang telah diatur kemudian digunakan oleh motor traksi, yang kondisinya direpresentasikan oleh Sensor 2, Sensor 4, Sensor 6, dan Sensor 12.

Energi mekanik dari motor diteruskan melalui *gearbox/final drive* menuju *axle* dan *wheelset*. Pada bagian keluaran mekanik tersebut, Sensor 8 merepresentasikan kecepatan roda dan Sensor 9 merepresentasikan gaya traksi. Sensor 11 berkaitan dengan pemantauan tahanan isolasi pada jalur elektrik antara inverter dan motor, sedangkan Sensor 12 berkaitan dengan arus motor selama proses pengereman regeneratif. Kemudian penjelasan rinci mengenai variabel operasi dan sensor beserta satuan pengukurannya ditampilkan pada tabel 3.1.

Tabel 3.1 Deskripsi Variabel Operasi dan Sensor pada Data Penelitian

Keterangan	Parameter	Besaran
Op_setting_1	Kecepatan	Kilometer per hour (km/h)
Op_setting_2	<i>Brake Force</i>	Kilonewton (kN)
Op_setting_3	Torsi Motor	Newton meter (Nm)
Sensor_1	Arus Input LAA	Ampere (A)
Sensor_2	Arus Input Motor	Ampere (A)
Sensor_3	Arus Output VVVF	Ampere (A)
Sensor_4	<i>Voltage Input Motor</i>	Volt (V)
Sensor_5	<i>Voltage Output VVVF</i>	Volt (V)
Sensor_6	<i>Temperature Motor</i>	Celsius (°C)
Sensor_7	<i>Temperature VVVF</i>	Celsius (°C)
Sensor_8	<i>Speed Wheel</i>	Radian per second (rad/s)
Sensor_9	<i>Traction Force</i>	Kilonewton (kN)
Sensor_10	<i>Power Consumption</i>	Watt (W)
Sensor_11	<i>Insulation Resistance</i>	Megaohm (MΩ)
Sensor_12	<i>Regenerative Motor Current</i>	Ampere (A)

Berdasarkan tabel 3.1, variabel pada dataset dapat dibedakan menjadi variabel operasi dan variabel sensor. Variabel operasi menggambarkan kondisi kerja sistem, sedangkan variabel sensor merekam respons elektrik, termal, dan mekanik yang terjadi selama sistem beroperasi. Kedua kelompok variabel perlu dianalisis secara terpisah karena perubahan nilai sensor tidak hanya dipengaruhi oleh degradasi, tetapi juga dapat disebabkan oleh perubahan kondisi operasi. Oleh karena itu, variabel operasi digunakan pada tahap *clustering*, sedangkan variabel sensor digunakan sebagai sumber informasi dalam pembentukan *Health Indicator*.

Dataset memiliki ukuran awal sebesar 32.845 baris dan 17 kolom serta mencakup 133 id pengamatan. Data kemudian dikelompokkan dan diurutkan berdasarkan id serta *TimeStamp* agar rangkaian pengamatan setiap unit tetap terjaga. Ringkasan karakteristik dataset disajikan pada tabel 3.2.

Tabel 3. 2 Ringkasan Data *Run-to-Failure*

Komponen	Nilai
Jumlah Baris	32.845
Jumlah Kolom	17
Jumlah id	133
Jumlah Variabel Operasi	3
Jumlah Variabel Sensor	12
Jenis Data	<i>Time-Based</i>

Berdasarkan tabel 3.2, satu id tidak merepresentasikan satu titik data, tetapi satu unit observasi yang terdiri atas rangkaian pengamatan sepanjang waktu operasi. Hubungan antarbaris pada id yang sama membentuk satu *trajectory* kondisi. Oleh karena itu, analisis dilakukan dengan mempertahankan struktur tiap unit agar pembentukan HI, pemodelan degradasi, dan estimasi RUL dapat dilakukan berdasarkan perkembangan kondisi secara utuh. Data historis tersebut tidak langsung digunakan dalam proses estimasi RUL, melainkan melalui berbagai proses hingga didapatkan model bentuk HI sehingga estimasi dapat dilakukan. Dengan demikian, tahap pengumpulan dan penyusunan data menjadi fondasi bagi seluruh proses analisis karena kualitas serta struktur data akan memengaruhi HI dan performa model estimasi RUL yang dihasilkan..

3.2 Pembagian Data *Train* dan *Test*

Setelah data historis dikumpulkan dan disusun berdasarkan unit, tahap berikutnya adalah membagi data menjadi kelompok *train* dan *test*. Penelitian ini menggunakan *fixed train-test split* dengan proporsi sekitar 80% data *train* dan 20% data *test*. Proporsi tersebut mengacu pada penelitian sebelumnya agar basis evaluasi tetap sebanding dan perbedaan performa lebih mencerminkan pengaruh metode yang digunakan, bukan perbedaan komposisi data.

Pembagian dilakukan pada level unit, bukan pada level baris pengamatan. Seluruh riwayat dari satu unit ditempatkan sepenuhnya pada data *train* atau data *test* agar *trajectory* degradasinya tetap utuh. Pembagian pada level baris berisiko menyebabkan sebagian riwayat unit yang sama masuk ke kedua kelompok sehingga menimbulkan *data leakage* dan menghasilkan evaluasi yang tidak valid. Data *train* digunakan untuk membangun seluruh komponen model, meliputi *clustering* kondisi operasional, penentuan parameter normalisasi sensor, pemilihan sensor, pembentukan *Health Indicator*, pemodelan degradasi polinomial, dan penyusunan referensi historis untuk estimasi *Remaining Useful Life*. Sementara itu, data *test* digunakan untuk mengevaluasi kemampuan model pada unit yang tidak dilibatkan dalam proses pembentukan model. Skema tersebut termasuk dalam *hold-out validation* karena data dibagi satu kali menggunakan komposisi yang telah ditetapkan. Meskipun identitas unit yang digunakan tidak harus sama dengan penelitian sebelumnya, proporsi pembagiannya dipertahankan agar hasil evaluasi dapat dibandingkan pada basis yang relatif sebanding.

Pembagian berbasis unit juga mempertahankan karakteristik data *run-to-failure*, yaitu setiap unit diperlakukan sebagai satu rangkaian kondisi dari awal hingga akhir pengamatan. Hal ini diperlukan karena pembentukan HI, pemodelan degradasi, dan estimasi RUL dilakukan berdasarkan perkembangan kondisi setiap unit, bukan berdasarkan titik pengamatan yang berdiri sendiri. Dengan demikian, pembagian data dirancang untuk menjaga keutuhan *trajectory*, mencegah *data leakage*, dan menyediakan data pengujian yang terpisah dari proses pembentukan model. Data *train* dan *test* yang diperoleh selanjutnya digunakan pada tahap pengujian variabel operasi dan pembentukan *cluster* kondisi operasional.

3.3 Pengujian *Mean Silhouette*

Sebelum dilakukan *clustering* kondisi operasional, variabel operasi terlebih dahulu diuji menggunakan *mean silhouette*. Tahap ini bertujuan menentukan kombinasi variabel yang mampu menghasilkan kelompok kondisi operasi paling kompak dan memiliki pemisahan yang jelas. Pemilihan variabel perlu dilakukan karena kualitas *clustering* akan memengaruhi proses normalisasi sensor dan pembentukan *Health Indicator* pada tahap berikutnya.

Variabel yang diuji terdiri atas Op_Setting1, Op_Setting2, dan Op_Setting3. Berdasarkan karakteristik dataset, Op_Setting1 merepresentasikan kecepatan, Op_Setting2 merepresentasikan *brake force*, dan Op_Setting3 merepresentasikan torsi motor. Ketiga variabel tersebut menggambarkan kondisi kerja sistem traksi, tetapi setiap kombinasi dapat menghasilkan kualitas pemisahan *cluster* yang berbeda. Pengujian dilakukan terhadap beberapa kombinasi variabel operasi menggunakan konfigurasi *clustering* yang sama. Penggunaan konfigurasi yang konsisten diperlukan agar perbedaan nilai *mean silhouette* yang diperoleh lebih mencerminkan pengaruh pemilihan variabel, bukan perubahan pengaturan metode *clustering*.

Nilai *silhouette* setiap data dan nilai *mean silhouette* masing-masing skenario dihitung berdasarkan Persamaan (2.1) dan Persamaan (2.2). Nilai yang lebih tinggi menunjukkan bahwa data dalam satu *cluster* memiliki kedekatan yang lebih baik dan lebih terpisah dari *cluster* lainnya. Sebaliknya, nilai yang lebih rendah menunjukkan bahwa pemisahan antar-*cluster* belum jelas atau masih terdapat tumpang tindih data. Skenario dengan nilai *mean silhouette* tertinggi dipilih sebagai masukan pada proses *K-Means clustering*. Hasil pengujian setiap skenario dan penetapan variabel operasi terpilih disajikan pada Bab IV, sedangkan prosedur *clustering* kondisi operasional dijelaskan pada subbab berikutnya.

3.4 *Clustering* Kondisi Operasional dengan *K-Means*

Setelah variabel operasi dipilih melalui pengujian *mean silhouette*, tahap berikutnya adalah melakukan *clustering* kondisi operasional menggunakan metode K-Means. Tahap ini bertujuan memisahkan variasi sinyal akibat perbedaan kondisi operasi sebelum dilakukan normalisasi sensor dan analisis degradasi. Pemisahan

diperlukan karena perubahan nilai sensor pada sistem traksi dapat berasal dari perubahan mode operasi maupun penurunan kondisi komponen.

Proses *clustering* dilakukan menggunakan variabel operasi terpilih dan jumlah *cluster* yang telah ditetapkan pada tahap sebelumnya. Metode K-Means mengelompokkan data berdasarkan kedekatannya terhadap *centroid* dengan mengacu pada fungsi objektif pada Persamaan (2.3). Data dalam *cluster* yang sama diharapkan memiliki karakteristik operasi yang lebih serupa dibandingkan data pada *cluster* lainnya.

Pembentukan *cluster* dan perhitungan *centroid* hanya dilakukan menggunakan data *train*. Setiap titik data *train* memperoleh label berdasarkan *centroid* terdekat, sedangkan data *test* tidak digunakan untuk memperbarui struktur *cluster*. Label kondisi operasional pada data *test* ditentukan berdasarkan kedekatannya terhadap *centroid* yang telah diperoleh dari data *train*. Prosedur ini digunakan untuk mencegah informasi dari data *test* masuk ke dalam proses pembentukan model.

Output tahap ini berupa label *cluster* untuk setiap titik pengamatan pada data *train* dan data *test*. Label tersebut menunjukkan kelompok kondisi operasional masing-masing data dan selanjutnya digunakan sebagai dasar normalisasi sensor per *cluster*. Dengan demikian, setiap nilai sensor dibandingkan terhadap karakteristik statistik dari kelompok operasi yang sejenis.

Tahap *clustering* ini tidak digunakan untuk mendeteksi degradasi secara langsung, tetapi untuk mengurangi tercampurnya pengaruh kondisi operasi dengan perubahan kesehatan sistem. Data yang telah dikelompokkan selanjutnya dinormalisasi berdasarkan parameter statistik masing-masing *cluster* agar proses pemilihan sensor dan pembentukan *Health Indicator* lebih berfokus pada pola degradasi.

3.5 Normalisasi Sensor Per Cluster

Setelah kondisi operasional dikelompokkan, tahap berikutnya adalah melakukan normalisasi sensor pada masing-masing *cluster*. Normalisasi diperlukan karena nilai sensor dipengaruhi oleh kondisi kesehatan komponen dan mode operasi saat pengukuran dilakukan. Jika seluruh data dinormalisasi secara global, perbedaan *baseline* antar kondisi operasi dapat tercampur dengan perubahan akibat

degradasi. Oleh karena itu, normalisasi dilakukan per *cluster* agar nilai sensor dibandingkan dalam konteks kondisi operasi yang sejenis (Davari et al., 2021; Xu et al., 2023).

Parameter normalisasi dihitung secara terpisah untuk setiap sensor pada masing-masing *cluster* menggunakan data *train*. Rata-rata dan simpangan baku yang diperoleh kemudian digunakan untuk mentransformasikan data *train* dan data *test* pada *cluster* yang sama. Pendekatan ini memungkinkan sensor dengan skala dan distribusi berbeda diubah ke bentuk yang lebih sebanding tanpa menghilangkan konteks operasionalnya.

Normalisasi dilakukan menggunakan transformasi *z-score* per *cluster*, sebagaimana ditunjukkan pada persamaan (3.1).

$$\widetilde{s}_{i,j}^{(k)} = \frac{s_{i,j} - \mu_{k,j}}{\sigma_{k,j}} \quad (3.1)$$

Pada persamaan (3.1), $\widetilde{s}_{i,j}^{(k)}$ merupakan nilai sensor ke- j pada sampel ke- i setelah dinormalisasi berdasarkan *cluster* ke- k . Nilai $s_{i,j}$ merupakan data sensor sebelum normalisasi, sedangkan $\mu_{k,j}$ dan $\sigma_{k,j}$ masing-masing merupakan rata-rata dan simpangan baku sensor ke- j yang dihitung dari data *train* pada *cluster* ke- k .

Perhitungan parameter hanya menggunakan data *train* untuk mencegah masuknya informasi data *test* ke dalam proses pembentukan model. Data *test* tetap dinormalisasi menggunakan rata-rata dan simpangan baku dari *cluster train* yang sesuai sehingga aturan transformasi diterapkan secara konsisten dan risiko *data leakage* dapat dihindari.

Output tahap ini berupa data sensor yang telah dinormalisasi berdasarkan kondisi operasionalnya. Data tersebut selanjutnya digunakan dalam proses pemilihan sensor sehingga analisis tren dilakukan pada variabel yang skalanya telah seragam dan pengaruh perbedaan *baseline* antar kondisi operasi telah dikurangi. Dengan demikian, tahap berikutnya dapat lebih berfokus pada pola perubahan yang berkaitan dengan degradasi sistem traksi.

3.6 Pemilihan Sensor *Trendable* untuk Pemodelan Degradasi

Setelah data sensor dinormalisasi berdasarkan kondisi operasional, tahap berikutnya adalah memilih sensor yang memiliki kecenderungan tren degradasi paling jelas. Tidak seluruh sensor digunakan dalam pembentukan *Health Indicator*

(HI) karena sebagian sensor dapat lebih dipengaruhi oleh fluktuasi operasi atau tidak menunjukkan perubahan yang konsisten terhadap waktu. Oleh karena itu, seleksi dilakukan agar HI dibentuk menggunakan sensor yang paling representatif terhadap penurunan kondisi sistem traksi.

Setiap sensor yang telah dinormalisasi dimodelkan menggunakan fungsi polinomial orde dua. Model ini dipilih karena tetap sederhana, tetapi mampu menangkap pola perubahan nonlinier pada *trajectory* sensor. Bentuk model polinomial tersebut ditunjukkan pada persamaan (3.2).

$$s_j(t) \approx at^2 + bt + c \quad (3.2)$$

Pada persamaan (3.2), $s_j(t)$ merupakan nilai sensor ke- j pada waktu t , sedangkan a , b , dan c merupakan koefisien model polinomial. Model tersebut digunakan untuk memperoleh representasi tren sensor yang lebih halus dibandingkan data pengamatan mentah.

Kecenderungan degradasi kemudian dinilai berdasarkan besar gradien kurva pada akhir *trajectory*. Bagian akhir pengamatan digunakan karena perubahan kondisi umumnya semakin terlihat ketika sistem mendekati akhir umur operasinya. Skor kecenderungan tren dinyatakan pada persamaan (3.3).

$$\text{TrendScore} \propto |2aT + b| \quad (3.3)$$

Pada persamaan (3.3), T merupakan waktu akhir pengamatan, sedangkan $2aT + b$ merupakan gradien model polinomial pada titik tersebut. Nilai absolut digunakan agar kekuatan tren dapat dinilai tanpa dipengaruhi arah perubahan. Semakin besar nilai *TrendScore*, semakin kuat kecenderungan perubahan sensor pada akhir *trajectory*.

Sensor kemudian diurutkan berdasarkan nilai *TrendScore*, dan sensor dengan peringkat tertinggi dipilih untuk tahap pembentukan HI. Setelah sensor terpilih, arah perubahannya diseragamkan karena sebagian sensor dapat meningkat ketika kondisi memburuk, sedangkan sensor lain menurun. Penyesuaian arah diperlukan agar seluruh sensor mempunyai orientasi degradasi yang konsisten saat digabungkan. Output tahap ini berupa kumpulan sensor *trendable* yang paling mampu merepresentasikan perubahan kondisi sistem traksi. Sensor tersebut selanjutnya digunakan dalam proses penentuan *health condition* dan pembentukan *Health Indicator*.

3.7 Pengujian dan Penetapan *Health Condition* Ideal

Setelah sensor *trendable* diperoleh, tahap berikutnya adalah menentukan *health condition* ideal sebagai kurva acuan dalam pembentukan *Health Indicator* (HI). Kurva ini digunakan untuk mengarahkan kombinasi sensor agar menghasilkan HI dengan pola perubahan yang jelas dan konsisten terhadap perkembangan degradasi. Tahap tersebut diperlukan karena kualitas HI akan memengaruhi representasi kondisi sistem, perbandingan pola antarunit, dan estimasi *Remaining Useful Life* (RUL).

Dasar penggunaan *health condition* ideal pada penelitian ini mengacu pada konsep pembentukan HI dalam penelitian terdahulu. Schwartz et al. (2022) menjelaskan bahwa HI digunakan untuk membentuk *degradation trajectory* yang disimpan sebagai referensi pada *Similarity-Based Model*. Hal tersebut menunjukkan bahwa HI perlu memiliki bentuk tren yang jelas agar kemiripan pola degradasi antarunit dapat dibandingkan secara bermakna. Penelitian tersebut juga menunjukkan bahwa HI lebih sesuai digunakan dibandingkan data sensor mentah karena mampu menggabungkan beberapa sumber informasi menjadi satu representasi kondisi.

Xu et al. (2023) menyatakan bahwa HI yang mampu menggambarkan perkembangan degradasi secara akurat merupakan salah satu prasyarat dalam membangun model prediksi RUL. HI yang memiliki kecenderungan lebih monoton dinilai lebih mampu menunjukkan perubahan kondisi secara bertahap. Berdasarkan konsep tersebut, kurva acuan pada penelitian ini dirancang memiliki bentuk sederhana dan arah perubahan yang dapat dievaluasi berdasarkan konsistensinya terhadap waktu.

Zhu et al. (2025) menjelaskan bahwa estimasi RUL tidak hanya memerlukan pemetaan data secara langsung menuju nilai RUL, tetapi juga dapat dilakukan melalui pembentukan HI untuk menangkap informasi degradasi secara lebih terstruktur. Pembentukan HI terlebih dahulu memungkinkan informasi dari beberapa sensor diringkas menjadi satu *trajectory* yang menggambarkan perkembangan kesehatan sistem. Oleh karena itu, bentuk acuan yang digunakan perlu mampu membantu menghasilkan HI dengan pola degradasi yang mudah dimodelkan dan dibandingkan.

Ketiga penelitian tersebut tidak menetapkan satu persamaan *health condition* yang harus digunakan. Namun, ketiganya mendukung prinsip bahwa HI perlu memiliki arah perubahan yang jelas, cukup monoton, dan mampu merepresentasikan perkembangan degradasi. Berdasarkan prinsip tersebut, penelitian ini menyusun beberapa kandidat kurva acuan untuk menguji bentuk yang paling sesuai terhadap karakteristik data sistem traksi.

Kandidat kurva dibentuk menggunakan fungsi polinomial sederhana karena mudah diinterpretasikan dan mampu memberikan variasi arah serta kelengkungan. HC-1 dan HC-2 dirancang sebagai kurva menurun dari kondisi awal menuju kondisi akhir, sesuai dengan konsep HI pada penelitian ini yang diharapkan menurun seiring terjadinya degradasi. Sementara itu, HC-3 dan HC-4 digunakan sebagai bentuk pembanding dengan arah perubahan yang berlawanan. Pengujian kedua kelompok tersebut dilakukan untuk memastikan bahwa arah kurva acuan yang dipilih benar-benar menghasilkan HI yang sesuai dengan karakteristik degradasi pada data, bukan hanya ditetapkan berdasarkan asumsi awal.

Seluruh kurva dibentuk pada domain waktu ternormalisasi $\tau \in [0,1]$. Nilai $\tau = 0$ merepresentasikan awal pengamatan dan $\tau = 1$ merepresentasikan akhir pengamatan. Normalisasi waktu memungkinkan unit dengan panjang *trajectory* berbeda dipetakan pada domain yang sama, sehingga setiap kandidat kurva dapat diterapkan secara konsisten pada seluruh data *train*. Kandidat *health condition* ideal yang digunakan ditunjukkan pada tabel 3.3.

Tabel 3. 3 Skema Pengujian Health Condition Ideal

Kode	Bentuk Kurva
HC-1	$(\tau) = 1 - \tau^2$
HC-2	$(\tau) = 1 - \tau^3$
HC-3	$(\tau) = 1 - (1 - \tau)^2$
HC-4	$(\tau) = 1 - (1 - \tau)^3$

Berdasarkan tabel 3.3, keempat kandidat *health condition* memiliki bentuk matematis dan kelengkungan yang berbeda. Setiap kandidat diterapkan secara terpisah pada data *train* untuk membentuk HI sementara. Selanjutnya, kurva HI yang dihasilkan dibandingkan untuk menentukan bentuk acuan yang paling sesuai dalam merepresentasikan degradasi sistem traksi. Dengan demikian, penetapan

health condition ideal tidak hanya didasarkan pada bentuk persamaannya, tetapi juga pada kualitas pola HI yang dihasilkan.

Kualitas HI sementara dievaluasi menggunakan jumlah unit dengan *slope* negatif, korelasi Spearman negatif, *delta end-start* negatif, dan rata-rata *monotonicity*. Nilai *slope* digunakan untuk menilai arah tren umum HI terhadap waktu, sedangkan korelasi Spearman digunakan untuk melihat konsistensi hubungan monoton antara HI dan waktu operasi. Nilai *delta end-start* menunjukkan perbedaan kondisi antara awal dan akhir pengamatan, sementara *monotonicity* digunakan untuk menilai konsistensi perubahan HI dalam mengikuti arah degradasi.

Selain evaluasi kuantitatif, kurva HI dari setiap kandidat juga divisualisasikan untuk mengamati bentuk *trajectory*, arah perubahan, dan kesesuaiannya dengan konsep degradasi pada penelitian ini. Evaluasi visual digunakan sebagai pendukung agar kandidat tidak hanya menghasilkan nilai metrik yang baik, tetapi juga membentuk pola HI yang dapat diinterpretasikan secara logis sebagai perkembangan kondisi sistem.

Seluruh proses pembentukan dan evaluasi kandidat dilakukan menggunakan data *train* agar data *test* tidak terlibat dalam penetapan kurva acuan. Kandidat dengan hasil evaluasi kuantitatif dan visual yang paling sesuai kemudian ditetapkan sebagai *health condition* ideal untuk pembentukan HI final. Hasil pengujian setiap kandidat dan penetapan kurva terpilih disajikan pada Bab IV, sedangkan kurva terpilih digunakan pada tahap pembentukan *Health Indicator* berikutnya.

3.8 Pembentukan *Health Indicator*

Setelah *health condition* ideal ditetapkan, tahap berikutnya adalah membentuk *Health Indicator* (HI) sebagai representasi kondisi kesehatan sistem traksi. HI digunakan untuk merangkum informasi dari beberapa sensor terpilih menjadi satu *trajectory* yang lebih ringkas dan mudah digunakan dalam pemodelan degradasi.

Sebelum digabungkan, arah degradasi setiap sensor diseragamkan. Penyesuaian ini diperlukan karena sebagian sensor dapat meningkat ketika kondisi sistem memburuk, sedangkan sensor lainnya menunjukkan arah menurun. Tanpa

penyeragaman tersebut, kontribusi antarsensor dapat saling berlawanan dan menghasilkan HI yang tidak konsisten. Sensor yang telah dinormalisasi dan disesuaikan arahnya kemudian diregresikan terhadap target *health condition* terpilih. Proses ini digunakan untuk memperoleh intersep dan koefisien yang mampu membentuk HI mendekati kurva acuan degradasi. Bentuk pembentukan HI pada penelitian ini dinyatakan pada persamaan (3.4).

$$h(t) \approx b_0 + \sum_{j=1}^p w_j z_j(t) \quad (3.4)$$

Pada persamaan (3.4), $h(t)$ merupakan nilai HI pada waktu t , b_0 merupakan intersep regresi, w_j merupakan koefisien sensor ke- j , $z_j(t)$ merupakan nilai sensor yang telah dinormalisasi dan disesuaikan arah degradasinya, sedangkan p merupakan jumlah sensor terpilih. Besarnya koefisien menunjukkan kontribusi masing-masing sensor dalam membentuk HI terhadap target *health condition*.

Parameter regresi dibentuk menggunakan data *train*, kemudian diterapkan dengan aturan yang sama pada data *test*. Prosedur ini menjaga agar informasi dari data *test* tidak digunakan dalam pembentukan HI sekaligus memastikan bahwa seluruh unit direpresentasikan menggunakan model yang konsisten. HI awal yang diperoleh selanjutnya melalui proses *smoothing* untuk mengurangi *noise* dan fluktuasi jangka pendek yang tidak berkaitan langsung dengan degradasi. Proses tersebut menghasilkan *trajectory* HI yang lebih stabil sehingga kecenderungan perubahan kondisi setiap unit dapat diamati dan dimodelkan dengan lebih jelas.

Output tahap ini berupa HI pada setiap unit data *train* dan data *test*. HI tersebut menjadi penghubung antara data sensor multidimensi dan tahap pemodelan degradasi polinomial. Dengan demikian, informasi dari beberapa sensor dapat direpresentasikan sebagai satu kurva kondisi yang selanjutnya digunakan dalam pembentukan model referensi dan estimasi *Remaining Useful Life*.

3.9 Polynomial Degradation Model

Setelah *Health Indicator* (HI) dibentuk untuk setiap unit, tahap berikutnya adalah memodelkan *trajectory* degradasi menggunakan fungsi polinomial. Pemodelan ini bertujuan mengubah deret nilai HI terhadap waktu menjadi representasi matematis yang lebih ringkas dan terstruktur. Representasi tersebut

diperlukan agar pola degradasi antarunit dapat dibandingkan secara konsisten pada tahap estimasi *Remaining Useful Life*.

Pada penelitian ini, HI setiap unit dimodelkan menggunakan polinomial orde dua. Orde tersebut dipilih karena cukup sederhana, tetapi tetap mampu merepresentasikan pola degradasi nonlinier. Bentuk model yang digunakan ditunjukkan pada persamaan (3.5).

$$\hat{h}_i(t) = a_i t^2 + b_i t + c_i \quad (3.5)$$

Pada persamaan (3.5), $\hat{h}_i(t)$ merupakan estimasi HI unit ke- i pada waktu t , sedangkan a_i , b_i , dan c_i merupakan koefisien polinomial yang menggambarkan karakteristik kurva degradasi unit tersebut. Koefisien a_i merepresentasikan kelengkungan kurva, b_i menunjukkan kecenderungan perubahan linier, dan c_i merupakan nilai intersep model.

Proses *fitting* dilakukan secara terpisah pada setiap unit agar perbedaan karakteristik degradasi antarunit tetap dipertahankan. Dengan demikian, setiap unit memiliki model degradasi sendiri dan tidak diasumsikan mengikuti satu pola global yang sama. Pendekatan ini sesuai dengan karakteristik data *run-to-failure* yang dapat menunjukkan laju dan bentuk penurunan kondisi yang berbeda.

Model polinomial juga membantu menangkap kecenderungan utama HI dan mengurangi pengaruh fluktuasi lokal yang masih tersisa setelah proses *smoothing*. Output tahap ini berupa model degradasi untuk setiap unit yang selanjutnya digunakan sebagai referensi pada estimasi RUL menggunakan pendekatan *residual-based weighted KNN*.

3.10 Estimasi RUL dengan *Residual-based Weighted KNN*

Setelah *trajectory Health Indicator* (HI) setiap unit direpresentasikan menggunakan *polynomial degradation model*, tahap berikutnya adalah mengestimasi *Remaining Useful Life* (RUL) menggunakan *residual-based weighted KNN*. Pendekatan berbasis kemiripan memanfaatkan pola degradasi historis sebagai referensi dengan asumsi bahwa unit yang menunjukkan perkembangan kondisi serupa cenderung memiliki kelanjutan degradasi dan umur sisa yang serupa. Oleh karena itu, unit *test* dibandingkan dengan unit-unit *run-to-failure* pada data *train* untuk memperoleh referensi historis yang paling relevan (Schwartz et al., 2022; Lin et al., 2023).

Pada penelitian ini, perbandingan tidak dilakukan langsung menggunakan data sensor mentah, tetapi menggunakan HI parsial unit *test* dan model degradasi polinomial unit *train*. Penggunaan representasi HI membuat informasi dari beberapa sensor dapat dibandingkan dalam bentuk *trajectory* kondisi yang lebih ringkas. Sementara itu, model polinomial menyediakan kurva referensi yang dapat dievaluasi pada rentang waktu pengamatan yang sama dengan unit *test*. Estimasi dilakukan pada tiga *breakpoint*, yaitu 50%, 70%, dan 90% dari panjang *trajectory* unit *test*. Penggunaan *breakpoint* bertujuan menyimulasikan kondisi ketika hanya sebagian riwayat operasi unit yang tersedia. Dengan skenario tersebut, kemampuan model dapat diamati pada tahap awal, menengah, dan mendekati akhir umur operasi.

Pada setiap *breakpoint*, HI unit *test* diambil sampai titik pengamatan tertentu, kemudian dibandingkan dengan kurva polinomial setiap unit *train* pada rentang pengamatan yang sama. Tingkat kemiripan dinilai menggunakan *residual distance*, yaitu jarak antara HI parsial unit *test* dan nilai HI yang dihasilkan oleh model degradasi unit *train*. Pendekatan berbasis jarak memungkinkan unit historis diurutkan berdasarkan kedekatan pola degradasinya terhadap unit yang diamati (Lin et al., 2023). Bentuk *residual distance* ditunjukkan pada persamaan (3.6).

$$d_i = \|h_{\text{now}} - h_{\text{ref},i}\|_2 \quad (3.6)$$

Pada persamaan (3.6), d_i merupakan *residual distance* terhadap unit *train* ke- i , h_{now} merupakan vektor HI parsial unit *test*, dan $h_{\text{ref},i}$ merupakan vektor HI referensi dari model polinomial unit *train* ke- i . Nilai *residual distance* yang lebih kecil menunjukkan bahwa pola degradasi unit *train* tersebut lebih mirip dengan unit *test*. Seluruh unit *train* kemudian diurutkan berdasarkan jaraknya untuk memperoleh *nearest neighbors*.

Metode KNN memanfaatkan informasi lokal dari sejumlah data referensi terdekat tanpa membentuk satu model global yang berlaku bagi seluruh unit. Karakteristik ini sesuai untuk data degradasi yang dapat memiliki pola berbeda antarunit karena prediksi dibentuk dari referensi historis yang paling mendekati kondisi unit *test* (Berthold Michael R., 2020; Lin et al., 2023). Pada penelitian ini, jumlah tetangga ditetapkan sebanyak $k = 5$.

Lima unit *train* dengan *residual distance* terkecil digunakan untuk membentuk kandidat RUL. Kandidat RUL masing-masing *neighbor* dihitung berdasarkan panjang penuh *trajectory* historisnya dan proporsi pengamatan unit *test* yang telah tersedia pada suatu *breakpoint*. Perhitungan kandidat RUL ditunjukkan pada persamaan (3.7).

$$RUL_m = N_{\text{train},m} \left(1 - \frac{k_{\text{obs}}}{N_{\text{test}}} \right) \quad (3.7)$$

Pada persamaan (3.7), RUL_m merupakan kandidat RUL dari *neighbor* ke- m , $N_{\text{train},m}$ merupakan panjang penuh *trajectory* unit *train* ke- m , k_{obs} merupakan jumlah titik unit *test* yang telah diamati sampai *breakpoint*, dan N_{test} merupakan panjang penuh *trajectory* unit *test*. Dengan pendekatan ini, setiap unit historis menghasilkan kandidat RUL berdasarkan panjang umur historis dan proporsi pengamatan unit *test*.

Kandidat RUL tidak langsung dirata-ratakan karena setiap *neighbor* memiliki tingkat kemiripan yang berbeda. Oleh karena itu, digunakan *weighted KNN* dengan metode *inverse distance weighting*. Pendekatan ini memberikan bobot lebih besar kepada *neighbor* dengan jarak lebih kecil sehingga referensi yang paling mirip memberikan pengaruh lebih besar terhadap prediksi (Seyghaly et al., 2024). Perhitungan bobot ditunjukkan pada persamaan (3.8).

$$w_m = \frac{\frac{1}{d_m + \varepsilon}}{\sum_{r=1}^k \frac{1}{d_r + \varepsilon}} \quad (3.8)$$

Pada persamaan (3.8), w_m merupakan bobot *neighbor* ke- m , d_m merupakan *residual distance* dari *neighbor* tersebut, dan ε merupakan konstanta kecil untuk mencegah pembagian dengan nol. Bobot dinormalisasi sehingga jumlah seluruh bobot dari lima *neighbors* sama dengan satu. Selain bobot dan kandidat RUL, kontribusi setiap *neighbor* terhadap hasil estimasi juga dihitung. Nilai kontribusi diperoleh dari perkalian antara bobot dan kandidat RUL masing-masing *neighbor* yang ditunjukkan pada persamaan berikut.

$$C_m = w_m \cdot RUL_m \quad (3.9)$$

Pada persamaan (3.9), C_m merupakan kontribusi *neighbor* ke- m . Nilai tersebut menunjukkan pengaruh setiap unit historis terhadap prediksi akhir. Dengan

demikian, kontribusi dipengaruhi oleh tingkat kemiripan pola degradasi dan kandidat RUL dari *neighbor* yang bersangkutan. Setelah kontribusi seluruh *neighbors* diperoleh, estimasi akhir RUL dihitung sebagai jumlah kontribusi dari lima unit terdekat, sebagaimana ditunjukkan pada Persamaan (3.10).

$$\widehat{RUL} = \sum_{m=1}^k C_m \quad (3.10)$$

atau secara ekuivalen dapat dituliskan sebagai berikut.

$$\widehat{RUL} = \sum_{m=1}^k w_m \cdot RUL_m \quad (3.11)$$

persamaan (3.10) dan (3.11) menunjukkan bahwa estimasi akhir merupakan agregasi berbobot dari kandidat RUL lima unit historis terdekat. Berbeda dari KNN regresi dengan rata-rata biasa, pendekatan berbobot mempertimbangkan perbedaan tingkat kedekatan setiap *neighbor* sehingga informasi dari unit yang paling mirip memperoleh penekanan lebih besar (Seyghaly et al., 2024).

Istilah *residual-based* digunakan karena kedekatan antarunit ditentukan berdasarkan selisih antara HI parsial unit *test* dan kurva degradasi polinomial unit *train*. Sementara itu, istilah *weighted KNN* menunjukkan bahwa kandidat RUL tidak diperlakukan secara sama, tetapi diberi bobot berdasarkan *residual distance*. Kombinasi tersebut memungkinkan model memanfaatkan beberapa pola historis sekaligus tanpa mengabaikan perbedaan tingkat kemiripannya (Schwartz et al., 2022; Lin et al., 2023).

Output tahap ini berupa estimasi RUL setiap unit *test* pada *breakpoint* 50%, 70%, dan 90%. Proses estimasi juga menghasilkan urutan *nearest neighbors*, *residual distance*, bobot, kandidat RUL, dan kontribusi setiap unit historis. Informasi tersebut digunakan untuk menjelaskan pembentukan prediksi secara rinci, sedangkan performa estimasi yang dihasilkan dievaluasi menggunakan metrik pada subbab berikutnya.

3.11 Evaluasi Model Estimator RUL

Setelah nilai estimasi *Remaining Useful Life* (RUL) diperoleh, tahap berikutnya adalah mengevaluasi performa model dengan membandingkan hasil prediksi terhadap RUL aktual pada data *test*. Evaluasi dilakukan pada *breakpoint*

50%, 70%, dan 90% untuk melihat kemampuan model ketika informasi degradasi yang tersedia masih terbatas hingga mendekati akhir umur operasi.

Metrik utama yang digunakan adalah *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE), dengan perhitungan mengacu pada Persamaan (2.11) sampai Persamaan (2.13). MAE menunjukkan rata-rata besar kesalahan dalam satuan RUL, RMSE memberikan penalti lebih besar terhadap prediksi dengan kesalahan tinggi, sedangkan MAPE menunjukkan kesalahan relatif terhadap RUL aktual dalam bentuk persentase.

Pada penelitian ini, MAPE digunakan sebagai dasar target performansi karena memberikan ukuran kesalahan relatif yang mudah diinterpretasikan pada unit dengan panjang umur berbeda. Model dinilai memenuhi target apabila menghasilkan MAPE kurang dari 30%. Sementara itu, MAE dan RMSE tetap digunakan untuk menunjukkan besarnya kesalahan absolut dan mengidentifikasi kemungkinan adanya prediksi yang menyimpang jauh.

Selain ketiga metrik utama tersebut, MAEfrac dan RMSEfrac dihitung sebagai metrik pendukung berdasarkan Persamaan (2.14) dan Persamaan (2.15). Kedua metrik tersebut menormalisasi kesalahan terhadap umur total masing-masing unit, sehingga digunakan untuk mendukung perbandingan dengan penelitian terdahulu yang menyajikan performa dalam bentuk fraksi umur total. MAEfrac dan RMSEfrac tidak menggantikan MAPE, tetapi memberikan sudut pandang tambahan terhadap besar kesalahan model.

Evaluasi pada beberapa *breakpoint* digunakan untuk mengamati perubahan performa seiring bertambahnya panjang *trajectory* yang tersedia. Pada *breakpoint* yang lebih tinggi, informasi pola degradasi menjadi lebih lengkap sehingga kemiripan terhadap unit historis diharapkan dapat ditentukan secara lebih akurat.

Output tahap ini berupa nilai MAE, RMSE, MAPE, MAEfrac, dan RMSEfrac pada setiap *breakpoint*. Ketiga metrik utama digunakan untuk menilai performa model, sedangkan MAEfrac dan RMSEfrac digunakan sebagai pendukung analisis perbandingan dengan penelitian sebelumnya. Hasil evaluasi tersebut disajikan dan dibahas lebih lanjut pada Bab IV.

3.12 Dashboard Prediksi RUL

Setelah model estimasi *Remaining Useful Life* (RUL) dibangun, tahap terakhir penelitian adalah merancang dashboard sebagai media antarmuka untuk mengintegrasikan proses analisis dalam satu sistem. Dashboard memungkinkan pengguna memuat data unit, menjalankan model, dan melihat hasil prediksi secara lebih terstruktur tanpa hanya bergantung pada tampilan lingkungan pemrograman.

Dashboard menerima masukan berupa data historis satu unit sistem traksi. Data tersebut kemudian diproses melalui tahapan identifikasi kondisi operasional, normalisasi sensor berdasarkan *cluster*, pembentukan *Health Indicator* (HI), pemodelan degradasi polinomial, dan estimasi RUL menggunakan *residual-based weighted KNN*. Dengan demikian, dashboard mengintegrasikan proses inferensi model dari data masukan hingga keluaran prediksi.

Fitur utama dashboard diawali dengan pemuatan file data unit yang akan dianalisis. Setelah data berhasil dibaca, dashboard menampilkan ringkasan yang mencakup identitas data, jumlah pengamatan, *cluster* operasional mayoritas, metode estimasi, dan nilai prediksi RUL. Informasi tersebut memberikan gambaran umum mengenai data masukan dan hasil analisis yang diperoleh. Selain keluaran numerik, dashboard menampilkan visualisasi HI dan *polynomial fit* dari unit masukan. Visualisasi ini menunjukkan *trajectory* kondisi unit serta representasi polinomial yang digunakan dalam proses estimasi. Dengan demikian, pengguna dapat mengamati pola degradasi yang mendasari nilai prediksi RUL.

Dashboard juga menyajikan informasi *nearest neighbors* dari data *train* yang menjadi referensi estimasi. Informasi tersebut meliputi identitas unit historis, *residual distance*, bobot, kandidat RUL, dan kontribusi setiap *neighbor* terhadap prediksi akhir. Untuk mendukung interpretasi, *trajectory* unit masukan dibandingkan secara visual dengan beberapa unit historis terdekat sehingga tingkat kemiripan pola degradasinya dapat diamati. Output tahap ini berupa dashboard yang mengintegrasikan pemuatan data, pemrosesan model, visualisasi degradasi, dan penyajian estimasi RUL. Dashboard tersebut digunakan sebagai bentuk implementasi model agar proses pembentukan prediksi dan referensi historis yang mendasarinya dapat ditampilkan secara lebih mudah dipahami.

BAB IV

PEMBAHASAN

4.1 Hasil *Pre-Processing* Data

Tahap *pre-processing* pada penelitian ini dilakukan untuk menyiapkan data sebelum masuk ke proses pembentukan *Health Indicator* dan estimasi *Remaining Useful Life* (RUL). Pada tahap ini, data dibagi menjadi data *train* dan data *test*, dilanjutkan dengan pengujian variabel operasi untuk *clustering*, segmentasi kondisi operasional menggunakan *K-Means*, serta normalisasi sensor per *cluster*. Tahapan ini dilakukan agar data yang digunakan pada proses selanjutnya menjadi lebih homogen dan lebih representatif terhadap pola degradasi sistem.

4.1.1 *Splitting* Data Menjadi Unit Train dan Test

Pada penelitian ini, tahap awal pada proses *pre-processing* adalah melakukan pembagian data menjadi data *train* dan data *test*. Pembagian ini dilakukan pada level unit, sehingga setiap unit *run-to-failure* tetap berada secara utuh dalam satu kelompok data dan tidak terpecah ke dalam data *train* dan data *test* secara bersamaan. Pendekatan ini digunakan agar *trajectory* degradasi tiap unit tetap terjaga dan proses pengujian model dapat dilakukan tanpa menimbulkan *data leakage*. Data *train* selanjutnya digunakan untuk membangun seluruh komponen model, sedangkan data *test* digunakan untuk mengevaluasi kemampuan model pada unit yang tidak dilibatkan dalam tahap pembentukan model. Hasil pembagian data tersebut ditunjukkan pada tabel 4.1.

Tabel 4.1 Pembagian ID Unit ke dalam Data *Train* dan Data *Test*

Data Train	Data Test
1, 2, 3, 4, 5, 6, 7, 9, 10, 12, 14, 15, 16, 17, 19, 20, 23, 24, 25, 26, 27, 29, 30, 31, 32, 34, 35, 36, 37, 38, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 55, 56, 57, 58, 59, 60, 61, 62, 63, 65, 66, 67, 68, 70, 71, 72, 73, 74, 75, 76, 77, 79, 82, 83, 84, 85, 86, 87, 88, 89, 90, 92, 93, 94, 96, 97, 99, 101, 102, 103, 104, 105, 106, 107, 108, 109, 112, 114, 115, 116, 117, 118, 119, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132	8, 11, 13, 18, 21, 22, 28, 33, 39, 40, 53, 54, 64, 69, 78, 80, 81, 91, 95, 98, 100, 110, 111, 113, 120, 133

Berdasarkan tabel 4.1, dapat dilihat bahwa seluruh unit telah teralokasi secara jelas ke dalam kelompok data *train* dan kelompok data *test*. Jumlah unit pada data *train* adalah 107 unit, sedangkan jumlah unit pada data *test* adalah 26 unit. Dengan demikian, proporsi pembagian data yang dihasilkan adalah sekitar 80,45% untuk data *train* dan 19,55% untuk data *test*. Hasil ini menunjukkan bahwa pembagian unit yang dilakukan telah menghasilkan rasio yang sebanding dengan proporsi pembagian data yang dirancang pada tahap metode. Pembagian ini selanjutnya menjadi dasar bagi seluruh tahapan analisis berikutnya, yaitu pembentukan model pada data *train* dan pengujian performa estimasi RUL pada data *test*.

4.1.2 Pengujian Variabel Operasi Untuk *Clustering*

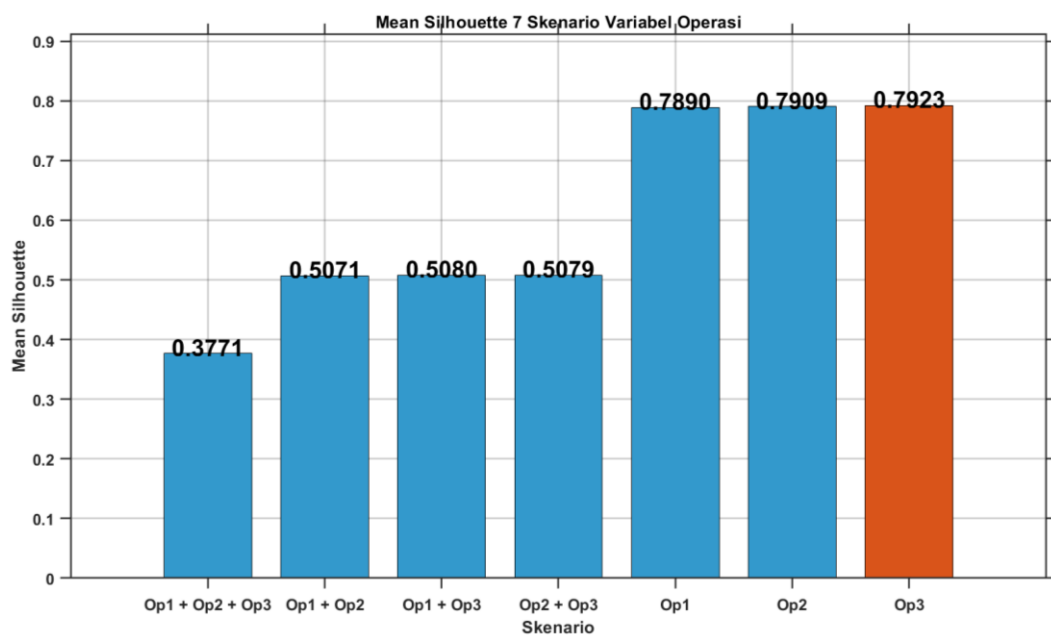
Setelah data *train* dan data *test* diperoleh, tahap berikutnya adalah melakukan pengujian variabel operasi yang akan digunakan pada proses *clustering* kondisi operasional. Tahap ini dilakukan karena kualitas *clustering* sangat dipengaruhi oleh variabel yang digunakan sebagai masukan. Pada penelitian ini, *clustering* kondisi operasional tidak dilakukan sekadar untuk membagi data ke dalam beberapa kelompok, tetapi untuk memisahkan pengaruh kondisi operasi dari perubahan sinyal sensor agar variasi akibat mode operasi tidak terbaca sebagai degradasi performa sistem traksi itu sendiri. Jika variabel operasi yang dipilih kurang mampu merepresentasikan perbedaan kondisi operasional, maka pemisahan *cluster* yang terbentuk menjadi kurang jelas dan dapat memengaruhi tahap analisis berikutnya, terutama pada proses normalisasi sensor per *cluster* dan pembentukan *Health Indicator*.

Pada penelitian ini, pengujian dilakukan terhadap beberapa skenario kombinasi variabel operasi, yaitu kombinasi tiga variabel operasi, kombinasi dua variabel operasi, dan masing-masing variabel operasi secara tunggal. Kualitas hasil *clustering* pada setiap skenario dievaluasi menggunakan nilai *mean silhouette*. Sesuai dengan dasar teori pada Bab II, nilai *silhouette* untuk setiap data dihitung dengan membandingkan dua komponen utama, yaitu *cohesion* dan *separation*. *Cohesion* menyatakan rata-rata jarak suatu data terhadap data lain dalam *cluster* yang sama, sedangkan *separation* menyatakan rata-rata jarak data tersebut terhadap *cluster* lain yang paling dekat. Dengan demikian, nilai *silhouette* menunjukkan

apakah suatu data lebih dekat dengan *cluster*-nya sendiri atau justru lebih dekat dengan *cluster* lain. Setelah nilai *silhouette* untuk seluruh data diperoleh, nilai tersebut dirata-ratakan sehingga menghasilkan *mean silhouette* sebagai indikator umum kualitas hasil *clustering*. Semakin tinggi nilai *mean silhouette*, maka secara umum *cohesion* antar data dalam satu *cluster* semakin baik dan *separation* terhadap *cluster* lain semakin jelas, sehingga pemisahan *cluster* dapat dinilai semakin baik. Melalui pendekatan ini, pemilihan variabel operasi dapat dilakukan secara lebih objektif sebelum tahap *K-Means clustering* diterapkan pada data *train*. Hasil pengujian variabel operasi berupa nilai *mean silhouette* pada penelitian ini ditampilkan pada tabel 4.2 dan divisualisasikan pada gambar 4.1.

Tabel 4.2 Hasil Pengujian Variabel Operasi Berdasarkan Nilai *Mean Silhouette*

No.	Skenario Variabel Operasi	Mean Silhouette
1	<i>Op1 + Op2 + Op3</i>	0,3771
2	<i>Op1 + Op2</i>	0,5071
3	<i>Op1 + Op3</i>	0,5080
4	<i>Op2 + Op3</i>	0,5079
5	<i>Op1</i>	0,7890
6	<i>Op2</i>	0,7909
7	<i>Op3</i>	0,7923



Gambar 4.1 Perbandingan *Mean Silhouette* Tujuh Skenario Variabel Operasi

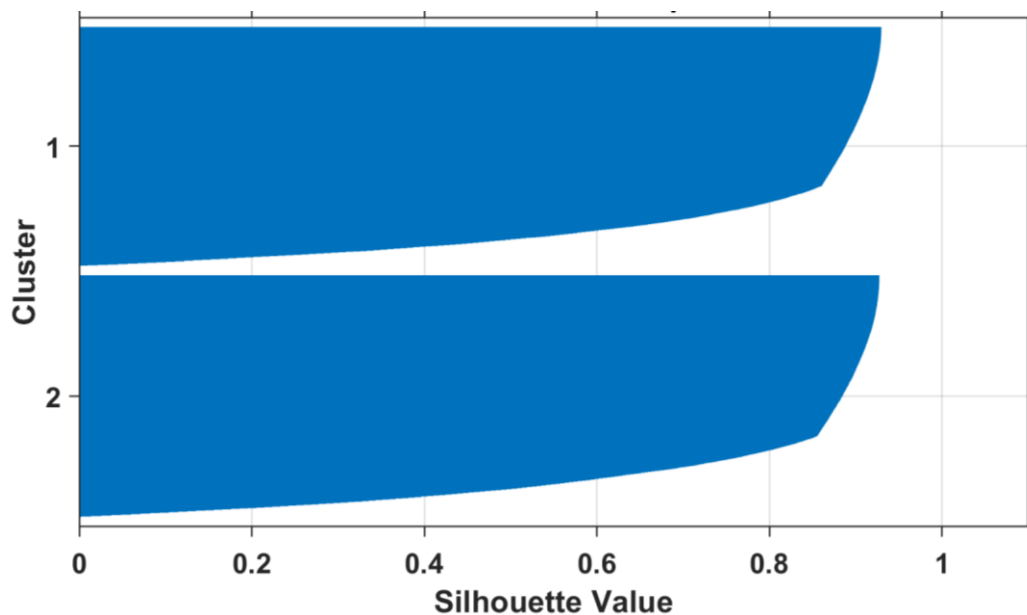
Berdasarkan tabel 4.2 dan gambar 4.1, dapat dilihat bahwa skenario yang menggunakan Op3 sebagai variabel tunggal menghasilkan nilai *mean silhouette* tertinggi, yaitu sebesar 0,7923. Nilai ini sedikit lebih tinggi dibandingkan skenario Op2 sebesar 0,7909 dan Op1 sebesar 0,7890. Sementara itu, seluruh skenario kombinasi dua variabel hanya menghasilkan nilai *mean silhouette* di sekitar 0,50, dan skenario kombinasi tiga variabel memberikan nilai paling rendah, yaitu 0,3771. Hasil ini menunjukkan bahwa pada dataset penelitian ini, penggunaan satu variabel operasi tunggal justru menghasilkan kualitas pemisahan *cluster* yang lebih baik dibandingkan penggunaan kombinasi dua atau tiga variabel operasi.

Jika ditinjau dari konsep *mean silhouette*, hasil tersebut menunjukkan bahwa pada skenario Op3, secara rata-rata setiap data memiliki *cohesion* yang lebih baik terhadap *cluster*-nya sendiri dan *separation* yang lebih jelas terhadap *cluster* lain dibandingkan skenario lainnya. Dengan kata lain, penggunaan Op3 memberikan kondisi di mana data dalam satu *cluster* cenderung lebih saling dekat, sedangkan jaraknya terhadap *cluster* terdekat lainnya cenderung lebih besar. Sebaliknya, pada skenario kombinasi dua atau tiga variabel, penambahan dimensi variabel operasi tidak selalu meningkatkan kualitas pemisahan *cluster*. Pada kondisi tertentu, kombinasi beberapa variabel justru dapat menyebabkan batas antar-*cluster* menjadi kurang tegas, sehingga nilai *silhouette* masing-masing data menurun dan *mean silhouette* keseluruhan menjadi lebih rendah. Hal ini menunjukkan bahwa semakin banyak variabel yang digunakan tidak selalu menghasilkan segmentasi yang lebih baik, terutama apabila variabel tambahan tersebut tidak memperkuat perbedaan kondisi operasional secara nyata.

Apabila dibandingkan lebih lanjut, perbedaan antara tiga skenario variabel tunggal memang tidak terlalu besar. Namun demikian, karena Op3 tetap memberikan nilai *mean silhouette* tertinggi di antara seluruh skenario yang diuji, maka variabel tersebut dipilih sebagai variabel operasi yang digunakan pada tahap *clustering* kondisi operasional. Pemilihan ini tidak hanya didasarkan pada hasil statistik berupa nilai *mean silhouette* tertinggi, tetapi juga pada makna fisik variabel tersebut terhadap sistem traksi. Pada penelitian ini, Op3 merepresentasikan torsi motor. Secara fisik, torsi motor berkaitan langsung dengan kebutuhan gaya gerak sistem traksi dalam menggerakkan rangkaian kereta, sehingga perubahan nilainya

lebih dekat dengan variasi beban kerja sistem dibandingkan variabel operasi lainnya. Dalam kondisi operasi riil LRT, kebutuhan torsi dapat berubah mengikuti tuntutan percepatan, perlambatan, tahanan gerak, serta variasi beban angkut, sehingga torsi motor menjadi variabel yang lebih representatif untuk membedakan kondisi operasional. Oleh karena itu, pemilihan Op3 sebagai variabel operasi terpilih pada penelitian ini memiliki dasar yang kuat, baik dari sisi kualitas *clustering* maupun dari sisi relevansi fisiknya terhadap kerja sistem traksi.

Untuk memperkuat interpretasi terhadap skenario terpilih, visualisasi *silhouette plot* untuk skenario Op3 dapat ditampilkan pada gambar 4.2. Visualisasi ini digunakan untuk menunjukkan distribusi nilai *silhouette* pada masing-masing *cluster* untuk variabel operasi terpilih, sehingga kualitas pemisahan data tidak hanya dilihat dari nilai rata-ratanya, tetapi juga dari pola distribusi nilai tiap data pada *cluster* yang terbentuk.



Gambar 4.2 *Silhouette Plot* Hasil *Clustering* Berdasarkan Op3

Berdasarkan visualisasi *silhouette plot* tersebut, sebagian besar data pada kedua *cluster* memiliki nilai *silhouette* positif. Kondisi ini menunjukkan bahwa secara umum data lebih dekat dengan *cluster*-nya sendiri dibandingkan dengan *cluster* lain yang paling dekat, sehingga hasil segmentasi yang terbentuk dapat dianggap cukup baik. Dominasi nilai *silhouette* positif ini juga konsisten dengan nilai *mean silhouette* Op3 yang tinggi pada Tabel 4.2. Dengan demikian, hasil pengujian variabel operasi pada tahap ini memberikan dasar yang lebih kuat dalam

penetapan variabel operasi yang digunakan pada proses *clustering* kondisi operasional selanjutnya. Secara keseluruhan, tahap pengujian ini memastikan bahwa variabel operasi yang dipilih benar-benar mampu memisahkan kondisi operasi dengan baik, sehingga pada tahap berikutnya pengaruh kondisi operasi dapat dikurangi dan informasi yang dipertahankan lebih merepresentasikan degradasi sistem traksi.l selanjutnya.

4.1.3 Clustering Kondisi Operasional dengan *K-Means*

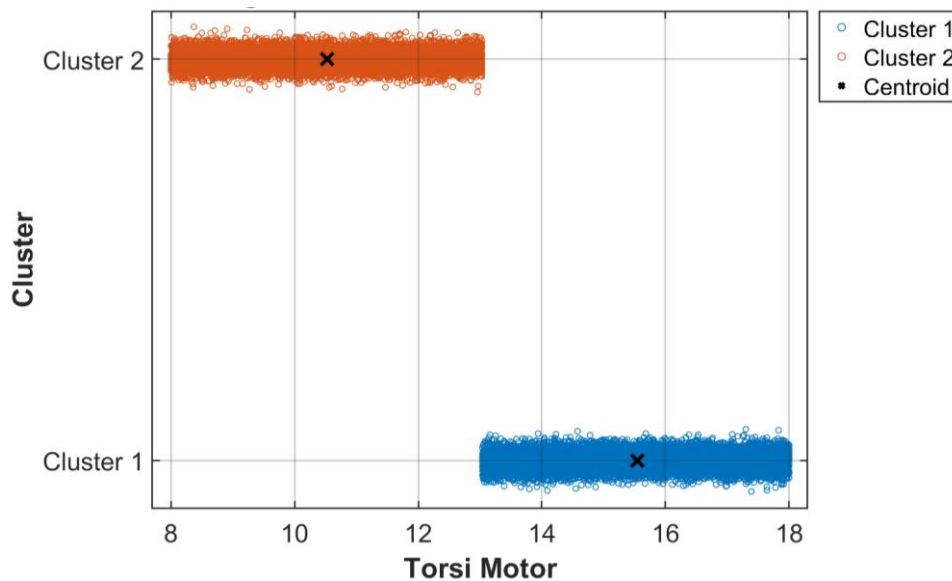
Setelah variabel operasi terpilih ditetapkan melalui pengujian *mean silhouette*, tahap berikutnya adalah melakukan *clustering* kondisi operasional menggunakan metode *K-Means*. Pada penelitian ini, proses *clustering* dilakukan pada data *train* dengan menggunakan variabel operasi terpilih, yaitu Op3. Tahap ini bertujuan untuk memisahkan data ke dalam beberapa kelompok kondisi operasional yang lebih homogen sebelum dilakukan proses normalisasi sensor dan pembentukan *Health Indicator*. Dengan demikian, data yang digunakan pada tahap berikutnya tidak lagi diperlakukan sebagai satu kumpulan data operasi yang bercampur, tetapi sebagai beberapa kelompok data yang memiliki karakteristik operasi yang lebih beragam.

Penggunaan *clustering* kondisi operasional pada penelitian ini dilakukan karena perubahan nilai sensor pada sistem traksi tidak hanya dipengaruhi oleh degradasi komponen, tetapi juga oleh variasi kondisi operasi. Jika seluruh data langsung digunakan tanpa pemisahan kondisi operasional, maka variasi sinyal akibat mode operasi yang berbeda berisiko tercampur dengan variasi sinyal akibat degradasi. Kondisi ini dapat menyebabkan perubahan sinyal yang sebenarnya berasal dari perbedaan beban kerja atau mode operasi justru terbaca sebagai penurunan performa sistem traksi itu sendiri. Oleh karena itu, proses segmentasi kondisi operasional menjadi tahap penting agar analisis pada tahap berikutnya lebih berfokus pada informasi degradasi, bukan pada variasi sinyal akibat perbedaan operasi. Pada penelitian ini, hasil *clustering* selanjutnya digunakan sebagai dasar normalisasi sensor per *cluster*, sehingga data dengan kondisi operasi tinggi dinormalisasi terhadap kelompok operasi tinggi, sedangkan data dengan kondisi operasi rendah dinormalisasi terhadap kelompok operasi rendah. Melalui

pendekatan ini, proses normalisasi *z-score* dapat menghasilkan data yang lebih homogen dalam masing-masing *cluster*.

Pada penelitian ini, metode *K-Means* digunakan karena mampu mengelompokkan data berdasarkan kedekatan karakteristik operasional secara relatif sederhana dan efisien. Sesuai dengan pembahasan pada Bab II, *K-Means* bekerja dengan menempatkan setiap data ke *cluster* yang memiliki *centroid* paling dekat, kemudian memperbarui posisi *centroid* secara iteratif hingga tercapai kondisi konvergen. Dalam konteks penelitian ini, proses tersebut digunakan untuk membentuk kelompok data operasi yang lebih homogen berdasarkan variabel operasi terpilih. Pada penelitian ini, proses *K-Means* dilakukan dengan jumlah *cluster* sebanyak dua. Penetapan dua *cluster* digunakan untuk membagi data operasi ke dalam dua kelompok utama berdasarkan kedekatan karakteristik variabel operasi terpilih, sehingga variasi kondisi operasi pada data *train* dapat direpresentasikan ke dalam dua kelompok utama yang berbeda sebelum proses normalisasi sensor dilakukan.

Hasil *clustering* kondisi operasional pada penelitian ini ditunjukkan pada gambar 4.3. Visualisasi tersebut memperlihatkan pembagian data *train* ke dalam dua *cluster* berdasarkan variabel operasi terpilih, yaitu Op3. Selain itu, titik pusat masing-masing *cluster* atau *centroid* juga ditampilkan untuk menunjukkan posisi representatif dari setiap kelompok data yang terbentuk.



Gambar 4.3 Hasil *Clustering* Kondisi Operasional Menggunakan Variabel *Op3*

Berdasarkan gambar 4.3, data *train* berhasil dipisahkan ke dalam dua kelompok kondisi operasional yang berbeda. Proses ini menunjukkan bahwa variabel Op3 mampu menghasilkan segmentasi data yang cukup jelas, sejalan dengan hasil pengujian pada subbab sebelumnya yang menunjukkan bahwa variabel ini memiliki nilai *mean silhouette* tertinggi dibandingkan skenario lain. Dengan demikian, pemilihan Op3 sebagai variabel operasi untuk *clustering* terbukti konsisten baik pada tahap evaluasi skenario maupun pada tahap implementasi segmentasi data. Dalam penelitian ini, Op3 merepresentasikan torsi motor, sehingga pemisahan *cluster* yang terbentuk dapat dipahami sebagai pemisahan dua kelompok operasi utama berdasarkan tingkat kebutuhan torsi. Secara fisik, hal ini penting karena torsi motor berkaitan langsung dengan beban kerja sistem traksi, sehingga perubahan kondisi operasi dapat lebih tepat dipisahkan sebelum analisis degradasi dilakukan.

Untuk memperjelas hasil segmentasi yang diperoleh, informasi ringkas mengenai hasil *K-Means clustering* disajikan pada tabel 4.3. Tabel ini memuat jumlah *cluster* yang digunakan, variabel operasi terpilih, nilai *centroid* masing-masing *cluster*, serta jumlah sampel pada tiap *cluster*.

Tabel 4.3 Ringkasan Hasil *K-Means Clustering* Kondisi Operasional

Komponen	Nilai
Variabel operasi terpilih	<i>Op3</i>
Jumlah <i>cluster</i>	2
<i>Centroid</i> Cluster 1	15,5465
<i>Centroid</i> Cluster 2	10,5300
Jumlah sampel Cluster 1	13.102 sampel
Jumlah sampel Cluster 2	13.279 sampel

Berdasarkan tabel 4.3, dapat dilihat bahwa proses *K-Means clustering* dengan variabel operasi Op3 berhasil membagi data *train* ke dalam dua *cluster* kondisi operasional. Hasil ini ditunjukkan oleh terbentuknya dua nilai *centroid*, yaitu 15,5465 untuk Cluster 1 dan 10,5300 untuk Cluster 2. Dalam konteks *K-Means*, nilai *centroid* merepresentasikan pusat kelompok atau nilai representatif dari variabel operasi pada masing-masing *cluster*. Dengan demikian, kedua nilai tersebut menunjukkan bahwa data pada Cluster 1 cenderung memiliki karakteristik

operasi Op3 yang berada di sekitar 15,5465, sedangkan data pada Cluster 2 cenderung berada di sekitar 10,5300. Perbedaan nilai *centroid* tersebut menunjukkan bahwa kedua *cluster* yang terbentuk memiliki karakteristik operasional yang berbeda, sehingga segmentasi yang dilakukan tidak membagi data secara acak, tetapi benar-benar memisahkan data ke dalam dua kelompok operasi berdasarkan kedekatan nilai torsi motor.

Selain ditinjau dari nilai *centroid*, interpretasi terhadap kedua *cluster* juga dapat diperkuat melalui rentang nilai Op3 pada data *train*. Pada penelitian ini, nilai Op3 pada data *train* berada pada rentang 8 sampai 18. Karena Op3 merepresentasikan torsi motor, maka rentang tersebut menunjukkan variasi kondisi operasi sistem traksi yang diamati pada dataset. Dalam konteks ini, *centroid* Cluster 1 sebesar 15,5465 berada pada bagian rentang yang relatif lebih tinggi, sedangkan *centroid* Cluster 2 sebesar 10,5300 berada pada bagian rentang yang relatif lebih rendah. Oleh karena itu, kedua *cluster* pada penelitian ini dapat diinterpretasikan sebagai kelompok operasi dengan level torsi motor yang relatif lebih tinggi dan relatif lebih rendah terhadap distribusi data yang diamati.

Interpretasi tersebut tidak dimaksudkan sebagai batas absolut teknis antara kondisi torsi tinggi dan torsi rendah menurut spesifikasi pabrikan, tetapi sebagai pemisahan relatif berdasarkan distribusi data *train* dan hasil segmentasi *K-Means*. Dengan demikian, istilah *cluster* operasi dengan torsi relatif lebih tinggi dan *cluster* operasi dengan torsi relatif lebih rendah pada penelitian ini digunakan untuk menunjukkan perbedaan level kondisi operasi dalam dataset, bukan untuk menetapkan klasifikasi operasional baku di luar data penelitian. Meskipun demikian, pemisahan relatif ini tetap penting karena menunjukkan bahwa data *train* memang memiliki dua kelompok utama kondisi operasi yang berbeda, sehingga analisis berikutnya dapat dilakukan pada kelompok yang lebih homogen.

Selain nilai *centroid*, Tabel 4.3 juga menunjukkan jumlah sampel pada masing-masing *cluster*, yaitu 13.102 sampel pada Cluster 1 dan 13.279 sampel pada Cluster 2. Jumlah sampel ini menunjukkan banyaknya titik data *train* yang masuk ke masing-masing kelompok setelah proses segmentasi dilakukan. Dari hasil tersebut, dapat dilihat bahwa distribusi jumlah sampel antara kedua *cluster* relatif seimbang. Kondisi ini menunjukkan bahwa proses *clustering* tidak menghasilkan

satu *cluster* yang terlalu dominan dibandingkan *cluster* lainnya, sehingga hasil segmentasi dapat dianggap cukup representatif untuk digunakan pada tahap analisis berikutnya.

Informasi mengenai nilai *centroid* dan jumlah sampel per *cluster* menjadi penting karena label *cluster* yang diperoleh pada tahap ini akan digunakan sebagai dasar pada proses normalisasi sensor per *cluster*. Pada tahap berikutnya, setiap data tidak dinormalisasi secara global, tetapi dinormalisasi sesuai dengan kelompok operasionalnya masing-masing. Dengan pendekatan ini, data yang berada pada kelompok operasi dengan torsi lebih tinggi akan dinormalisasi menggunakan parameter statistik dari *cluster* operasi tinggi, sedangkan data yang berada pada kelompok operasi dengan torsi lebih rendah akan dinormalisasi menggunakan parameter statistik dari *cluster* operasi rendah. Tujuan pendekatan ini adalah agar perbedaan *baseline* akibat kondisi operasi tidak ikut terbawa ke dalam proses normalisasi. Dengan demikian, hasil normalisasi *z-score* yang diperoleh menjadi lebih homogen dalam masing-masing *cluster* dan lebih mencerminkan penyimpangan relatif terhadap kondisi operasi sejenis, bukan terhadap seluruh data operasi yang bercampur. Secara keseluruhan, tahap *clustering* kondisi operasional pada penelitian ini bukan ditujukan untuk mendeteksi degradasi secara langsung, tetapi untuk memisahkan pengaruh kondisi operasi agar informasi yang digunakan pada tahap prognostik selanjutnya lebih benar-benar merepresentasikan degradasi sistem traksi.

4.1.4 Normalisasi Sensor Per Cluster

Setelah data *train* berhasil dipisahkan ke dalam dua *cluster* kondisi operasional, tahap berikutnya adalah melakukan normalisasi sensor per *cluster*. Tahap ini dilakukan karena nilai sensor pada sistem traksi tidak hanya dipengaruhi oleh degradasi komponen, tetapi juga oleh perbedaan kondisi operasi. Jika seluruh data sensor langsung digunakan tanpa mempertimbangkan hasil *clustering*, maka perbedaan level sinyal akibat mode operasi yang berbeda masih dapat terbawa ke tahap berikutnya dan berisiko memengaruhi pembentukan *Health Indicator*. Dalam kondisi tersebut, perubahan sinyal yang sebenarnya berasal dari variasi operasi dapat salah diinterpretasikan sebagai penurunan performa sistem traksi. Oleh karena itu, pada penelitian ini normalisasi dilakukan secara terpisah untuk setiap

cluster operasional agar data sensor yang dihasilkan menjadi lebih homogen dan lebih merepresentasikan penyimpangan relatif terhadap kondisi operasi sejenis.

Pada penelitian ini, parameter normalisasi dihitung menggunakan data *train* pada masing-masing *cluster*, yaitu berupa nilai rata-rata dan simpangan baku untuk setiap sensor. Nilai rata-rata digunakan sebagai pusat data pada masing-masing *cluster*, sedangkan simpangan baku digunakan untuk menunjukkan tingkat sebaran nilai sensor terhadap rata-ratanya. Kedua parameter tersebut kemudian digunakan sebagai dasar untuk mentransformasikan data sensor pada tahap normalisasi, sehingga setiap sinyal disesuaikan terhadap karakteristik operasional *cluster*-nya masing-masing. Dengan pendekatan ini, data yang berada pada kelompok operasi dengan torsi lebih tinggi dinormalisasi menggunakan parameter statistik dari *cluster* operasi tinggi, sedangkan data yang berada pada kelompok operasi dengan torsi lebih rendah dinormalisasi menggunakan parameter statistik dari *cluster* operasi rendah. Melalui cara ini, proses normalisasi *z-score* dilakukan pada kelompok data yang lebih seragam, sehingga hasil normalisasi menjadi lebih sesuai dengan konteks operasional masing-masing. Berikut nilai rata-rata dan simpangan baku sensor pada masing-masing *cluster* disajikan pada tabel 4.4 dan 4.5.

Tabel 4.4 Nilai Rata-Rata dan Simpangan Baku Sensor pada *Cluster* 1

Sensor	Rata-rata	Simpangan baku
Sensor1	579,05	37,4280
Sensor2	42,8400	3,2793
Sensor3	74,5020	2,0245
Sensor4	579,05	37,4280
Sensor5	92986	7615,6
Sensor6	52,8000	3,6783
Sensor7	1416,8	106,76
Sensor8	1201,1	119,89
Sensor9	2226,1	147,47
Sensor10	8520,1	338,95
Sensor11	266,03	138,49
Sensor12	472,76	26,3910

Tabel 4.5 Nilai Rata-Rata dan Simpangan Baku Sensor pada *Cluster 2*

Sensor	Rata-rata	Simpangan baku
Sensor1	579,66	37,2120
Sensor2	42,9120	3,2179
Sensor3	74,4910	2,0115
Sensor4	579,66	37,2120
Sensor5	92800	7420,5
Sensor6	52,8090	3,6925
Sensor7	1418,8	105,62
Sensor8	1202,9	118,84
Sensor9	2230,3	143,93
Sensor10	8526,9	335,02
Sensor11	267,25	137,98
Sensor12	472,94	26,4000

Parameter statistik pada tabel 4.4 dan 4.5 selanjutnya digunakan pada proses normalisasi sensor dengan transformasi *z-score* per *cluster*, sebagaimana dijelaskan pada Bab III. Pada pendekatan ini, setiap nilai sensor disesuaikan terhadap rata-rata dan simpangan baku dari *cluster* operasional tempat data tersebut berada, bukan terhadap seluruh data secara global. Dengan demikian, data dari *cluster* operasi tinggi dinormalisasi menggunakan parameter statistik *cluster* operasi tinggi, sedangkan data dari *cluster* operasi rendah dinormalisasi menggunakan parameter *cluster* operasi rendah. Melalui cara ini, perbedaan *baseline* sinyal akibat variasi kondisi operasi dapat dikurangi, sehingga hasil normalisasi menjadi lebih homogen dan lebih sesuai digunakan pada tahap pembentukan *Health Indicator*.

Berdasarkan Tabel 4.4 dan Tabel 4.5, masing-masing *cluster* memiliki parameter statistik sensor yang tidak sepenuhnya sama. Perbedaan ini tampak pada beberapa sensor yang menunjukkan nilai rata-rata berbeda antara *Cluster 1* dan *Cluster 2*, seperti Sensor5, Sensor7, Sensor8, Sensor9, Sensor10, dan Sensor11. Hal ini menunjukkan bahwa data pada masing-masing *cluster* memiliki karakteristik operasional yang berbeda, sehingga penggunaan satu parameter normalisasi global untuk seluruh data menjadi kurang sesuai. Jika seluruh data dinormalisasi menggunakan satu rata-rata dan satu simpangan baku yang sama, variasi akibat

operasi tinggi dan operasi rendah masih dapat tercampur dalam hasil transformasi data.

Selain nilai rata-rata, simpangan baku pada masing-masing sensor juga menunjukkan bahwa tingkat sebaran data di tiap *cluster* tidak identik. Perbedaan simpangan baku ini menunjukkan bahwa variasi sinyal sensor pada satu *cluster* dapat lebih besar atau lebih kecil dibandingkan *cluster* lainnya. Kondisi tersebut memperkuat alasan bahwa proses normalisasi perlu dilakukan berdasarkan *cluster* operasional, sehingga data sensor tidak hanya disesuaikan terhadap skala statistiknya, tetapi juga terhadap konteks operasi tempat data tersebut berasal.

Makna penting dari hasil ini adalah bahwa normalisasi pada penelitian ini tidak hanya bertujuan untuk menyamakan skala antar-sensor, tetapi juga untuk menyesuaikan transformasi data terhadap kondisi operasional masing-masing *cluster*. Dengan menggunakan rata-rata dan simpangan baku yang dihitung secara khusus pada tiap *cluster*, sinyal sensor yang telah dinormalisasi diharapkan menjadi lebih sebanding antarunit tanpa menghilangkan pola perubahan dasarnya terhadap waktu. Pendekatan ini membuat penyimpangan nilai sensor setelah normalisasi lebih mencerminkan kondisi relatif terhadap operasi sejenis, bukan sekadar akibat perbedaan mode operasi.

Dengan demikian, normalisasi sensor per *cluster* menjadi tahap penting sebelum pembentukan *Health Indicator*. Melalui penggunaan parameter normalisasi yang sesuai untuk masing-masing *cluster*, pengaruh variasi mode operasi terhadap sinyal sensor dapat dikurangi, sehingga informasi yang dipertahankan pada tahap berikutnya lebih merepresentasikan degradasi komponen daripada sekadar perbedaan kondisi operasional. Tahap ini berperan sebagai jembatan antara segmentasi kondisi operasional dan pembentukan HI, sehingga HI yang dihasilkan lebih berorientasi pada pola degradasi sistem traksi.

4.2 Hasil Pembentukan *Health Indicator*

Setelah proses *pre-processing* data selesai dilakukan, tahap berikutnya adalah membentuk *Health Indicator* (HI) sebagai representasi kondisi kesehatan sistem traksi. Tahap ini penting karena data sensor mentah masih terdiri atas banyak variabel dan belum secara langsung menunjukkan kondisi degradasi sistem dalam bentuk yang ringkas. Oleh karena itu, pada penelitian ini pembentukan HI

dilakukan melalui beberapa tahapan, yaitu pemilihan sensor yang paling *trendable*, penetapan *health condition* ideal, dan pemodelan HI dari sensor-sensor terpilih.

Secara umum, hasil pada tahap ini diharapkan sejalan dengan dasar teori pada Bab II, yaitu bahwa HI yang baik harus mampu merepresentasikan degradasi sistem dalam bentuk yang lebih sederhana, memiliki kecenderungan perubahan yang relatif konsisten, dan cukup stabil untuk digunakan pada tahap estimasi *Remaining Useful Life* (RUL). Dengan demikian, pembahasan pada subbab ini tidak hanya menunjukkan hasil pembentukan HI, tetapi juga menilai apakah karakteristik HI yang diperoleh sudah sesuai dengan konsep teoritis yang telah dijelaskan sebelumnya.

Pada penelitian ini, hasil pembentukan HI dibahas melalui tiga tahapan utama. Tahap pertama adalah pemilihan sensor *trendable* untuk menentukan sensor-sensor yang paling relevan terhadap kecenderungan degradasi. Tahap kedua adalah penetapan *health condition* ideal sebagai target pembentukan HI. Tahap ketiga adalah pemodelan HI dari sensor-sensor terpilih serta evaluasi kualitas HI yang dihasilkan. Uraian hasil pada masing-masing tahapan tersebut disajikan pada Subbab 4.2.1 sampai dengan Subbab 4.2.3.

4.2.1 Pemilihan Sensor *Trendable Polynomial*

Setelah data sensor dinormalisasi berdasarkan *cluster* kondisi operasional, tahap berikutnya adalah memilih sensor dengan kecenderungan tren degradasi paling jelas. Seleksi diperlukan karena tidak seluruh sensor menunjukkan perubahan yang relevan terhadap penurunan kondisi sistem. Penggunaan seluruh sensor tanpa seleksi berpotensi memasukkan fluktuasi yang kurang berkaitan dengan degradasi ke dalam pembentukan *Health Indicator* (HI).

Pemilihan sensor dilakukan menggunakan pendekatan *polynomial trendability* dengan polinomial orde dua. Setiap sensor dievaluasi berdasarkan kekuatan gradien kurva pada bagian akhir *trajectory*, karena perubahan kondisi cenderung semakin terlihat ketika unit mendekati akhir pengamatan. Sensor dengan nilai *trend score* lebih tinggi dipandang memiliki kecenderungan perubahan lebih kuat dan lebih layak digunakan dalam pembentukan HI. Hasil pemeringkatan seluruh sensor disajikan pada tabel 4.6.

Tabel 4.6 Peringkat Sensor Berdasarkan Nilai *Trend Score*

Peringkat	Sensor	Trend Score
1	<i>Sensor3</i>	0,0035026
2	<i>Sensor2</i>	0,0032931
3	<i>Sensor6</i>	0,0030546
4	<i>Sensor5</i>	0,0029421
5	<i>Sensor8</i>	0,0029083
6	<i>Sensor10</i>	0,0028481
7	<i>Sensor7</i>	0,0027917
8	<i>Sensor9</i>	0,0026734
9	<i>Sensor12</i>	0,0026164
10	<i>Sensor11</i>	0,0025375
11	<i>Sensor1</i>	0,0025227
12	<i>Sensor4</i>	0,0025227

Berdasarkan Berdasarkan Tabel 4.6, Sensor3 memiliki *trend score* tertinggi, diikuti oleh Sensor2, Sensor6, Sensor5, dan Sensor8. Kelima sensor tersebut dipilih untuk pembentukan HI karena menunjukkan kecenderungan perubahan paling kuat pada bagian akhir *trajectory*. Sensor terpilih masing-masing merepresentasikan arus keluaran VVVF, arus masukan motor, temperatur motor, tegangan keluaran VVVF, dan kecepatan roda. Urutan peringkat tersebut menunjukkan bahwa sensor yang terpilih tidak hanya berasal dari satu bagian sistem, tetapi tersebar pada beberapa bagian penting dalam rantai kerja sistem traksi. Hal ini menjadi penting karena perubahan kondisi sistem traksi dapat tercermin melalui beberapa parameter, baik dari sisi kelistrikan, termal, maupun mekanik.

Pemilihan lima sensor dilakukan dengan mempertimbangkan keseimbangan antara kecukupan informasi dan pengendalian variasi sinyal. Penggunaan sensor terlalu banyak berpotensi memasukkan parameter dengan *trend score* lebih rendah, sedangkan penggunaan sensor terlalu sedikit berisiko menghilangkan informasi penting. Oleh karena itu, lima sensor teratas dipilih untuk mempertahankan informasi dari beberapa bagian sistem traksi tanpa memasukkan terlalu banyak variabel dengan tren yang lebih lemah. Sensor di luar lima peringkat teratas tetap merekam perilaku sistem, tetapi kecenderungan perubahannya

terhadap degradasi relatif lebih rendah berdasarkan metode seleksi yang digunakan. Dengan pendekatan ini, pembentukan HI diarahkan pada sensor yang paling informatif terhadap perubahan kondisi, tetapi tetap mempertahankan keterwakilan dari beberapa fungsi utama sistem traksi. Selain itu, pemilihan sensor yang lebih terbatas juga membantu menjaga agar HI yang dibentuk tidak terlalu dipengaruhi oleh sinyal yang fluktuatif atau kurang konsisten terhadap arah degradasi.

Selain didukung oleh hasil *trend score*, pemilihan kelima sensor tersebut juga berkaitan dengan prinsip fisika pada rantai konversi energi sistem traksi. Feng et al. (2017) menjelaskan bahwa tegangan, arus, temperatur, dan kecepatan merupakan kelompok parameter utama yang umum dipantau pada sistem traksi kereta. Kelima sensor terpilih mencakup kelompok tersebut sehingga memberikan representasi dari sisi elektrik, termal, dan mekanik sistem traksi. Hal ini menunjukkan bahwa hasil seleksi sensor tidak hanya relevan secara statistik, tetapi juga masih dapat dikaitkan dengan proses fisik yang terjadi pada sistem. Dengan demikian, sensor yang dipilih tidak hanya memiliki nilai *trend score* tinggi, tetapi juga memiliki makna terhadap alur kerja sistem traksi.

Sensor3 dan Sensor5 masing-masing merepresentasikan arus dan tegangan keluaran VVVF, sedangkan Sensor2 merepresentasikan arus yang diterima motor traksi. Ketiga parameter tersebut berkaitan dengan daya listrik yang disalurkan dari inverter menuju motor. Secara fisik, perubahan tegangan dan arus dapat memengaruhi daya yang diterima motor, sehingga ketiga sensor tersebut relevan untuk menggambarkan perubahan kebutuhan daya dan respons elektrik selama operasi. Hubungan antara ketiga sensor ini juga menunjukkan bahwa perubahan pada sisi inverter dan motor dapat diamati secara berurutan melalui sinyal arus dan tegangan yang tersedia pada dataset.

Arus motor juga berkaitan dengan pembentukan medan elektromagnetik dan torsi yang dihasilkan motor. Pada kondisi fluks terkendali, peningkatan kebutuhan torsi umumnya diikuti peningkatan komponen arus pembentuk torsi. Namun, pengaruh utama kondisi operasi telah dikurangi melalui *clustering* berdasarkan torsi motor dan normalisasi sensor per *cluster*. Dengan demikian, tren Sensor2, Sensor3, dan Sensor5 setelah normalisasi lebih diarahkan pada perubahan

relatif terhadap kondisi operasi yang sejenis, bukan sekadar perbedaan tingkat torsi antaroperasi.

Sensor6 merepresentasikan temperatur motor dan berkaitan dengan rugi-rugi energi yang berubah menjadi panas. Temperatur motor dapat dipengaruhi oleh arus, rugi inti, gesekan, kondisi pendinginan, dan durasi pembebanan. Oleh karena itu, Sensor6 dapat merepresentasikan dampak termal dari beban elektrik motor. Pemantauan temperatur penting karena beban termal tinggi dapat menekan lilitan dan sistem isolasi motor. Penelitian motor traksi menunjukkan bahwa temperatur bagian kritis perlu dijaga karena kelebihan beban termal dapat meningkatkan risiko kegagalan. Kajian isolasi mesin listrik juga menunjukkan bahwa tegangan termal berperan dalam degradasi dan umur isolasi, sehingga tren Sensor6 memiliki hubungan fisik dengan kondisi kesehatan motor traksi (Mehboob et al., 2024; Zhou et al., 2024). Dengan demikian, Sensor6 tidak hanya menggambarkan kondisi sesaat, tetapi juga dapat menunjukkan akumulasi pengaruh termal selama sistem bekerja.

Sensor8 merepresentasikan kecepatan sudut roda sebagai keluaran mekanik dari rantai sistem traksi. Kecepatan roda berkaitan dengan hasil akhir konversi energi listrik menjadi energi mekanik melalui motor traksi dan sistem transmisi. Dengan demikian, Sensor8 dapat memberikan informasi mengenai respons mekanik sistem terhadap daya dan torsi yang diberikan. Perubahan pola kecepatan pada kondisi operasi yang sejenis dapat menjadi indikasi tambahan dalam membaca perubahan perilaku sistem traksi, khususnya pada sisi keluaran mekanik. Parameter ini juga menjadi pelengkap bagi sensor elektrik dan termal karena menunjukkan bagaimana perubahan pada sisi masukan sistem akhirnya tercermin pada keluaran gerak.

Secara fisik, Secara fisik, kelima sensor terpilih membentuk urutan pemantauan yang saling terhubung. Sensor3 dan Sensor5 menggambarkan keluaran elektrik VVVF, Sensor2 menggambarkan arus yang diterima motor, Sensor6 merepresentasikan konsekuensi termal, dan Sensor8 menggambarkan keluaran mekanik pada roda. Dengan demikian, pemilihannya tidak hanya didasarkan pada nilai *trend score*, tetapi juga mencakup rantai energi dari inverter, motor traksi, hingga roda. Hasil tersebut memperkuat bahwa lima sensor terpilih memiliki

relevansi statistik dan fisik dalam merepresentasikan perubahan kondisi sistem traksi. Sensor-sensor tersebut selanjutnya digunakan pada pengujian *health condition ideal* dan pembentukan HI agar indikator kesehatan yang dibentuk tetap mencakup sisi elektrik, termal, dan mekanik sistem.

4.2.2 Pengujian dan Penetapan *Health Condition Ideal*

Setelah sensor *trendable* terpilih, tahap berikutnya adalah melakukan pengujian beberapa bentuk *health condition ideal* yang akan digunakan sebagai target dalam pembentukan *Health Indicator* (HI). Tahap ini dilakukan karena bentuk kurva acuan akan memengaruhi arah dan karakteristik HI yang dihasilkan. Jika kurva acuan yang digunakan kurang sesuai, maka HI yang terbentuk dapat memiliki arah perubahan yang tidak sejalan dengan konsep degradasi sistem, sehingga menjadi kurang representatif untuk digunakan pada tahap pemodelan degradasi dan estimasi *Remaining Useful Life* (RUL). Oleh karena itu, pada penelitian ini dilakukan pengujian terhadap empat kandidat *health condition ideal* untuk menentukan bentuk acuan yang paling sesuai bagi pembentukan HI.

Pada penelitian ini, evaluasi terhadap masing-masing kandidat dilakukan menggunakan beberapa indikator kualitas HI, yaitu jumlah unit dengan *slope* negatif, jumlah unit dengan korelasi Spearman negatif, jumlah unit dengan *delta end-start* negatif, serta rata-rata nilai *monotonicity*. Seperti yang sudah dijelaskan pada Bab III, fungsi Indikator-indikator tersebut digunakan untuk mendapatkan HI yang arah penurunannya konsisten, hubungan monoton terhadap waktu operasi, dan perubahan global yang sesuai antara awal dan akhir pengamatan. Berikut hasil evaluasi *health condition ideal* yang ditampilkan pada tabel 4.7 dan 4.8.

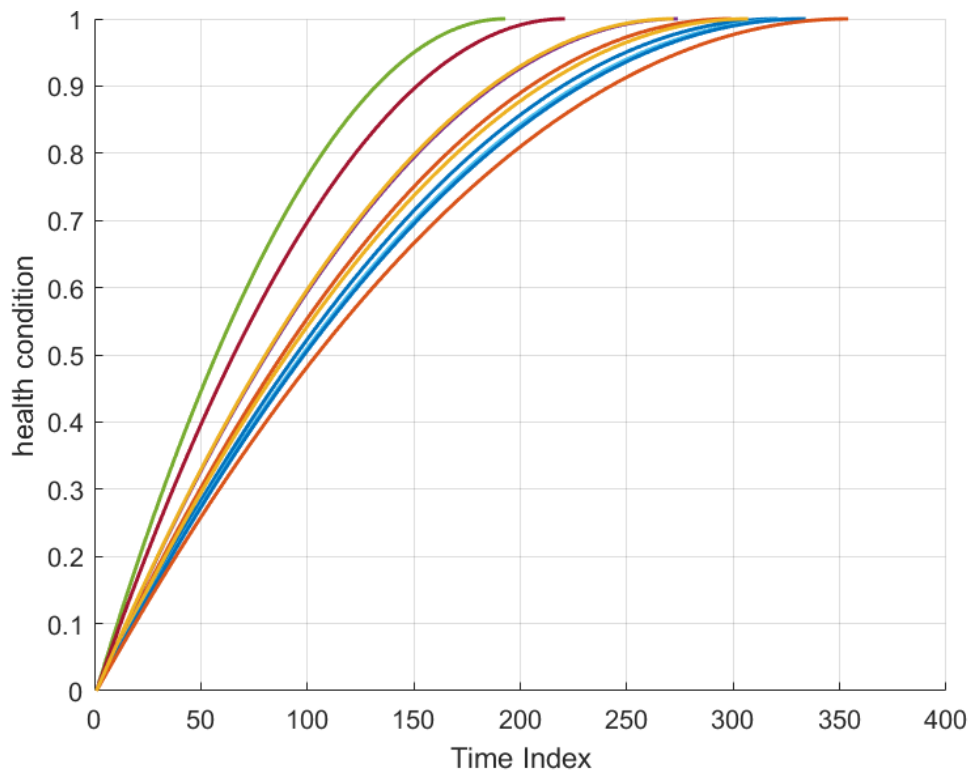
Tabel 4.7 Hasil Pengujian Empat Bentuk *Health Condition Ideal* (Bagian 1)

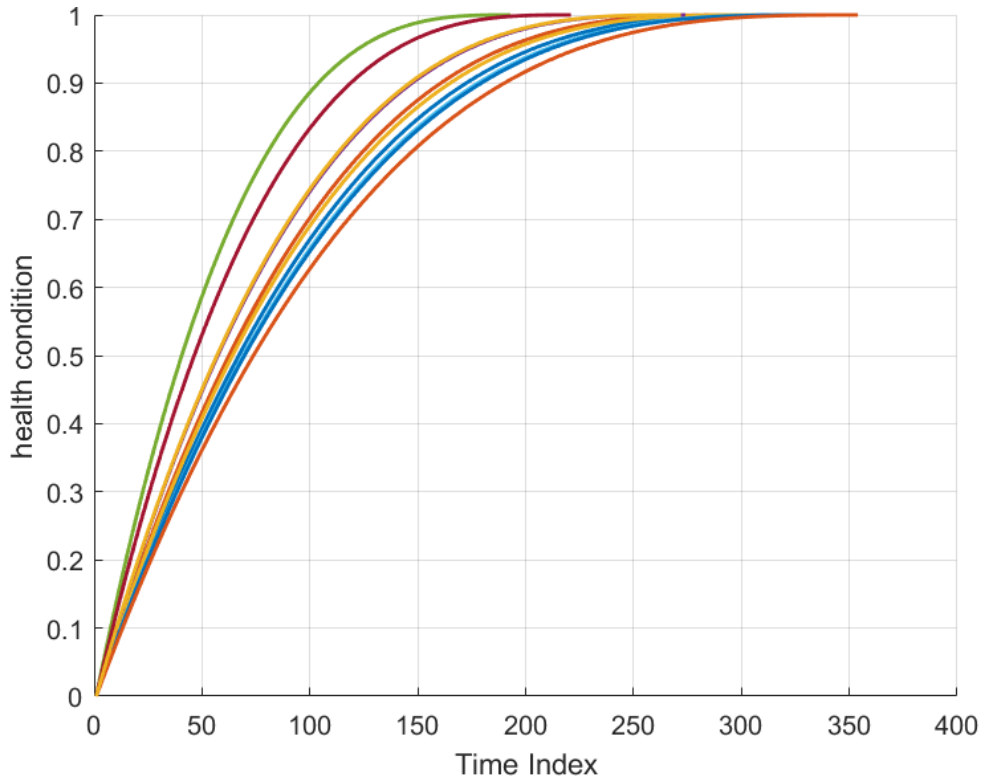
<i>Health Condition Ideal</i>	Jumlah Unit <i>Train</i>	<i>Slope</i> Negatif	<i>Slope</i> Positif / Nol	<i>Spearman</i> Negatif
HC-1 (τ) = $1 - \tau^2$	107	101	6	101
HC-2 (τ) = $1 - \tau^3$	107	101	6	101
HC-3 (τ) = $1 - (1 - \tau)^2$	107	6	101	6
HC-4 (τ) = $1 - (1 - \tau)^3$	107	6	101	6

Tabel 4.8 Hasil Pengujian Empat Bentuk *Health Condition* Ideal (Bagian 2)

<i>Health Condition</i> Ideal	<i>Spearman</i> Positif	<i>Delta</i> <i>End-Start</i> Negatif	<i>Delta</i> <i>End-Start</i> Positif	Rata-rata <i>Monotonicity</i>
HC-1 $(\tau) = 1 - \tau^2$	6	98	9	0,56792
HC-2 $(\tau) = 1 - \tau^3$	6	98	9	0,56797
HC-3 $(\tau) = 1 - (1 - \tau)^2$	101	10	97	0,43249
HC-4 $(\tau) = 1 - (1 - \tau)^3$	101	9	98	0,43206

Berdasarkan tabel 4.7 dan 4.8, HC-3 dan HC-4 didominasi tren positif, dengan 101 unit memiliki *slope* positif atau nol dan hanya 6 unit memiliki *slope* negatif. Pola yang sama terlihat pada korelasi Spearman, yaitu 101 unit bernilai positif dan 6 unit bernilai negatif. Indikator *delta end-start* juga menunjukkan dominasi perubahan positif, yaitu 97 unit pada HC-3 dan 98 unit pada HC-4. Hal ini menunjukkan bahwa HI yang dibentuk cenderung naik, sehingga tidak sesuai dengan konsep HI untuk estimasi RUL yang diharapkan menurun seiring bertambahnya degradasi sistem. Visualisasi kurva HC-3 dan HC-4 disajikan pada gambar 4.4 dan 4.5.

**Gambar 4.4** Kurva HC-3 $(\tau) = 1 - (1 - \tau)^2$

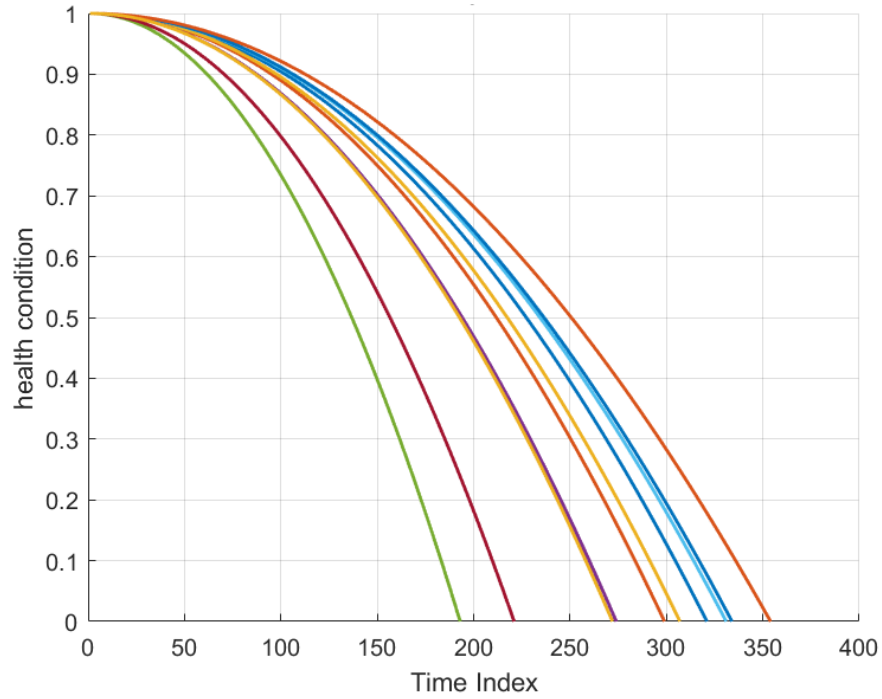


Gambar 4.5 Kurva HC-4 ($\tau = 1 - (1 - \tau)^3$)

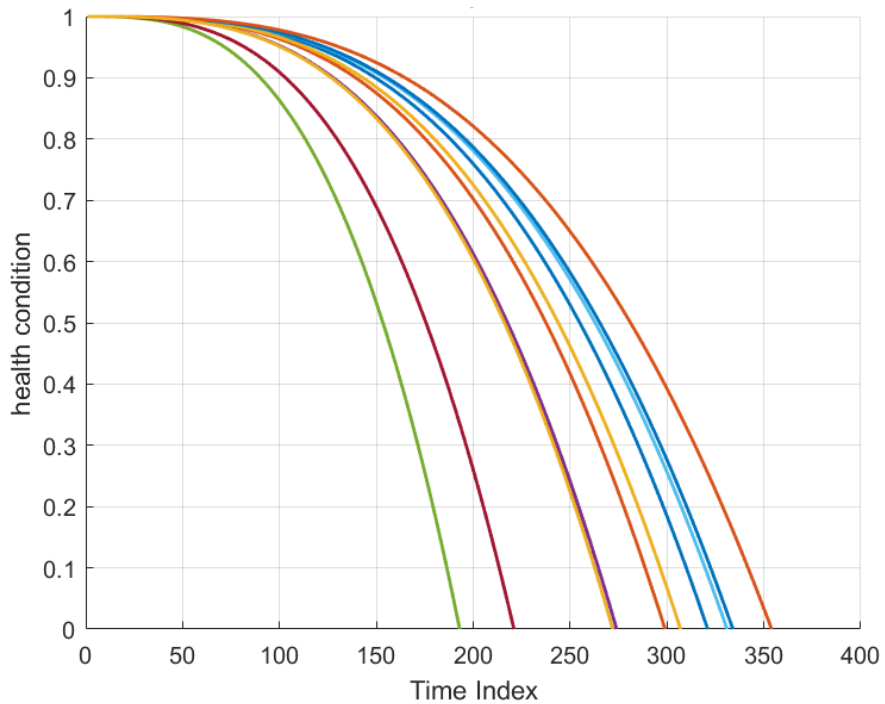
Berdasarkan gambar 4.4 dan 4.5, dapat dilihat bahwa HC-3 dan HC-4 sama-sama membentuk kurva yang meningkat terhadap waktu. Orientasi kurva ini berlawanan dengan konsep dasar HI pada penelitian ini, yaitu indikator kesehatan yang diharapkan menurun ketika kondisi sistem semakin terdegradasi. Hasil visualisasi ini konsisten dengan evaluasi numerik pada Tabel 4.7 dan Tabel 4.8, sehingga memperkuat bahwa HC-3 dan HC-4 tidak sesuai digunakan sebagai kurva acuan pembentukan HI. Setelah dua kandidat tersebut dieliminasi, tersisa HC-1 dan HC-2 sebagai dua kandidat yang sama-sama menghasilkan karakteristik HI yang sesuai dengan konsep degradasi. Berdasarkan Tabel 4.7 dan Tabel 4.8, kedua kandidat tersebut sama-sama menghasilkan 101 unit dengan *slope* negatif, 101 unit dengan korelasi Spearman negatif, dan 98 unit dengan *delta end-start* negatif dari total 107 unit *train*. Hasil ini menunjukkan bahwa baik HC-1 maupun HC-2 mampu menghasilkan HI dengan arah perubahan menurun yang dominan dan cukup konsisten terhadap waktu.

Meskipun demikian, pemilihan *health condition ideal* pada penelitian ini tidak hanya didasarkan pada dominasi tren menurun, tetapi juga pada konsistensi perubahan HI yang dihasilkan. Dalam hal ini, HC-2 memiliki rata-rata *monotonicity*

tertinggi, yaitu sebesar 0,56797, sedikit lebih tinggi dibandingkan HC-1 yang sebesar 0,56792. Oleh karena itu, HC-1 dan HC-2 dibandingkan lebih lanjut melalui visualisasi bentuk pada gambar 4.6 dan 4.7.



Gambar 4.6 Kurva HC-1 ($\tau = 1 - \tau^2$)



Gambar 4.7 Kurva HC-2 ($\tau = 1 - \tau^3$)

Berdasarkan gambar 4.6 dan 4.7, dapat dilihat bahwa HC-1 dan HC-2 sama-sama membentuk kurva yang menurun terhadap waktu, sehingga keduanya sesuai dengan konsep dasar HI yang diharapkan mengikuti arah degradasi sistem. Namun demikian, terdapat perbedaan pada karakter kelengkungan kedua kurva tersebut. HC-1 menunjukkan penurunan yang lebih bertahap sejak awal pengamatan, sedangkan HC-2 cenderung lebih landai pada bagian awal dan menjadi lebih tajam pada bagian akhir pengamatan. Dalam konteks penelitian ini, bentuk kurva HC-2 dinilai lebih sesuai untuk menggambarkan kondisi sistem yang pada awal pengamatan masih relatif stabil, kemudian menunjukkan penurunan kondisi yang lebih jelas ketika mendekati akhir masa operasinya. Jika dikaitkan dengan hasil evaluasi numerik sebelumnya, visualisasi ini juga mendukung hasil bahwa HC-2 memberikan rata-rata *monotonicity* tertinggi di antara keempat kandidat yang diuji.

Dengan demikian, hasil pengujian pada subbab ini menunjukkan bahwa HC-3 dan HC-4 tidak dipilih karena menghasilkan tren kurva yang meningkat, sehingga tidak sesuai dengan konsep degradasi HI untuk estimasi RUL. Sementara itu, HC-1 dan HC-2 sama-sama menghasilkan tren menurun yang sesuai, tetapi HC-2 dipilih sebagai *health condition ideal* karena memiliki nilai *monotonicity* tertinggi dan karakter bentuk kurva yang lebih sesuai untuk menggambarkan degradasi sistem traksi pada penelitian ini. Kurva HC-2 yang terpilih selanjutnya digunakan sebagai acuan pada tahap pembentukan HI final yang dibahas pada subbab berikutnya.

4.2.3 Pemodelan *Health Indicator*

Setelah sensor *trendable* dan *health condition ideal* terpilih ditetapkan, tahap berikutnya adalah melakukan pemodelan *Health Indicator* (HI) final. Pada penelitian ini, HI dibentuk untuk merangkum informasi dari beberapa sensor terpilih menjadi satu indikator kondisi yang lebih sederhana dan lebih mudah diamati sebagai *trajectory* degradasi. Tahap ini penting karena data sensor yang masih terpisah-pisah akan lebih sulit digunakan secara langsung untuk membaca pola penurunan kondisi sistem, sedangkan HI memungkinkan informasi tersebut direpresentasikan dalam satu kurva kesehatan yang lebih ringkas.

Pada penelitian ini, pemodelan HI dilakukan menggunakan lima sensor terpilih, yaitu Sensor3, Sensor2, Sensor6, Sensor5, dan Sensor8. Sebelum digabungkan, arah degradasi masing-masing sensor terlebih dahulu diseragamkan agar seluruh sensor mengikuti orientasi penurunan kondisi yang sama. Selanjutnya, sensor-sensor tersebut dikombinasikan untuk membentuk HI *raw*, kemudian dilakukan normalisasi ke rentang 0–1, dilanjutkan dengan proses *smoothing*, dan akhirnya level awal HI disejajarkan ke sekitar 1. Dengan tahapan ini, HI yang dihasilkan diharapkan menjadi lebih stabil dan lebih mudah dibaca sebagai representasi degradasi sistem.

Informasi mengenai hasil pemodelan HI pada penelitian ini ditampilkan pada tabel 4.9. Tabel ini memuat sensor-sensor yang digunakan dalam pembentukan HI, koefisien pembobotan, intersep, serta arah penyeragaman degradasi untuk masing-masing sensor.

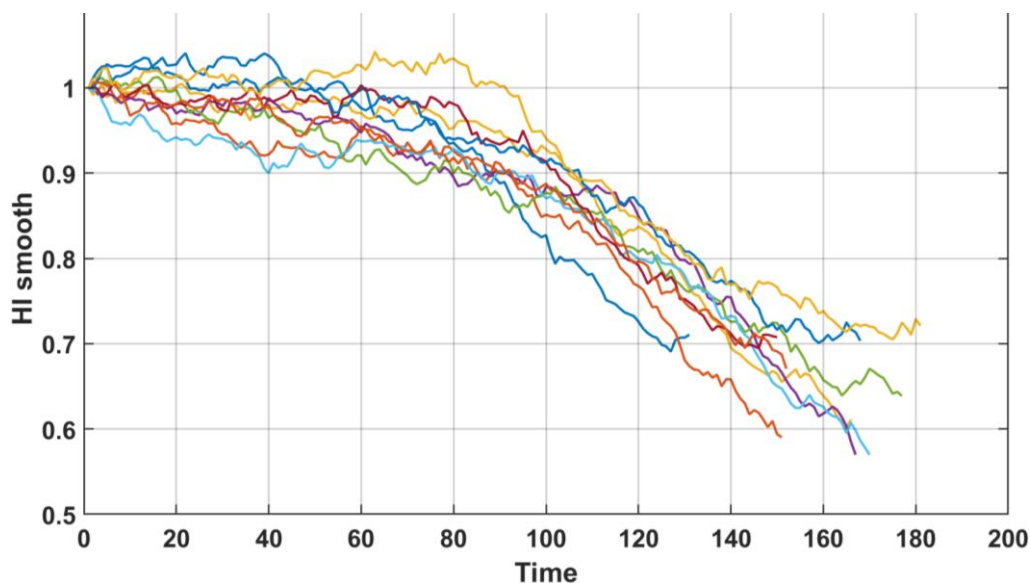
Tabel 4.9 Ringkasan Model Pembentukan Health Indicator

Komponen	Nilai
Sensor terpilih	Sensor3, Sensor2, Sensor6, Sensor5, Sensor8
Koefisien HI	0,0001; 0,4864; 0,0280; 0,3368; -0,1403
Intersep	0,7490
Arah sensor setelah penyeragaman degradasi	-1, -1, +1, -1, -1
Tahap lanjutan HI	Normalisasi 0–1, <i>smoothing</i> , dan penyelarasan level awal

Berdasarkan tabel 4.9, dapat dilihat bahwa pembentukan HI final pada penelitian ini tidak dilakukan menggunakan seluruh sensor secara langsung, tetapi hanya menggunakan lima sensor yang sebelumnya telah dipilih berdasarkan kemampuan trennya dalam merepresentasikan degradasi. Penggunaan sensor terpilih ini menunjukkan bahwa pembentukan HI difokuskan pada variabel-variabel yang paling relevan terhadap perubahan kondisi sistem. Selain itu, adanya arah penyeragaman degradasi juga menunjukkan bahwa sebelum dikombinasikan, masing-masing sensor terlebih dahulu disesuaikan agar memiliki orientasi perubahan yang konsisten terhadap penurunan kesehatan sistem.

Koefisien pembobotan dan intersep pada Tabel 4.9 menunjukkan bahwa HI final dibentuk sebagai kombinasi linear dari sensor-sensor terpilih. Dalam konteks penelitian ini, koefisien tersebut merepresentasikan kontribusi relatif tiap sensor terhadap HI yang dibangun. Setelah HI *raw* diperoleh, proses normalisasi dan *smoothing* dilakukan agar kurva HI menjadi lebih stabil, tidak terlalu dipengaruhi fluktuasi lokal, dan lebih sesuai dengan karakteristik HI yang dijelaskan pada dasar teori, yaitu memiliki arah perubahan yang cukup konsisten terhadap degradasi sistem.

Untuk melihat bentuk HI final yang dihasilkan pada beberapa ID unit pada data *train*, visualisasi *full trajectory* HI ditunjukkan pada gambar 4.6. Gambar ini menampilkan kurva HI *smooth* dari beberapa unit pada data *train*, sehingga pola umum penurunan HI data *train*, sehingga pola umum penurunan HI dapat diamati secara lebih langsung.

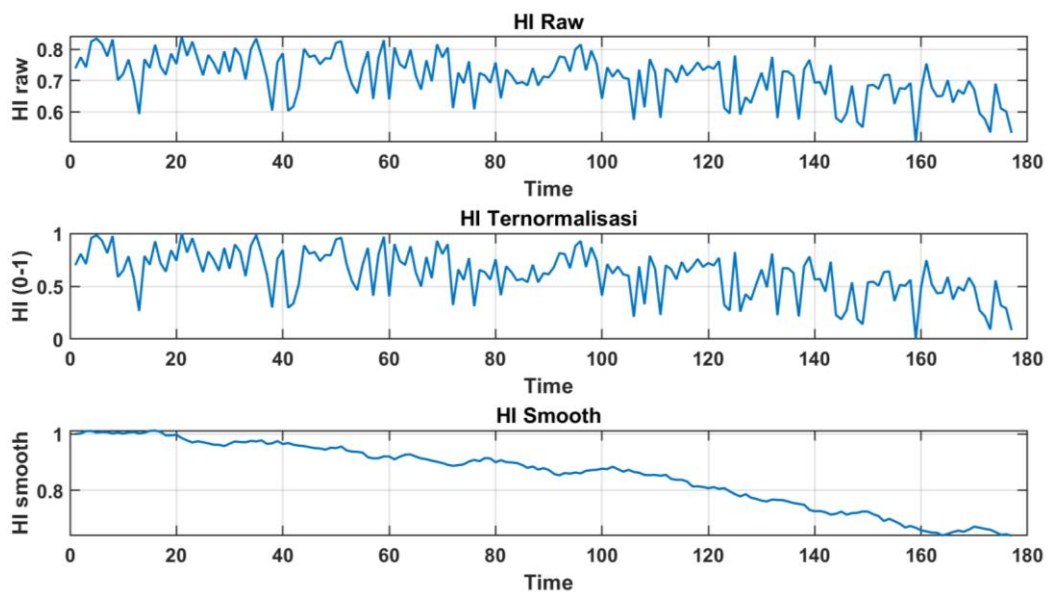


Gambar 4.8 *Full Trajectory Health Indicator* pada Beberapa ID Unit Data *Train*

Berdasarkan gambar 4.6, dapat dilihat bahwa sebagian besar *trajectory* HI dimulai dari level yang berada di sekitar 1 pada awal pengamatan, kemudian secara bertahap menurun seiring bertambahnya waktu operasi. Pola ini menunjukkan bahwa HI yang dibentuk telah mengikuti orientasi degradasi yang diharapkan, yaitu menurun ketika kondisi sistem semakin memburuk. Meskipun masih terdapat variasi bentuk kurva antarunit, secara umum seluruh kurva tetap menunjukkan kecenderungan penurunan yang relatif konsisten.

Perbedaan bentuk antar-*trajectory* pada Gambar 4.6 juga menunjukkan bahwa setiap unit memiliki karakteristik degradasi yang tidak sepenuhnya identik. Hal ini sejalan dengan konsep dasar prognosis bahwa pola degradasi antarunit dapat berbeda walaupun berasal dari sistem yang sama. Namun demikian, dengan HI yang telah dibentuk, variasi tersebut tetap dapat ditampilkan dalam bentuk yang lebih seragam dan lebih mudah dibandingkan antarunit, sehingga tahap selanjutnya, yaitu pemodelan degradasi *polynomial* dan estimasi RUL, dapat dilakukan dengan dasar representasi kondisi yang lebih baik.

Untuk memperjelas tahapan pembentukan HI pada satu unit contoh, gambar 4.7 menampilkan perbandingan HI *raw*, HI ternormalisasi, dan HI *smooth* pada ID 59 yang termasuk ke dalam data *train*. Gambar ini digunakan untuk menunjukkan perubahan bentuk kurva sejak masih berupa hasil gabungan awal sensor sampai menjadi HI final yang lebih halus dan lebih mudah diinterpretasikan.



Gambar 4.9 Tahapan Pembentukan *Health Indicator* pada ID 59 Data *Train*

Berdasarkan gambar 4.7, dapat dilihat bahwa HI *raw* masih menunjukkan fluktuasi yang cukup besar sepanjang waktu pengamatan. Setelah dilakukan normalisasi ke rentang 0–1, pola perubahan HI menjadi lebih terstandarkan, tetapi fluktuasi lokal masih tetap terlihat. Selanjutnya, setelah melalui proses *smoothing*, kurva HI menjadi jauh lebih halus dan memperlihatkan kecenderungan penurunan yang lebih jelas dari awal hingga akhir pengamatan. Hasil ini menunjukkan bahwa

proses *smoothing* berperan penting dalam mengurangi gangguan fluktuasi lokal sehingga pola degradasi sistem dapat terbaca dengan lebih baik.

Jika dikaitkan dengan dasar teori pada Bab II, hasil pemodelan HI pada penelitian ini telah menunjukkan karakteristik yang sesuai, yaitu mampu merepresentasikan kondisi sistem dalam bentuk indikator yang lebih sederhana dan memiliki kecenderungan perubahan yang mengikuti arah degradasi. Dengan demikian, HI yang diperoleh tidak hanya berfungsi sebagai hasil olahan data sensor, tetapi juga sebagai dasar representasi kondisi yang digunakan pada tahap berikutnya, yaitu pemodelan degradasi berbasis *polynomial* dan estimasi *Remaining Useful Life* (RUL).

Secara keseluruhan, hasil pada subbab ini menunjukkan bahwa proses pemodelan HI berhasil mengubah beberapa sensor terpilih menjadi satu indikator kesehatan yang lebih ringkas, lebih stabil, dan lebih mudah diamati sebagai *trajectory* degradasi. Oleh karena itu, HI yang telah dibentuk pada penelitian ini dapat digunakan sebagai masukan utama pada tahap estimasi RUL yang dibahas pada subbab berikutnya.

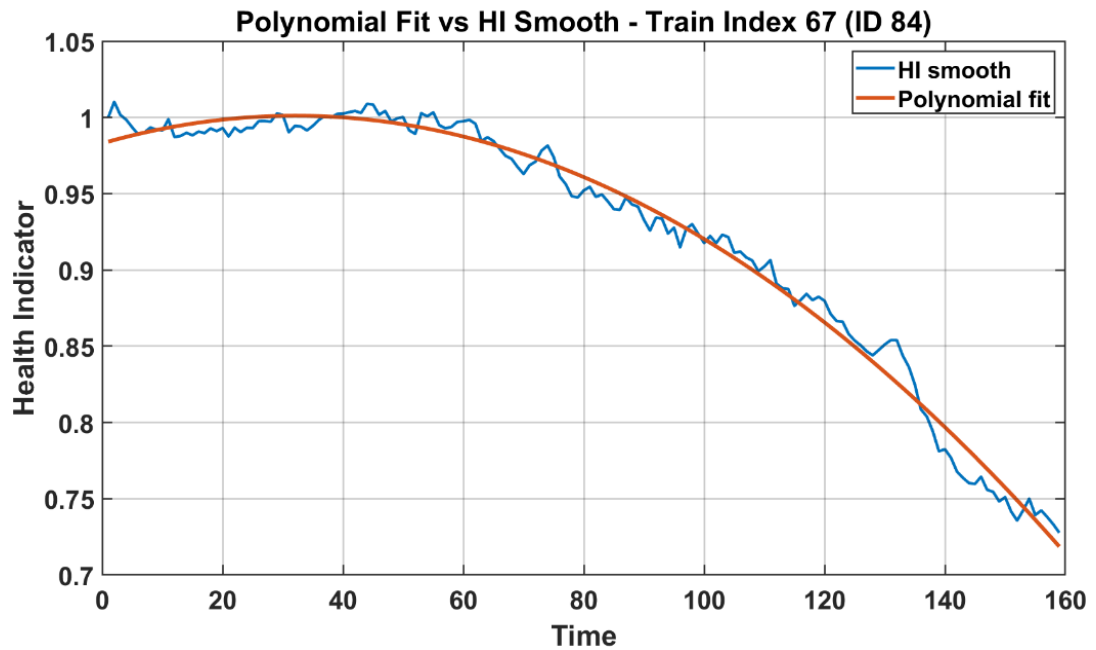
4.3 Hasil Estimasi *Remaining Useful Life* (RUL)

4.3.1 *Polynomial Degradation Model*

Setelah *Health Indicator* (HI) final diperoleh, tahap berikutnya adalah membentuk *polynomial degradation model* untuk merepresentasikan pola degradasi setiap unit dalam bentuk fungsi yang lebih ringkas. Pada penelitian ini, pemodelan degradasi dilakukan dengan melakukan *fitting polynomial* terhadap kurva HI *smooth* pada masing-masing unit pada data *train*. Tahap ini diperlukan karena kurva HI yang masih berbentuk deret waktu diskret perlu diubah menjadi representasi matematis yang lebih sederhana agar dapat digunakan pada proses perbandingan kemiripan antarunit pada tahap berikutnya.

Pada penelitian ini, orde *polynomial* yang digunakan adalah orde dua. Pemilihan orde dua dilakukan karena bentuk ini cukup sederhana untuk digunakan, tetapi tetap mampu menangkap kecenderungan perubahan nonlinier pada kurva degradasi. Dengan demikian, model yang dihasilkan tidak terlalu kompleks, namun tetap dapat mengikuti bentuk umum penurunan HI sepanjang *trajectory* pengamatan.

Untuk memperjelas hasil pemodelan degradasi yang diperoleh, contoh *polynomial fit* pada salah satu unit pada data *train* disajikan pada gambar 4.8. Gambar ini menampilkan perbandingan antara kurva *HI smooth* aktual dan kurva hasil *approximate polynomial* pada ID unit 84 yang termasuk ke dalam data *train*.



Gambar 4.10 Perbandingan Kurva *HI Smooth* dan *Polynomial Fit* pada ID 84 Data *Train*

Berdasarkan gambar 4.8, dapat dilihat bahwa kurva *polynomial fit* mampu mengikuti bentuk umum kurva *HI smooth* dengan cukup baik. Meskipun kurva hasil *fitting* tampak lebih halus dibandingkan kurva *HI* aktual, pola penurunan kondisi secara keseluruhan tetap dapat direpresentasikan dengan baik. Hal ini menunjukkan bahwa pendekatan *polynomial* orde dua yang digunakan pada penelitian ini cukup mampu menangkap tren degradasi utama tanpa terlalu dipengaruhi oleh fluktuasi lokal pada *HI*.

Untuk memperjelas hasil pemodelan degradasi pada unit contoh tersebut, ringkasan parameter *polynomial fit* disajikan pada tabel 4.10. Tabel ini memuat identitas unit contoh pada data *train*, nilai *RMSE polynomial fit*, nilai *monotonicity HI*, *slope linear HI*, serta koefisien *polynomial* yang diperoleh.

Tabel 4.10 Ringkasan Hasil *Polynomial Fit*

Komponen	Nilai
Unit <i>train</i> terpilih untuk visualisasi	ID 84 pada data <i>train</i>
<i>Polynomial RMSE</i>	0,009506
<i>Monotonicity</i> HI	0,6076
<i>Slope</i> HI linear	-0,001680
Koefisien <i>polynomial</i> [<i>a b c</i>]	(-0,0000; 0,0011; 0,9830)

Berdasarkan tabel 4.10, dapat dilihat bahwa unit contoh yang digunakan memiliki nilai *RMSE polynomial fit* sebesar 0,009506. Nilai ini menunjukkan bahwa selisih antara kurva HI *smooth* dan kurva hasil *approximately polynomial* relatif kecil, sehingga *fit* yang diperoleh dapat dianggap cukup baik. Selain itu, nilai *monotonicity* HI sebesar 0,6076 dan *slope* linear yang bernilai negatif menunjukkan bahwa kurva HI pada unit tersebut memiliki kecenderungan menurun yang cukup konsisten, sehingga sesuai untuk digunakan sebagai contoh representatif dalam visualisasi *polynomial fit*.

Koefisien *polynomial* [*a b c*] yang diperoleh juga menunjukkan bahwa kurva degradasi pada unit contoh dapat direpresentasikan dalam bentuk fungsi kuadratik. Dalam konteks penelitian ini, koefisien-koefisien tersebut tidak hanya berfungsi sebagai hasil *fitting*, tetapi juga menjadi dasar representasi degradasi tiap unit yang selanjutnya akan dibandingkan pada tahap *similarity*. Dengan kata lain, setiap unit pada data *train* tidak lagi direpresentasikan hanya sebagai deret HI, tetapi sebagai bentuk *polynomial degradation model* yang lebih ringkas.

Untuk melihat kualitas *fitting polynomial* secara umum pada seluruh data *train*, ringkasan nilai *RMSE polynomial fit* ditampilkan pada tabel 4.11. Tabel ini memuat nilai minimum, maksimum, dan rata-rata *RMSE* hasil *fitting polynomial* pada seluruh unit pada data *train*.

Tabel 4.11 Ringkasan *RMSE Polynomial Fit* pada Data *Train*

Komponen	Nilai
<i>RMSE</i> minimum	0,009506
<i>RMSE</i> maksimum	0,039484
<i>RMSE</i> rata-rata	0,021931

Berdasarkan tabel 4.11, dapat dilihat bahwa nilai *RMSE polynomial fit* pada data *train* secara umum berada pada tingkat yang relatif kecil. Nilai *RMSE* rata-rata sebesar 0,021931 menunjukkan bahwa *approximately polynomial* orde dua yang digunakan pada penelitian ini cukup mampu merepresentasikan bentuk kurva *HI smooth* pada sebagian besar unit pada data *train*. Meskipun terdapat variasi kualitas *fit* antarunit, secara umum hasil ini menunjukkan bahwa *polynomial degradation model* yang dibentuk telah cukup memadai untuk digunakan sebagai dasar perbandingan kemiripan pada tahap berikutnya.

Jika dikaitkan dengan dasar teori pada Bab II dan metode pada Bab III, hasil ini menunjukkan bahwa representasi degradasi berbasis *polynomial* berhasil menyederhanakan kurva *HI* menjadi bentuk matematis yang lebih ringkas tanpa kehilangan tren degradasi utamanya. Oleh karena itu, *polynomial degradation model* yang diperoleh pada penelitian ini dapat digunakan sebagai dasar pada tahap selanjutnya, yaitu analisis *Residual-Based KNN Similarity* dan estimasi *Remaining Useful Life*.

4.3.2 Estimasi RUL Per *Breakpoint*

Setelah *polynomial degradation model* dibentuk untuk setiap unit pada data *train*, tahap berikutnya adalah melakukan pengukuran kemiripan antara unit pada data *test* dan referensi historis menggunakan pendekatan *residual-based KNN similarity*. Pada penelitian ini, estimasi *Remaining Useful Life* (RUL) dilakukan berdasarkan lima *nearest neighbors* ($k = 5$) yang memiliki jarak residual paling kecil terhadap unit *test*. Pendekatan ini digunakan karena unit *test* diprediksi dengan memanfaatkan informasi dari unit-unit pada data *train* yang memiliki pola degradasi paling mirip, sehingga estimasi RUL dapat dilakukan berdasarkan referensi historis yang paling relevan.

Pada penelitian ini, pengujian dilakukan pada tiga *breakpoint*, yaitu 50%, 70%, dan 90% dari *life trajectory*. Penggunaan beberapa *breakpoint* dilakukan untuk melihat bagaimana performa estimasi berubah ketika panjang data pengamatan pada unit *test* semakin bertambah. Secara umum, semakin besar *breakpoint*, maka informasi degradasi yang tersedia pada unit *test* menjadi semakin lengkap, sehingga proses pencarian *nearest neighbors* diharapkan menjadi lebih akurat dan hasil estimasi RUL menjadi lebih baik.

Ringkasan hasil estimasi RUL pada unit contoh, yaitu ID unit 13 yang termasuk ke dalam data *test*, disajikan pada tabel 4.12. Tabel ini memuat nilai RUL aktual, RUL prediksi, *absolute error* pada unit contoh, serta nilai MAE dan RMSE keseluruhan model pada masing-masing *breakpoint*.

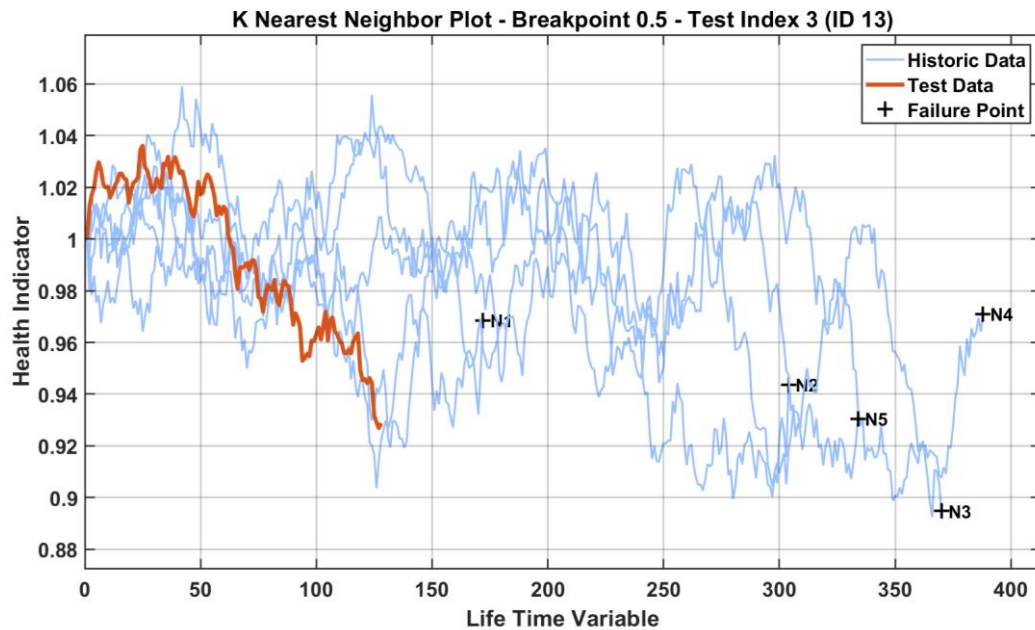
Tabel 4.12 Ringkasan Hasil Estimasi RUL pada ID 13 Data *Test*

Breakpoint	True RUL	Predicted RUL	Absolute Error
50%	128,00	150,82	22,82
70%	77,00	80,09	3,09
90%	26,00	18,64	7,36

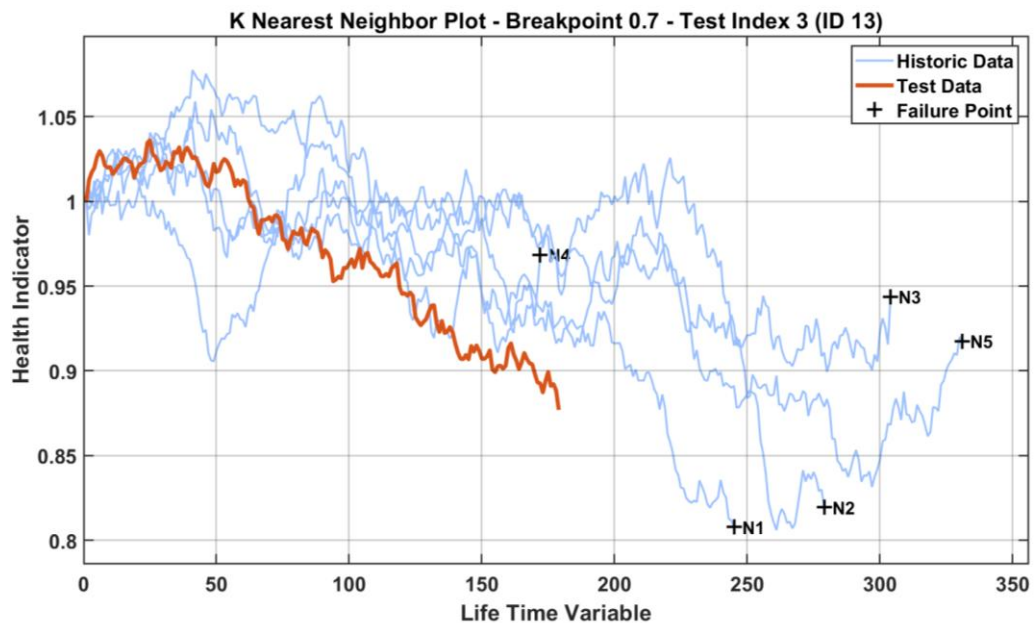
Apabila ditinjau pada unit contoh ID 13, hasil estimasi terbaik diperoleh pada *breakpoint* 70%, dengan *absolute error* sebesar 3,09. Pada *breakpoint* 50%, prediksi RUL masih cenderung terlalu tinggi dengan *absolute error* sebesar 22,82, yang menunjukkan bahwa ketika data pengamatan baru tersedia sampai sekitar setengah umur unit, informasi degradasi yang terbaca belum cukup kuat untuk menghasilkan estimasi yang sangat presisi. Pada kondisi tersebut, *trajectory* degradasi unit masih relatif pendek sehingga pola penurunan HI yang terbentuk belum sepenuhnya mampu merepresentasikan kecenderungan kerusakan secara lebih jelas. Akibatnya, proses pencarian *nearest neighbors* masih berpotensi menghasilkan referensi historis yang belum sepenuhnya paling sesuai dengan pola unit contoh.

Sementara itu, pada *breakpoint* 90%, meskipun performa keseluruhan model paling baik, pada unit contoh ini prediksi justru sedikit lebih rendah dari nilai aktual dengan *absolute error* sebesar 7,36. Kondisi ini menunjukkan bahwa performa agregat model dapat membaik, tetapi pada unit tertentu masih mungkin muncul deviasi prediksi akibat karakteristik *nearest neighbors* yang terpilih. Dengan demikian, hasil pada unit contoh ID 13 menunjukkan bahwa peningkatan panjang data pengamatan memang cenderung memperbaiki akurasi estimasi, tetapi kesesuaian pola unit historis yang digunakan sebagai referensi tetap menjadi faktor penting dalam menentukan besar kecilnya error prediksi pada masing-masing *breakpoint*.

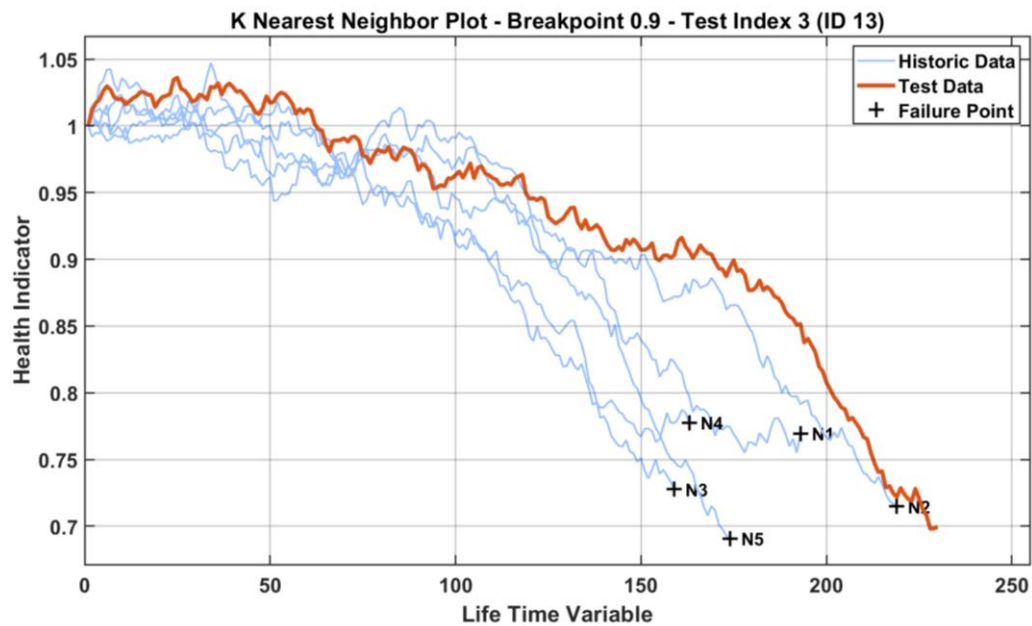
Untuk memperjelas hasil pencarian kemiripan pada masing-masing breakpoint, visualisasi K Nearest Neighbor Plot pada unit contoh ID 13 disajikan pada gambar 4.9 sampai dengan 4.11. Ketiga gambar tersebut menunjukkan posisi trajectory unit test terhadap lima unit historis terdekat, beserta titik kegagalan masing-masing neighbor.



Gambar 4.11 *K Nearest Neighbor Plot* pada Breakpoint 50% untuk ID 13 Data Test



Gambar 4.12 *K Nearest Neighbor Plot* pada Breakpoint 70% untuk ID 13 Data Test



Gambar 4.13 *K Nearest Neighbor Plot* pada Breakpoint 90% untuk ID 13 Data Test

Berdasarkan gambar 4.9, dapat dilihat bahwa pada *breakpoint* 50%, *trajectory* unit test masih relatif pendek sehingga pola degradasi yang tersedia belum cukup lengkap. Akibatnya, lima *nearest neighbors* yang terpilih masih menunjukkan variasi bentuk kurva dan posisi titik kegagalan yang cukup lebar. Kondisi ini sejalan dengan hasil pada Tabel 4.12, di mana prediksi RUL pada *breakpoint* 50% masih memiliki *absolute error* yang cukup besar.

Pada gambar 4.10, ketika *breakpoint* meningkat menjadi 70%, *trajectory* unit test sudah memperlihatkan pola degradasi yang lebih jelas. Pada kondisi ini, unit test dapat dibandingkan dengan unit-unit historis yang lebih mirip, sehingga hasil estimasi menjadi lebih akurat. Hal tersebut juga tercermin pada Tabel 4.12, di mana *absolute error* untuk ID 13 turun menjadi 3,09. Dengan demikian, *breakpoint* 70% pada unit contoh ini memberikan titik kompromi yang baik antara panjang data pengamatan dan akurasi prediksi.

Sementara itu, pada gambar 4.11, *trajectory* unit test pada *breakpoint* 90% sudah sangat panjang dan pola degradasinya terlihat jauh lebih lengkap. Secara umum, kondisi ini membuat proses pencarian kemiripan menjadi lebih kuat, yang juga didukung oleh penurunan MAE dan RMSE keseluruhan model. Namun demikian, pada unit contoh ID 13, prediksi yang dihasilkan justru sedikit lebih rendah dari nilai aktual. Hal ini menunjukkan bahwa beberapa *neighbor* terpilih

memiliki titik kegagalan yang relatif lebih awal dibandingkan umur sisa aktual unit *test*, sehingga hasil pembobotan cenderung menghasilkan prediksi yang sedikit *underestimate*.

Untuk memperjelas komposisi *nearest neighbors* pada tiap *breakpoint*, hasil ID unit pada data *train*, jarak residual, bobot *inverse-distance*, *Candidate RUL*, dan *Contribution* ditampilkan pada tabel 4.13 sampai dengan 4.15.

Tabel 4.13 *Nearest Neighbors* pada Breakpoint 50% untuk ID 13 Data *Test*

Neighbor	ID	Distance	Weight	Candidate RUL	Contribution
N1	50	0,1461	0,2649	86	22,78
N2	108	0,1954	0,1981	152	30,11
N3	131	0,2110	0,1835	185	33,94
N4	88	0,2113	0,1832	194	35,54
N5	9	0,2272	0,1703	167	28,45
Total			1,0000		150,82

Tabel 4.14 *Nearest Neighbors* pada Breakpoint 70% untuk ID 13 Data *Test*

Neighbor	ID	Distance	Weight	Candidate RUL	Contribution
N1	20	0,2879	0,2187	73	15,96
N2	41	0,3115	0,2021	84	16,97
N3	108	0,3222	0,1954	91	17,78
N4	50	0,3273	0,1923	52	10,00
N5	6	0,3285	0,1916	99	18,97
Total			1,0000		79,68

Tabel 4.15 *Nearest Neighbors* pada Breakpoint 90% untuk ID 13 Data *Test*

Neighbor	ID	Distance	Weight	Candidate RUL	Contribution
N1	5	0,2932	0,2094	19	3,98
N2	32	0,2997	0,2049	22	4,51
N3	84	0,3059	0,2008	16	3,21
N4	27	0,3069	0,2001	16	3,20
N5	122	0,3321	0,1849	17	3,14
Total			1,0000		18,04

Nilai *Candidate RUL* menunjukkan sisa umur masing-masing unit historis pada titik *breakpoint* yang sama, sedangkan *Contribution* menunjukkan kontribusi setiap *neighbor* terhadap estimasi RUL akhir. Nilai *Contribution* diperoleh dari hasil perkalian antara bobot dan *Candidate RUL*. Dengan demikian, total *Contribution* dari seluruh *neighbor* menjadi dasar pembentukan nilai prediksi RUL.

Berdasarkan tabel 4.13, pada *breakpoint* 50%, *neighbor* dengan jarak residual terkecil adalah N1, yaitu ID unit 50 dengan distance sebesar 0,1461 dan weight sebesar 0,2649. Karena jaraknya paling kecil, unit ini memperoleh bobot terbesar dibandingkan *neighbor* lainnya. Namun demikian, nilai *Candidate RUL* dari lima *neighbor* yang terpilih masih cukup bervariasi, yaitu 86, 152, 185, 194, dan 167. Nilai *Contribution* terbesar berasal dari N4, yaitu sebesar 35,54, karena meskipun bobotnya bukan yang paling besar, unit tersebut memiliki *Candidate RUL* paling tinggi sebesar 194. Total *Contribution* pada *breakpoint* 50% menghasilkan estimasi RUL sebesar 150,82. Nilai ini lebih tinggi dibandingkan RUL aktual sebesar 128,00, sehingga prediksi pada *breakpoint* ini cenderung *overestimate*.

Berdasarkan tabel 4.14, pada *breakpoint* 70%, nilai distance antar-*neighbor* berada pada rentang 0,2879 sampai 0,3285, sedangkan nilai weight berada pada rentang 0,1916 sampai 0,2187. Distribusi bobot yang relatif lebih seimbang menunjukkan bahwa tidak ada satu *neighbor* yang terlalu mendominasi hasil estimasi. Nilai *Candidate RUL* pada *breakpoint* ini juga lebih terkonsentrasi pada rentang yang lebih dekat dengan RUL aktual unit *test*, yaitu 73, 84, 91, 52, dan 99. Total *Contribution* yang dihasilkan sebesar 79,68, sedangkan nilai prediksi model sebesar 80,09 dan RUL aktual sebesar 77,00. Kondisi ini menunjukkan bahwa kombinasi antara distance, weight, *Candidate RUL*, dan *Contribution* pada *breakpoint* 70% mampu menghasilkan prediksi yang paling dekat terhadap nilai aktual.

Berdasarkan tabel 4.15, pada *breakpoint* 90%, kelima *neighbor* memiliki nilai weight yang juga relatif seimbang, yaitu berada pada rentang 0,1849 sampai 0,2094. Akan tetapi, nilai *Candidate RUL* dari unit historis yang terpilih relatif kecil, yaitu 19, 22, 16, 16, dan 17. Hal ini menyebabkan nilai *Contribution* dari masing-masing *neighbor* juga kecil, dengan total *Contribution* sebesar 18,04. Nilai

tersebut mendekati hasil prediksi model sebesar 18,64, tetapi masih lebih rendah dibandingkan RUL aktual unit *test* sebesar 26,00. Dengan demikian, prediksi pada *breakpoint* 90% mengalami *underestimation* karena unit-unit historis yang terpilih memiliki sisa umur yang lebih pendek dibandingkan unit *test*.

4.3.3 Evaluasi Model Estimator RUL

Setelah proses estimasi *Remaining Useful Life* (RUL) dilakukan pada beberapa *breakpoint*, tahap berikutnya adalah mengevaluasi kedekatan hasil prediksi terhadap RUL aktual pada seluruh data *test*. Evaluasi dilakukan pada *breakpoint* 50%, 70%, dan 90% untuk mengamati performa model ketika informasi degradasi yang tersedia masih terbatas hingga mendekati akhir umur operasi.

Metrik utama yang digunakan meliputi *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE). MAE menunjukkan rata-rata kesalahan absolut dalam satuan RUL, RMSE memberikan penalti lebih besar terhadap prediksi dengan kesalahan tinggi, sedangkan MAPE menunjukkan kesalahan relatif terhadap RUL aktual dalam bentuk persentase. Selain itu, MAEfrac dan RMSEfrac digunakan sebagai metrik pendukung untuk menunjukkan besar kesalahan relatif terhadap umur total unit dan mendukung perbandingan dengan penelitian sebelumnya. Hasil evaluasi model estimator RUL pada setiap *breakpoint* ditampilkan pada tabel 4.16.

Tabel 4.16 Hasil Evaluasi Error Model Estimator RUL

Breakpoint	MAE	RMSE	MAPE	MAEfrac	RMSEfrac
50%	31,1294	38,7466	27,1163%	13,5278%	16,8350%
70%	17,1567	22,7426	24,7927%	7,4326%	9,7830%
90%	5,0832	6,8757	20,6347%	2,0686%	2,6400%

Berdasarkan tabel 4.16, seluruh metrik error mengalami penurunan seiring meningkatnya *breakpoint*. Pada *breakpoint* 50%, model menghasilkan MAE sebesar 31,1294 dan RMSE sebesar 38,7466. Nilai tersebut menunjukkan bahwa ketika hanya setengah *trajectory* unit yang tersedia, pola degradasi yang diamati masih terbatas sehingga penentuan kemiripan terhadap unit historis belum dapat dilakukan secara optimal. Pada *breakpoint* 70%, MAE menurun menjadi 17,1567 dan RMSE menjadi 22,7426. Penurunan tersebut menunjukkan bahwa tambahan informasi degradasi membantu model memperoleh unit referensi yang lebih sesuai.

MAEfrac dan RMSEfrac juga menurun menjadi 7,4326% dan 9,7830%, yang menunjukkan berkurangnya kesalahan relatif terhadap umur total unit.

Performa terbaik diperoleh pada *breakpoint* 90%, dengan MAE sebesar 5,0832 dan RMSE sebesar 6,8757. Pada tahap ini, sebagian besar *trajectory* degradasi telah tersedia sehingga bentuk HI parsial unit *test* dapat dibandingkan dengan model degradasi unit *train* secara lebih akurat. Hal tersebut juga ditunjukkan oleh MAEfrac sebesar 2,0686% dan RMSEfrac sebesar 2,6400%. Nilai MAPE menurun dari 27,1163% pada *breakpoint* 50% menjadi 20,6347% pada *breakpoint* 90%. Penurunannya tidak sebesar metrik lainnya karena MAPE menggunakan RUL aktual sebagai penyebut. Ketika unit mendekati akhir umur operasi, nilai RUL aktual semakin kecil sehingga kesalahan absolut yang kecil tetap dapat menghasilkan persentase kesalahan yang relatif besar. Meskipun demikian, seluruh nilai MAPE berada di bawah batas performansi sebesar 30%, sehingga model memenuhi target yang telah ditetapkan.

Untuk menilai posisi performa model terhadap penelitian terdahulu, hasil penelitian ini dibandingkan dengan penelitian Husain. Penelitian tersebut dipilih sebagai pembanding kuantitatif karena menggunakan *breakpoint* serta metrik MAE, RMSE, MAEfrac, dan RMSEfrac yang sama. Sementara itu, penelitian Natasya tidak dimasukkan ke dalam tabel karena menggunakan *error mean*, *error median*, dan *error standard deviation* pada lima data validasi. Perbedaan definisi metrik tersebut menyebabkan hasilnya tidak dapat dibandingkan secara langsung dengan metrik pada penelitian ini. MAPE juga tidak dimasukkan dalam tabel perbandingan karena tidak dilaporkan pada penelitian Husain. Perbandingan hasil evaluasi ditampilkan pada tabel 4.17.

Tabel 4.17 Perbandingan Hasil Evaluasi terhadap Penelitian Terdahulu

Breakpoint	Penelitian	MAE	RMSE	MAEfrac	RMSEfrac
50%	Al Husaini (2025)	60,67	71,15	28,48%	35,76%
	Penelitian ini	31,1294	38,7466	13,5278%	16,8350%
70%	Al Husaini (2025)	38,06	47,18	17,26%	22,95%
	Penelitian ini	17,1567	22,7426	7,4326%	9,7830%
90%	Al Husaini (2025)	14,10	19,00	6,39%	9,08%
	Penelitian ini	5,0832	6,8757	2,0686%	2,6400%

Berdasarkan tabel 4.17, penelitian ini menghasilkan nilai MAE, RMSE, MAEfrac, dan RMSEfrac yang lebih rendah pada seluruh *breakpoint*. Pada *breakpoint* 50%, MAE menurun dari 60,67 menjadi 31,1294 dan RMSE menurun dari 71,15 menjadi 38,7466. Pada *breakpoint* 70%, MAE penelitian ini sebesar 17,1567 dibandingkan 38,06 pada penelitian Husain. Perbedaan semakin terlihat pada *breakpoint* 90%, ketika MAE dan RMSE penelitian ini masing-masing sebesar 5,0832 dan 6,8757, sedangkan penelitian Husain menghasilkan 14,10 dan 19,00.

Pola yang sama terlihat pada metrik fraksional. MAEfrac penelitian ini pada *breakpoint* 50%, 70%, dan 90% masing-masing sebesar 13,5278%, 7,4326%, dan 2,0686%, sedangkan penelitian Husain menghasilkan 28,48%, 17,26%, dan 6,39%. RMSEfrac penelitian ini juga lebih rendah pada seluruh *breakpoint*, terutama pada *breakpoint* 90% yang menurun dari 9,08% menjadi 2,6400%.

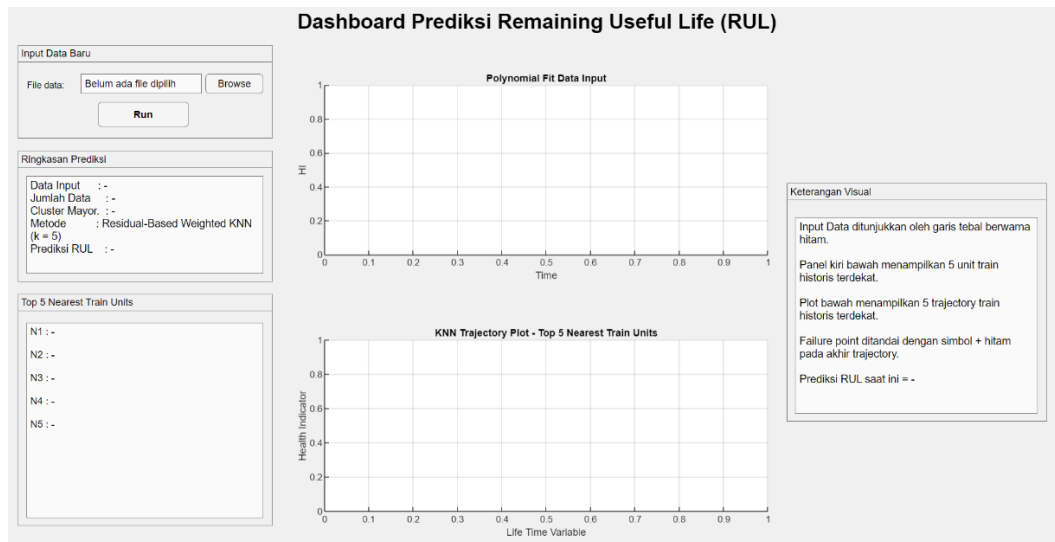
Perbedaan hasil tersebut berkaitan dengan perbedaan tahapan pengolahan data dan metode estimasi yang digunakan. Penelitian Husain menggunakan model degradasi *similarity-based exponential* dengan segmentasi K-Means dan *Isolation Forest*, sedangkan penelitian ini menggunakan model polinomial dan *residual-based weighted KNN*. Pada penelitian ini, kemiripan ditentukan berdasarkan *residual distance* antara HI parsial unit *test* dan model degradasi unit *train*, kemudian kandidat RUL digabungkan menggunakan pembobotan jarak. Meskipun hasil penelitian ini menunjukkan error yang lebih rendah pada metrik yang dibandingkan, perbedaan performa tetap perlu dipahami dalam konteks perbedaan metode pembentukan HI, pengolahan data, model degradasi, dan estimator yang digunakan.

Secara keseluruhan, model estimator RUL menunjukkan peningkatan performa seiring bertambahnya informasi degradasi. Seluruh nilai MAPE telah memenuhi target di bawah 30%, sedangkan hasil perbandingan terhadap penelitian Husain menunjukkan nilai MAE, RMSE, MAEfrac, dan RMSEfrac yang lebih rendah pada seluruh *breakpoint*.

4.4 Hasil Visualisasi Dashboard

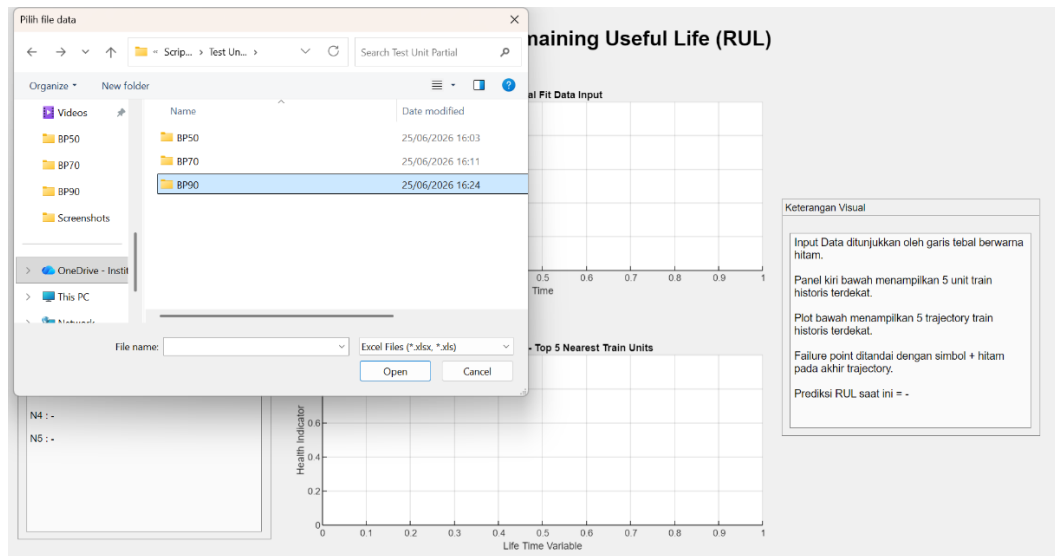
Setelah dashboard prediksi *Remaining Useful Life* (RUL) dikembangkan, tahap berikutnya adalah melakukan uji coba menggunakan file input dari data *test*. File tersebut tersedia dalam bentuk *full trajectory* dan data parsial pada *breakpoint* 50%, 70%, dan 90%. Pengujian ini dilakukan untuk menunjukkan bahwa dashboard dapat menjalankan proses estimasi pada panjang data pengamatan yang berbeda, mulai dari kondisi ketika informasi degradasi masih terbatas hingga mendekati akhir pengamatan.

Tahapan penggunaan dashboard ditunjukkan pada gambar 4.14 sampai dengan gambar 4.17. Rangkaian gambar tersebut memperlihatkan proses dari tampilan awal dashboard, pemilihan folder *breakpoint*, pemilihan file input, hingga penyajian hasil estimasi dan visualisasi pendukung.



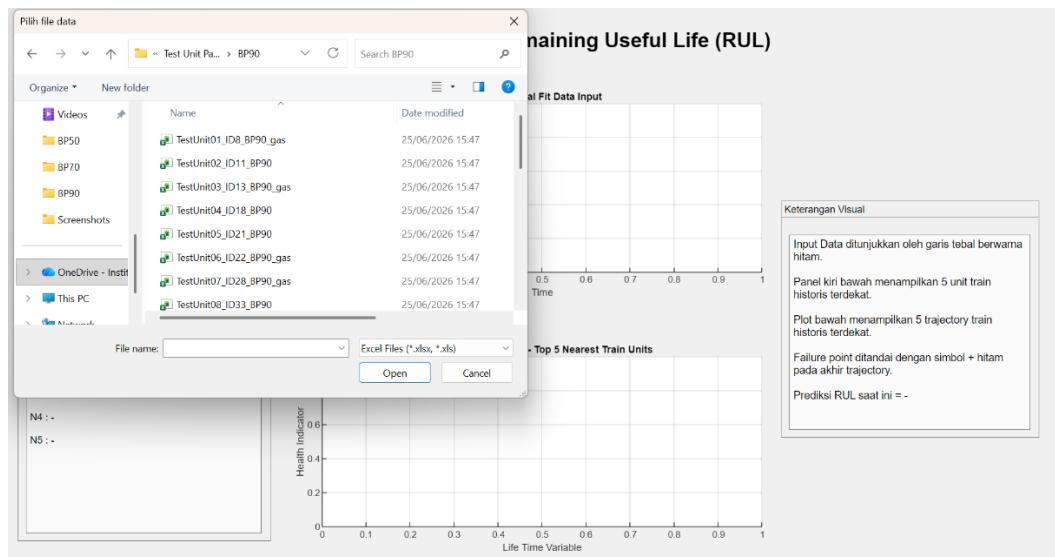
Gambar 4.14 Tampilan Awal Dashboard Prediksi RUL

Berdasarkan gambar 4.14, dashboard dijalankan melalui MATLAB hingga menampilkan antarmuka dalam kondisi siap digunakan. Pada tampilan awal, panel ringkasan prediksi, grafik *polynomial fit*, dan *KNN trajectory plot* masih kosong karena data belum dimasukkan. Pengguna dapat menekan tombol *Browse* untuk memilih file data historis yang akan dianalisis tanpa perlu mengubah atau menjalankan bagian *script* secara manual.



Gambar 4.15 Pemilihan Folder File Input pada Dashboard

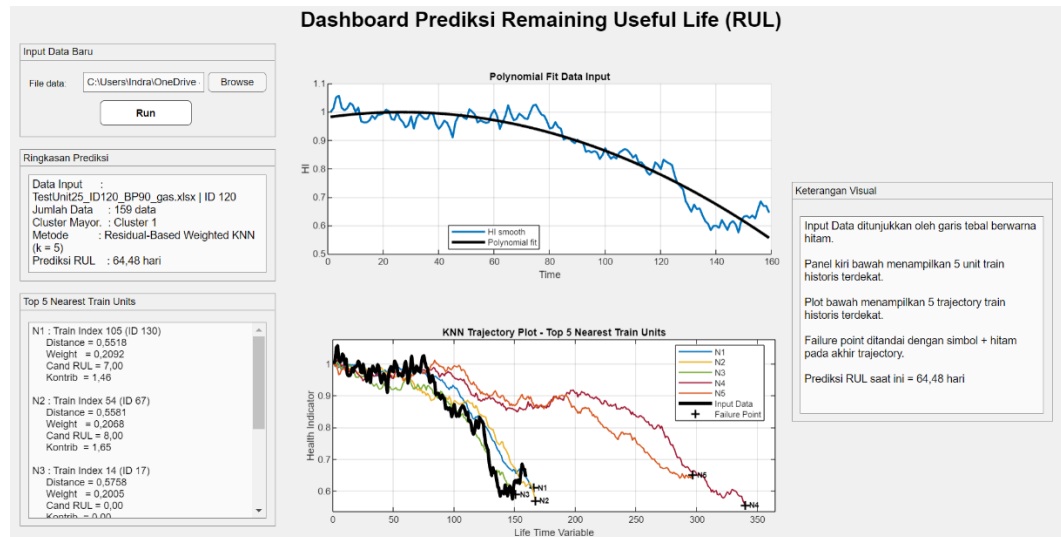
Gambar 4.15 menunjukkan proses pemilihan folder yang berisi data input sesuai skenario pengamatan, yaitu BP50, BP70, atau BP90. Penggunaan folder tersebut dilakukan untuk simulasi menggunakan data *test*. Dalam penerapan selanjutnya, pengguna dapat memasukkan data unit baru dengan struktur yang sesuai untuk memperoleh estimasi RUL.



Gambar 4.16 Pemilihan File Input Data *Test* pada Salah Satu Folder *Breakpoint*

Setelah folder dipilih, pengguna menentukan file unit yang akan dianalisis sebagaimana ditunjukkan pada gambar 4.16. File input tersimpan dalam format Excel dan berisi data historis satu unit sampai dengan *breakpoint* tertentu. Pemilihan file dilakukan sesuai skenario pengamatan yang akan diuji agar jumlah

data yang diproses sesuai dengan kondisi simulasi. File tersebut kemudian dibaca dan diperiksa oleh dashboard sebelum digunakan sebagai masukan bagi model estimator untuk menjalankan tahapan analisis dan estimasi RUL



Gambar 4.17 Tampilan Hasil Prediksi pada Dashboard

Berdasarkan gambar 4.17, setelah file dipilih dan tombol *Run* dijalankan, *dashboard* menampilkan identitas data, jumlah pengamatan, *cluster* operasional mayoritas, metode estimasi, dan nilai prediksi RUL. *Dashboard* juga menampilkan grafik *polynomial fit* serta *KNN trajectory plot* yang membandingkan unit masukan dengan lima unit historis terdekat. Informasi ini membantu pengguna melihat hasil prediksi sekaligus dasar pola degradasi dan unit referensi yang digunakan.

Pada contoh pengujian, file input berasal dari unit ID 120 pada data *test* dengan skenario *breakpoint* 90%. *Dashboard* menghasilkan estimasi RUL sebesar 64,48 hari. Tampilan tersebut menunjukkan bahwa pembacaan data, pembentukan representasi degradasi, pemilihan unit historis terdekat, dan penyajian hasil prediksi telah terintegrasi dalam satu antarmuka.

Secara keseluruhan, hasil pengujian menunjukkan bahwa *dashboard* dapat digunakan untuk menjalankan estimasi RUL pada berbagai panjang data pengamatan, baik *full trajectory* maupun data parsial pada *breakpoint* 50%, 70%, dan 90%. Dengan demikian, *dashboard* tidak hanya menampilkan nilai akhir estimasi, tetapi juga menyediakan visualisasi pola degradasi dan unit historis sebagai dasar prediksi. Video simulasi penggunaan *dashboard* secara lengkap disajikan pada Lampiran C.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan mengenai estimasi *Remaining Useful Life* (RUL) sistem traksi LRT Jabodebek menggunakan pendekatan *similarity based polynomial*, maka diperoleh beberapa kesimpulan sebagai berikut.

- Tujuan pembentukan *Health Indicator* (HI) berhasil dicapai melalui segmentasi kondisi operasional, normalisasi sensor per *cluster*, pemilihan sensor *trendable*, dan pengujian *health condition* ideal. Op3 atau torsi motor dipilih sebagai variabel *clustering*, sedangkan Sensor3, Sensor2, Sensor6, Sensor5, dan Sensor8 digunakan dalam pembentukan HI karena memiliki *trend score* tertinggi. HI dibentuk menggunakan kurva acuan $(\tau) = 1 - \tau^3$, kemudian melalui proses *smoothing* dan penyesuaian level awal sehingga dapat merepresentasikan kecenderungan degradasi sistem traksi.
- Estimator RUL berhasil dirancang menggunakan model degradasi polinomial orde dua dan *residual-based weighted KNN*. Lima unit historis dengan jarak residual terkecil digunakan sebagai *nearest neighbors* dan diberi bobot berdasarkan *inverse distance weighting*. Pada *breakpoint* 50%, model menghasilkan MAE 31,1294, RMSE 38,7466, MAPE 27,1163%, MAEfrac 13,5278%, dan RMSEfrac 16,8350%. Pada *breakpoint* 70%, nilainya menurun menjadi 17,1567, 22,7426, 24,7927%, 7,4326%, dan 9,7830%. Performa terbaik diperoleh pada *breakpoint* 90% dengan MAE 5,0832, RMSE 6,8757, MAPE 20,6347%, MAEfrac 2,0686%, dan RMSEfrac 2,6400%. Seluruh nilai MAPE berada di bawah target 30%.
- Pendekatan yang dikembangkan menghasilkan nilai MAE, RMSE, MAEfrac, dan RMSEfrac yang lebih rendah dibandingkan penelitian sebelumnya pada metrik dan *breakpoint* yang sebanding. Hasil tersebut menunjukkan bahwa pemodelan polinomial dan pembobotan berdasarkan jarak residual dapat meningkatkan performa estimasi RUL ketika informasi degradasi semakin lengkap. Selain itu, dashboard yang dikembangkan

mampu menampilkan data masukan, *trajectory* HI, *polynomial fit*, unit historis terdekat, dan hasil estimasi RUL dalam satu antarmuka untuk mendukung pemantauan kondisi sistem traksi.

5.2 Saran

Berdasarkan hasil penelitian, maka saran untuk pengembangan selanjutnya adalah sebagai berikut.

- Formulasi HI dapat dikembangkan menggunakan metode pembobotan fitur, ekstraksi fitur, atau model pembelajaran lain untuk memperoleh representasi degradasi yang lebih stabil dan sensitif.
- Estimator RUL perlu diuji menggunakan variasi pembagian data dan metode pembandingan yang lebih luas. Informasi ketidakpastian seperti *confidence interval*, variasi antar-*neighbor*, atau tingkat keyakinan model juga dapat ditambahkan.
- Dashboard dapat dikembangkan dengan integrasi penuh terhadap *pipeline* model agar proses pemuatan data, prediksi, dan penyajian hasil berjalan secara otomatis tanpa penyesuaian manual.
- Pengujian selanjutnya dapat menggunakan data yang lebih beragam, periode operasi lebih panjang, dan subsistem lain agar kemampuan generalisasi serta potensi implementasi model di lingkungan industri dapat dievaluasi lebih kuat.

DAFTAR PUSTAKA

- Ahmed, M., Seraj, R., & Islam, S. M. S. (2020). The K-Means Algorithm: A Comprehensive Survey And Performance Evaluation. In *Electronics (Switzerland)* (Vol. 9, Number 8, Pp. 1–12). MDPI AG. <https://doi.org/10.3390/Electronics9081295>
- Berghout, T., & Benbouzid, M. (2022). A Systematic Guide For Predicting Remaining Useful Life With Machine Learning. In *Electronics (Switzerland)* (Vol. 11, Number 7). MDPI. <https://doi.org/10.3390/Electronics11071125>
- Berthold Michael R. (2020). *Advances In Intelligent Data Analysis XVIII* (M. R. Berthold, A. Feelders, & G. Kreml, Eds.; 18th Ed., Vol. 12080). Springer International Publishing. <https://doi.org/10.1007/978-3-030-44584-3>
- Davari, N., Veloso, B., Costa, G. De A., Pereira, P. M., Ribeiro, R. P., & Gama, J. (2021). A Survey On Data-Driven Predictive Maintenance For The Railway Industry. In *Sensors* (Vol. 21, Number 17). MDPI AG. <https://doi.org/10.3390/S21175739>
- De Pater, I., & Mitici, M. (2022). Novel Metrics To Evaluate Probabilistic Remaining Useful Life Prognostics With Applications To Turbofan Engines. *Proceedings Of The European Conference Of The PHM Society 2022*, 96–109.
- Gere, A. (2023). Recommendations For Validating Hierarchical Clustering In Consumer Sensory Projects. *Current Research In Food Science*, 6. <https://doi.org/10.1016/J.Crfs.2023.100522>
- Halder, R. K., Uddin, M. N., Uddin, M. A., Aryal, S., & Khraisat, A. (2024). Enhancing K-Nearest Neighbor Algorithm: A Comprehensive Review And Performance Analysis Of Modifications. *Journal Of Big Data*, 11(1). <https://doi.org/10.1186/S40537-024-00973-Y>
- Iqbal Arjiansah, H. (2021). Penggunaan Suplai Daya Rel Ketiga Dan Rugi-Rugi Sistem Propulsi Pada Light Rail Transit (LRT). *Simposium Nasional RAPI XX*, 167–174.
- Lin, Y. H., Ding, Z. Q., & Li, Y. F. (2023). Similarity Based Remaining Useful Life Prediction Based On Gaussian Process With Active Learning. *Reliability Engineering And System Safety*, 238. <https://doi.org/10.1016/J.Res.2023.109461>
- Paudel, M., Lim, L. J., Yap, F. F., & Kho, K. (2021). Vibration Analysis Of The Third Rail Structure Of A Mass Rapid Transit System With Structural Defects. *Applied Sciences (Switzerland)*, 11(18). <https://doi.org/10.3390/App11188410>

- Schwartz, S., Montero Jiménez, J. J., Vingerhoeds, R., & Salaün, M. (2022). An Unsupervised Approach For Health Index Building And For Similarity-Based Remaining Useful Life Estimation. *Computers In Industry*, 141. <https://doi.org/10.1016/j.compind.2022.103716>
- Seyghaly, R., Garcia, J., Masip-Bruin, X., & Kuljanin, J. (2024). SBNNR: Small-Size Bat-Optimized KNN Regression. *Future Internet*, 16(11). <https://doi.org/10.3390/fi16110422>
- Shutaywi, M., & Kachouie, N. N. (2021). Silhouette Analysis For Performance Evaluation In Machine Learning With Applications To Clustering. *Entropy*, 23(6). <https://doi.org/10.3390/E23060759>
- Wani, A. A. (2024). Comprehensive Analysis Of Clustering Algorithms: Exploring Limitations And Innovative Solutions. *Peerj Computer Science*, 10, 1–45. <https://doi.org/10.7717/PEERJ-CS.2286>
- Xu, Z., Bashir, M., Liu, Q., Miao, Z., Wang, X., Wang, J., & Ekere, N. (2023). A Novel Health Indicator For Intelligent Prediction Of Rolling Bearing Remaining Useful Life Based On Unsupervised Learning Model. *Computers And Industrial Engineering*, 176. <https://doi.org/10.1016/j.cie.2023.108999>
- Zhu, T., Chen, Z., Zhou, D., Chen, Z., & Pan, E. (2025). Remaining Useful Life Prediction By Degradation Distribution Transport Health Indicator And Consolidated Memory Stabilized LSTM. *Mechanical Systems And Signal Processing*, 236. <https://doi.org/10.1016/j.ymssp.2025.113039>
- Feng, J., Xu, J., Liao, W., & Liu, Y. (2017). Review on the traction system sensor technology of a rail transit train. *Sensors*, 17(6), 1356. <https://doi.org/10.3390/s17061356>
- Mehboob, F., Fattouh, A., & Sahoo, S. (2024). Synergizing transfer learning and multi-agent systems for thermal parametrization in induction traction motors. *Applied Sciences*, 14(11), 4455. <https://doi.org/10.3390/app14114455>
- Zhou, X., Giangrande, P., Ji, Y., Zhao, W., Ijaz, S., & Galea, M. (2024). Insulation for rotating low-voltage electrical machines: Degradation, lifetime modeling, and accelerated aging tests. *Energies*, 17(9), 1987. <https://doi.org/10.3390/en17091987>

LAMPIRAN A
HASIL PREDIKSI RUL

ID Data Test	True RUL 50%	Est RUL 50%	Abs Error 50%	True RUL 70%	Est RUL 70%	Abs Error 70%	True RUL 90%	Est RUL 90%	Abs Error 90%
8	115	157,91	42,91	69	85,28	16,28	23	23,78	0,78
11	150	164,94	14,94	90	97,88	7,88	30	21,61	8,39
13	128	150,82	22,82	77	80,09	3,09	26	18,64	7,36
18	171	131,02	39,98	103	101,61	1,39	34	33,54	0,46
21	113	127,77	14,77	68	68,29	0,29	23	30,55	7,55
22	105	127,25	22,25	63	59,52	3,48	21	25,14	4,14
28	101	137,43	36,43	61	83,94	22,94	20	20,94	0,94
33	100	131,19	31,19	60	79,16	19,16	20	27,47	7,47
39	101	133,97	32,97	61	81,15	20,15	20	26,70	6,70
40	103	156,07	53,07	62	87,42	25,42	21	21,34	0,34
53	152	142,46	9,54	91	94,04	3,04	30	31,28	1,28
54	98	132,84	34,84	59	78,19	19,19	20	34,31	14,31
64	92	116,28	24,28	55	98,55	43,55	18	21,86	3,86
69	113	114,88	1,88	68	85,54	17,54	23	26,75	3,75
78	125	142,79	17,79	75	90,34	15,34	25	33,73	8,73
80	152	132,39	19,61	91	72,25	18,75	30	30,66	0,66
81	93	145,43	52,43	56	101,60	45,60	19	22,47	3,47
91	177	165,47	11,53	106	103,02	2,98	35	29,07	5,93
95	119	138,35	19,35	72	87,42	15,42	24	29,21	5,21
98	122	114,45	7,55	73	99,00	26,00	24	19,40	4,60
100	130	146,68	16,68	78	96,63	18,63	26	29,33	3,33
110	93	130,69	37,69	56	57,31	1,31	19	25,99	6,99
111	148	143,10	4,90	89	87,84	1,16	30	29,94	0,06
113	92	164,25	72,25	55	67,67	12,67	18	21,22	3,22
120	88	149,33	61,33	53	74,48	21,48	18	19,45	1,45
133	244	137,62	106,38	147	83,69	63,31	49	27,83	21,17

Keterangan tabel ditampilkan pada halaman berikutnya.

Keterangan Tabel :

ID Data Test menunjukkan nomor identitas unit pada data *test* yang digunakan dalam proses evaluasi model. **True RUL** merupakan nilai RUL aktual pada masing-masing *breakpoint*, yaitu 50%, 70%, dan 90% dari umur operasi unit. **Est RUL** merupakan nilai RUL hasil estimasi model menggunakan pendekatan *similarity-based polynomial* dengan *residual-based weighted KNN*. Sementara itu, **Abs Error** merupakan nilai absolut dari selisih antara Est RUL dan True RUL. Nilai Abs Error digunakan untuk menunjukkan besar kesalahan prediksi pada setiap ID data *test* tanpa memperhatikan arah kesalahan prediksi.

LAMPIRAN B
HASIL NILAI ERROR

ID Data	Error 50	Error frac50	APE 50	Error 70	Error frac70	APE 70	Error 90	Error frac90	APE 90
8	42,91	0,19	37,31	16,29	0,07	23,60	0,78	0,00	3,41
11	14,94	0,05	9,96	7,88	0,03	8,75	-8,39	-0,03	27,96
13	22,82	0,09	17,83	3,09	0,01	4,02	-7,36	-0,03	28,32
18	-39,98	-0,12	23,38	-1,39	0,00	1,35	-0,46	0,00	1,36
21	14,77	0,07	13,07	0,29	0,00	0,43	7,55	0,03	32,81
22	22,25	0,11	21,19	-3,48	-0,02	5,53	4,14	0,02	19,69
28	36,43	0,18	36,07	22,94	0,11	37,61	0,94	0,00	4,70
33	31,19	0,16	31,19	19,16	0,10	31,93	7,47	0,04	37,36
39	32,97	0,16	32,64	20,15	0,10	33,03	6,70	0,03	33,51
40	53,07	0,26	51,52	25,42	0,12	41,00	0,34	0,00	1,64
53	-9,54	-0,03	6,27	3,04	0,01	3,35	1,28	0,00	4,27
54	34,84	0,18	35,55	19,19	0,10	32,53	14,31	0,07	71,53
64	24,28	0,13	26,39	43,55	0,24	79,19	3,86	0,02	21,46
69	1,89	0,01	1,67	17,54	0,08	25,79	3,75	0,02	16,31
78	17,79	0,07	14,23	15,34	0,06	20,46	8,73	0,03	34,92
80	-19,61	-0,06	12,90	-18,75	-0,06	20,60	0,66	0,00	2,21
81	52,43	0,28	56,38	45,60	0,24	81,43	3,47	0,02	18,29
91	-11,53	-0,03	6,51	-2,98	-0,01	2,81	-5,93	-0,02	16,93
95	19,35	0,08	16,26	15,42	0,06	21,42	5,21	0,02	21,71
98	-7,55	-0,03	6,18	26,00	0,11	35,61	-4,60	-0,02	19,17
100	16,68	0,06	12,83	18,63	0,07	23,89	3,33	0,01	12,83
110	37,69	0,20	40,53	1,31	0,01	2,34	6,99	0,04	36,80
111	-4,90	-0,02	3,31	-1,16	0,00	1,30	-0,06	0,00	0,20
113	72,25	0,39	78,53	12,67	0,07	23,04	3,22	0,02	17,86
120	61,33	0,35	69,70	21,48	0,12	40,54	1,45	0,01	8,06
133	-106,38	-0,22	43,60	-63,31	-0,13	43,07	-21,17	-0,04	43,20

Keterangan tabel ditampilkan pada halaman berikutnya.

Keterangan Tabel :

Error merupakan selisih antara nilai estimasi RUL dan nilai RUL aktual, sehingga nilainya dapat positif maupun negatif. Nilai positif menunjukkan model melakukan *overestimate*, sedangkan nilai negatif menunjukkan model melakukan *underestimate*.

ErrorFrac merupakan nilai error yang dinormalisasi terhadap umur total unit. **APE** atau *Absolute Percentage Error* merupakan nilai error absolut relatif terhadap RUL aktual pada setiap ID data test. Nilai MAPE diperoleh dari rata-rata APE seluruh ID data test pada masing-masing *breakpoint*.

LAMPIRAN C

TAUTAN DAN QR CODE LAMPIRAN DIGITAL

Link Lampiran Digital:

<https://its.id/m/TaRulUwais>

QR Code Lampiran Digital:



Keterangan:

Lampiran C berisi tautan dan **QR code** menuju folder lampiran digital penelitian. Folder tersebut memuat kode program MATLAB, video simulasi penggunaan *dashboard* prediksi RUL, serta file output pendukung yang digunakan dalam penelitian ini. Kode program MATLAB mencakup tahapan pembacaan data historis, pembagian data *train* dan *test*, pengujian variabel operasi, *clustering*, normalisasi sensor, pembentukan *Health Indicator*, pemodelan degradasi polinomial, estimasi RUL menggunakan *residual-based weighted KNN*, serta evaluasi model. Video simulasi disertakan untuk menunjukkan penggunaan *dashboard* prediksi RUL, mulai dari pemilihan file input hingga tampilan hasil estimasi RUL dan visualisasi pendukungnya.

Halaman ini sengaja dikosongkan

BIODATA PENULIS



Penulis bernama Muhammad Jamil Uwais Saragih Sumbayak dilahirkan di Jakarta, 21 Februari 2003, merupakan anak pertama dari 2 bersaudara. Penulis telah menempuh pendidikan formal yaitu di TK An-Najjah Jakarta Timur, SDS Krishna Jakarta timur, SMPN 115 Jakarta Selatan, dan SMAN 54 Jakarta Timur . Setelah lulus dari SMA tahun 2021, Penulis diterima di Departemen Teknik Fisika FTIRS - ITS pada tahun 2021 dan terdaftar dengan NRP 5009211123. Selama masa perkuliahan, penulis aktif berkegiatan didalam maupun diluar lingkup ITS seperti mengikuti berbagai kepanitiaan acara kampus dan juga ikut serta menjadi tim dari aksi pengabdian masyarakat ke daerah 3T di Labuan Bajo, Wakatobi, dan Sumba. Pengalaman di Industri, penulis melaksanakan kerja praktik di PT. Kilang Pertamina Balikpapan pada proyek *Refinery Development Master Plan* RU V Balikpapan divisi *Construction* subdivisi *Instrument*. Penulis melaksanakan penelitian tugas akhir di Laboratorium *Embedded and Cyber-Physical Systems* Teknik Fisika ITS dengan topik “*Remaining Useful Life*”