



## A K-NN Based Model for Predicting Electrical Grid Failures from Maintenance Records

Devina Harfia<sup>1</sup>, Mahdi<sup>2</sup>, Umri Erdiansyah<sup>3</sup>

<sup>1,2,3</sup> Jurusan Teknologi Informasi dan Komputer, Prodi Teknik Informatika, Politeknik Negeri Lhokseumawe, Jln. B. Aceh - Medan  
Km. 280 Buketrata 24301, Indonesia

\*Penulis Korespondensi : mahdi@pnl.ac.id

### INFORMASI ARTIKEL

*Riwayat artikel:*

Diajukan pada 03 November 2025

Direvisi pada 21 November 2025

Publikasi pada 20 Desember 2025

*Kata kunci:*

*K-Nearest Neighbor* (K-NN)

Klasifikasi

Pemeliharaan

Listrik

PLN.

*Keywords:*

*K-Nearest Neighbor* (K-NN)

Classification

Maintenance

Electricity

PLN.

### ABSTRAK

Gangguan pada jaringan distribusi listrik dapat menyebabkan pemadaman dan menurunkan keandalan sistem. Namun, pemeliharaan rutin PLN belum dilengkapi sistem prediktif yang mampu memanfaatkan data historis secara optimal untuk mengidentifikasi potensi risiko. Penelitian ini bertujuan mengembangkan sistem klasifikasi risiko gangguan jaringan listrik menggunakan metode *K-Nearest Neighbor* (K-NN). Data yang digunakan berasal dari pemeliharaan gardu dan Saluran Udara Tegangan Menengah (SUTM), yang kemudian diproses melalui tahapan *preprocessing*, pelabelan tingkat risiko dengan clustering, serta klasifikasi menggunakan K-NN. Pengujian dilakukan pada 13 penyulang, dengan hasil klasifikasi 8 penyulang berisiko RENDAH, 3 berisiko SEDANG, dan 2 berisiko TINGGI. Nilai parameter K ditentukan melalui *cross-validation*, di mana K = 5 dipilih sebagai nilai optimal. Hasil evaluasi menunjukkan akurasi sebesar 83,75%, dengan performa terbaik pada kelas RENDAH (*precision*, *recall*, dan *F1-score* masing-masing 93,41%). Kelas TINGGI memperoleh *precision* 82,76% dan *F1-score* 76,19%, sedangkan kelas SEDANG masih lebih rendah dengan *precision* 62,5% dan *F1-score* 66,67%. Secara keseluruhan, model menghasilkan rata-rata *precision* 79,6%, *recall* 78,5%, dan *F1-score* 78,8%. Temuan ini membuktikan bahwa metode K-NN cukup efektif dalam memprediksi risiko gangguan jaringan listrik dan dapat dijadikan alat bantu pengambilan keputusan pemeliharaan preventif di lingkungan PLN.

### ABSTRACT

Disturbances in distribution networks can cause outages and reduce system reliability. However, PLN's routine maintenance is not yet supported by predictive systems that optimally utilize historical data to identify potential risks. This study aims to develop a risk classification system for distribution network disturbances using the *K-Nearest Neighbor* (K-NN) method. The data were collected from substation and Medium Voltage Overhead Line (SUTM) maintenance, then processed through *preprocessing*, risk labeling using clustering, and classification with K-NN. The system was tested on 13 feeders, resulting in 8 categorized as LOW risk, 3 as MEDIUM, and 2 as HIGH. The parameter K was determined through *cross-validation*, where K = 5 was selected as the optimal value. Evaluation results showed an accuracy of 83.75%, with the best performance obtained in the LOW class (*precision*, *recall*, and *F1-score* of 93.41%). The HIGH class achieved 82.76% *precision* and 76.19% *F1-score*, while the MEDIUM class was relatively lower with 62.5% *precision* and 66.67% *F1-score*. Overall, the model achieved an average *precision* of 79.6%, *recall* of 78.5%, and *F1-score* of 78.8%. These findings indicate that the K-NN method is fairly effective in predicting distribution network disturbance risks and can serve as a decision-support tool for preventive maintenance in PLN's operations.

## 1. Pendahuluan

Energi listrik merupakan kebutuhan mendasar bagi masyarakat di Indonesia. Penggunaannya tersebar di berbagai sektor, termasuk industri, komersial, dan rumah tangga [1]. Secara umum, energi listrik yang dihasilkan dari pusat pembangkit akan disalurkan melalui saluran transmisi tegangan tinggi (high voltage) ke pusat-pusat beban, seperti gardu induk. Penggunaan tegangan tinggi ini bertujuan untuk meminimalkan kehilangan daya selama proses transmisi [2][3]. Pemeliharaan jaringan listrik, khususnya pada jaringan tegangan menengah, merupakan langkah penting untuk menjaga agar infrastruktur tetap berfungsi dengan baik dan untuk mencegah kerusakan yang dapat menimbulkan gangguan [4]. Saluran distribusi, yang mayoritas berbentuk saluran udara, sangat rentan terhadap berbagai gangguan, baik yang disebabkan oleh faktor internal maupun eksternal, seperti kondisi cuaca atau vegetasi [2][5].

Pemeliharaan gardu dan Saluran Udara Tegangan Menengah (SUTM) merupakan langkah penting untuk menjaga keandalan jaringan listrik [6]. Kegiatan ini bertujuan mengatasi penurunan efisiensi dan kerusakan pada perlengkapan jaringan, seperti ausnya isolator, ketidakseimbangan beban gardu, dan gangguan akibat pertumbuhan vegetasi yang terlalu dekat dengan saluran [7]. Penurunan efisiensi tersebut dapat menyebabkan lonjakan arus, panas berlebih, hingga pemadaman mendadak, yang berpotensi menimbulkan kerugian bagi masyarakat maupun penyedia layanan [3][8]. PT PLN ULP Lhoksukon secara rutin melakukan pemeliharaan gardu dan SUTM di wilayah kerjanya, seperti Cot Girek, Lhoksukon, Matang Kuli, Bukit Hagu, dan sekitarnya. Kegiatan ini meliputi inspeksi, perbaikan, serta pemasangan alat pengaman, yang dilakukan untuk meminimalkan risiko gangguan dan memastikan perangkat bekerja optimal sesuai fungsinya. Namun, efektivitas pemeliharaan masih memerlukan pengelolaan yang lebih terarah guna mengurangi risiko kerusakan secara lebih signifikan [4][9]. Meskipun pemeliharaan jaringan listrik telah dilakukan secara rutin, gangguan masih kerap terjadi akibat proses identifikasi dan analisis risiko yang belum sepenuhnya optimal [10]. Gangguan seperti pemadaman mendadak, kerusakan isolator, serta distribusi beban yang tidak seimbang sering kali tidak terdeteksi lebih awal. Hal ini disebabkan oleh metode pemeliharaan yang masih bergantung pada inspeksi manual dan analisis konvensional, tanpa pemanfaatan maksimal terhadap data historis yang tersedia [11][12]. Padahal, data hasil inspeksi, perbaikan, serta pemasangan alat pengaman menyimpan informasi penting yang dapat digunakan untuk memperkirakan potensi gangguan [13]. Ketidakefektifan dalam memanfaatkan data ini menyebabkan tindakan preventif menjadi kurang tepat sasaran dan risiko gangguan tetap tinggi [14][15].

Penelitian ini memanfaatkan pendekatan klasifikasi berbasis algoritma sebagai solusi untuk menganalisis potensi risiko gangguan pada jaringan listrik. Salah satu metode yang sesuai untuk tugas klasifikasi ini adalah algoritma *K-Nearest Neighbor* (K-NN), yang bekerja dengan menentukan kelas suatu data berdasarkan kedekatan jarak terhadap data lain yang telah dikategorikan sebelumnya [16][17]. Algoritma ini efektif dalam mengelompokkan data ke dalam kategori tertentu berdasarkan kemiripan atribut, khususnya pada data yang bersifat numerik dan memiliki hubungan spasial antar fitur [18][19]. Penggunaan K-NN memungkinkan sistem untuk mengenali pola dari data historis dan menerapkannya pada data baru dengan proses yang relatif sederhana namun akurat [20].

Penelitian ini bertujuan untuk menyelesaikan permasalahan tersebut dengan mengimplementasikan *K-Nearest Neighbor* (K-NN) sebagai pendekatan klasifikasi dalam menganalisis risiko gangguan pada jaringan listrik. Dengan memanfaatkan data historis pemeliharaan, algoritma ini membangun model prediksi yang mampu mengklasifikasikan tingkat risiko gangguan ke dalam tiga kategori, yaitu tinggi, sedang, dan rendah. Klasifikasi tersebut digunakan untuk memetakan kondisi jaringan serta memberikan gambaran mengenai prioritas perbaikan dan tindakan pencegahan yang diperlukan. Informasi yang dihasilkan dapat mendukung proses penjadwalan pemeliharaan dan pengalokasian sumber daya secara lebih efisien, sehingga membantu mengurangi potensi gangguan berulang dan meningkatkan keandalan jaringan distribusi listrik.

## 2. Metode

### 2.1 Deskripsi Data

Data yang digunakan dalam penelitian ini diperoleh dari dokumentasi historis kegiatan pemeliharaan jaringan distribusi listrik di PT PLN (Persero) ULP Lhoksukon. Data mencakup dua komponen utama infrastruktur, yaitu gardu distribusi dan Saluran Udara Tegangan Menengah (SUTM). Pengumpulan data dilakukan secara sistematis melalui arsip internal dan sistem informasi PLN selama periode Maret 2023 hingga Februari 2025, sehingga mencakup 24 bulan pengamatan terhadap 12 penyulang aktif di wilayah kerja ULP Lhoksukon.

Seluruh data tersebut kemudian diproses dan digunakan sebagai dataset utama dalam pembangunan model klasifikasi risiko gangguan. Secara substansi, data pemeliharaan gardu terdiri atas kegiatan inspeksi rutin (Tier 1), pemeriksaan teknis dengan alat bantu (Tier 2), serta *Condition-Based Maintenance* (CBM) yang meliputi pengukuran parameter operasional transformator, penggantian komponen yang mengalami degradasi, hingga relokasi gardu akibat perubahan beban atau kondisi lingkungan. Sementara itu, data pemeliharaan SUTM mencakup inspeksi visual jalur distribusi, pemangkasan vegetasi atau pembersihan *right of way*, serta perbaikan perangkat pelindung seperti arrester dan *fuse cut-out* (FCO). Selain itu, tindakan pengamanan terhadap gangguan fauna juga dicatat, termasuk pemasangan pelindung hewan pada konduktor dan transformator distribusi. Dalam penelitian ini, risiko gangguan jaringan diklasifikasikan ke dalam empat dimensi utama dengan bobot sesuai tingkat kontribusinya terhadap potensi gangguan, yaitu:

1. Skor Inspeksi (35%),
2. Skor Pemeliharaan (30%),
3. Skor Perbaikan Kritis (25%), dan
4. Skor Kondisi Infrastruktur (10%).

Pembobotan tersebut dirancang untuk merepresentasikan proporsi pengaruh masing-masing faktor terhadap potensi risiko gangguan pada sistem distribusi listrik. Setelah proses pembersihan (data *cleaning*) dari entri yang tidak relevan, tindakan minor dengan frekuensi sangat rendah, serta data duplikat, diperoleh 791 entri data akhir yang siap digunakan dalam tahap pra-pemrosesan (*preprocessing*) dan implementasi

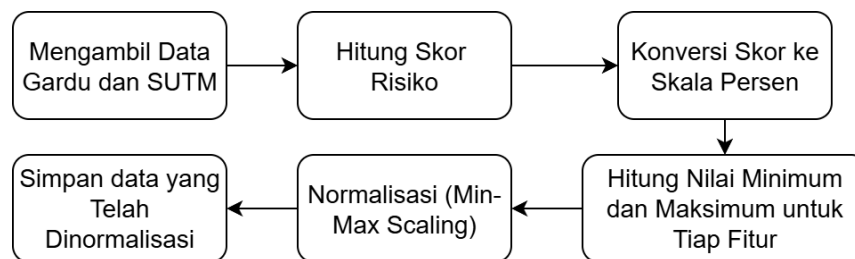
algoritma *K-Nearest Neighbor* (K-NN). Dataset akhir ini selanjutnya diklasifikasikan ke dalam tiga kategori tingkat risiko, yaitu:

1. Risiko Rendah: 454 entri,
2. Risiko Sedang: 171 entri, dan
3. Risiko Tinggi: 166 entri.

Pembagian kategori tersebut menjadi dasar dalam proses pelatihan dan pengujian model K-NN untuk menentukan pola klasifikasi risiko gangguan jaringan listrik berdasarkan data historis pemeliharaan PLN.

## 2.2 Tahapan *preprocessing* dan *labelling*

Proses *preprocessing* merupakan tahap awal yang krusial dalam penelitian ini, karena memastikan bahwa seluruh fitur (atribut) dalam dataset memiliki rentang nilai yang seragam sehingga tidak ada satu fitur pun yang mendominasi proses perhitungan jarak pada algoritma *K-Nearest Neighbor* (K-NN). Tahapan *preprocessing* ditunjukkan secara konseptual pada Gambar 1 Rancangan *Preprocessing*.

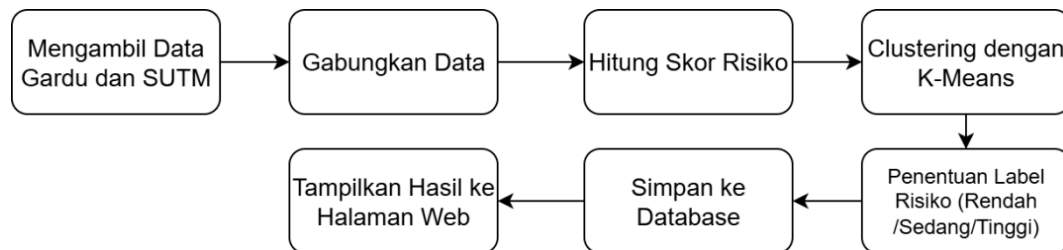


Gambar 1. Rancangan *Preprocessing*

Proses dimulai dengan pengambilan data mentah hasil pemeliharaan gardu distribusi dan jaringan Saluran Udara Tegangan Menengah (SUTM) yang masih dalam bentuk asli dan belum memiliki nilai risiko numerik. Masing-masing skor dihitung melalui fungsi evaluasi terpisah sesuai jenis kegiatan pemeliharaan, dan hasilnya dikonversi ke skala 0–100 persen untuk memudahkan normalisasi. Selanjutnya, sistem menghitung nilai minimum dan maksimum pada setiap fitur untuk menjadi acuan penerapan metode *Min-Max Scaling*. Metode ini digunakan agar seluruh nilai fitur berada pada rentang 0 hingga 1, dengan rumus sebagai berikut:

Apabila suatu fitur memiliki nilai minimum dan maksimum yang sama, maka sistem secara otomatis menetapkan nilai normalisasi sebesar 0 guna menghindari pembagian dengan nol (*division by zero*). Proses ini dilakukan secara berulang untuk setiap atribut dalam dataset hingga menghasilkan data baru yang telah ternormalisasi secara konsisten. Dataset hasil normalisasi ini selanjutnya digunakan dalam tahap clustering (*K-Means*) dan klasifikasi (K-NN). Dengan demikian, skala nilai antar fitur menjadi seimbang dan proses perhitungan jarak Euclidean dapat menghasilkan hasil klasifikasi yang lebih akurat. Tahap labeling

berfungsi untuk memberikan kategori risiko gangguan jaringan listrik pada setiap unit data berdasarkan hasil pengolahan data pemeliharaan. Alur lengkap proses labeling ditunjukkan pada Gambar 2 Rancangan Labeling.



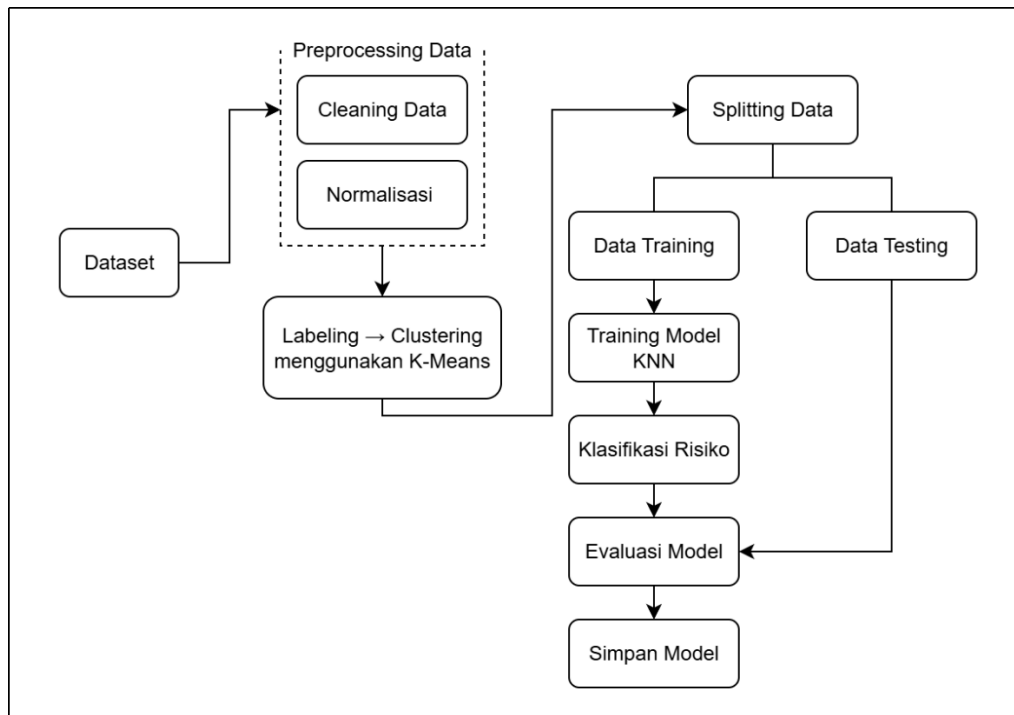
Gambar 2. Rancangan Labeling

Data yang digunakan dalam proses ini diperoleh dari dua sumber utama, yaitu pemeliharaan gardu distribusi dan pemeliharaan jaringan SUTM. Kedua data tersebut mencakup aspek inspeksi visual, aktivitas pemeliharaan rutin, tindakan perbaikan kritis, dan penilaian kondisi infrastruktur fisik. Setelah kedua sumber data dikumpulkan, sistem melakukan penggabungan (*merging*) berdasarkan identitas lokasi seperti ID penyulang atau koordinat jaringan, agar setiap unit data mewakili kondisi jaringan tertentu secara komprehensif. Setelah data tergabung, sistem menghitung skor risiko total menggunakan bobot yang telah ditentukan sebelumnya.

Proses pelabelan ini dilakukan secara otomatis, tanpa intervensi manual, untuk menjaga objektivitas dan konsistensi hasil. Label yang dihasilkan disimpan ke dalam basis data, dan juga ditampilkan melalui antarmuka web aplikasi dalam bentuk tabel serta diagram visualisasi distribusi risiko per penyulang. Visualisasi ini membantu pihak teknis PLN dalam melakukan monitoring kondisi jaringan dan pengambilan keputusan pemeliharaan preventif secara berbasis data. Hasil akhir proses labeling kemudian digunakan sebagai dataset berlabel untuk tahap pelatihan model klasifikasi K-NN pada bagian selanjutnya.

### 2.3 Algoritma K-NN

Metode penelitian yang digunakan dalam studi ini adalah *K-Nearest Neighbor* (K-NN), yaitu algoritma klasifikasi berbasis kemiripan (*similarity-based*) yang menentukan kelas suatu data baru berdasarkan kedekatannya terhadap data latih. Algoritma ini bekerja dengan mencari sejumlah tetangga terdekat (*nearest neighbors*) dari data baru, kemudian menetapkan kelas berdasarkan mayoritas label dari tetangga tersebut. Arsitektur sistem klasifikasi risiko gangguan jaringan listrik ditunjukkan pada Gambar 3 Arsitektur Umum Metode K-NN.



Gambar 3. Arsitektur Umum Metode K-NN

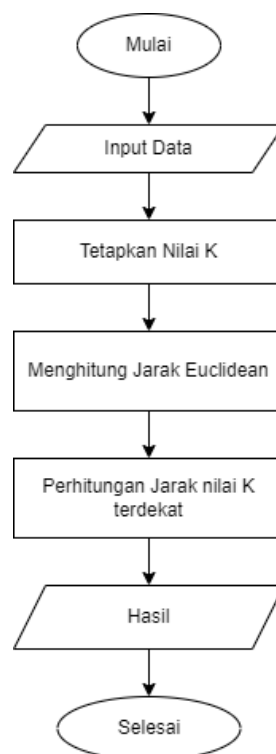
Secara keseluruhan, sistem terdiri atas beberapa tahapan yang terintegrasi dan saling berurutan, dimulai dari pengumpulan data pemeliharaan jaringan listrik, yang mencakup aktivitas inspeksi, perbaikan, dan kondisi infrastruktur gardu serta Saluran Udara Tegangan Menengah (SUTM). Data mentah hasil kegiatan tersebut selanjutnya melewati tahap *preprocessing*, yang meliputi pembersihan data dari nilai kosong atau tidak valid, serta normalisasi skala numerik menggunakan *Min-max Scaling* untuk memastikan kesetaraan bobot antar fitur. Tahap berikutnya adalah labeling otomatis menggunakan algoritma *K-Means Clustering*, di mana data dikelompokkan ke dalam tiga kategori risiko rendah, sedang, dan tinggi berdasarkan skor risiko gabungan dari empat dimensi utama (inspeksi, pemeliharaan, perbaikan kritis, dan infrastruktur). Setelah setiap data memiliki label risiko, dataset dibagi menjadi dua subset, yaitu data latih (*training*) dan data uji (*testing*), dengan teknik *stratified splitting* untuk menjaga proporsi distribusi kelas. Proses klasifikasi dilakukan menggunakan algoritma K-NN, yang menghitung jarak antara data uji dan seluruh data latih dengan rumus Euclidean Distance sebagai berikut:

$$d_{iS}(x_1, x_2) = \sqrt{\sum_{i=0}^n (x_{1i} - x_{2i})^2}$$

Secara keseluruhan, dengan masing-masing menyatakan nilai fitur ke-*i* dari data uji dan data latih. Sistem memilih *K* tetangga terdekat berdasarkan nilai jarak terkecil, kemudian menentukan kelas risiko akhir berdasarkan voting mayoritas label tetangga tersebut. Seluruh proses dilakukan secara terintegrasi di sisi server, menggunakan bahasa pemrograman PHP dengan basis data MySQL tanpa arsitektur layanan API terpisah. Pendekatan ini menghasilkan sistem yang sederhana, efisien, dan mudah diimplementasikan

pada lingkungan operasional PLN. Hasil klasifikasi yang dihasilkan ditampilkan melalui antarmuka web dan disimpan dalam basis data sebagai referensi untuk analisis lanjutan dan pelaporan periodik kondisi jaringan distribusi listrik.

Desain metode K-NN divisualisasikan yang menggambarkan alur kerja sistem klasifikasi risiko gangguan jaringan listrik secara keseluruhan. Proses dimulai dengan inisialisasi sistem dan pemanggilan dataset historis pemeliharaan jaringan listrik. Dataset ini mencakup berbagai parameter teknis seperti frekuensi gangguan, jenis tindakan pemeliharaan (preventif/korektif), hasil inspeksi visual lapangan, dan kondisi infrastruktur fisik. Seluruh parameter tersebut menjadi fitur utama dalam proses pembelajaran algoritma. Berikut Gambar 4 merupakan Flowchart Metode K-NN.



Gambar 4. Flowchart Metode K-NN

Setelah dataset siap digunakan, sistem menentukan nilai parameter  $K$ , yaitu jumlah tetangga terdekat yang dipertimbangkan dalam proses klasifikasi. Penentuan nilai  $K$  dilakukan melalui proses *cross-validation*, di mana beberapa kandidat nilai  $K$  (3, 5, 7, ...) diuji untuk memperoleh akurasi rata-rata terbaik. Berdasarkan hasil pengujian, diperoleh nilai  $K = 5$  sebagai parameter optimal, karena menghasilkan akurasi tertinggi dan stabilitas klasifikasi terbaik. Langkah selanjutnya adalah menghitung jarak Euclidean antara setiap data uji dengan seluruh data latih. Data-data ini kemudian diurutkan berdasarkan jarak terkecil, dan sistem memilih lima data terdekat sebagai referensi klasifikasi. Kelas risiko ditentukan melalui mekanisme voting mayoritas, di mana kelas yang paling sering muncul di antara tetangga terdekat menjadi hasil klasifikasi akhir.

Proses klasifikasi menghasilkan label risiko akhir (*low*, *medium*, *high*) untuk setiap unit jaringan distribusi. Hasil ini ditampilkan pada antarmuka web dalam bentuk tabel dan visualisasi grafik, sehingga

memudahkan pengguna dalam meninjau distribusi risiko per penyulang. Informasi yang dihasilkan dapat dimanfaatkan oleh pihak teknis PLN untuk menentukan prioritas pemeliharaan preventif, seperti:

1. Penanganan cepat pada penyulang berisiko tinggi,
2. Pemantauan rutin pada penyulang berisiko sedang, dan
3. Pemeliharaan periodik untuk penyulang berisiko rendah.

Dengan demikian, rancangan metode K-NN yang diimplementasikan tidak hanya berperan dalam klasifikasi teknis, tetapi juga menjadi alat bantu pengambilan keputusan strategis untuk meningkatkan efisiensi, keandalan, dan keberlanjutan operasi sistem distribusi listrik.

## 2.4 Proses *training* dan evaluasi

Data yang telah melewati tahap *preprocessing* akan dibagi menjadi dua kategori utama: data latih dan data uji. Pengujian dilakukan untuk mengevaluasi performa model dalam mengklasifikasikan data berdasarkan metode *K-Nearest Neighbor* (K-NN). Data kemudian dikelompokkan ke dalam dua kategori seperti yang ditunjukkan pada Tabel 1.

Tabel 1. Pembagian Label Data

No	Label	Jumlah Data
1	Data Latih	80%
2	Data Uji	20%

Data pada Tabel 1 menunjukkan jumlah data yang diklasifikasikan berdasarkan kategori risiko gangguan jaringan listrik. Pembagian data dilakukan dengan proporsi 80% sebagai data latih dan 20% sebagai data uji untuk memastikan model dapat belajar dengan optimal sebelum dilakukan pengujian.

## 3. Hasil Dan Pembahasan

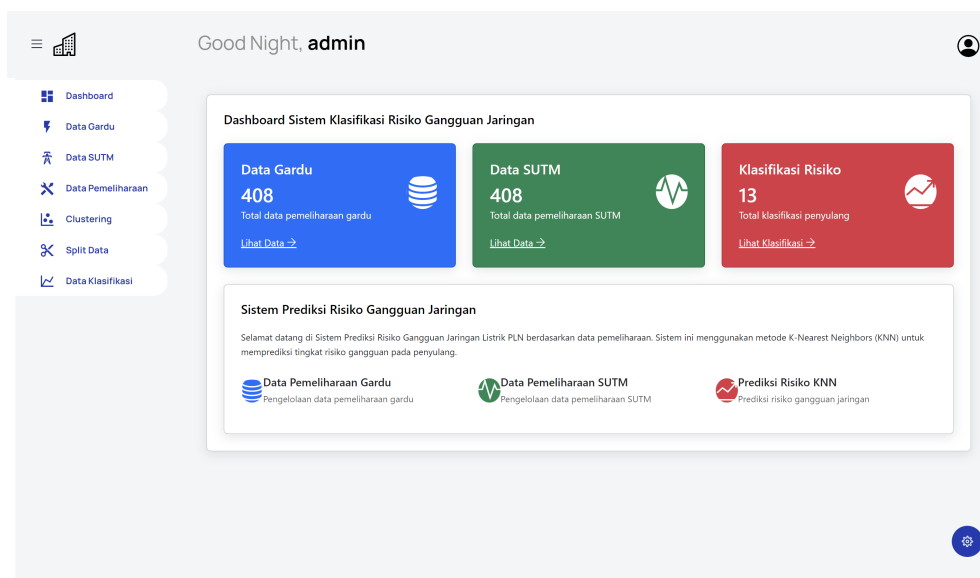
### 3.1 Tampilan dan hasil implementasi

Bagian ini menjelaskan hasil implementasi antarmuka pengguna (user interface) dari sistem klasifikasi risiko gangguan jaringan listrik berbasis metode *K-Nearest Neighbor* (K-NN). Implementasi antarmuka ini bertujuan untuk mempermudah pengguna (admin PLN) dalam mengelola data, menjalankan proses analisis, serta menampilkan hasil klasifikasi secara interaktif dan informatif. Setiap tampilan dirancang dengan memperhatikan aspek kemudahan penggunaan (*usability*), keterbacaan (*readability*), dan efisiensi proses, sehingga seluruh proses mulai dari input data hingga evaluasi hasil dapat dilakukan secara terintegrasi.

#### 1. Tampilan *Dashboard*

Halaman *Dashboard* merupakan pusat kendali utama sistem setelah pengguna berhasil login. Pada halaman ini ditampilkan ringkasan informasi penting seperti jumlah data gardu dan SUTM yang telah diunggah, status terkini hasil klasifikasi, serta rekap aktivitas pemeliharaan dalam bentuk kartu informasi dan grafik ringkas. *Dashboard* dirancang dengan antarmuka yang sederhana dan responsif, memudahkan

pengguna memantau data secara real-time tanpa perlu membuka menu tambahan. Tampilan *Dashboard* disajikan pada Gambar 5.



Gambar 5. Hasil Tampilan *Dashboard*

## 2. Tampilan Data Gardu

Halaman data gardu digunakan untuk menampilkan dan mengelola informasi terkait kegiatan pemeliharaan pada gardu distribusi. Data mencakup identitas gardu, nama penyulang, serta jenis tindakan teknis seperti inspeksi, penggantian komponen, dan pengamanan lingkungan. Sistem menyediakan fitur tambah, ubah, hapus, serta unggah data secara massal melalui file Excel. Dilengkapi pula fitur pencarian dan pengurutan otomatis untuk mempermudah pengelolaan data. Tampilan halaman data gardu dapat dilihat pada Gambar 6.

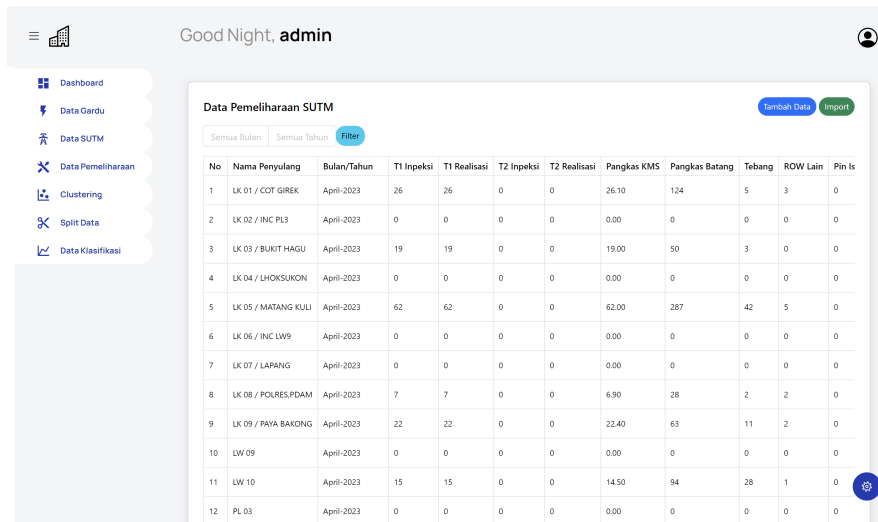
The table displays maintenance data for various substations (Gardu) in March 2023. The columns include: No, Nama Penyulang, Bulan/Tahun, T1 Inspeksi, T1 Realisasi, T2 Inspeksi, T2 Realisasi, Pengukuran, Pergantian Arrester, Pergantian FCO, and Relokas.

No	Nama Penyulang	Bulan/Tahun	T1 Inspeksi	T1 Realisasi	T2 Inspeksi	T2 Realisasi	Pengukuran	Pergantian Arrester	Pergantian FCO	Relokas
1	LK 01 / COT GIREK	Maret-2023	297	297	0	0	24	0	0	0
2	LK 02 / INC PL3	Maret-2023	0	0	0	0	0	0	0	0
3	LK 03 / BUKIT HAGU	Maret-2023	115	115	0	0	9	0	0	0
4	LK 04 / LHOXSUKON	Maret-2023	18	18	0	0	11	0	0	0
5	LK 05 / MATANG KULU	Maret-2023	173	173	0	0	24	0	0	0
6	LK 06 / INC LW9	Maret-2023	0	0	0	0	0	0	0	0
7	LK 07 / LAPANG	Maret-2023	0	0	0	0	11	0	0	0
8	LK 08 / POLRES.PDAM	Maret-2023	47	47	0	0	17	0	0	0
9	LK 09 / PAVIA BAKONG	Maret-2023	64	64	0	0	24	0	0	0
10	LW 09	Maret-2023	0	0	0	0	0	0	0	0
11	LW 10	Maret-2023	2	2	0	0	25	0	0	0

Gambar 6. Hasil Tampilan Data Gardu

### 3. Tampilan Data SUTM

Halaman ini berfungsi menampilkan data pemeliharaan pada Saluran Udara Tegangan Menengah (SUTM), mencakup aktivitas seperti inspeksi visual, pemangkasan vegetasi, perbaikan komponen pelindung (arrester, FCO), dan tindakan pengamanan jalur. Data SUTM dapat dimasukkan secara manual atau diimpor dari file Excel untuk efisiensi. Informasi ini menjadi sumber utama pada tahap clustering dan pembentukan dataset model K-NN. Tampilan halaman SUTM dapat dilihat pada Gambar 7.

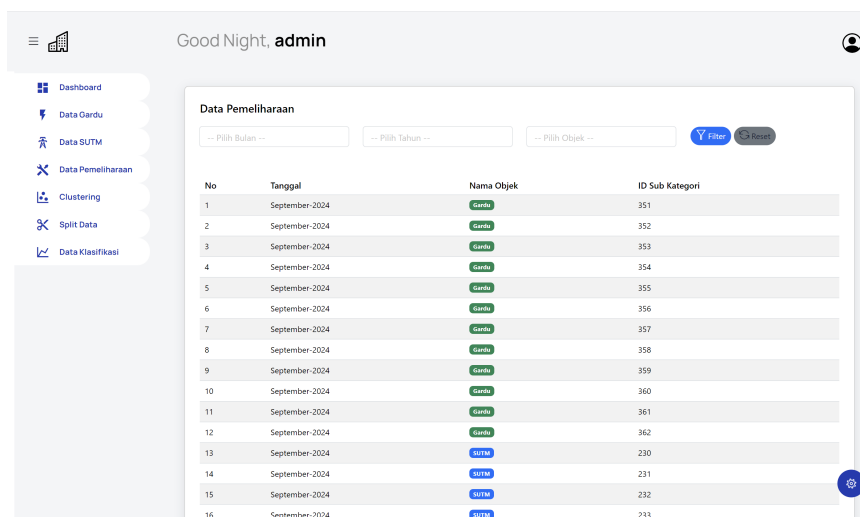


No	Nama Penyulang	Bulan/Tahun	T1 Inspeksi	T1 Realisasi	T2 Inspeksi	T2 Realisasi	Pangkas KMS	Pangkas Batang	Tebang	ROW Lain	Pin Is
1	LK 01 / COT GIREK	April-2023	26	26	0	0	26.10	124	5	3	0
2	LK 02 / INC PL3	April-2023	0	0	0	0	0.00	0	0	0	0
3	LK 03 / BUKIT HAGU	April-2023	19	19	0	0	19.00	50	3	0	0
4	LK 04 / LHOXSUNON	April-2023	0	0	0	0	0.00	0	0	0	0
5	LK 05 / MATANG KULU	April-2023	62	62	0	0	62.00	287	42	5	0
6	LK 06 / INC LW9	April-2023	0	0	0	0	0.00	0	0	0	0
7	LK 07 / LAPANG	April-2023	0	0	0	0	0.00	0	0	0	0
8	LK 08 / POLRES.PDAM	April-2023	7	7	0	0	6.90	28	2	2	0
9	LK 09 / PAHA BAKONG	April-2023	22	22	0	0	22.40	63	11	2	0
10	LW 09	April-2023	0	0	0	0	0.00	0	0	0	0
11	LW 10	April-2023	15	15	0	0	14.50	94	28	1	0
12	PL 03	April-2023	0	0	0	0	0.00	0	0	0	0

Gambar 7. Hasil Tampilan Data SUTM

### 4. Tampilan Data Pemeliharaan

Halaman data pemeliharaan menyajikan rekap aktivitas pemeliharaan gabungan dari gardu dan SUTM. Informasi ditampilkan berdasarkan tanggal pelaksanaan, jenis tindakan, dan objek pemeliharaan. Rekapitulasi ini berfungsi sebagai basis evaluasi dan validasi data sebelum proses *preprocessing* serta menjadi acuan dalam proses labeling risiko. Tampilan ini ditunjukkan pada Gambar 8.

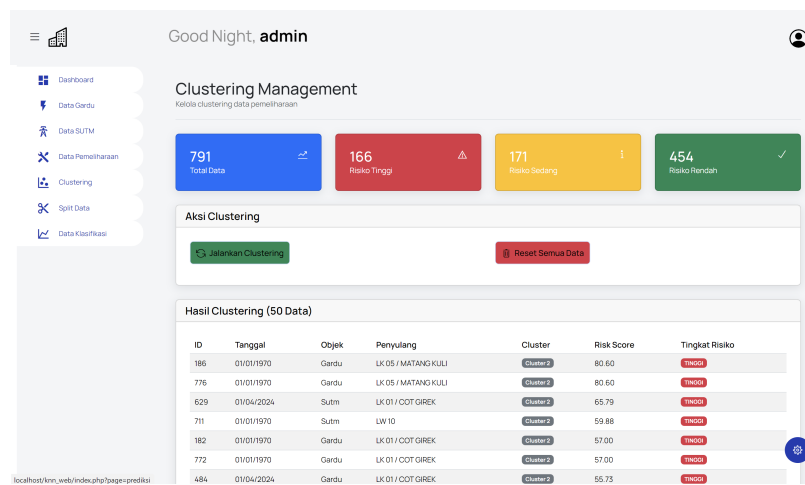


No	Tanggal	Nama Objek	ID Sub Kategori
1	September-2024	Gardu	351
2	September-2024	Gardu	352
3	September-2024	Gardu	353
4	September-2024	Gardu	354
5	September-2024	Gardu	355
6	September-2024	Gardu	356
7	September-2024	Gardu	357
8	September-2024	Gardu	358
9	September-2024	Gardu	359
10	September-2024	Gardu	360
11	September-2024	Gardu	361
12	September-2024	Gardu	362
13	September-2024	SUTM	230
14	September-2024	SUTM	231
15	September-2024	SUTM	232
16	September-2024	SUTM	233

Gambar 8. Hasil Tampilan Data Pemeliharaan

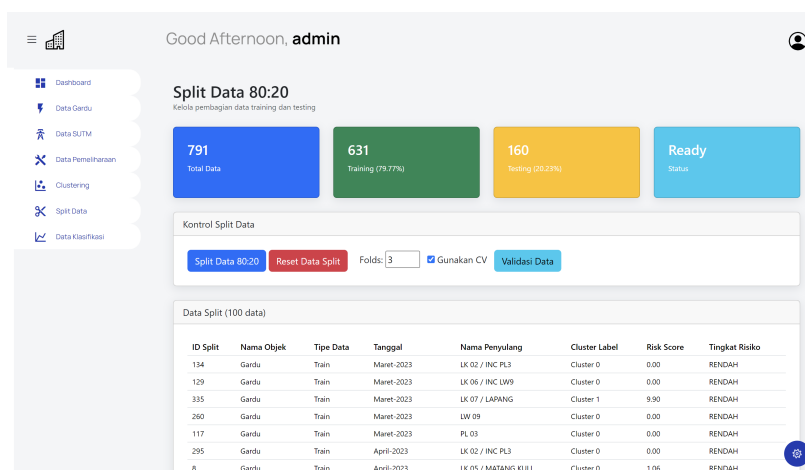
### 5. Tampilan Clustering

Halaman clustering memungkinkan pengguna menjalankan proses pengelompokan data pemeliharaan menggunakan algoritma *K-Means* secara otomatis. Sistem telah menetapkan parameter jumlah cluster dan batas iterasi secara default, sehingga admin cukup menekan tombol “Proses”. Data kemudian dikelompokkan menjadi tiga kategori risiko: tinggi, sedang, dan rendah. Hasil pengelompokan divisualisasikan dalam bentuk tabel dan diagram, serta digunakan sebagai dasar pembentukan label risiko untuk tahap pelatihan model K-NN. Tampilan ini dapat dilihat pada Gambar 9.



Gambar 9. Hasil Tampilan Clustering

### 6. Tampilan Split Data



Gambar 10. Hasil Tampilan Split Data

Setelah proses clustering selesai, sistem menyediakan fitur split data untuk membagi dataset menjadi dua bagian, yaitu data latih (*training*) dan data uji (*testing*). Proporsi pembagian yang digunakan adalah 80% data latih dan 20% data uji, untuk menjamin proses pelatihan dan evaluasi berjalan seimbang dan tidak bias.

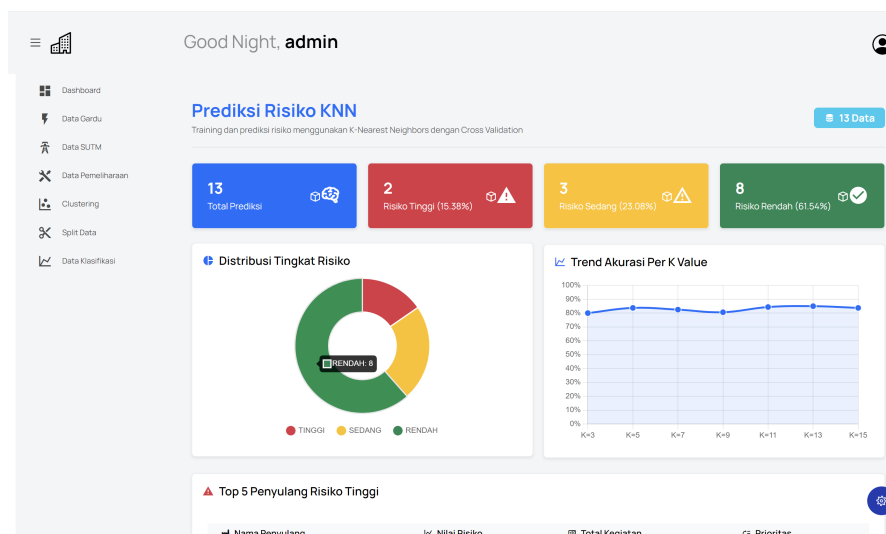
Halaman ini menampilkan rasio pembagian dan jumlah data pada masing-masing subset seperti ditunjukkan pada Gambar 10.

## 7. Tampilan Klasifikasi

Halaman klasifikasi merupakan komponen utama yang menjalankan dan menampilkan hasil prediksi risiko gangguan jaringan listrik menggunakan algoritma *K-Nearest Neighbor* (K-NN). Admin dapat menentukan nilai **K** sebelum proses dimulai. Setelah klasifikasi dilakukan, sistem menampilkan hasil dalam bentuk tabel berisi nama penyulang, kategori risiko (tinggi, sedang, rendah), nilai risiko numerik, dan tingkat kepercayaan model. Selain hasil tabel, sistem juga menampilkan *Confusion matrix* beserta empat metrik evaluasi utama:

1. *Accuracy* – rasio prediksi benar terhadap total data uji,
2. *Precision* – ketepatan model dalam meminimalkan prediksi positif palsu,
3. *Recall* – kemampuan model mendeteksi semua kasus aktual, dan
4. *F1-Score* – rata-rata harmonis antara *precision* dan *recall*.

Hasil ini membantu menilai apakah model telah mencapai performa optimal atau masih memerlukan penyesuaian parameter **K**. Tampilan klasifikasi disajikan pada Gambar 11.



Gambar 11. Hasil Tampilan Data Klasifikasi

## 9. Tampilan Laporan Hasil Klasifikasi

Halaman laporan hasil klasifikasi merupakan antarmuka akhir yang menyajikan ringkasan keseluruhan hasil analisis. Bagian atas halaman menampilkan statistik ringkas seperti jumlah total klasifikasi, distribusi kategori risiko, rata-rata nilai risiko, serta jumlah kegiatan pemeliharaan. Bagian utama berisi tabel hasil klasifikasi tiap penyulang lengkap dengan nilai risiko, kategori kelas, dan waktu klasifikasi. Sistem juga menerapkan warna latar berbeda untuk tiap tingkat risiko agar memudahkan identifikasi visual. Seluruh hasil

klasifikasi dapat diekspor dalam format PDF atau Excel, untuk keperluan pelaporan teknis dan dokumentasi resmi. Tampilan laporan hasil klasifikasi ditunjukkan pada Gambar 12.

**LAPORAN HASIL PREDIKSI RISIKO KNN**

Tanggal Cetak: Rabu, 27 Agustus 2025  
Waktu Cetak: 23.45.14  
Format: Landscape Detail

---

**RINGKASAN STATISTIK**

Total Data Prediksi: 13  
Rata-rata Risk Score: 13.29  
Rata-rata Total Kegiatan: 63

Distribusi Risiko:  
• Tinggi: 2 data (15.38%)  
• Sedang: 3 data (23.08%)  
• Rendah: 8 data (61.54%)

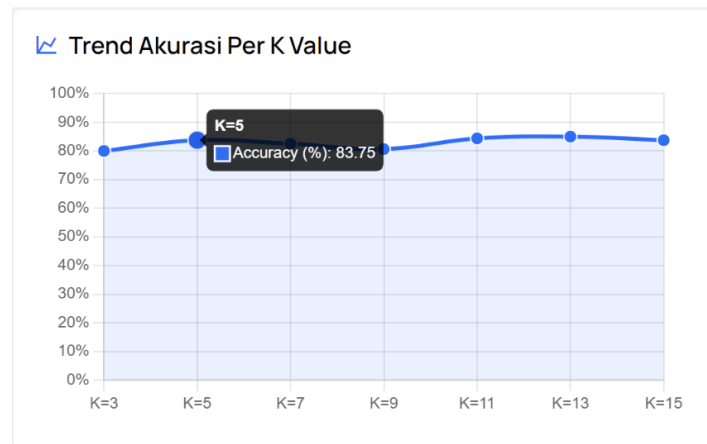
---

**DATA PREDIKSI DETAIL**

No	Nama Penyulang	Tingkat Risiko	Nilai Risiko	Total Kegiatan	K Value	Saran Perbaikan	Tanggal Prediksi
1	LK 01 / COT GIREK	SEDANG	23.98	163	K=5	Terapkan prioritas medium dengan pemeriksaan komponen; Lakukan penyeimbang beban gardu untuk distribusi stabil; Evaluasi penyebab tingginya aktivitas pemeliharaan	27/08/2025 23.08
2	LK 02 / INC PL3	RENDAH	1.38	5	K=5	Terapkan preventive maintenance dengan inspeksi berkala; Pertahankan jadwal pemeliharaan rutin	27/08/2025 23.08
3	LK 03 / BUKIT HAGU	TINGGI	23.14	126	K=5	Terapkan corrective & predictive maintenance fokus perbaikan darurat; Gunakan alat monitoring real-time seperti sensor suhu; Pertimbangkan relisasi atau pembangunan gardu baru; Segera lakukan shutdown maintenance untuk pencegahan gangguan massal	27/08/2025 23.08
4	LK 04 / LHOKSUKON	SEDANG	13.85	29	K=5	Terapkan prioritas medium dengan pemeriksaan komponen; Lakukan penyeimbang beban gardu untuk distribusi stabil; Evaluasi penyebab tingginya aktivitas pemeliharaan	27/08/2025 23.08
5	LK 05 / MATANG KULI	TINGGI	26.64	164	K=5	Terapkan corrective & predictive maintenance fokus perbaikan darurat; Gunakan alat monitoring real-time seperti sensor suhu; Pertimbangkan relisasi atau pembangunan gardu baru; Segera lakukan shutdown maintenance untuk pencegahan gangguan massal	27/08/2025 23.08
6	LK 05 / MATANG KULI	RENDAH	18.63	84	K=5	Terapkan preventive maintenance dengan inspeksi berkala; Pertahankan jadwal pemeliharaan rutin	27/08/2025 23.08

Gambar 12. Laporan Hasil Klasifikasi

### 3.2 Analisis hasil

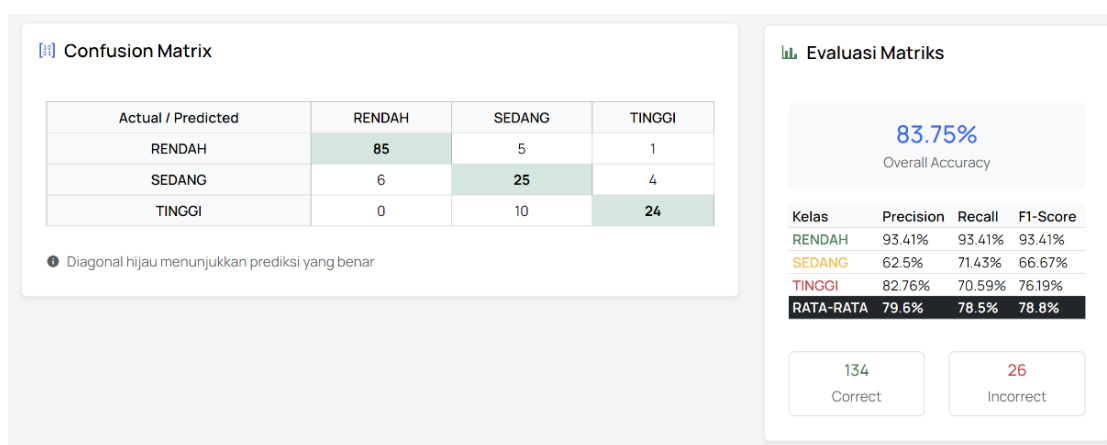


Gambar 13. Hasil Pengujian *Cross-validation*

Model K-NN dievaluasi menggunakan pendekatan *confusion matrix*, yang menggambarkan jumlah prediksi benar dan salah untuk masing-masing kategori risiko (rendah, sedang, dan tinggi). Proses evaluasi ini dilaksanakan pada data uji hasil pembagian dataset dengan rasio 80:20 antara data latih dan data uji. Sistem mendefinisikan tiga kelas risiko, yaitu rendah, sedang, dan tinggi, kemudian menghitung jumlah prediksi benar (true positive) dan salah (false positive, false negative) pada masing-masing kelas. Nilai-nilai ini selanjutnya digunakan untuk menghitung metrik evaluasi *accuracy*, *precision*, *recall*, dan *F1-score* guna menilai kinerja model secara menyeluruh. Sebelum evaluasi utama dilakukan, sistem melaksanakan proses *cross-validation* untuk menentukan parameter nilai K optimal pada algoritma K-NN. Proses ini bertujuan

mencari keseimbangan antara kompleksitas model dan kemampuan generalisasi. Hasil uji *cross-validation* terhadap berbagai nilai  $K$  ditampilkan pada Gambar 13.

Model K-NN dievaluasi menggunakan pendekatan *confusion matrix*, yang menggambarkan jumlah prediksi benar dan salah untuk masing-masing kategori risiko (rendah, sedang, dan tinggi). Proses evaluasi ini dilaksanakan pada data Berdasarkan hasil pengujian, terlihat bahwa nilai akurasi model bervariasi pada setiap percobaan  $K$ . Akurasi tertinggi diperoleh pada  $K = 5$ , dengan capaian 83,75%, sehingga nilai ini dipilih sebagai parameter terbaik. Nilai  $K = 5$  dianggap paling optimal karena mampu menjaga keseimbangan antara bias dan variansi, serta menghasilkan stabilitas performa yang konsisten di berbagai skenario pengujian. Setelah parameter  $K$  ditetapkan, sistem melanjutkan proses evaluasi utama menggunakan *confusion matrix* untuk menilai detail performa klasifikasi per kelas risiko. Hasil pengujian ditunjukkan pada Gambar 14.



Gambar 14. Hasil Pengujian *Confusion matrix*

Berdasarkan hasil pengujian, diperoleh akurasi keseluruhan sebesar 83,75%, dengan 134 prediksi benar dan 26 prediksi salah dari total data uji. Model menunjukkan performa yang sangat baik pada kelas risiko rendah, dengan:

- Precision*: 93,41%
- Recall*: 93,41%
- F1-score*: 93,41%

Namun, pada kelas risiko sedang, performa model relatif lebih rendah dengan:

- Precision*: 62,50%
- Recall*: 71,43%
- F1-score*: 66,67%

Hal ini menunjukkan bahwa terdapat tumpang tindih karakteristik data antara kelas sedang dengan kelas lain yang menyebabkan beberapa kesalahan klasifikasi. Sementara itu, pada kelas risiko tinggi, model menunjukkan performa yang cukup baik, dengan:

- Precision*: 82,76%

- b. *Recall*: 70,59%
- c. *F1-score*: 76,19%

Secara rata-rata, model mencatat nilai metrik sebagai berikut:

- a. *Precision* rata-rata: 79,6%
- b. *Recall* rata-rata: 78,5%
- c. *F1-score* rata-rata: 78,8%

Hasil ini menunjukkan bahwa model K-NN dengan  $K = 5$  memiliki kinerja klasifikasi yang baik dan stabil di semua kategori risiko. Meski masih terdapat kelemahan pada pembedaan kelas sedang, model mampu memberikan prediksi yang cukup akurat dan dapat diandalkan untuk keperluan operasional PLN dalam memprioritaskan pemeliharaan jaringan listrik berdasarkan tingkat risiko.

#### 4. Kesimpulan

Berdasarkan hasil implementasi metode *K-Nearest Neighbor* (K-NN), sistem yang dikembangkan berhasil memprediksi tingkat risiko gangguan jaringan listrik dengan menggunakan empat aspek utama, yaitu inspeksi, pemeliharaan rutin, perbaikan kritis, dan kondisi infrastruktur. Data yang digunakan terdiri dari 791 entri hasil dokumentasi pemeliharaan pada 13 penyulang aktif di wilayah kerja PLN ULP Lhoksukon. Hasil klasifikasi menunjukkan bahwa sistem mampu mengelompokkan penyulang menjadi 8 risiko rendah, 3 risiko sedang, dan 2 risiko tinggi. Proses klasifikasi dilakukan dengan pembagian data latih dan data uji secara stratifikasi serta pemilihan parameter  $K=5$  melalui *cross-validation*, sehingga diperoleh akurasi sebesar 83,75%. Selain itu, hasil pengujian black box menunjukkan bahwa seluruh fungsi sistem berjalan sesuai kebutuhan pengguna, sedangkan pengujian white box memastikan bahwa logika program dan alur algoritma telah sesuai dengan rancangan yang diimplementasikan.

Berdasarkan *confusion matrix*, performa terbaik diperoleh pada kelas risiko rendah dengan *precision*, *recall*, dan *F1-score* masing-masing sebesar 93,41%, sementara kelas tinggi mencapai *precision* 82,76% dan *F1-score* 76,19%, serta kelas sedang memperoleh *precision* 62,5% dan *F1-score* 66,67%. Secara keseluruhan, model menghasilkan rata-rata *precision* 79,6%, *recall* 78,5%, dan *F1-score* 78,8%. Hasil ini menunjukkan bahwa metode K-NN efektif dalam memprediksi risiko gangguan jaringan listrik berbasis data historis pemeliharaan, dengan performa sistem yang stabil dan dapat diandalkan untuk mendukung pengambilan keputusan pemeliharaan preventif.

#### Ucapan Terimakasih

Penulis menyampaikan terima kasih kepada PT PLN (Persero) Unit Layanan Pelanggan (ULP) Lhoksukon yang telah memberikan izin, dukungan data, serta informasi teknis terkait kegiatan pemeliharaan jaringan distribusi listrik yang digunakan dalam penelitian ini. Ucapan terima kasih juga disampaikan kepada Politeknik Negeri Lhokseumawe khususnya Jurusan Teknologi Informasi dan Komputer, atas dukungan fasilitas, bimbingan akademik, dan sarana penelitian yang memungkinkan terlaksananya implementasi dan pengujian sistem ini.

Tidak lupa, apresiasi diberikan kepada seluruh rekan dan pihak yang telah berkontribusi dalam proses pengumpulan data, analisis, serta pengembangan sistem klasifikasi risiko berbasis *K-Nearest Neighbor* (K-NN) ini, sehingga penelitian dapat terselesaikan dengan baik dan memberikan manfaat bagi pengembangan sistem prediksi gangguan jaringan listrik di masa mendatang.

## Referensi

- [1] K. Octavianus Bachri, E. Jonathan, and A. D. Soewono, "Analisis Gangguan Pada Sistem Gardu Listrik Berbasis Gas Insulated Switchgear 150 Kv Dengan Metode High Voltage Test," *Cylind. J. Ilm. Tek. Mesin*, vol. 9, pp. 2252–925X, 2023.
- [2] T. Andriani, M. Hidayatullah, S. Esabella, and P. Studi Teknik Elektro, "Pemeliharaan Jaringan Distribusi di PT. PLN ULP 2 Mawasangka," *Hexagon*, vol. 2, no. 1, 2021.
- [3] S. Amalia and E. Saputra, "Pemeliharaan Jaringan Saluran Udara Tegangan Menengah (SUTM) 20 kV Feeder Mata Air," *J. Tek. Elektro Inst. Teknol. Padang*, vol. 9, no. 2, 2020, doi: 10.21063/JTE.2020.3133911.
- [4] S. Arena, G. Manca, S. Murru, P. F. Orrù, R. Perna, and D. Reforgiato Recupero, "Data Science Application for Failure Data Management and Failure Prediction in the Oil and Gas Industry: A Case Study," *Appl. Sci.*, vol. 12, no. 20, Oct. 2022, doi: 10.3390/app122010617.
- [5] S. Rezig, Z. Achour, and N. Rezg, "RETRACTED: Using Data Mining Methods for Predicting Sequential Maintenance Activities," *Appl. Sci.*, vol.8, no. 11, Nov. 2018, doi: 10.3390/app8112184.
- [6] M. Kafil, "Penerapan Metode *K-Nearest Neighbor* Untuk Prediksi Penjualan Berbasis Web Pada Boutiq Dealove Bondowoso," 2019.
- [7] L. Abd, R. Hakim, A. A. Rizal, and D. Ratnasari, "Aplikasi Prediksi Kelulusan Mahasiswa Berbasis *K-Nearest Neighbor* (K-NN)," 2019.
- [8] Siska Ayu Widiiana, Iqbal Firdaus, Edwin Tenda, and Eliasta Ketaren, "Sistem Informasi Prediksi Penjualan Produk Thrift Di Toko Manado Menggunakan Algoritma *K-Nearest Neighbor* (K-NN)," *J. TIMES*, vol. 12, no. 2, pp. 52–57, 2023, doi: 10.51351/jim.12.2.2023708.
- [9] Roziana, E. Ramadhani, and S. Anwar, "Analisis Gangguan Listrik Di Pltu Nagan Raya Menggunakan Diagram Kendali Exponentially Weighted Moving Average Dan Decision On Belief," *INTERVAL J. Ilm. Mat.*, vol. 2, no. 2, pp. 104–117, 2022.
- [10] W. Pratama et al., "Prediksi Pemeliharaan Transformator Distribusi Berbasis Artificial Neural Network," *SNESTIK*, p. 112, 2023, doi: 10.31284/p.snestik.2023.4004.
- [11] M. Yunus and N. K. A. Pratiwi, "Prediksi Status Gizi Balita Dengan Algoritma *K-Nearest Neighbor* (K-NN) di Puskesmas Cakranegara," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 4, no. 4, pp. 221–231, Feb. 2023, doi: 10.35746/jtim.v4i4.328.
- [12] N. Kwintarini Suparman, B. Arif Dermawan, and T. Nur Padilah, "Prediksi Barang Keluar TB. Wijaya Bangunan Menggunakan Algoritma K-NN Regression dengan RStudio," *MEI*, 2021.
- [13] Rachmadhany Iman, Basuki Rahmat, and Achmad Junaidi, "Implementasi Algoritma *K-Means* dan Knearest Neighbors (K-NN) Untuk Identifikasi Penyakit Tuberkulosis Pada Paru-Paru," *Repeater Publ. Tek. Inform. dan Jar.*, vol. 2, no. 3, pp. 12–25, 2024, doi: 10.62951/repeater.v2i3.77.
- [14] G. Dimas Prenata, "Klasifikasi Keandalan Sistem Distribusi Tenaga Listrik Di Pt. Pln (Persero) Up3 Surabaya Selatan Menggunakan Metode *K-Nearest Neighbor* (K-NN)," *Jl. Semolowaru No*, vol. 11, no. 3, p. 60118, 2023, doi:10.23960/jitet.v11i3%20s1.3397.
- [15] M. A. Harriz and H. Setiyowati, "Komparasi Algoritma Decision Tree Dan K-NN Dalam Mengklasifikasi Daerah Berdasarkan Produksi Listrik," *JIKO (Jurnal Inform. dan Komputer)*, vol. 7, no. 2, p. 167, 2023, doi: 10.26798/jiko.v7i2.787.
- [16] Retno Wahyusari, "Perbandingan *K-Nearest Neighbor* (K-NN) Dan Support Vector Regression (Svr) Untuk Prediksi Konsumsi Energi Listrik," *Pros. Snast*, no. November, pp. G44-50, 2024, doi: 10.34151/prosidingsnast.v1i1.5034.
- [17] P. Hidayat, R. Kurniawan, Y. A. Wijaya, and T. Suprapti, "Optimasi Algoritma K-Nearest (K-NN) Neighbors Pada Prediksi Risiko Penyakit Kardiovaskular," *J. Inform. dan Tek. Elektro Terap.*, vol. 13, no. 1, 2025, doi: 10.23960/jitet.v13i1.5864.
- [18] R. Joto and M. Urfan Barran Rusyda Marzuq, "Analisis Perencanaan Sistem Jaringan Distribusi Listrik dan Perkembangan Beban Pada Perumahan The Grand Kenjeran Surabaya," *ELPOSYS J. Sist. Kelistrikan*, vol. 09, no. 3, 2022.
- [19] R. J. Alfirdausy and S. Bahri, "Implementasi Algoritma *K-Nearest Neighbor* untuk Klasifikasi Diagnosis Penyakit Alzheimer," *Techno.Com*, vol. 22, no.3, pp. 635–642, 2023, doi: 10.33633/tc.v22i3.8393.
- [20] A. Putri, C. Syaficha Hardiana, E. Novfuja, F. Try Puspa Siregar, Y. Fatma, and R. Wahyuni, "Komparasi Algoritma K-NN, Naive Bayes dan SVM untuk Prediksi Kelulusan Mahasiswa Tingkat Akhir," *MALCOM Indones. J. Mach. Learn. Comput. Sci. J. Homepage*, vol. 3, no. 1, pp. 20–26, 2023, [Online]. Available: <https://journal.irpi.or.id/index.php/malcom/article/view/610>