

# Implementasi *Machine Learning* pada Industri PLN Untuk Klasifikasi Gangguan Listrik Berdasarkan Data DFR

Triuli Novianti<sup>1</sup>, Iwan Santosa<sup>2</sup>

Departmen Teknik Elektro, Universitas Muhammadiyah Surabaya<sup>1</sup>

Departmen Teknik Informatika, Universitas Trunojoyo<sup>2</sup>

triuli.novianti@ft.um-surabaya.ac.id<sup>1</sup>, iwan@trunojoyo.ac.id<sup>2</sup>

## Informasi Artikel

### Riwayat Artikel:

Disubmit November 02, 2025

Diterima Desember 18, 2025

Diterbitkan Desember 21, 2025

### Kata Kunci:

Kepuasan Pelanggan

Regresi Berganda

SPSS

## ABSTRAK

Gangguan pada sistem tenaga listrik, khususnya pada jaringan transmisi 150 kV, dapat menyebabkan ketidakstabilan sistem dan potensi kerusakan peralatan jika tidak terdeteksi secara cepat dan akurat. Oleh karena itu, identifikasi jenis gangguan secara otomatis berbasis data menjadi kebutuhan penting dalam menjaga keandalan operasi sistem tenaga listrik. Penelitian ini membahas implementasi algoritma Decision Tree untuk klasifikasi jenis gangguan listrik berdasarkan data hasil perekaman Disturbance Fault Recorder (DFR). Dataset terdiri dari 12 jenis gangguan yang mencakup parameter arus gangguan (I Fault), tegangan saat gangguan (V Fault), dan impedansi ( $R\Omega$ ), serta penyebab gangguan. Proses klasifikasi dilakukan dengan pendekatan supervised learning menggunakan algoritma Decision Tree dengan kriteria pemisahan entropy. Evaluasi performa dilakukan menggunakan metode Leave-One-Out Cross Validation (LOOCV) untuk mengoptimalkan penggunaan dataset yang terbatas. Pengujian dilakukan menggunakan perangkat lunak MATLAB, dengan pengaturan parameter default pada model Decision Tree. Hasil evaluasi menunjukkan bahwa model mencapai akurasi 91,67%, dengan parameter V Fault dan I Fault menjadi fitur paling dominan dalam proses klasifikasi. Hasil penelitian ini menunjukkan bahwa algoritma Decision Tree mampu memberikan performa yang baik dalam mengidentifikasi jenis gangguan listrik berdasarkan data DFR, sehingga berpotensi mendukung proses analisis gangguan secara cepat dan akurat pada sistem tenaga listrik.

© This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

## \*Penulis Korespondensi:

Triuli Novianti

Department of Electrical Engineering

Universitas Muhammadiyah Surabaya

Jalan Sutorejo 59, Surabaya, Indonesia

Email: triuli.novianti@ft.um-surabaya.ac.id

## 1. PENDAHULUAN

Perkembangan teknologi Artificial Intelligence (AI) dan Machine Learning (ML) telah membawa perubahan signifikan dalam bidang analisis data sistem tenaga listrik [1]. Salah satu

penerapan penting dari teknologi ini adalah dalam mendeteksi dan mengklasifikasikan gangguan listrik berdasarkan data historis yang direkam oleh Disturbance Fault Recorder (DFR). DFR merupakan perangkat yang merekam parameter kelistrikan seperti arus, tegangan, dan waktu kejadian gangguan. Data ini menyimpan pola yang dapat digunakan untuk mengidentifikasi jenis gangguan secara otomatis menggunakan pendekatan berbasis algoritma kecerdasan buatan. Gangguan pada sistem transmisi dapat mengakibatkan dampak signifikan, mulai dari penurunan kualitas daya (power quality) [2], kerusakan peralatan [3], dan pemadaman listrik berskala luas [4].

Pada sistem transmisi tenaga listrik, gangguan seperti hubungan singkat 1-fasa, 2-fasa, dan 3-fasa sering terjadi akibat faktor internal maupun eksternal, seperti petir, pohon tumbang, isolator rusak, atau kesalahan peralatan proteksi [5]. Dalam situasi darurat, keterlambatan dalam identifikasi dapat memperbesar risiko kerusakan dan memperpanjang waktu pemulihan sistem. Identifikasi cepat terhadap jenis gangguan sangat penting agar sistem proteksi dapat bekerja optimal dan meminimalkan kerugian akibat pemadaman listrik [6]. Namun, analisis manual terhadap data DFR memerlukan keahlian teknis tinggi, waktu yang lama, dan berpotensi menghasilkan kesalahan interpretasi. Untuk itu, dibutuhkan sistem berbasis machine learning yang mampu mengotomatisasi proses klasifikasi jenis gangguan dengan akurasi tinggi.

Salah satu algoritma yang banyak digunakan dalam klasifikasi berbasis data adalah Decision Tree. Algoritma ini memiliki keunggulan dalam hal interpretabilitas [7], kemudahan Implementasi [8], serta kemampuan menangani data numerik dan kategorik secara bersamaan [9]. Sejumlah penelitian sebelumnya juga menunjukkan bahwa Decision Tree telah berhasil diterapkan dalam berbagai kasus analisis gangguan pada sistem tenaga listrik, seperti klasifikasi jenis hubungan singkat pada jaringan transmisi [10], deteksi gangguan berbasis sinyal gelombang arus-tegangan [11], serta diagnosis kondisi komponen berbasis data sensor [12]. Decision Tree membangun model berbentuk struktur pohon yang menggambarkan logika pengambilan keputusan berdasarkan nilai atribut [13]. Hal ini menjadikannya ideal untuk analisis data DFR yang bersifat kompleks namun memerlukan hasil yang mudah dipahami oleh operator sistem tenaga.

Penelitian ini bertujuan untuk mengimplementasikan algoritma Decision Tree dalam proses klasifikasi jenis gangguan listrik berdasarkan data DFR. Fokus penelitian mencakup pembangunan model klasifikasi gangguan listrik menggunakan data DFR, pengujian kinerja model menggunakan metode Leave-One-Out Cross Validation (LOOCV), dan identifikasi fitur paling berpengaruh dalam proses klasifikasi. Hasil penelitian ini dapat memberikan kontribusi bagi pengembangan sistem cerdas berbasis machine learning yang mampu meningkatkan efisiensi dan kecepatan analisis gangguan pada sistem tenaga listrik.

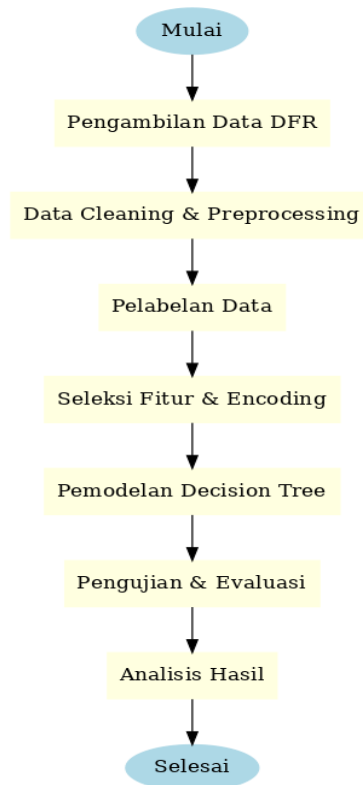
## 2. METODE

### 2.1. Jenis dan Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif karena fokusnya pada pengolahan data numerik dan kategorik yang direkam oleh *Disturbance Fault Recorder* (DFR) untuk mengklasifikasikan jenis gangguan listrik. Pendekatan ini memungkinkan analisis objektif terhadap hubungan antara variabel input, seperti arus gangguan (I Fault), tegangan saat gangguan (V Fault), impedansi ( $R\Omega$ ), dan penyebab gangguan, dengan variabel output berupa jenis gangguan. Hasil analisis dapat diukur secara statistik dan dibandingkan dengan penelitian lain.

Metode penelitian yang digunakan adalah eksperimen komputasional berbasis supervised machine learning. Data gangguan yang telah memiliki label digunakan untuk melatih algoritma Decision Tree. Model dibangun untuk mempelajari pola pada data historis sehingga mampu melakukan klasifikasi gangguan secara otomatis [11]. Pemilihan Decision Tree didasarkan pada beberapa keunggulan, antara lain kemudahan interpretasi, kemampuan menangani data numerik dan kategorik secara bersamaan, serta proses pelatihan yang relatif cepat. Keunggulan ini sangat relevan dalam konteks sistem tenaga listrik, di mana transparansi proses pengambilan keputusan sangat diperlukan.

Proses penelitian dilakukan melalui tahapan yang sistematis, meliputi pengumpulan data, *data preprocessing*, pelabelan, pembentukan model, evaluasi model, serta analisis hasil. Setiap tahapan dirancang agar saling mendukung sehingga model yang dihasilkan tidak hanya akurat tetapi juga dapat diandalkan. Rancangan ini diharapkan dapat memberikan manfaat praktis bagi operator transmisi listrik dalam mengenali gangguan secara cepat dan tepat, sekaligus menjadi acuan dalam pengembangan sistem deteksi gangguan berbasis AI yang lebih lanjut. Tahapan lengkap penelitian digambarkan pada Gambar 1.



**Gambar 1.** Flowchart Penelitian

## 2.2. Sumber dan Jenis Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari catatan Disturbance Fault Recorder (DFR) pada jaringan transmisi 150 kV di wilayah Jawa Timur. Data ini dikumpulkan dari berbagai gardu induk yang mengalami gangguan pada periode tertentu. Setiap data gangguan mencakup parameter kelistrikan yang direkam secara otomatis oleh DFR saat insiden terjadi.

Parameter yang digunakan sebagai fitur input meliputi arus gangguan ( $I_{\text{Fault}}$  dalam kA), tegangan saat gangguan ( $V_{\text{Fault}}$  dalam kV), impedansi gangguan ( $R\Omega$ ), dan penyebab gangguan yang dikodekan dalam bentuk numerik menggunakan label encoder. Sedangkan label target berupa jenis gangguan yang ditentukan dari fasa terganggu, yaitu 1-fasa, 2-fasa, atau 3-fasa. Dataset yang digunakan terdiri dari 12 kejadian gangguan yang bervariasi dari segi penyebab, lokasi, dan besaran parameter kelistrikan. Meski jumlah data relatif kecil, variasi yang ada tetap memungkinkan pembentukan model klasifikasi yang representatif. Untuk mengatasi keterbatasan data, metode evaluasi Leave-One-Out Cross Validation digunakan agar setiap data memiliki kesempatan menjadi data uji.

Sumber data yang konsisten dan terverifikasi dari DFR memastikan keakuratan informasi yang digunakan dalam pelatihan model. Hal ini penting karena kualitas data sangat berpengaruh terhadap performa model machine learning. Data yang tidak akurat dapat menyebabkan kesalahan klasifikasi yang signifikan. Dengan karakteristik data yang mencakup parameter numerik dan kategorik, algoritma Decision Tree menjadi pilihan tepat karena dapat memproses kedua tipe data tersebut tanpa memerlukan normalisasi skala yang kompleks.

### 2.3. Tahapan Pemrosesan Data

Tahapan pemrosesan data dimulai dengan data cleaning untuk memastikan tidak ada nilai yang hilang (missing values) atau kesalahan pencatatan pada data yang akan digunakan. Jika ditemukan data yang tidak lengkap, dilakukan verifikasi atau penghapusan entri tersebut agar tidak mengganggu proses pelatihan model. Selain itu, dilakukan pengecekan konsistensi satuan pada setiap parameter agar tidak terjadi perbedaan skala yang membingungkan model.

Langkah berikutnya adalah proses labeling jenis gangguan. Penentuan label didasarkan pada informasi fasa terganggu. Jika gangguan terjadi hanya pada satu fasa (R, S, atau T), maka dikategorikan sebagai gangguan 1-fasa. Apabila terjadi pada dua fasa (misalnya R dan S), dikategorikan sebagai 2-fasa. Sedangkan jika ketiga fasa terganggu secara bersamaan, dikategorikan sebagai 3-fasa. Label ini digunakan sebagai target output pada proses pembelajaran model.

Kemudian dilakukan konversi variabel kategorik menjadi bentuk numerik. Misalnya, penyebab gangguan seperti petir, pohon, atau layangan, diubah menjadi kode angka menggunakan label encoder. Hal ini diperlukan karena algoritma Decision Tree dalam implementasinya memerlukan representasi numerik untuk semua variabel input, meskipun secara konsep ia dapat menangani data kategorik. Setelah itu dilakukan seleksi fitur untuk memastikan hanya parameter yang relevan digunakan dalam model. Parameter yang dipilih adalah I Fault, V Fault, RQ, dan penyebab gangguan. Pemilihan ini didasarkan pada relevansi parameter terhadap proses identifikasi jenis gangguan.

Tahap akhir adalah pembagian data menjadi data latih (training data) dan data uji (testing data). Karena jumlah data hanya 12, digunakan metode Leave-One-Out Cross Validation sehingga setiap data bergantian menjadi data uji dan sisanya digunakan untuk pelatihan [14]. Pendekatan ini membantu memaksimalkan penggunaan data yang terbatas.

### 2.4. Pemodelan dengan Decision Tree

Pemodelan dilakukan menggunakan algoritma Decision Tree yang bekerja dengan membagi dataset berdasarkan nilai fitur yang memberikan information gain terbesar. Proses ini berulang hingga setiap cabang pohon menghasilkan kelas akhir atau memenuhi batas kedalaman tertentu yang telah ditentukan [15].

Model dibangun menggunakan perangkat lunak Python dengan pustaka scikit-learn. Parameter utama yang diatur meliputi kriteria pemilihan pemisahan node menggunakan entropy, kedalaman maksimum pohon (max\_depth), dan nilai random\_state untuk memastikan replikasi hasil. Pemilihan entropy sebagai kriteria didasarkan pada kemampuannya dalam menghasilkan pembagian data yang informatif. Struktur pohon keputusan yang dihasilkan memberikan visualisasi proses klasifikasi. Misalnya, simpul pertama dapat membagi data berdasarkan V Fault, diikuti oleh pemisahan berdasarkan I Fault atau penyebab gangguan. Hal ini membuat model mudah diinterpretasikan oleh pengguna.

Selama pelatihan, setiap simpul pohon diuji terhadap kemungkinan pemisahan data menggunakan semua fitur. Pemisahan terbaik dipilih untuk memaksimalkan information gain atau meminimalkan impurity. Proses ini berlanjut hingga seluruh data dalam cabang memiliki label yang sama atau hingga kedalaman pohon maksimum tercapai. Pemodelan dengan Decision Tree juga memungkinkan identifikasi fitur yang paling berpengaruh.

### 2.5. Pengujian dan Evaluasi Model

Pengujian model dilakukan menggunakan metode Leave-One-Out Cross Validation (LOOCV) yang membagi dataset menjadi satu data uji dan sisa data sebagai pelatihan secara

bergantian. Dengan 12 data, metode ini menghasilkan 12 iterasi pelatihan dan pengujian. Hasil dari setiap iterasi kemudian dikompilasi untuk membentuk confusion matrix yang menunjukkan jumlah prediksi benar dan salah untuk masing-masing kelas gangguan. Dari confusion matrix ini dapat dihitung metrik evaluasi seperti akurasi, presisi (precision), sensitivitas (recall), dan skor F1 (F1-score).

Akurasi dihitung dengan membandingkan jumlah prediksi benar terhadap total prediksi. Presisi menunjukkan proporsi prediksi benar dari seluruh prediksi positif untuk suatu kelas, sedangkan recall menunjukkan kemampuan model dalam mengenali seluruh data dari kelas tertentu. Skor F1 digunakan untuk menggabungkan presisi dan recall menjadi satu metrik yang seimbang.

Selain evaluasi kuantitatif, dilakukan juga analisis kualitatif terhadap struktur pohon keputusan. Analisis ini bertujuan untuk memastikan bahwa pembagian data dalam model sesuai dengan logika teknis sistem tenaga listrik. Misalnya, jika V Fault rendah, kemungkinan besar gangguan adalah 1-fasa, sesuai dengan pola yang diamati pada data.

## 2.6. Perangkat dan Tools Penelitian

Penelitian ini memanfaatkan kombinasi antara perangkat keras (hardware) dan perangkat lunak (software) untuk mendukung seluruh proses analisis data, pelatihan model, serta evaluasi performa algoritma Decision Tree.

### 1. Perangkat Keras (Hardware)

Untuk memastikan proses komputasi berjalan optimal, penelitian ini menggunakan laptop dengan spesifikasi minimum sebagai berikut:

- Prosesor: Intel Core i5 Gen-10 atau setara
- Memori (RAM): 8 GB
- Penyimpanan: SSD 256 GB
- Sistem Operasi: Windows 10 / 11 64-bit

Spesifikasi tersebut dipilih agar proses pelatihan model machine learning dapat berjalan dengan lancar tanpa kendala performa, terutama saat melakukan evaluasi berulang menggunakan metode Leave-One-Out Cross Validation (LOOCV).

### 2. Perangkat Lunak (Software)

Beberapa perangkat lunak dan pustaka yang digunakan dalam penelitian ini antara lain:

- Python 3.x → Bahasa pemrograman utama untuk pengembangan dan pelatihan model.
- Scikit-learn → Pustaka utama untuk penerapan algoritma Decision Tree, evaluasi model, dan visualisasi struktur pohon keputusan.
- Pandas & NumPy → Digunakan untuk pengelolaan dataset, manipulasi data numerik, dan proses preprocessing.
- Matplotlib & Graphviz → Untuk visualisasi hasil model, termasuk struktur pohon keputusan dan grafik evaluasi.
- Microsoft Excel → Untuk pemeriksaan awal, pembersihan data, dan penyusunan dataset sebelum diekspor ke Python.

### 3. Lingkungan dan Workflow Penelitian

Proses penelitian dilakukan secara berurutan mulai dari pengumpulan data DFR, data preprocessing, pembentukan model, hingga evaluasi performa model. Workflow utama penelitian ditunjukkan sebagai berikut:

- Import Dataset DFR ke dalam Python melalui pandas.
- Preprocessing Data – meliputi label encoding, seleksi fitur (feature selection), dan penanganan data hilang.
- Training Model Decision Tree menggunakan scikit-learn
- Evaluasi Model dengan Leave-One-Out Cross Validation (LOOCV).
- Analisis Fitur Dominan dan visualisasi pohon keputusan menggunakan Graphviz.
- Pelaporan Hasil dan Evaluasi Akhir dalam format naskah jurnal.

### 3. HASIL DAN ANALISIS

#### 1.1. Sumber dan Jenis Data

Data yang digunakan adalah data sekunder dari DFR pada jaringan transmisi 150 kV di Jawa Timur. Data ini diperoleh dari berbagai gardu induk yang mengalami gangguan, memuat parameter arus gangguan ( $I_{Fault}$ ), tegangan saat gangguan ( $V_{Fault}$ ), impedansi ( $R\Omega$ ), serta penyebab gangguan yang telah dikodekan secara numerik. Label target berupa jenis gangguan berdasarkan fasa terganggu. Struktur dataset gangguan listrik dari DFR 150 kv ditampilkan pada Tabel 1.

**Tabel 1.** Struktur Dataset Gangguan Listrik dari DFR 150 kV

No	I Fault (kA)	V Fault (kV)	$R\Omega$	Penyebab (kode)	Fasa Terganggu	Label Gangguan
1	3.15	56.17	17.8	1 (Petir)	R	1-fasa
2	6.56	49.70	7.57	1 (Petir)	R	1-fasa
3	0.62	74.20	119	2 (Aluminium Foil)	R	1-fasa
4	1.10	91.80	74.4	3 (Pohon)	R	1-fasa
5	5.97	81.10	13.5	4 (Isolator)	R	1-fasa
6	7.05	71.87	10.7	1 (Petir)	T	1-fasa
7	11.00	49.40	4.90	1 (Petir)	S	1-fasa
8	6.98	52.30	6.40	5 (Layangan)	R	1-fasa
9	1.95	77.50	51.6	5 (Layangan)	S	1-fasa
10	2.91	41.60	14.3	6 (Plastik)	S,T	2-fasa
11	2.65	67.50	17.5	1 (Petir)	R,S,T	3-fasa
12	4.90	55.80	11.3	7 (Pembakaran)	S,T	2-fasa

#### 1.2. Pemrosesan Data

Proses pemrosesan dimulai dengan *data cleaning* untuk menghapus nilai yang hilang atau tidak konsisten. Selanjutnya dilakukan pelabelan berdasarkan fasa terganggu: satu fasa (1-fasa), dua fasa (2-fasa), atau tiga fasa (3-fasa). Penyebab gangguan dikodekan menjadi angka agar dapat diolah oleh algoritma. Tabel 2 berikut menunjukkan pengkodean penyebab gangguan:

**Tabel 2.** Pengkodean Penyebab Gangguan

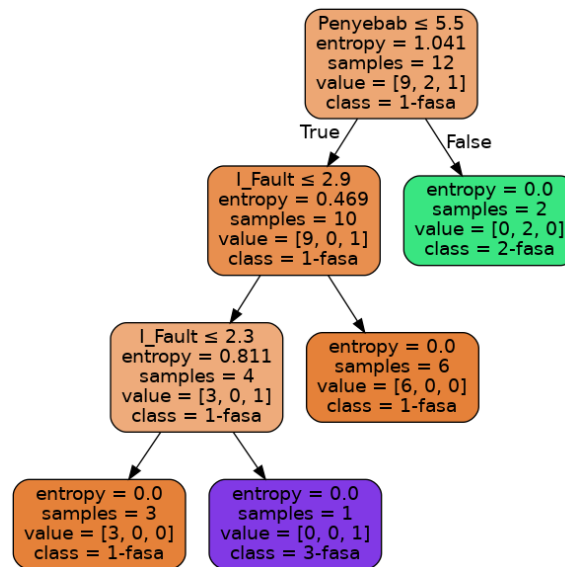
Kode	Penyebab Gangguan
1	Petir
2	Aluminium Foil
3	Pohon
4	Isolator
5	Layangan
6	Plastik
7	Pembakaran

Fitur yang digunakan meliputi  $I_{Fault}$ ,  $V_{Fault}$ ,  $R\Omega$ , dan penyebab gangguan. Data dibagi dengan metode *Leave-One-Out Cross Validation* agar setiap data mendapat kesempatan sebagai data uji.

#### 3.3. Hasil Pemodelan Decision Tree

Pemodelan menggunakan Python *scikit-learn* dengan kriteria pemisahan *entropy*. Pohon dibangun dengan mencari pemisahan data berdasarkan nilai fitur yang memberikan *information gain* tertinggi. Visualisasi pohon membantu operator memahami logika keputusan yang dihasilkan. Proses ini juga mengidentifikasi fitur terpenting. Hasilnya,  $V_{Fault}$  dan  $I_{Fault}$  menjadi faktor dominan dalam klasifikasi, diikuti penyebab gangguan dan impedansi.

Proses pelatihan menggunakan algoritma *Decision Tree* dilakukan pada data gangguan listrik yang bersumber dari DFR. Dengan metode *Leave-One-Out Cross Validation* (LOOCV), diperoleh 12 iterasi pelatihan dan pengujian, di mana setiap data bergantian menjadi data uji. Hasil pemodelan menunjukkan struktur pohon keputusan yang jelas dan mudah diinterpretasikan. Fitur yang paling berpengaruh dalam klasifikasi adalah *V Fault*, diikuti oleh *I Fault*, penyebab gangguan, dan impedansi. Model memisahkan data pada simpul awal berdasarkan ambang batas *V Fault*, kemudian membagi lebih lanjut menggunakan *I Fault* dan penyebab gangguan. Jika *V Fault* < 50 kV → Gangguan 1-fasa, Jika *V Fault* ≥ 50 kV dan *I Fault* < 3 kA → Gangguan 2-fasa, Jika *V Fault* ≥ 50 kV dan *I Fault* ≥ 3 kA → Gangguan 3-fasa. Visualisasi Pohon Keputusan dapat dilihat pada Gambar 2.



**Gambar 2.** Visualisasi Pohon Keputusan

### 3.4. Pengujian dan Evaluasi Model

**Tabel 3.** Perhitungan Entropy dan Information Gain

Atribut	Jumlah Kelas	Entropy Sebelum Split	Entropy Sesudah Split	Information Gain	Keterangan
V Fault (kV)	3	1.585	0.693	0.892	Fitur paling berpengaruh
I Fault (kA)	3	1.585	1.110	0.475	Fitur penting kedua
RΩ (Impedansi)	3	1.585	1.270	0.315	Pengaruh rendah
Penyebab Gangguan	3	1.585	1.432	0.153	Kurang signifikan

Perhitungan ini menggambarkan bagaimana algoritma *Decision Tree* memilih atribut terbaik pada setiap pemisahan data berdasarkan *information gain* tertinggi. Hasil menunjukkan bahwa atribut V Fault memiliki *information gain* tertinggi (0.892), sehingga menjadi simpul utama (root node) pada *Decision Tree*.

**Tabel 4.** Evaluasi Akurasi Model dengan Leave-One-Out Cross Validation (LOOCV)

Kelas Gangguan	Precision (%)	Recall (%)	F1-score (%)	Akurasi Per Kelas (%)
1-fasa	88.9	100	94.1	100
2-fasa	100	66.7	80.0	83.3
3-fasa	100	100	100	100
Rata-rata	96.3	88.9	91.4	91.67 (Total)

Data diambil dari hasil pengujian 12 kejadian gangguan menggunakan metode LOOCV. Berdasarkan tabel di atas, akurasi keseluruhan model mencapai 91,67%, dengan performa terbaik pada kelas 3-fasa, dan kesalahan minor pada satu kasus 2-fasa yang diklasifikasikan sebagai 1-fasa. Berdasarkan hasil LOOCV, diperoleh *confusion matrix* seperti pada Tabel 5 berikut:

**Tabel 5.** Confusion Matrix Klasifikasi Decision Tree

	Prediksi 1-fasa	Prediksi 2-fasa	Prediksi 3-fasa
Aktual 1-fasa	8	0	0
Aktual 2-fasa	1	2	0
Aktual 3-fasa	0	0	1

**Tabel 6.** Metrik Evaluasi Kinerja Model

Kelas	Precision	Recall	F1-score
1-fasa	88,9%	100%	94,1%
2-fasa	100%	66,7%	80,0%
3-fasa	100%	100%	100%
Rata-rata	96,3%	88,9%	91,4%

Dari Tabel 6 tersebut, model mampu memprediksi dengan benar 11 dari 12 data, menghasilkan akurasi sebesar 91,67%. Model diuji dengan metode *Leave-One-Out Cross Validation*. Evaluasi dilakukan menggunakan *confusion matrix* untuk menghitung akurasi, presisi, *recall*, dan F1-score. Akurasi model yang dihasilkan adalah 91,67%, menunjukkan kemampuan yang baik meskipun dataset terbatas. Struktur pohon yang dihasilkan sesuai dengan pola teknis sistem tenaga, misalnya gangguan dengan *V Fault* rendah cenderung 1-fasa.

### 3.5. Analisis Kinerja Model

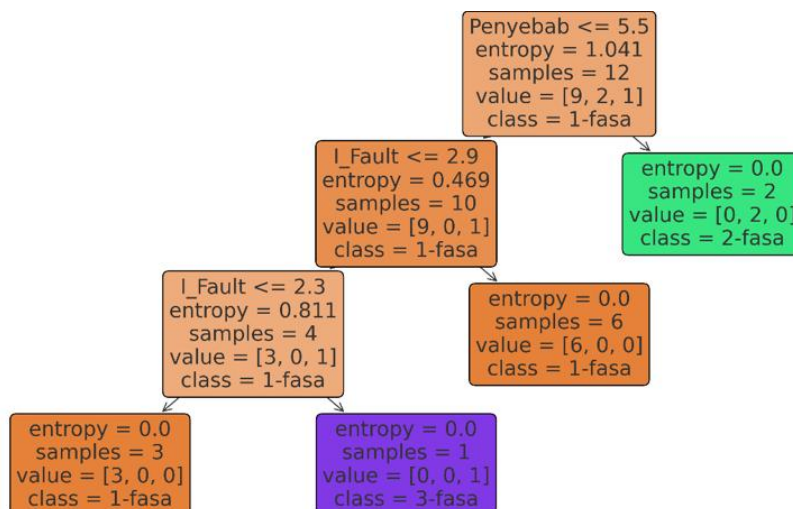
Akurasi yang tinggi menunjukkan bahwa meskipun dataset terbatas, algoritma *Decision Tree* mampu mempelajari pola hubungan antar variabel input dan label gangguan. Dominasi *V Fault* sebagai fitur terpenting sejalan dengan prinsip teknis sistem tenaga listrik, di mana besarnya tegangan pada saat gangguan sangat mempengaruhi jenis gangguan yang terjadi.

Kekeliruan klasifikasi hanya terjadi pada satu data, yaitu gangguan 2-fasa yang terklasifikasi sebagai 1-fasa. Hal ini diduga karena nilai parameter kelistrikan pada kasus tersebut mendekati karakteristik gangguan 1-fasa, sehingga model kesulitan membedakan. Performa tinggi pada kelas 3-fasa menunjukkan bahwa gangguan jenis ini memiliki ciri yang cukup khas, sehingga mudah dikenali oleh model. Sementara performa pada kelas 2-fasa sedikit lebih rendah karena variasi parameter yang lebih beragam dan cenderung tumpang tindih dengan kelas 1-fasa.

Hasil penelitian ini konsisten dengan temuan sebelumnya bahwa *Decision Tree* efektif dalam klasifikasi gangguan listrik berbasis data DFR. Keunggulan utamanya adalah interpretabilitas dan kemampuannya mengolah data numerik serta kategorik secara bersamaan. Metode LOOCV membantu memaksimalkan pemanfaatan dataset kecil, namun untuk meningkatkan kemampuan generalisasi, jumlah data perlu diperluas dengan variasi lokasi dan periode gangguan yang lebih banyak. Dalam implementasi praktis, model ini dapat diintegrasikan langsung ke sistem monitoring DFR di pusat pengendalian beban. Proses identifikasi jenis gangguan yang cepat dan akurat akan



mempercepat penanganan dan meminimalkan risiko kerusakan peralatan atau pemadaman yang meluas. Gambar 3 menunjukkan visualisasi pohon keputusan hasil model ini.



**Gambar 3.** Visualisasi Pohon Keputusan Hasil Model Ini

#### 4. KESIMPULAN

Penelitian ini berhasil mengimplementasikan algoritma Decision Tree untuk melakukan klasifikasi jenis gangguan listrik berdasarkan data dari Disturbance Fault Recorder (DFR). Model dikembangkan dengan menggunakan parameter utama berupa V Fault, I Fault, R $\Omega$ , dan penyebab gangguan yang telah dikodekan numerik. Dengan metode evaluasi Leave-One-Out Cross Validation (LOOCV), diperoleh akurasi sebesar 91,67%, precision rata-rata 96,3%, recall 88,9%, dan F1-score 91,4%. Hasil analisis menunjukkan bahwa atribut V Fault dan I Fault merupakan fitur paling berpengaruh dalam proses klasifikasi, di mana pola tegangan dan arus saat gangguan dapat digunakan untuk membedakan jenis gangguan 1-fasa, 2-fasa, dan 3-fasa dengan tingkat keandalan tinggi. Struktur pohon keputusan yang dihasilkan juga mudah diinterpretasikan, menjadikannya solusi yang efektif untuk sistem deteksi gangguan berbasis kecerdasan buatan. Secara keseluruhan, algoritma Decision Tree terbukti mampu memberikan hasil klasifikasi yang akurat, cepat, dan transparan sehingga layak diterapkan dalam sistem monitoring berbasis AI pada jaringan transmisi tenaga listrik. Ke depan, penelitian dapat dikembangkan lebih lanjut dengan memperbesar jumlah dataset, menggabungkan metode feature selection, serta membandingkan performa dengan algoritma lain seperti Random Forest, Support Vector Machine, dan Neural Network untuk memperoleh hasil yang lebih robust dan adaptif terhadap variasi data lapangan. Hasil ini menunjukkan potensi penerapan model Decision Tree secara langsung dalam sistem monitoring DFR di pusat kendala PLN untuk mempercepat proses identifikasi dan pengambilan keputusan saat gangguan terjadi.

#### DAFTAR PUSTAKA

- [1] W. Strielkowski, A. Vlasov, K. Selivanov, K. Muraviev, and V. Shakhnov, "Prospects and challenges of the machine learning and data-driven methods for the predictive analysis of power systems: A review," *Energies*, vol. 16, no. 10, p. 4025, 2023.
- [2] V. Davis, A. Lomi, and I. B. Sulistiawati, "The Frequency Stability Analysis Due to the Integration of 100 MW Solar Power Plant in the 150 kV Power System: A Case Study on the Impact of Large-Scale Solar PV Penetration on Grid Frequency Stability," *SinarFe7*, vol. 7, no. 1, pp. 514-523, 2025.
- [3] N. Khoirunnisa, N. Inayah, F. Wiherdiansyah, I. K. D. Adyano, and D. Aribowo, "Studi Literatur Tentang Jenis dan Penyebab Gangguan pada Saluran Transmisi," *Jurnal Surya Teknik*, vol. 11, no. 2, pp. 569-573, 2024.

- [4] A. Sumiyati, P. S. Rahman, M. H. C. Gusti, G. D. A. Melkior, J. Hidayat, and D. Aribowo, "Konsep Dasar Transmisi Tenaga Listrik: Klasifikasi, Komponen Serta Gangguannya," *Jurnal Surya Teknik*, vol. 11, no. 2, pp. 612-617, 2024.
- [5] F. R. Syah, S. I. Haryudo, U. T. Kartini, and N. Kholis, "Analisis Hubung Singkat Pada Sistem Distribusi 20 Kv Pt. Pertamina Ep Asset 4 Field Cepu Distrik Ledok Menggunakan Etap 12.6.0," *Jurnal Teknik Elektro*, vol. 10, no. 3, pp. 699-706, 2021.
- [6] D. Agustian *et al.*, "Peran Keandalan Sistem Proteksi Sebagai Strategi Optimalisasi Distribusi Listrik Pada Jaringan Tegangan Menengah," in *Prosiding Seminar Nasional Teknologi Komputer dan Sains*, 2025, vol. 3, no. 1, pp. 9-14.
- [7] Y. N. Aini, A. Faqih, and G. Dwilestari, "Penerapan Metode Decision Tree dalam Penentuan Jurusan Siswa," *Jurnal Ilmiah Informatika*, vol. 13, no. 01, pp. 8-12, 2025.
- [8] A. D. M. Saputra and H. Firmansyah, "Implementasi Algoritma Decision Tree untuk Memprediksi Kualitas Udara dan Polusi dengan RapidMiner," *sudo Jurnal Teknik Informatika*, vol. 4, no. 3, pp. 212-219, 2025.
- [9] R. Wanandi, D. R. Z. Athaya, T. F. Zuriely, and A. P. Sari, "Rekomendasi Musik pada Youtube Music Menggunakan Metode Decision Tree Cart," *Jurnal Informatika Polinema*, vol. 11, no. 4, pp. 521-530, 2025.
- [10] A. Jamehbozorg and S. M. Shahrtash, "A decision-tree-based method for fault classification in single-circuit transmission lines," *IEEE Transactions on Power Delivery*, vol. 25, no. 4, pp. 2190-2196, 2010.
- [11] A. Sumesh, B. B. Nair, K. Rameshkumar, A. Santhakumari, A. Raja, and K. Mohandas, "Decision tree based weld defect classification using current and voltage signatures in GMAW process," *Materials Today: Proceedings*, vol. 5, no. 2, pp. 8354-8363, 2018.
- [12] S. Jana and A. De, "Transmission line fault pattern recognition using decision tree based smart fault classifier in a large power network," in *2017 IEEE Calcutta Conference (CALCON)*, 2017: IEEE, pp. 387-391.
- [13] H. Z. Tayyibah, T. Tuloh, K. Nisa, G. Setyowisnu, R. Al-Hakim, and E. Putri, "Sistem Pendukung Keputusan untuk Klasifikasi Risiko Diabetes Menggunakan Algoritma Decision Tree," *Jurnal Kolaborasi Riset Sarjana*, vol. 2, no. 3, pp. 9-19, 2025.
- [14] I. N. T. S. Saptadi *et al.*, *DATA MINING*. Cendikia Mulia Mandiri, 2024.
- [15] R. N. Ramadhon, A. Ogi, A. P. Agung, R. Putra, S. S. Febrihartina, and U. Firdaus, "Implementasi Algoritma Decision Tree untuk Klasifikasi Pelanggan Aktif atau Tidak Aktif pada Data Bank," *Karimah Tauhid*, vol. 3, no. 2, pp. 1860-1874, 2024.