

Analisis Algoritma Prediksi Sistem Tenaga Listrik Menggunakan Machine Learning dan Implikasi Aksiologis

Sri Farida Utami^{1*}; Syaad Patmanthara¹

1. Department of Electrical Engineering and Informatics, Universitas Negeri Malang
Jl. Cakrawala No.5, Sumber Sari, Lowokwaru, Malang, Jawa Timur 65145 Indonesia

**Email: sri.farida.2505349@students.um.ac.id*

Received: 1 Desember 2025 | Accepted: 18 Desember 2025 | Published: 7 Januari 2026

ABSTRACT

Electrical fault detection is an important aspect in ensuring the stability and security of electric power systems. This study aims to classify the types of electrical faults using data mining algorithms with a machine learning approach. Data were collected through simulations of electrical circuits under normal and fault conditions, then analyzed using the Decision Tree, Random Forest, Naive Bayes, and K-Nearest Neighbors algorithms. The test results show that the Decision Tree and Random Forest algorithms achieve 100% accuracy and performance, while Naive Bayes also shows optimal performance with 100% accuracy. In contrast, K-Nearest Neighbors shows lower performance with 82.20% accuracy. These findings indicate that the Decision Tree and Random Forest algorithms are very effective in detecting and classifying electrical fault types accurately. The results of this study provide an important contribution to the development of automatic protection systems for more reliable and adaptive electric power systems.

Keywords: *electrical fault detection, machine learning, decision tree, random forest, naive bayes, classification*

ABSTRAK

Deteksi gangguan listrik merupakan aspek penting dalam memastikan stabilitas dan keamanan sistem tenaga listrik. Penelitian ini bertujuan untuk mengklasifikasikan jenis gangguan listrik menggunakan algoritma data mining dengan pendekatan machine learning. Data dikumpulkan melalui simulasi rangkaian listrik dalam kondisi normal dan gangguan, kemudian dianalisis menggunakan algoritma Decision Tree, Random Forest, Naive Bayes, dan K-Nearest Neighbors. Hasil pengujian menunjukkan bahwa algoritma Decision Tree dan Random Forest mencapai tingkat akurasi dan performa 100%, sedangkan Naive Bayes juga menunjukkan performa optimal dengan akurasi 100%. Sebaliknya, K-Nearest Neighbors menunjukkan performa yang lebih rendah dengan akurasi 82,20%. Temuan ini memperlihatkan bahwa algoritma Decision Tree dan Random Forest sangat efektif dalam mendeteksi dan mengklasifikasi tipe gangguan listrik secara akurat. Hasil penelitian ini memberikan kontribusi penting dalam pengembangan sistem proteksi otomatis untuk sistem tenaga listrik yang lebih andal dan adaptif.

Kata kunci: *deteksi gangguan listrik, machine learning, decision tree, random forest, Naive Bayes, klasifikasi*

1. PENDAHULUAN

Deteksi gangguan listrik (electrical fault detection) merupakan aspek penting dalam sistem tenaga listrik untuk menjaga stabilitas dan keamanan operasional. Gangguan seperti hubung singkat, gangguan tanah, dan beban berlebih dapat menyebabkan kerusakan peralatan serta potensi kecelakaan jika tidak terdeteksi dengan cepat dan tepat. [1] Dengan pesatnya perkembangan teknologi, teknik data mining klasifikasi menjadi solusi efektif dalam mengidentifikasi dan mengklasifikasikan jenis gangguan listrik berdasarkan data besar dan kompleks yang berasal dari sistem kelistrikan. Dengan menggunakan algoritma klasifikasi proses deteksi dapat berlangsung secara otomatis dengan peningkatan akurasi dan kecepatan, sehingga mendukung terbentuknya sistem proteksi yang lebih adaptif dan andal. [2] Berikut adalah beberapa metode klasifikasi yang digunakan:

1. Decision Tree

Decision tree adalah algoritma klasifikasi yang menyusun model dalam bentuk pohon keputusan, dimana setiap node mewakili pertanyaan atau kondisi berdasarkan fitur data yang diklasifikasikan. Algoritma ini populer karena hasilnya mudah dipahami dan diinterpretasi, serta mampu menangani data yang bersifat kategorikal maupun numerik. Dalam konteks deteksi gangguan listrik, decision tree dapat mengklasifikasikan jenis gangguan berdasarkan ciri khas sinyal listrik yang terekam, memberikan keputusan yang cepat dan transparan. [3]

2. Random Forest

Random forest merupakan ensemble learning yang menggabungkan banyak pohon keputusan (decision trees) secara acak untuk meningkatkan akurasi prediksi dan mengurangi risiko overfitting. Dalam deteksi gangguan listrik, random forest dapat secara efektif menangkap pola kompleks dan variabilitas data, sehingga menghasilkan klasifikasi yang lebih stabil dan andal dibandingkan satu pohon keputusan tunggal. [4]

3. Naive Bayes

Naive Bayes adalah algoritma klasifikasi probabilistik yang menggunakan teorema Bayes dengan asumsi independensi antar fitur. Meski asumsi ini sederhana, Naive Bayes sering kali memberikan performa yang baik dan efisien pada dataset besar. Algoritma ini cocok untuk deteksi gangguan listrik karena mampu menghitung probabilitas terjadinya jenis gangguan berdasarkan ciri-ciri sinyal dari data historis dengan kecepatan tinggi. [5]

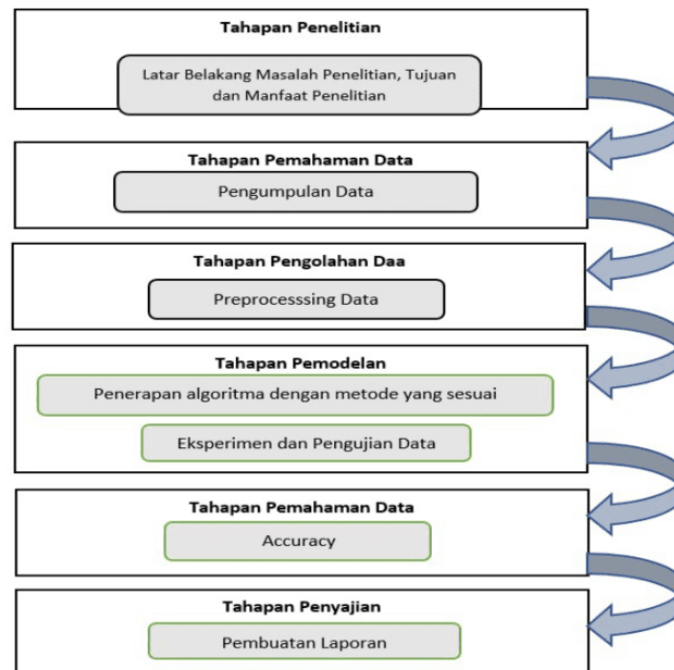
4. K-Nearest Neighbors (KNN)

KNN adalah algoritma berbasis instance yang mengklasifikasikan data baru berdasarkan kedekatannya dengan data contoh di dalam ruang fitur. KNN mengelompokkan data sesuai dengan mayoritas kelas dari k tetangga terdekatnya. Dalam aplikasi deteksi gangguan, KNN dapat mengidentifikasi jenis gangguan dengan membandingkan pola sinyal saat ini dengan pola sinyal gangguan yang sudah dikenal sebelumnya, menjadikannya metode yang intuitif namun efektif. [6]

2. METODE/PERANCANGAN PENELITIAN

Penelitian yang digunakan adalah model penelitian eksperimen. Bertujuan untuk melakukan prediksi status gangguan listrik, berdasarkan nilai akurasi dan evaluasi pada

algoritma klasifikasi data mining. Penelitian ini menekankan pada teori-teori yang sudah ada dengan dilandasi oleh kerangka pemikiran pemecahan masalah. [7]



Gambar 1. Kerangka Penelitian

Penelitian ini dilakukan dengan menjalankan beberapa langkah proses penelitian yaitu:

1. Pengumpulan Data.
2. Pengolahan Data Awal.
3. Pengukuran Penelitian.
4. Analisa hasil Penerapan Algoritma.

3. HASIL DAN PEMBAHASAN

3.1. Pengumpulan Data

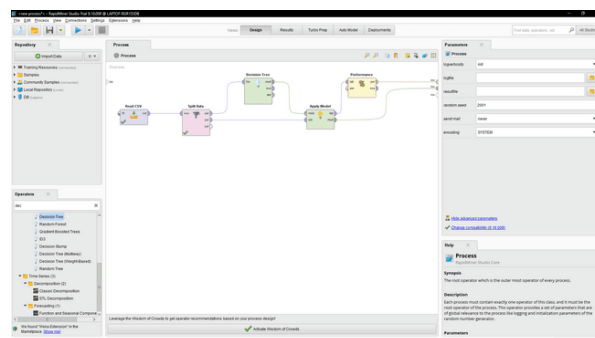
Pengumpulan data dilakukan dengan mensimulasikan rangkaian listrik dalam kondisi normal maupun kondisi gangguan. Data yang dikumpulkan meliputi tegangan dan arus pada garis keluaran sistem tenaga serta dilabeli berdasarkan tipe gangguan yang terjadi. Jumlah data yang berhasil dikumpulkan mendekati 12.000 data point, sehingga dataset ini cocok digunakan untuk pengujian dan pengembangan algoritma deteksi otomatis gangguan listrik berbasis machine learning atau data mining. Yang bersumber dari website Kaggle yang di upload oleh E Sathya Prakash. [8]

G	C	B	A	Ia	Ib	Ic	Va	Vb	Vc	Status
1	0	0	1	-151.292	-9.67745	85.80016	0.40075	-0.13293	-0.26781	LG Fault
1	0	0	1	-336.186	-76.2833	18.3289	0.312732	-0.12363	-0.1891	LG Fault
1	0	0	1	-502.892	-174.648	-80.9247	0.265728	-0.1143	-0.15143	LG Fault
1	0	0	1	-593.942	-217.703	-124.892	0.235511	-0.10494	-0.13057	LG Fault
1	0	0	1	-643.664	-224.159	-132.283	0.209537	-0.09555	-0.11398	LG Fault
1	0	0	1	-632.313	-181.715	-90.7955	0.193116	-0.08614	-0.10697	LG Fault
1	0	0	1	-557.392	-119.469	-29.5294	0.210004	-0.07671	-0.13329	LG Fault
1	0	0	1	-458.8	-96.3189	-7.38185	0.273652	-0.06726	-0.20639	LG Fault
1	0	0	1	-385.669	-97.9898	-10.0768	0.334649	-0.0578	-0.27685	LG Fault
1	0	0	1	-359.929	-87.3195	-0.45222	0.34742	-0.04831	-0.29911	LG Fault
1	0	0	1	-370.627	-66.7202	19.07983	0.323889	-0.03882	-0.28507	LG Fault
1	0	0	1	-418.537	-49.0224	35.68937	0.274574	-0.02932	-0.24525	LG Fault
1	0	0	1	-501.663	-55.2636	28.33883	0.217809	-0.01981	-0.198	LG Fault
1	0	0	1	-597.931	-96.5978	-14.1252	0.179217	-0.01029	-0.16892	LG Fault
1	0	0	1	-668.715	-136.772	-55.4499	0.160162	-0.00078	-0.15938	LG Fault
1	0	0	1	-695.776	-141.532	-61.3804	0.147641	0.00874	-0.15638	LG Fault
1	0	0	1	-689.362	-118.659	-39.697	0.139482	0.018256	-0.15774	LG Fault
1	0	0	1	-657.62	-87.1873	-9.43491	0.143773	0.027766	-0.17154	LG Fault
1	0	0	1	-608.736	-64.9491	11.57459	0.167518	0.03727	-0.20479	LG Fault
1	0	0	1	-562.586	-60.9178	14.3582	0.201832	0.046765	-0.2486	LG Fault
1	0	0	1	-540.049	-61.8202	12.1897	0.220405	0.056248	-0.27665	LG Fault
1	0	0	1	-547.016	-53.8706	18.8548	0.208581	0.065717	-0.2743	LG Fault
1	0	0	1	-576.131	-42.9498	28.47321	0.177478	0.07517	-0.25265	LG Fault
1	0	0	1	-620.141	-42.1509	27.95211	0.1433	0.084605	-0.2279	LG Fault
1	0	0	1	-671.491	-57.2031	11.56258	0.11633	0.094019	-0.21035	LG Fault
1	0	0	1	-716.567	-79.8524	-12.4411	0.100397	0.103409	-0.20381	LG Fault
1	0	0	1	-739.987	-90.9179	-24.8775	0.092266	0.112774	-0.20504	LG Fault
1	0	0	1	-738.156	-81.5928	-16.9396	0.087626	0.122111	-0.20974	LG Fault
1	0	0	1	-719.674	-62.4816	0.768475	0.087876	0.131418	-0.21929	LG Fault
1	0	0	1	-694.319	-47.9211	13.91023	0.096317	0.140693	-0.23701	LG Fault
1	0	0	1	-669.601	-43.3099	17.08731	0.11113	0.149932	-0.26106	LG Fault
1	0	0	1	-654.106	-44.7845	14.16385	0.122625	0.159135	-0.28176	LG Fault
1	0	0	1	-654.554	-44.5347	12.95013	0.119844	0.168299	-0.28814	LG Fault
1	0	0	1	-670.702	-40.4641	15.54311	0.101783	0.177421	-0.2792	LG Fault
1	0	0	1	-696.301	-38.6123	15.90337	0.07834	0.186499	-0.26484	LG Fault
1	0	0	1	-724.595	-44.5065	8.504262	0.059088	0.195531	-0.25462	LG Fault
1	0	0	1	-749.564	-55.9119	-4.41908	0.046813	0.204516	-0.25133	LG Fault
1	0	0	1	-765.009	-64.4135	-14.4514	0.039569	0.213449	-0.25302	LG Fault

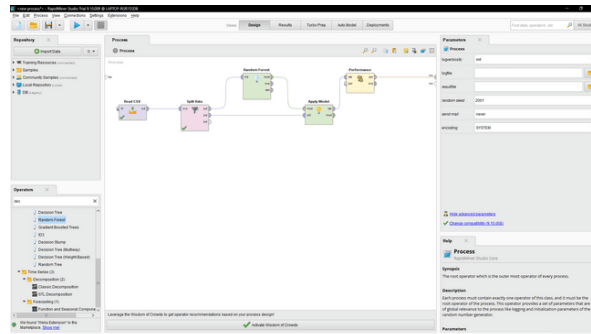
Gambar 2. Dataset Deteksi dan Klasifikasi Kesalahan Listrik

3.2. Pengolahan Data

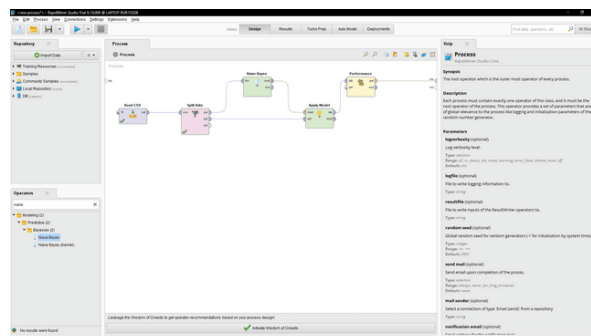
Dalam pengujian ini menggunakan rapid miner dengan operator 10-fold cross-validation untuk mendapatkan hasil accuracy yang tinggi pada setiap algoritma yang diuji menggunakan dataset. Berikut adalah proses Preprocessing. [9]



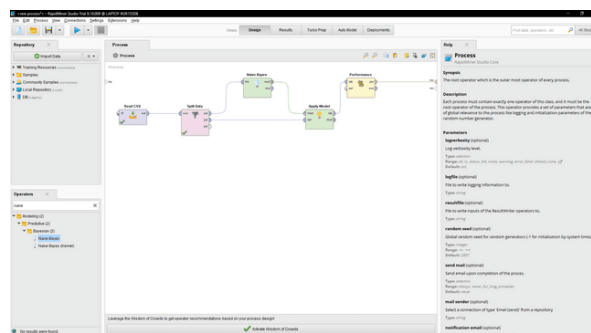
Gambar 3. Preprocessing Decision Tree



Gambar 4. Preprocessing Random Forest



Gambar 5. Preprocessing Naive Bayes



Gambar 6. Preprocessing K-Nearest Neighbors

3.3. Pengukuran Penelitian

Berdasarkan hasil olah data melalui aplikasi Rapid Miner, maka diperoleh Confution Matrix guna mengukur tingkat akurasi, dari algoritma Decision Tree, yaitu sebesar akurasi sebesar 72.41% dan menghasilkan sebuah model pohon keputusan seperti berikut.

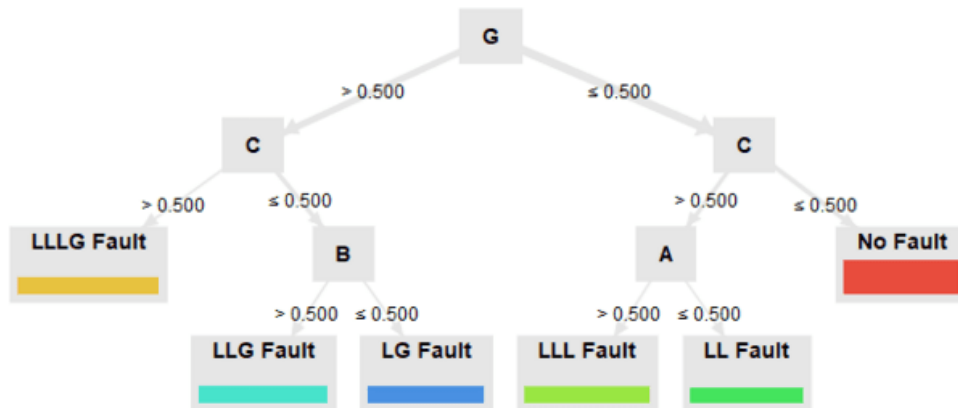
1. Decision Tree

Matriks menunjukkan performa sempurna dalam klasifikasi gangguan listrik untuk keenam tipe kelas, dengan 100% akurasi, precision, dan recall di seluruh kelas model mengklasifikasikan semua data secara tepat pada seluruh typenya, tanpa error satupun. Ini sangat ideal untuk aplikasi sistem proteksi atau diagnosis otomatis dalam lingkungan yang serupa dengan dataset pengujian tersebut.

accuracy: 100.00%

	true LG Fault	true LLG Fault	true LL Fault	true LLL Fault	true LLLG Fault	true No Fault	class precision
pred. LG Fault	226	0	0	0	0	0	100.00%
pred. LLG Fault	0	227	0	0	0	0	100.00%
pred. LL Fault	0	0	201	0	0	0	100.00%
pred. LLL Fault	0	0	0	219	0	0	100.00%
pred. LLLG Fault	0	0	0	0	227	0	100.00%
pred. No Fault	0	0	0	0	0	473	100.00%
class recall	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	

Gambar 7. Hasil Uji Decision Tree



Gambar 8. Model Decision Tree

2. Random Forest

Model klasifikasi ini bekerja sempurna pada data uji yang diberikan, baik untuk seluruh tipe gangguan maupun kondisi tanpa gangguan, tanpa ada error satu pun. Hasil ini menunjukkan model sangat andal pada data ini, cocok untuk sistem deteksi otomatis gangguan listrik dalam lingkungan yang serupa.

Namun, hasil 100% juga menandakan dataset bisa jadi bersih, sangat terstruktur, atau model telah sangat menyesuaikan diri pada data (kemungkinan overfitting perlu diuji lebih jauh pada data lain yang berbeda).

accuracy: 100.00%

	true LG Fault	true LLG Fault	true LL Fault	true LLL Fault	true LLLG Fault	true No Fault	class precision
pred. LG Fault	226	0	0	0	0	0	100.00%
pred. LLG Fault	0	227	0	0	0	0	100.00%
pred. LL Fault	0	0	201	0	0	0	100.00%
pred. LLL Fault	0	0	0	219	0	0	100.00%
pred. LLLG Fault	0	0	0	0	227	0	100.00%
pred. No Fault	0	0	0	0	0	473	100.00%
class recall	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	

Gambar 9. Hasil Uji Random Forest

3. Naive Bayes

Berdasarkan confusion matrix, model klasifikasi gangguan listrik menunjukkan performa yang sangat baik dengan akurasi, presisi, dan recall 100% pada seluruh kelas gangguan (LG, LL, LLG, LLL, LLLG) maupun kondisi normal. Hal ini

menandakan sistem mampu mengenali setiap pola gangguan dengan tepat tanpa kesalahan klasifikasi.

accuracy: 100.00%

	true LG Fault	true LLG Fault	true LL Fault	true LLL Fault	true LLLG Fault	true No Fault	class precision
pred. LG Fault	226	0	0	0	0	0	100.00%
pred. LLG Fault	0	227	0	0	0	0	100.00%
pred. LL Fault	0	0	201	0	0	0	100.00%
pred. LLL Fault	0	0	0	219	0	0	100.00%
pred. LLLG Fault	0	0	0	0	227	0	100.00%
pred. No Fault	0	0	0	0	0	473	100.00%
class recall	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	

Gambar 10. Hasil Uji Naive Bayes

4. K-Nearest Neighbors

Berdasarkan confusion matrix dengan akurasi 82.20%, model klasifikasi mampu mengenali sebagian besar jenis gangguan listrik dengan baik, terutama untuk kelas LG, LLG, LL, dan No Fault, yang menunjukkan precision di atas 95%. Namun, model masih mengalami kesalahan signifikan pada kelas LLL dan LLLG Fault, yang memiliki tingkat precision dan recall di bawah 50%. Kesamaan karakteristik sinyal antar kedua jenis gangguan menjadi faktor utama penurunan performa, sehingga diperlukan optimalisasi lebih lanjut dalam tahap ekstraksi fitur dan pelatihan model.

accuracy: 82.20%

	true LG Fault	true LLG Fault	true LL Fault	true LLL Fault	true LLLG Fault	true No Fault	class precision
pred. LG Fault	225	10	0	0	0	0	95.74%
pred. LLG Fault	1	215	0	1	5	0	95.85%
pred. LL Fault	0	0	198	1	2	0	98.51%
pred. LLL Fault	0	0	0	79	117	0	40.31%
pred. LLLG Fault	0	2	0	138	103	0	42.39%
pred. No Fault	0	0	3	0	0	473	99.37%
class recall	99.55%	94.71%	98.51%	36.97%	45.37%	100.00%	

Gambar 11. Hasil Uji K-Nearest Neighbors

3.4. Analisa Hasil Penelitian

Tabel hasil pengolahan data dapat dilihat pada tabel dibawah ini :

Tabel 1. Keseluruhan Hasil Uji

Metode	Accuracy	Rata-rata Precision	Rata-Rata Recall
Decision Tree	100%	100%	100%
Random Forest	100%	100%	100%
Naive Bayes	100%	100%	100%
K-Nearest Neighbors.	82,20%	78,53%	78,37%

4. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma decision tree, random forest dan naive bayes menunjukkan performa yang optimal dengan tingkat akurasi mencapai 100% dalam klasifikasi berbagai jenis gangguan listrik. Hal ini menunjukkan bahwa ketiga algoritma ini sangat efektif dan dapat diandalkan untuk digunakan dalam sistem deteksi otomatis gangguan listrik. Di sisi lain, K-Nearest Neighbors memiliki tingkat akurasi yang lebih rendah, yaitu 82,20%, dan mengalami kesulitan dalam membedakan jenis gangguan yang memiliki karakteristik sinyal yang serupa. Hasil ini menegaskan pentingnya pemilihan algoritma yang tepat untuk meningkatkan keakuratan dalam sistem deteksi otomatis gangguan listrik. Disarankan untuk melakukan pengujian lebih lanjut pada dataset yang berbeda guna memastikan keandalan model dan mengoptimalkan proses klasifikasi.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada LPDP atas dukungan beasiswa yang telah membantu dalam pelaksanaan dan penyelesaian penelitian ini. Penulis juga mengucapkan terima kasih kepada semua pihak promotor, dosen matakuliah, rekan-rekan doktoral Departemen Teknik Elektro dan Informatika, Universitas Negeri Malang yang telah memberi dukungan yang membantu pelaksanaan penelitian dan atau penulisan artikel.

DAFTAR PUSTAKA

- [1] S. S. Barani K, K. S, H. S, and M. S. N, "Fault Detection and Fault Classification in Transmission Lines using Machine Learning," in *2023 International Conference on Intelligent Technologies for Sustainable Electric and Communications Systems (iTech SECOM)*, 2023, pp. 158–162. doi: 10.1109/iTechSECOM59882.2023.10435012.
- [2] H. Marsadi, S. Amalia, A. M. N. Putra, A. Syofian, E. Erhaneli, and A. Antonov, "Perancangan Klasifikasi Kelainan Penyaluran Tenaga Listrik Jaringan Distribusi 20 kV Kota Sabang Menggunakan Neural Network," *J. Multidisiplin Teknol. dan Arsit.*, vol. 2, no. 2, pp. 823–835, 2024, doi: 10.57235/motekar.v2i2.3947.
- [3] Oscario Oscario, "Penerapan Data Mining Dengan Metode Decision Tree Menggunakan Algoritma C4.5 Untuk Memprediksi Gaya Belajar Yang Cocok Untuk Siswa Siswi (Studi Kasus : SD Sariputra Jambi)," vol. lim, no. 2009, pp. 1–25, 2019.
- [4] A. Seva, S. N. Tirumala Rao, and M. Sireesha, "Prediction of Liver Disease with Random Forest Classifier Through SMOTE-ENN Balancing," in *2024 IEEE 13th International Conference on Communication Systems and Network Technologies (CSNT)*, 2024, pp. 928–933. doi: 10.1109/CSNT60213.2024.10546170.
- [5] M. Langarizadeh and F. Moghbeli, "Applying Naive Bayesian Networks to Disease Prediction: a Systematic Review.," *Acta Inform. medica AIM J. Soc. Med. Informatics Bosnia Herzegovina Cas. Drus. za Med. Inform. BiH*, vol. 24, no. 5, pp. 364–369, Oct. 2016, doi: 10.5455/aim.2016.24.364-369.
- [6] M. A. jabbar, B. L. Deekshatulu, and P. Chandra, "Classification of Heart Disease Using K- Nearest Neighbor and Genetic Algorithm," *Procedia Technol.*, vol. 10, pp. 85–94, 2013, doi: <https://doi.org/10.1016/j.protcy.2013.12.340>.
- [7] A. F. Riany and G. Testiana, "Penerapan Data Mining untuk Klasifikasi Penyakit

- Stroke Menggunakan Algoritma Naïve Bayes,” *J. SAINTEKOM*, vol. 13, no. 1, pp. 42–54, 2023, doi: 10.33020/saintekom.v13i1.352.
- [8] E Sathya Prakash, “Electrical Fault detection and classification,” 2021. https://www.kaggle.com/datasets/esathyaprakash/electrical-fault-detection-and-classification/data?select=detect_dataset.csv (accessed Nov. 11, 2025).
- [9] R. Oktafiani, A. Hermawan, and D. Avianto, “Pengaruh Komposisi Split data Terhadap Performa Klasifikasi Penyakit Kanker Payudara Menggunakan Algoritma Machine Learning,” *J. Sains dan Inform.*, vol. 9, no. April, pp. 19–28, 2023, doi: 10.34128/jsi.v9i1.622.
- [10] K. Dedes, A. Wibawa, and L. Budiarto, “Sistematika Filsafat Menurut Ontologi, Epistemologi, dan Aksiologi dalam Artificial Intelligence,” *J. Inov. Teknol. dan Edukasi Tek.*, vol. 1, no. 8, pp. 584–591, 2021, doi: 10.17977/um068v1i82021p584-591.
- [11] H. Yulianto, “Filsafat Statistika Bisnis Di Era Big Data Dan Ai : Relasi Triadik Ontologi , Etika , Dan Epistemologis the Philosophy of Business Statistics in the Era of Big Data and Ai : a Triadic Interrelation of Ontological , Ethical , and Epistemological,” no. April, pp. 7460–7474, 2025.
- [12] F. Rio Naibaho, “Tinjauan Ontologi, Epistemologi, Dan Aksiologi Dalam Bidang Ilmu Steganografi,” *J. Sci. Soc. Res.*, vol. 4307, no. 1, pp. 534–542, 2025, [Online]. Available: <http://jurnal.goretanpena.com/index.php/JSSR534>
- [13] E. O. Ayala, M. M. Altamirano-Bustamante, and A. de Hoyos-Bermea, “Axiology and dynamics of contemporary research groups: a systematic review and hermeneutic meta-analysis of knowledge, values, and social elements.,” *Front. Res. metrics Anal.*, vol. 10, p. 1525587, 2025, doi: 10.3389/frma.2025.1525587.
- [14] H. Herdianto, M. Zarlis, and Z. Nasution, “Filsafat Ilmu Sebagai Landasan Pengembangan Komputer Menuju Komputer Vision,” *J. Ilm. Abdi Ilmu*, vol. 14, no. 2, pp. 171–184, 2021, [Online]. Available: <https://journal.pancabudi.ac.id/index.php/abdiilmu/article/view/4054>
- [15] M. S. S. Uqba, “Philosophy of Science in the Era of Big Data and Artificial Intelligence: Challenges and Opportunities,” *Al-Iqro’ J. Islam. Stud. Vol. 1, No. 1, Januari 2024*, DOI 10.54622/aijis.v1i1.271 ISSN 3032-3436 ISSN 3031-3732, vol. 1, no. Al-Iqro’: Journal of Islamic Studies, pp. 102–111, 2024, doi: 10.54622/aijis.v1i1.271.