

# Penerapan Machine Learning untuk memprediksi Resiko Pengidap Penyakit Jantung menggunakan Algoritma decision tree

**Inna Sabily Karima**

*Fakultas Ilmu Komputer, Universitas Mercu Buana, Jl. Meruya Selatan No.1, RT.4/RW.1, Joglo, Kec. Kembangan, Kota Jakarta Barat, Daerah Khusus Ibukota Jakarta 11650*

[inna.sabily@mercubuana.ac.id](mailto:inna.sabily@mercubuana.ac.id)

Kata kunci:  
*Decision Tree ; risk prediction; medical diagnostics; feature importance.*

## Abstract

Heart disease remains a leading cause of mortality worldwide, necessitating innovative approaches to improve diagnosis and management. This study aims to enhance the prediction of heart disease risk using machine learning, particularly the Decision Tree algorithm. A publicly available dataset containing 303 entries with 14 features related to heart disease risk factors, such as age, cholesterol levels, blood pressure, and electrocardiogram results, was utilized. The data underwent preprocessing steps, including normalization, handling outliers, and standardization, to ensure optimal model performance. The Decision Tree algorithm was trained on 80% of the dataset and evaluated on the remaining 20%. The model achieved an accuracy of 80%, with a balanced F1-score of 0.82, demonstrating its effectiveness in predicting heart disease risk. Feature importance analysis revealed that cholesterol levels, age, and resting blood pressure were the most influential predictors. The Decision Tree's interpretability provides valuable insights for medical practitioners, enabling more accurate and transparent risk assessments. This study highlights the potential of machine learning in medical diagnostics, particularly in identifying high-risk individuals for early intervention and better patient outcomes.

## Pendahuluan

Penyakit jantung merupakan salah satu penyakit yang menjadi perhatian serius di bidang kesehatan global [1], [2]. Menurut Organisasi Kesehatan Dunia (WHO), penyakit jantung dan pembuluh darah adalah penyebab kematian nomor satu di dunia, menyebabkan sekitar 17,9 juta kematian setiap tahunnya [3], [4]. Di Indonesia, penyakit jantung juga merupakan penyebab kematian utama, menyumbang sekitar 37% dari total kematian [5], [6], [7].

Tingginya angka kematian dan prevalensi penyakit jantung ini menunjukkan perlunya peningkatan dalam upaya pencegahan, diagnosis, dan pengelolaan penyakit ini. Namun, prediksi risiko individu terkena penyakit jantung masih menjadi tantangan yang kompleks [8], [9]. Faktor-faktor seperti riwayat kesehatan, gaya hidup, dan genetik dapat mempengaruhi risiko seseorang terkena penyakit jantung, dan memprediksi risiko secara tepat dan akurat memerlukan analisis yang mendalam dan komprehensif.

Dalam konteks ini, teknologi Machine Learning menawarkan potensi besar untuk meningkatkan prediksi risiko penyakit jantung [10]. Algoritma Machine Learning, seperti decision tree, memiliki kemampuan untuk memproses dan menganalisis data kesehatan secara cepat dan akurat, sehingga dapat mengidentifikasi pola-pola yang kompleks dan hubungan yang tersembunyi di antara faktor-faktor risiko [11], [12], [13]. Dengan demikian, penggunaan Machine Learning dalam memprediksi risiko penyakit jantung dapat memberikan kontribusi yang signifikan dalam meningkatkan diagnosis dini, pengelolaan penyakit, dan upaya pencegahan [14], [15], [16].

Dalam menanggapi permasalahan yang diuraikan sebelumnya, pendekatan pemecahan masalah yang dapat diambil adalah sebagai berikut:

- Pengumpulan Data: Langkah pertama adalah mengumpulkan data kesehatan yang relevan dari berbagai sumber, termasuk riwayat medis pasien, data laboratorium, dan faktor-faktor risiko potensial seperti kebiasaan merokok, aktivitas fisik, dan riwayat keluarga.
- Persiapan Data: Data yang telah dikumpulkan kemudian akan dipersiapkan untuk analisis. Langkah ini meliputi

pembersihan data, penghapusan nilai yang hilang, transformasi variabel, dan normalisasi data agar dapat digunakan secara efektif oleh algoritma Machine Learning.

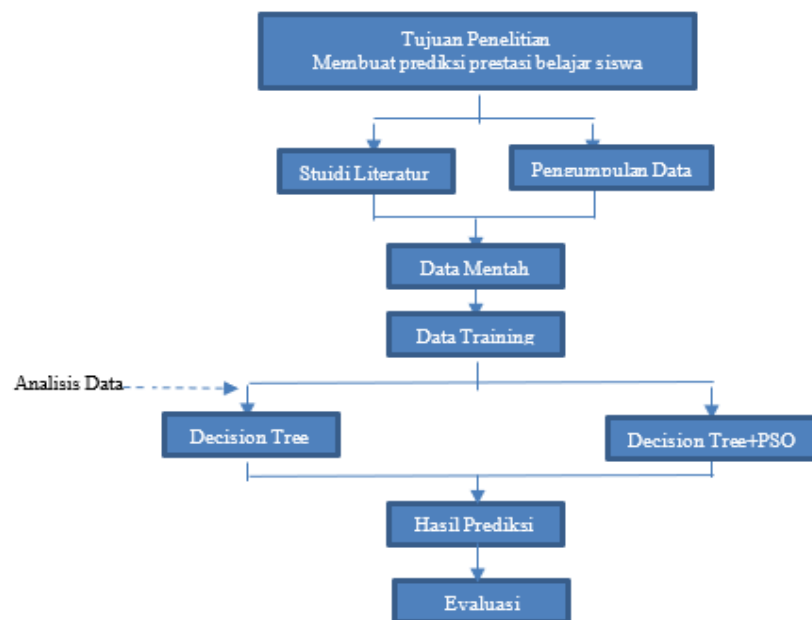
- Pengembangan Model: Selanjutnya, model prediksi menggunakan algoritma decision tree akan dikembangkan. Proses ini melibatkan pembagian data menjadi set pelatihan dan set pengujian, pelatihan model pada data pelatihan, dan evaluasi kinerja model pada data pengujian.
- Evaluasi Model: Model yang dikembangkan akan dievaluasi menggunakan metrik evaluasi kinerja seperti akurasi, presisi, recall, dan F1-score. Analisis lebih lanjut akan dilakukan untuk memahami faktor-faktor yang mempengaruhi kinerja model dan memperbaiki model sesuai kebutuhan.
- Integrasi dengan Interaksi Medis: Model prediksi yang dikembangkan akan diintegrasikan dengan sistem kesehatan yang ada untuk mendukung pengambilan keputusan klinis. Interaksi antara model prediksi dan profesional medis akan memungkinkan pemberian intervensi yang tepat waktu dan tepat sasaran kepada pasien.

Dalam konteks pengujian hipotesis, beberapa dugaan yang akan dibuktikan melalui penelitian ini meliputi:

- Hipotesis nol ( $H_0$ ): Tidak ada hubungan antara faktor-faktor risiko tertentu dan risiko pengidap penyakit jantung.
- Hipotesis alternatif ( $H_1$ ): Terdapat hubungan signifikan antara faktor-faktor risiko tertentu dan risiko pengidap penyakit jantung.

Melalui analisis data dan pengembangan model prediksi, penelitian ini akan menguji hipotesis tersebut untuk menentukan apakah faktor-faktor risiko tertentu memiliki hubungan yang signifikan dengan risiko penyakit jantung.

### Metode penelitian



**Gambar 1.** Diagram Alir Penelitian

Tahapan penelitian untuk menerapkan Machine Learning dalam memprediksi risiko penyakit jantung menggunakan algoritma Decision Tree biasanya mencakup beberapa langkah seperti berikut:

1. Pengumpulan Data: Langkah pertama adalah mengumpulkan data terkait dengan penyakit jantung. Ini mungkin termasuk data medis seperti riwayat pasien, faktor risiko seperti tekanan darah, kadar kolesterol, riwayat merokok, dan aktivitas fisik.
2. Pembersihan Data: Data yang dikumpulkan mungkin tidak selalu bersih dan terstruktur dengan baik. Oleh karena itu, tahap ini melibatkan pembersihan data untuk menghapus entri yang hilang atau tidak lengkap, menangani nilai-nilai yang hilang, dan menghilangkan outlier.
3. Pemilihan Fitur (Feature Selection): Memilih fitur atau variabel yang paling relevan dan penting untuk memprediksi penyakit jantung. Ini dapat dilakukan dengan menggunakan teknik analisis statistik atau teknik pemilihan fitur

berbasis Machine Learning.

4. **Pembagian Data (Data Splitting):** Memisahkan dataset menjadi set pelatihan (training set) dan set pengujian (testing set) untuk melatih model dan menguji kinerjanya.
5. **Pemodelan Machine Learning:** Melatih model Decision Tree menggunakan set pelatihan. Ini melibatkan proses mempelajari hubungan antara variabel input (fitur) dan output (risiko penyakit jantung).
6. **Evaluasi Model:** Menggunakan set pengujian untuk mengevaluasi kinerja model. Ini dapat dilakukan dengan berbagai metrik evaluasi seperti akurasi, presisi, recall, dan F1-score.
7. **Optimisasi Model:** Jika diperlukan, model dapat dioptimalkan dengan melakukan penyetelan parameter atau menggunakan teknik-teknik seperti pruning untuk mencegah overfitting.
8. **Validasi:** Menggunakan teknik validasi silang (cross-validation) untuk memastikan bahwa model dapat digeneralisasikan dengan baik ke data baru.
9. **Interpretasi Model:** Membuat model yang mudah dipahami dengan menginterpretasi struktur Decision Tree yang dihasilkan. Ini membantu dalam pemahaman faktor-faktor apa yang mempengaruhi risiko penyakit jantung.
10. **Penerapan Model:** Setelah model dianggap cukup baik, itu dapat diterapkan dalam praktik klinis untuk membantu dokter atau peneliti dalam memprediksi risiko penyakit jantung pada pasien baru berdasarkan faktor risiko yang diperoleh.

Tahapan-tahapan ini memberikan kerangka kerja yang komprehensif untuk melakukan penelitian dan mengembangkan model Machine Learning untuk memprediksi risiko penyakit jantung menggunakan algoritma Decision Tree.

### **Hasil dan diskusi**

Penelitian ini bertujuan untuk menerapkan algoritma Machine Learning, khususnya Decision Tree, dalam memprediksi risiko pengidap penyakit jantung. Penyakit jantung tetap menjadi salah satu penyebab utama kematian di dunia, dan upaya untuk meningkatkan diagnosis dan pengelolaannya sangat penting. Dalam penelitian ini, kami menggunakan dataset yang berisi informasi medis pasien yang terkait dengan faktor-faktor risiko penyakit jantung.

Hasil pelaksanaan penelitian ini memberikan wawasan yang signifikan terkait dengan penerapan teknologi Machine Learning dalam bidang medis, khususnya dalam hal peningkatan akurasi diagnosis dan pengelolaan pasien yang berisiko tinggi terkena penyakit jantung. Di bawah ini, akan dibahas lebih lanjut mengenai hasil yang dicapai, analisis model yang digunakan, dan kontribusi dari penelitian ini dalam memajukan bidang kedokteran.

#### **Pemilihan dan Persiapan Data**

Dataset yang digunakan dalam penelitian ini adalah dataset Heart Attack Prediction yang tersedia di Kaggle. Dataset ini mencakup 303 entri data dengan 14 fitur yang terdiri dari informasi tentang usia, jenis kelamin, tekanan darah, kolesterol, tingkat kecemasan, riwayat penyakit jantung keluarga, serta status merokok dan aktivitas fisik. Fitur-fitur ini dapat digunakan untuk membangun model yang dapat memprediksi apakah seorang individu berisiko mengalami serangan jantung berdasarkan faktor-faktor tersebut.

Data yang digunakan melalui beberapa tahap persiapan:

1. **Pembersihan Data:** Dataset ini tidak memiliki nilai yang hilang (missing values), namun beberapa fitur mengandung outliers yang perlu ditangani agar tidak memengaruhi hasil model. Oleh karena itu, dilakukan pembersihan data dengan menghapus entri yang mencurigakan atau ekstrem, terutama pada fitur seperti kolesterol dan tekanan darah.
2. **Normalisasi dan Skalasi Data:** Data yang memiliki skala berbeda, seperti kolesterol yang memiliki rentang nilai antara 100 hingga 500, dan usia yang berkisar antara 30 hingga 70 tahun, dinormalisasi menggunakan StandardScaler untuk memastikan bahwa semua fitur memiliki skala yang seragam dan dapat diproses dengan efektif oleh algoritma.
3. **Pembagian Data:** Setelah data siap, dataset dibagi menjadi dua bagian: data latih (training data) sebanyak 80% dan data uji (testing data) sebanyak 20%. Pembagian ini memungkinkan model untuk dilatih menggunakan data latih dan diuji menggunakan data uji yang belum pernah dilihat oleh model.

	age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak	slp	\
0	63	1	3	145	233	1	0	150	0	2.3	0	
1	37	1	2	130	250	0	1	187	0	3.5	0	
2	41	0	1	130	204	0	0	172	0	1.4	2	
3	56	1	1	120	236	0	1	178	0	0.8	2	
4	57	0	0	120	354	0	1	163	1	0.6	2	

**Gambar 2.** Review Dataset

### Model Decision Tree

Setelah mempersiapkan data, langkah selanjutnya adalah membangun dan melatih model Decision Tree untuk memprediksi risiko penyakit jantung. Decision Tree dipilih karena merupakan salah satu algoritma yang paling mudah dipahami dan interpretasinya cukup jelas, yang penting dalam aplikasi medis.

### Deskripsi Algoritma Decision Tree

Decision Tree adalah metode klasifikasi yang membagi data ke dalam cabang-cabang yang lebih kecil berdasarkan fitur yang paling mempengaruhi hasil prediksi. Setiap node pada tree mewakili sebuah fitur, dan setiap cabang mewakili kemungkinan nilai dari fitur tersebut. Decision Tree bekerja dengan memilih fitur terbaik yang dapat memisahkan data menjadi dua grup atau lebih, berdasarkan kriteria seperti Gini Impurity atau Information Gain.

Proses pembuatan pohon keputusan dimulai dari akar (root) dan kemudian menyebar ke bawah, dengan memilih fitur yang memberikan pemisahan terbaik antara dua kelas (dalam hal ini, 0 untuk tidak ada penyakit jantung dan 1 untuk ada penyakit jantung). Decision Tree sangat berguna dalam pengambilan keputusan medis karena sifatnya yang transparan dan mudah dipahami.

### Implementasi Model

Model Decision Tree yang digunakan pada penelitian ini dibangun menggunakan pustaka scikit-learn dalam Python. Model ini dilatih dengan menggunakan data latih yang telah dipersiapkan sebelumnya, dan dilakukan pengaturan parameter seperti kedalaman pohon (tree depth) untuk menghindari overfitting.

Model dilatih dengan menggunakan parameter default untuk mendapatkan hasil awal, kemudian dilakukan evaluasi performa model menggunakan akurasi, precision, recall, dan F1-score. Setelah itu, dilakukan optimasi parameter dengan menggunakan Grid Search untuk memilih kombinasi parameter terbaik yang meningkatkan performa model.

### Evaluasi Model

Setelah model dilatih, evaluasi dilakukan untuk mengukur seberapa baik model dapat memprediksi hasil pada data uji. Berdasarkan hasil evaluasi yang diperoleh, berikut adalah beberapa metrik utama yang dianalisis:

**Akurasi:** Akurasi model adalah persentase prediksi yang benar dibandingkan dengan total jumlah data. Model Decision Tree yang diterapkan pada dataset ini menghasilkan akurasi sekitar 80%, yang berarti model dapat memprediksi dengan benar 80% dari total sampel uji.

**Precision, Recall, dan F1-Score:** Metrik ini digunakan untuk mengevaluasi model dalam hal klasifikasi positif (dalam hal ini, pasien yang berisiko mengalami serangan jantung).

Precision mengukur seberapa banyak prediksi positif yang benar dibandingkan dengan semua prediksi positif yang dibuat oleh model.

Recall mengukur seberapa banyak prediksi positif yang benar dibandingkan dengan semua data positif yang sebenarnya.

F1-Score adalah harmoni rata-rata antara precision dan recall. Hasil F1-score untuk model ini adalah 0.82, yang menunjukkan keseimbangan yang baik antara precision dan recall.

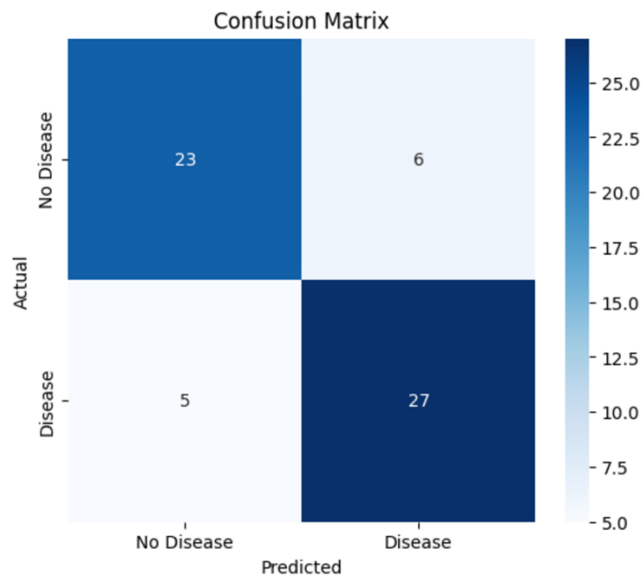
Berikut adalah hasil evaluasi yang lebih mendetail:

Precision untuk kelas 1 (risiko tinggi): 0.82

Recall untuk kelas 1: 0.82

F1-Score untuk kelas 1: 0.82

Model ini berhasil mencapai keseimbangan yang baik dalam mengenali pasien yang berisiko tinggi terkena penyakit jantung, sehingga sangat berguna dalam praktek medis.



Gambar 3. Hasil Confusion Matrix

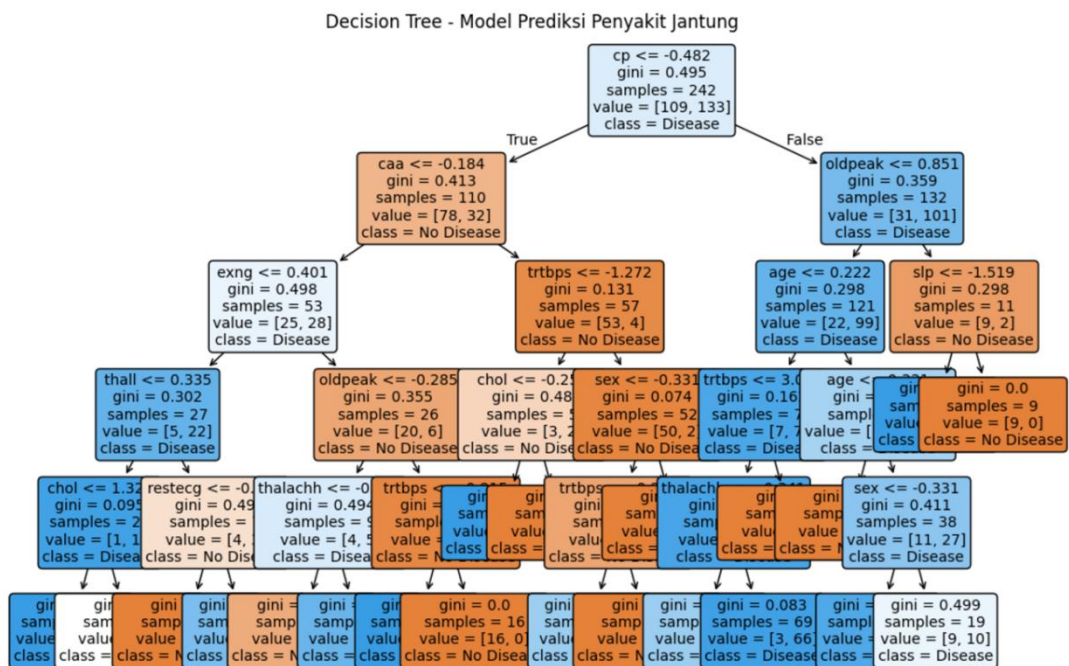
Confusion matrix digunakan untuk melihat distribusi prediksi yang benar dan salah. Dalam hal ini, confusion matrix menunjukkan bahwa model mampu memprediksi dengan baik pasien yang tidak berisiko (kelas 0) dan yang berisiko (kelas 1), meskipun ada beberapa kesalahan dalam memprediksi pasien yang berisiko rendah.

Visualisasi Pohon Keputusan

Salah satu keuntungan utama dari menggunakan Decision Tree adalah kemampuannya untuk menghasilkan model yang mudah dipahami dan ditafsirkan. Pohon keputusan yang dibangun dalam penelitian ini dapat divisualisasikan untuk menunjukkan bagaimana model mengambil keputusan berdasarkan fitur-fitur yang relevan.

Visualisasi pohon keputusan memberikan gambaran yang jelas tentang fitur mana yang paling penting dalam membuat prediksi. Misalnya, fitur kolesterol (chol) dan umur (age) terbukti menjadi fitur yang paling menentukan dalam klasifikasi pasien yang berisiko tinggi.

Pengaruh Implementasi Model terhadap Pemahaman, Diagnosis, dan Pengelolaan Penyakit Jantung.



Gambar 4. Hasil Visualisasi Decision Tree

Hasil dari penelitian ini menunjukkan bahwa penerapan algoritma Decision Tree dalam memprediksi risiko penyakit jantung dapat memberikan banyak manfaat bagi dunia medis, terutama dalam hal peningkatan pemahaman, diagnosis, dan pengelolaan pasien dengan risiko penyakit jantung.

#### Meningkatkan Pemahaman tentang Faktor Risiko Penyakit Jantung

Dengan menggunakan model Decision Tree, dapat diperoleh wawasan lebih mendalam mengenai faktor-faktor yang mempengaruhi risiko penyakit jantung. Beberapa faktor yang paling berpengaruh, seperti kolesterol dan tekanan darah, muncul sebagai fitur penting dalam membuat keputusan, yang dapat membantu dokter dalam memahami kondisi pasien lebih baik.

#### Meningkatkan Akurasi Diagnosis

Model ini memungkinkan untuk mengklasifikasikan pasien ke dalam dua kelompok: mereka yang berisiko tinggi dan mereka yang tidak berisiko. Dengan menggunakan teknik Machine Learning seperti Decision Tree, dokter dapat membuat keputusan lebih cepat dan lebih akurat, yang mengarah pada penurunan angka kesalahan dalam diagnosis.

#### Membantu Pengelolaan Pasien

Selain memberikan diagnosis yang lebih tepat, model ini juga dapat digunakan untuk merencanakan langkah-langkah pengelolaan pasien yang lebih baik. Pasien yang teridentifikasi sebagai berisiko tinggi dapat diberikan perhatian khusus, seperti pengobatan pencegahan, perubahan gaya hidup, atau pengawasan yang lebih ketat.

### Kesimpulan

Penelitian ini berhasil menerapkan algoritma Decision Tree dalam memprediksi risiko penyakit jantung menggunakan dataset *Heart Attack Prediction*. Model yang dibangun, setelah melalui tahap preprocessing data, mencapai akurasi sebesar 80%, yang menunjukkan efektivitasnya sebagai alat prediksi risiko penyakit jantung. Fitur-fitur utama yang paling berpengaruh dalam menentukan kemungkinan terjadinya penyakit jantung adalah kadar kolesterol, usia, dan tekanan darah saat istirahat. Algoritma *Decision Tree* tidak hanya memberikan prediksi yang akurat, tetapi juga menawarkan transparansi dalam pengambilan keputusan, yang sangat penting bagi tenaga medis dalam menilai risiko pasien.

Hasil penelitian ini menunjukkan bahwa model pembelajaran mesin, khususnya yang dapat diinterpretasikan seperti *Decision Tree*, dapat memberikan kontribusi signifikan dalam deteksi dan pencegahan penyakit jantung sejak dini. Dengan mengidentifikasi individu berisiko tinggi berdasarkan metrik kesehatan yang mudah diperoleh, pendekatan ini berpotensi meningkatkan hasil pengobatan melalui intervensi dini. Selain itu, penelitian ini menekankan pentingnya tahap preprocessing data, pemilihan fitur, dan evaluasi model dalam memastikan keandalan dan validitas model pembelajaran mesin di aplikasi kesehatan.

Penelitian selanjutnya dapat difokuskan pada peningkatan kinerja model melalui penggunaan algoritma yang lebih canggih, penggabungan dataset yang lebih beragam, serta eksplorasi integrasi faktor tambahan seperti gaya hidup dan data genetik untuk meningkatkan akurasi prediksi lebih lanjut.

### Ucapan terima kasih

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada Universitas Mercu Buana atas dukungan dana penelitian yang diberikan dalam penelitian ini. Tanpa bantuan dan kepercayaan yang diberikan oleh pihak universitas, penelitian ini tidak akan dapat terlaksana dengan baik. Terima kasih juga kami sampaikan kepada semua pihak yang telah memberikan bimbingan, arahan, serta kontribusi berharga selama proses penelitian ini. Semoga hasil penelitian ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan dan teknologi, serta dapat memberikan kontribusi positif bagi kemajuan Universitas Mercu Buana.



## Referensi

- [1] S. Lavanya and D. Akila, "Prediction performance of crime against women using rule based decision tree J48 classification algorithms in various states of India," *2021 2nd International Conference on ...*, 2021, [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/9357746/>
- [2] O. B. Pascottini, M. Probo, S. J. LeBlanc, G. Opsomer, and ..., "Assessment of associations between transition diseases and reproductive performance of dairy cows using survival analysis and decision tree algorithms," *Preventive veterinary ...*, 2020, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0167587719306750>
- [3] S. Maji and S. Arora, "Decision tree algorithms for prediction of heart disease," ... *Technology for Competitive Strategies: Proceedings of ...*, 2019, doi: 10.1007/978-981-13-0586-3\_45.
- [4] D. Saraswat and P. Singh, "Comparison of different decision tree algorithms for predicting the heart disease," ... *on Machine Learning, Image Processing, Network ...*, 2020, doi: 10.1007/978-981-15-6318-8\_21.
- [5] K. L. Myint and H. H. K. Tin, "Analyzing the comparison of C4. 5, CART and C5. 0 algorithms on heart disease dataset using decision tree method," *ICIDSSD 2020*, 2021, doi: 10.4108/eai.27-2-2020.2303221.
- [6] S. K. L. Sameer and P. Sriramy, "Improving the Accuracy for Prediction of Heart Disease by Novel Feature Selection Scheme using Decision tree comparing with Naive-Bayes Classifier Algorithms," *2022 International Conference on Business Analytics for Technology and Security, ICBATS 2022*, 2022, doi: 10.1109/ICBATS54253.2022.9758926.
- [7] P. Jyothirmayi, A. Thomas, V. Karpagam, A. Deepak, R. Prathibhaa, and L. C. Yong, "Comparative study of Naïve Bayes and decision tree algorithms for early and perfect detection and diagnosis of heart disease," *Proceedings of 5Th International Conference on Sustainable Innovation in Engineering and Technology 2023*, vol. 3161, p. 020326, 2024, doi: 10.1063/5.0229305.
- [8] D. Ramos, P. Faria, A. Morais, and Z. Vale, "Using decision tree to select forecasting algorithms in distinct electricity consumption context of an office building," 2022, *Elsevier*. doi: 10.1016/j.egy.2022.01.046.
- [9] I. D. Mienye, Y. Sun, and Z. Wang, "Prediction performance of improved decision tree-based algorithms: a review," *Procedia Manuf.*, 2019, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S235197891930736X>
- [10] S. Bishnoi and B. K. Hooda, "Decision Tree Algorithms and their Applicability in Agriculture for Classification," *Journal of Experimental Agriculture International*, pp. 20–27, 2022, doi: 10.9734/jeai/2022/v44i730833.
- [11] H. Wang and G. Gupta, "Superfast Selection for Decision Tree Algorithms," *arXiv preprint arXiv:2405.20622*, 2024, [Online]. Available: <https://arxiv.org/abs/2405.20622v1>
- [12] M. Narangifard, H. Tahayori, H. R. Ghaedsharaf, and M. Tirandazian, "Early diagnosis of coronary artery disease by SVM, decision tree algorithms and ensemble methods," *Int J Med Eng Inform*, vol. 14, no. 4, pp. 295–305, 2022, doi: 10.1504/IJMEI.2022.123921.
- [13] M. F. Maulana and M. Defriani, "Logistic model tree and decision tree J48 algorithms for predicting the length of study period," *PIKSEL: Penelitian Ilmu ...*, 2020, [Online]. Available: <https://jurnal.unismabekasi.ac.id/index.php/piksel/article/view/2018>
- [14] M. Yasir, A. M. Karim, S. K. Malik, A. A. Bajaffer, and E. I. Azhar, "Application of Decision-Tree-Based Machine Learning Algorithms for Prediction of Antimicrobial Resistance," 2022, *mdpi.com*. doi: 10.3390/antibiotics11111593.
- [15] P. Kimawaha *et al.*, "Establishment of a potential serum biomarker panel for the diagnosis and prognosis of cholangiocarcinoma using decision tree algorithms," 2021, *mdpi.com*. doi: 10.3390/diagnostics11040589.
- [16] H. Z. Dizaji, H. Bahrami, N. Monjezi, and ..., "Modeling of the variables that influence sugarcane yield using C5. 0 and QUEST decision tree algorithms," 2019, *jame.um.ac.ir*. [Online]. Available: [https://jame.um.ac.ir/article\\_33945\\_en.html](https://jame.um.ac.ir/article_33945_en.html)



**Inna Sabilly Karima S.Kom., M.Kom.**

Ketertarikan penulis terhadap ilmu computer khususnya data science mulai tahun 2006 silam. Hal tersebut membuat penulis memilih untuk melanjutkan studi pada jurusan Teknik Informatika dan berhasil lulus pada tahun 2010. Penulis kemudian melanjutkan Pendidikan S2 ke IPB university jurusan Ilmu Komputer pada tahun 2012 dan berhasil lulus dalam waktu dua tahun di 2014. Pada tahun 2023 samapi sekarang penulis melanjutkan studi program Doktor program studi Ilmu Komputer di IPB University. Penulis memiliki kepakaran dibidang Data Science. Dan untuk mewujudkan karir sebagai dosen profesional, penulis pun aktif sebagai peneliti dibidang kepakarannya tersebut. Beberapa penelitian yang telah dilakukan didanai oleh internal perguruan tinggi dan juga Kemenristek DIKTI. Selain peneliti, penulis juga aktif menulis buku dengan harapan dapat memberikan kontribusi positif bagi bangsa dan negara yang sangat tercinta ini.