



Department of Digital Business

Journal of Artificial Intelligence and Digital Business (RIGGS)

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 5 No. 2 (2026) pp: 22738- 22744

P-ISSN: 2963-9298, e-ISSN: 2963-914X

Klasifikasi Penyakit Jantung Berdasarkan Algoritma Random Forest pada Dataset Heart Disease

Yeheskiel Ravena Damanik¹, Albert Adolf Angie MS², Esther Laura Rumahorbo³, Elisabeth Jessica Sitanggang⁴
Program Studi Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas HKBP Nommensen
Pematangsiantar

yeheskielravenadamanik@gmail.com¹, albertmanalu121102@gmail.com², esther.rumahorbo17@gmail.com³,
elisabethjessicastg@gmail.com⁴

Abstrak

Penyakit jantung merupakan salah satu penyebab utama kematian di dunia dan menjadi masalah kesehatan yang memerlukan deteksi dini secara cepat serta akurat untuk mengurangi risiko komplikasi maupun kematian. Seiring berkembangnya teknologi, penerapan data mining dalam bidang kesehatan dapat dimanfaatkan untuk membantu proses klasifikasi penyakit berdasarkan data medis pasien. Penelitian ini bertujuan untuk mengimplementasikan algoritma **Random Forest** dalam mengklasifikasikan penyakit jantung menggunakan dataset **Heart Disease** yang terdiri dari 918 data dengan berbagai atribut klinis yang berhubungan dengan kondisi pasien. Proses pengolahan data dilakukan menggunakan perangkat lunak **Orange Data Mining** melalui beberapa tahapan, yaitu pemuatan dataset, seleksi atribut, pembagian data menjadi data latih dan data uji, pembangunan model klasifikasi, serta evaluasi kinerja model menggunakan **Test & Score** dan **Confusion Matrix**. Evaluasi dilakukan untuk mengetahui kemampuan model dalam mengklasifikasikan pasien yang terindikasi maupun tidak terindikasi penyakit jantung berdasarkan nilai prediksi yang dihasilkan. Hasil penelitian menunjukkan bahwa algoritma Random Forest mampu memberikan performa klasifikasi yang baik dengan tingkat akurasi yang tinggi serta mampu mengenali kelas positif dan kelas negatif secara efektif berdasarkan hasil evaluasi. Hal ini menunjukkan bahwa algoritma Random Forest memiliki potensi yang baik untuk diterapkan sebagai sistem pendukung keputusan dalam proses identifikasi awal penyakit jantung. Penelitian ini diharapkan dapat menjadi referensi bagi penelitian selanjutnya dalam pengembangan model klasifikasi berbasis data mining pada bidang kesehatan guna meningkatkan efektivitas dan efisiensi proses deteksi penyakit jantung.

Kata Kunci: Data Mining, Random Forest, Klasifikasi Penyakit Jantung, Orange Data Mining, Heart Disease.

1. Latar Belakang

Penyakit jantung merupakan salah satu penyebab utama kematian di dunia dan menjadi masalah kesehatan yang terus meningkat setiap tahunnya. Tingginya angka penderita penyakit jantung menyebabkan kebutuhan terhadap sistem pendukung diagnosis yang cepat, akurat, dan efisien semakin diperlukan. Kemajuan teknologi informasi, khususnya pada bidang **Machine Learning** dan **Data Mining**, memberikan peluang dalam membantu tenaga medis melakukan identifikasi dini terhadap risiko penyakit jantung melalui analisis data pasien. Dengan memanfaatkan data historis pasien, model klasifikasi dapat dibangun untuk menghasilkan prediksi yang lebih cepat sehingga mendukung proses pengambilan keputusan klinis [1], [2].

Data Mining merupakan proses menemukan pola atau informasi yang tersembunyi dari sekumpulan data berukuran besar menggunakan teknik statistik, pembelajaran mesin, maupun kecerdasan buatan. Salah satu teknik Data Mining yang banyak diterapkan dalam bidang kesehatan adalah klasifikasi, yaitu proses mengelompokkan data berdasarkan karakteristik tertentu untuk menghasilkan prediksi terhadap data baru. Pada kasus penyakit jantung, atribut seperti usia, jenis kelamin, tekanan darah, kadar kolesterol, kadar gula darah puasa, hasil elektrokardiogram, detak jantung maksimum, serta riwayat nyeri dada merupakan faktor yang berpengaruh terhadap kondisi kesehatan pasien [3], [4].

Berbagai algoritma Machine Learning telah diterapkan dalam klasifikasi penyakit jantung, di antaranya Logistic Regression, Decision Tree, Support Vector Machine (SVM), K-Nearest Neighbor (KNN), dan Random Forest. Di antara algoritma tersebut, **Random Forest** menjadi salah satu metode yang banyak digunakan karena mampu menghasilkan performa klasifikasi yang tinggi melalui pembentukan sejumlah *decision tree* yang kemudian

digabungkan menggunakan mekanisme *majority voting*. Pendekatan tersebut mampu meningkatkan stabilitas model, mengurangi risiko *overfitting*, serta memberikan hasil klasifikasi yang lebih baik dibandingkan penggunaan satu pohon keputusan saja [1], [5].

Penelitian yang dilakukan oleh **Felix et al.** menunjukkan bahwa algoritma Random Forest mampu memberikan performa yang baik dalam memprediksi penyakit jantung menggunakan pendekatan Data Mining. Hasil penelitian tersebut menunjukkan bahwa Random Forest dapat dimanfaatkan sebagai metode klasifikasi yang efektif dalam mendukung proses diagnosis penyakit jantung [3]. Penelitian lain yang dilakukan oleh **Khofiyah dan Kuncoro** mengombinasikan algoritma Random Forest dengan teknik **Synthetic Minority Oversampling Technique (SMOTE)** sehingga mampu meningkatkan kemampuan model dalam mengatasi ketidakseimbangan data dan menghasilkan performa klasifikasi yang lebih baik [2]. Selain itu, penelitian **Nasution et al.** membandingkan beberapa algoritma Machine Learning seperti Logistic Regression, Random Forest, SVM, dan KNN, dimana Random Forest menunjukkan performa yang kompetitif pada proses klasifikasi penyakit jantung [5].

Perkembangan penelitian terbaru juga menunjukkan bahwa optimasi terhadap algoritma Random Forest mampu meningkatkan kualitas prediksi penyakit jantung. Narasimhan dan Victor mengembangkan pendekatan optimasi berbasis **Metaheuristic Optimization** yang dikombinasikan dengan Random Forest sehingga menghasilkan peningkatan performa klasifikasi pada dataset penyakit jantung [1]. Penelitian lain oleh **Al-Waeli et al.** juga menunjukkan bahwa Random Forest memiliki kemampuan yang baik dalam melakukan klasifikasi pada **UCI Heart Disease Dataset**, sehingga algoritma ini layak digunakan sebagai model prediksi penyakit jantung [8].

Selain algoritma yang digunakan, perangkat lunak yang mendukung proses analisis data juga menjadi faktor penting dalam penelitian Data Mining. Salah satu perangkat lunak yang banyak digunakan adalah **Orange Data Mining**, karena menyediakan lingkungan visual (*visual programming*) yang memudahkan pengguna dalam melakukan prapemrosesan data, pembangunan model klasifikasi, hingga evaluasi performa model tanpa memerlukan penulisan kode program secara kompleks. Orange Data Mining juga menyediakan berbagai metode evaluasi seperti **Test & Score**, **Confusion Matrix**, serta **Receiver Operating Characteristic (ROC)** yang mampu memberikan informasi mengenai kemampuan model dalam melakukan klasifikasi [6], [7].

Berdasarkan penelitian-penelitian sebelumnya, masih diperlukan implementasi algoritma Random Forest menggunakan **Heart Disease Dataset** pada Orange Data Mining untuk mengetahui kemampuan model dalam mengklasifikasikan pasien yang menderita maupun tidak menderita penyakit jantung berdasarkan atribut-atribut klinis yang tersedia. Dataset yang digunakan dalam penelitian ini terdiri dari **918 data** dengan **12 atribut**, sedangkan evaluasi model dilakukan menggunakan **Accuracy**, **Precision**, **Recall**, **F1-Score**, **Area Under Curve (AUC)**, serta **Confusion Matrix**. Hasil penelitian diharapkan dapat memberikan gambaran mengenai performa algoritma Random Forest dalam klasifikasi penyakit jantung sekaligus menjadi referensi bagi penelitian selanjutnya pada bidang Data Mining di sektor kesehatan [1], [3], [9].

2. Metode Penelitian

Penelitian ini menggunakan metode **Knowledge Discovery in Database (KDD)** dengan pendekatan klasifikasi menggunakan algoritma **Random Forest**. Proses penelitian dilakukan menggunakan perangkat lunak **Orange Data Mining** untuk membangun model klasifikasi penyakit jantung berdasarkan dataset *Heart Disease*. Tahapan penelitian meliputi pengumpulan data, prapemrosesan data, pembangunan model klasifikasi, dan evaluasi kinerja model. Evaluasi dilakukan menggunakan *Test & Score*, *Confusion Matrix*, dan *ROC Analysis* untuk mengetahui kemampuan algoritma dalam mengklasifikasikan data penyakit jantung.

2.1 Dataset Penelitian

Dataset yang digunakan dalam penelitian ini adalah **Heart Disease Dataset** yang terdiri dari **918 data** pasien dengan **12 atribut**, yaitu *Age*, *Sex*, *ChestPainType*, *RestingBP*, *Cholesterol*, *FastingBS*, *RestingECG*, *MaxHR*, *ExerciseAngina*, *Oldpeak*, *ST_Slope*, dan *HeartDisease*. Atribut **HeartDisease** digunakan sebagai kelas (*class label*) yang menunjukkan kondisi pasien, yaitu **0** untuk tidak menderita penyakit jantung dan **1** untuk menderita penyakit jantung.

Sebagai ilustrasi, sebagian data yang digunakan pada penelitian ditampilkan pada Tabel 1.

DOI: <https://doi.org/10.31004/riggs.v5i2.12186>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

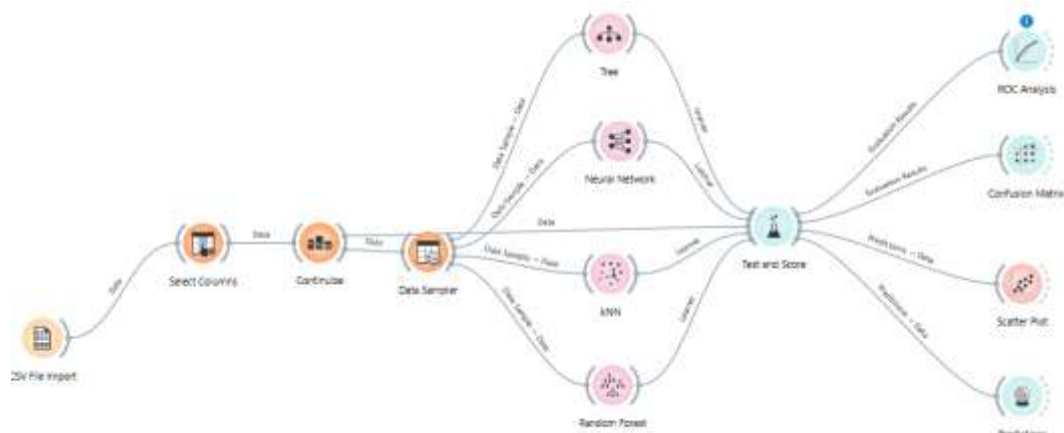
Tabel 1. Contoh Dataset Heart Disease

Age	Sex	ChestPainType	RestingBP	Cholesterol	MaxHR	HeartDisease
40	M	ATA	140	289	172	0
49	F	NAP	160	180	156	1
37	M	ATA	130	283	98	0
48	F	ASY	138	214	108	1
54	M	NAP	150	195	122	0

2.2 Tahapan Penelitian

Tahapan penelitian dilakukan secara sistematis menggunakan Orange Data Mining. Dataset terlebih dahulu dimuat ke dalam aplikasi menggunakan widget **File**, kemudian dilakukan pemilihan atribut menggunakan **Select Columns** untuk menentukan atribut prediktor dan atribut target. Selanjutnya data diproses menggunakan **Data Sampler** untuk membagi data menjadi data pelatihan (*training data*) dan data pengujian (*testing data*). Model klasifikasi kemudian dibangun menggunakan algoritma **Random Forest**. Setelah model terbentuk, dilakukan evaluasi menggunakan widget **Test & Score**, sedangkan hasil klasifikasi dianalisis lebih lanjut menggunakan **Confusion Matrix** dan **ROC Analysis**. Alur penelitian yang digunakan ditunjukkan pada Gambar 1.

Gambar 1. Workflow Penelitian Menggunakan Orange Data Mining



2.3

Prapemrosesan Data

Tahap prapemrosesan dilakukan untuk memastikan data siap digunakan dalam proses klasifikasi. Pada penelitian ini dilakukan pemilihan atribut menggunakan widget **Select Columns**, sehingga atribut **HeartDisease** ditetapkan sebagai variabel target (*target variable*), sedangkan atribut lainnya digunakan sebagai variabel prediktor. Selanjutnya dilakukan pembagian data menggunakan **Data Sampler** sehingga data pelatihan dan data pengujian dapat digunakan dalam proses pembentukan dan pengujian model klasifikasi.

2.4 Pembangunan Model

Model klasifikasi dibangun menggunakan algoritma **Random Forest** yang tersedia pada Orange Data Mining. Algoritma ini membentuk sejumlah pohon keputusan (*decision tree*) dari data pelatihan, kemudian hasil prediksi setiap pohon digabungkan menggunakan mekanisme *majority voting* untuk menghasilkan keputusan akhir.

Pendekatan tersebut bertujuan meningkatkan akurasi klasifikasi dan mengurangi kemungkinan terjadinya *overfitting* pada model.

2.5 Evaluasi Model

Evaluasi model dilakukan menggunakan widget **Test & Score** pada Orange Data Mining. Pengujian menghasilkan beberapa metrik evaluasi, yaitu **Accuracy**, **Area Under Curve (AUC)**, **Precision**, **Recall**, dan **F1-Score**. Selain itu, digunakan **Confusion Matrix** untuk mengetahui jumlah prediksi benar dan prediksi salah pada setiap kelas, serta **ROC Analysis** untuk mengevaluasi kemampuan model dalam membedakan kelas positif dan kelas negatif. Hasil evaluasi tersebut digunakan sebagai dasar dalam menganalisis performa algoritma Random Forest pada proses klasifikasi penyakit jantung.

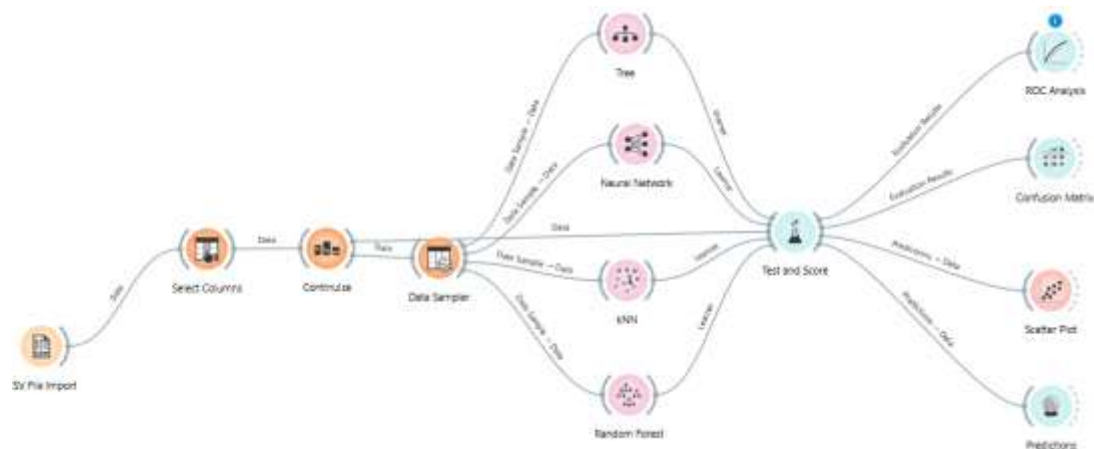
3. Hasil dan Diskusi

Penelitian ini bertujuan untuk mengetahui performa algoritma Random Forest dalam melakukan klasifikasi penyakit jantung menggunakan Heart Disease Dataset melalui perangkat lunak Orange Data Mining. Evaluasi dilakukan menggunakan metode 5-Fold Cross Validation dengan metrik Area Under Curve (AUC), Classification Accuracy (CA), Precision, Recall, F1-Score, Matthews Correlation Coefficient (MCC), Confusion Matrix, dan Receiver Operating Characteristic (ROC).

3.1 Hasil Implementasi Model

Proses klasifikasi diawali dengan mengimpor dataset ke dalam Orange Data Mining, kemudian dilakukan pemilihan atribut menggunakan widget **Select Columns**, transformasi data menggunakan **Continue**, pembagian data menggunakan **Data Sampler**, serta proses pelatihan model menggunakan empat algoritma klasifikasi yaitu **k-Nearest Neighbor (kNN)**, **Decision Tree**, **Neural Network**, dan **Random Forest**. Selanjutnya seluruh model dievaluasi menggunakan widget **Test and Score**, **ROC Analysis**, dan **Confusion Matrix**.

Gambar 2. Workflow Klasifikasi Menggunakan Orange Data Mining



Workflow pada Gambar 2 menunjukkan bahwa seluruh algoritma memperoleh dataset pelatihan yang sama sehingga proses evaluasi dilakukan pada kondisi yang identik. Dengan demikian, hasil evaluasi setiap algoritma dapat dibandingkan secara objektif.

3.2 Hasil Evaluasi Model

Hasil evaluasi model menggunakan widget **Test and Score** ditunjukkan pada Tabel 2.

Tabel 2. Hasil Evaluasi Algoritma Klasifikasi

Algoritma	AUC	Accuracy	F1-Score	Precision	Recall	MCC
Neural Network	0,920	0,853	0,853	0,853	0,853	0,702

Decision Tree	0,782	0,794	0,794	0,795	0,794	0,584
Random Forest	0,915	0,849	0,848	0,848	0,849	0,693

Berdasarkan Tabel 2, algoritma **Random Forest** memperoleh nilai **Accuracy sebesar 84,9%**, **AUC sebesar 0,915**, **Precision sebesar 84,8%**, **Recall sebesar 84,9%**, **F1-Score sebesar 84,8%**, dan **MCC sebesar 0,693**. Nilai tersebut menunjukkan bahwa Random Forest memiliki kemampuan yang baik dalam mengklasifikasikan data pasien berdasarkan atribut yang tersedia.

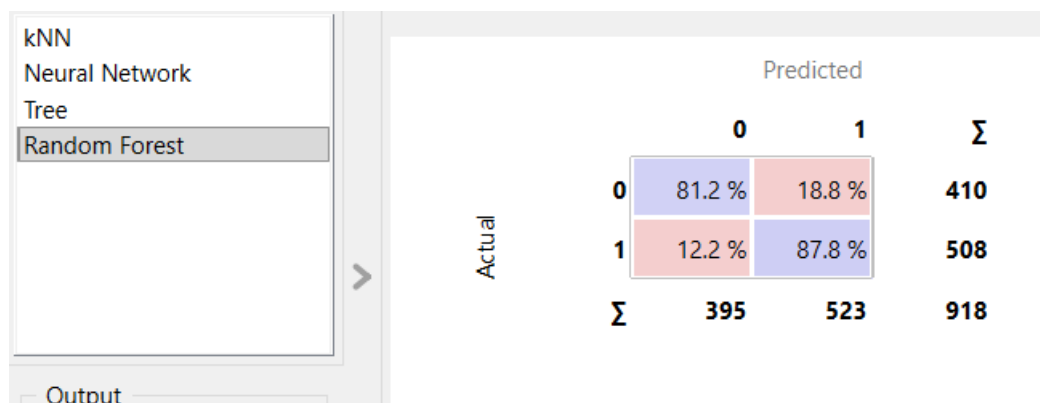
Apabila dibandingkan dengan Decision Tree, Random Forest menghasilkan peningkatan performa pada seluruh metrik evaluasi. Hal ini menunjukkan bahwa pendekatan **ensemble learning** yang digunakan Random Forest mampu meningkatkan kemampuan generalisasi model dibandingkan penggunaan satu pohon keputusan.

Sementara itu, Neural Network memperoleh nilai Accuracy sedikit lebih tinggi dibandingkan Random Forest. Namun demikian, perbedaan performa kedua algoritma relatif kecil sehingga Random Forest masih menjadi salah satu metode klasifikasi yang kompetitif pada dataset penyakit jantung.

3.3 Analisis Confusion Matrix

Evaluasi lebih lanjut dilakukan menggunakan **Confusion Matrix** untuk mengetahui kemampuan model dalam mengklasifikasikan masing-masing kelas.

Gambar 3. Confusion Matrix Algoritma Random Forest



Berdasarkan Gambar 3 diketahui bahwa dataset terdiri atas **410 data** pada kelas **tidak menderita penyakit jantung (kelas 0)** dan **508 data** pada kelas **menderita penyakit jantung (kelas 1)**.

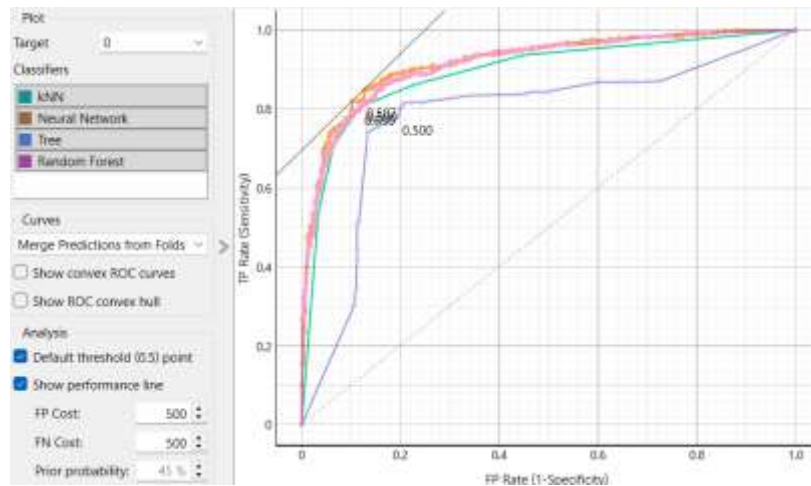
Model Random Forest berhasil mengklasifikasikan **81,2%** data pada kelas **0** secara benar, sedangkan **18,8%** sisanya salah diklasifikasikan sebagai kelas **1**. Pada kelas **1**, model berhasil mengidentifikasi **87,8%** data secara benar dan hanya **12,2%** data yang salah diklasifikasikan menjadi kelas **0**.

Hasil tersebut menunjukkan bahwa Random Forest memiliki kemampuan yang lebih baik dalam mendeteksi pasien yang benar-benar menderita penyakit jantung dibandingkan dalam mengidentifikasi pasien yang tidak menderita penyakit jantung. Tingginya tingkat keberhasilan pada kelas positif menjadi nilai penting karena kesalahan dalam mendeteksi pasien yang berisiko dapat berdampak terhadap proses diagnosis medis.

3.4 Analisis ROC Curve

Evaluasi performa model juga dilakukan menggunakan **Receiver Operating Characteristic (ROC)**.

Gambar 4. Kurva ROC Seluruh Algoritma



Kurva ROC menunjukkan hubungan antara **True Positive Rate (Sensitivity)** dan **False Positive Rate (1-Specificity)** pada berbagai nilai ambang klasifikasi. Berdasarkan Gambar 4 terlihat bahwa kurva **Random Forest** berada jauh di atas garis diagonal acak dan mendekati sudut kiri atas grafik. Hal tersebut menunjukkan bahwa model memiliki kemampuan diskriminasi yang baik dalam membedakan pasien yang menderita dan tidak menderita penyakit jantung.

Nilai **AUC sebesar 0,915** semakin memperkuat bahwa Random Forest memiliki kualitas klasifikasi yang tinggi. Menurut interpretasi nilai AUC, model dengan nilai di atas **0,90** termasuk dalam kategori **excellent classification**, sehingga model yang dibangun pada penelitian ini memiliki performa yang sangat baik dalam melakukan klasifikasi penyakit jantung.

3.5 Diskusi

Hasil penelitian menunjukkan bahwa algoritma **Random Forest** mampu memberikan performa klasifikasi yang baik pada **Heart Disease Dataset** dengan nilai **Accuracy sebesar 84,9%** dan **AUC sebesar 0,915**. Nilai tersebut menunjukkan bahwa metode Random Forest efektif dalam mempelajari pola hubungan antar atribut seperti usia, tekanan darah, kadar kolesterol, detak jantung maksimum, dan karakteristik klinis lainnya.

Performa yang diperoleh sejalan dengan penelitian **Felix et al. [3]** yang melaporkan bahwa Random Forest mampu menghasilkan performa klasifikasi yang tinggi pada prediksi penyakit jantung menggunakan pendekatan Data Mining. Hasil penelitian ini juga mendukung penelitian **Khofiyah dan Kuncoro [2]** yang menunjukkan bahwa Random Forest memiliki kemampuan klasifikasi yang baik bahkan ketika dikombinasikan dengan teknik penyeimbangan data SMOTE.

Selain itu, penelitian **Nasution et al. [5]** menunjukkan bahwa Random Forest memiliki performa yang kompetitif dibandingkan Logistic Regression, Support Vector Machine, dan K-Nearest Neighbor. Temuan tersebut konsisten dengan hasil penelitian ini, dimana Random Forest mampu menghasilkan performa yang lebih baik dibandingkan Decision Tree dan memiliki performa yang mendekati Neural Network.

Kemampuan Random Forest dalam membangun banyak pohon keputusan (*multiple decision tree*) serta menggabungkan hasil prediksi menggunakan mekanisme **majority voting** menyebabkan model menjadi lebih stabil dan mampu mengurangi risiko **overfitting**. Karakteristik tersebut menjadikan Random Forest sebagai salah satu algoritma yang sesuai digunakan pada dataset kesehatan yang memiliki atribut heterogen dan jumlah data yang relatif besar.

Secara keseluruhan, hasil penelitian menunjukkan bahwa Random Forest merupakan metode yang layak digunakan sebagai model klasifikasi penyakit jantung menggunakan Orange Data Mining. Model yang dihasilkan mampu memberikan performa yang baik dan berpotensi digunakan sebagai sistem pendukung keputusan dalam proses deteksi dini penyakit jantung.

4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, algoritma Random Forest mampu digunakan untuk memprediksi penyakit jantung menggunakan Heart Disease Dataset melalui perangkat lunak Orange Data Mining. Hasil evaluasi model menunjukkan bahwa Random Forest memperoleh nilai Accuracy sebesar 84,9%, Area Under Curve (AUC) sebesar 0,915, Precision sebesar 84,8%, Recall sebesar 84,9%, F1-Score sebesar 84,8%, dan Matthews Correlation Coefficient (MCC) sebesar 0,693. Selain itu, hasil Confusion Matrix menunjukkan bahwa model mampu mengklasifikasikan 81,2% data pada kelas tidak menderita penyakit jantung dan 87,8% data pada kelas menderita penyakit jantung dengan benar. Nilai AUC yang berada di atas 0,90 menunjukkan bahwa model memiliki kemampuan klasifikasi yang sangat baik dalam membedakan pasien yang berisiko dan tidak berisiko mengalami penyakit jantung. Hasil penelitian ini menunjukkan bahwa algoritma Random Forest memiliki performa yang baik dan dapat dimanfaatkan sebagai salah satu metode klasifikasi dalam mendukung proses deteksi dini penyakit jantung. Penggunaan Orange Data Mining juga mempermudah proses prapemrosesan data, pembangunan model, hingga evaluasi hasil sehingga dapat menjadi alternatif perangkat lunak yang efektif dalam penelitian berbasis data mining. Untuk penelitian selanjutnya, disarankan melakukan optimasi parameter Random Forest, menerapkan teknik penyeimbangan data seperti SMOTE, serta membandingkan performa model dengan algoritma ensemble learning lainnya, misalnya XGBoost, LightGBM, atau AdaBoost, menggunakan dataset yang lebih besar dan lebih beragam sehingga diperoleh model prediksi yang memiliki tingkat akurasi dan kemampuan generalisasi yang lebih tinggi.

Referensi

- [1] G. Narasimhan and A. Victor, "A Hybrid Approach with Metaheuristic Optimization and Random Forest in Improving Heart Disease Prediction," *Scientific Reports*, vol. 15, 2025. doi: 10.1038/s41598-024-73867-x.
- [2] S. N. Khofiyah and A. P. Kuncoro, "Heart Disease Prediction Using Random Forest Algorithm and Synthetic Minority Oversampling Technique (SMOTE)," *Komputa: Jurnal Ilmiah Komputer dan Informatika*, vol. 14, no. 1, 2025. doi: 10.34010/komputa.v14i1.14678.
- [3] F. Felix, D. Sitanggang, Y. Laia, A. Amalia, M. Radhi, and E. S. Barus, "Application of Data Mining Using the Random Forest Method to Predict Heart Disease," *Jurnal Sistem Informasi dan Ilmu Komputer*, vol. 7, no. 2, pp. 35–47, 2024. doi: 10.34012/jurnal sistem informasi dan ilmu komputer.v7i2.4801.
- [4] A. Alariyibi, M. El-Jarai, and A. Maatuk, "Evaluating the Accuracy of Classification Algorithms for Detecting Heart Disease Risk," *International Journal of Advanced Computer Science and Applications*, 2023.
- [5] N. Nasution, M. A. Hasan, and F. B. Nasution, "Predicting Heart Disease Using Machine Learning: An Evaluation of Logistic Regression, Random Forest, SVM, and KNN Models on the UCI Heart Disease Dataset," *IT Journal Research and Development*, vol. 9, no. 2, 2025. doi: 10.25299/itjrd.2025.17941.
- [6] M. Al Amin, A. Rahman, and M. Iqbal, "Performance Evaluation of Machine Learning Algorithms Using Orange Data Mining," *CogITO Smart Journal*, vol. 11, no. 1, 2025. doi: 10.31154/cogito.v11i1.782.80-90.
- [7] A. Soumya and A. Itten, "Heart Disease Prediction Using Random Forest Algorithm: A Comprehensive Analysis," *International Journal of Engineering Research and Science & Technology*, vol. 21, no. 2, 2025. doi: 10.62643/ijerst.2025.v21.i2.p119-123.
- [8] N. K. Al-Waeli et al., "Early Detection of Heart Disease Using a Random Forest Classifier on the UCI Heart Disease Dataset," *International Journal of Computer Applications*, vol. 187, no. 67, 2025. doi: 10.5120/ijca2025926161.
- [9] S. Rimal et al., "Artificial Intelligence-Based Heart Disease Prediction Using Ensemble Learning Techniques," *Scientific Reports*, vol. 15, 2025. doi: 10.1038/s41598-025-93675-1.
- [10] G. Narasimhan and A. Victor, "A Hybrid Approach with Metaheuristic Optimization and Random Forest in Improving Heart Disease Prediction," *Scientific Reports*, vol. 15, 2025.
- [11] F. Felix et al., "Application of Data Mining Using the Random Forest Method to Predict Heart Disease," *Jurnal Sistem Informasi dan Ilmu Komputer*, vol. 7, no. 2, 2024.
- [12] N. Nasution et al., "Predicting Heart Disease Using Machine Learning," *IT Journal Research and Development*, vol. 9, no. 2, 2025.
- [13] S. N. Khofiyah and A. P. Kuncoro, "Heart Disease Prediction Using Random Forest Algorithm and SMOTE," *Komputa*, vol. 14, no. 1, 2025.
- [14] N. K. Al-Waeli et al., "Early Detection of Heart Disease Using a Random Forest Classifier," *International Journal of Computer Applications*, vol. 187, no. 67, 2025.
- [15] S. Rimal et al., "Artificial Intelligence-Based Heart Disease Prediction Using Ensemble Learning Techniques," *Scientific Reports*, vol. 15, 2025.