



Pendekatan *Ensemble Learning* Untuk Klasifikasi Penyakit Jantung Koroner

Rangga Roris^{1*}, M Farrel Aziz², M Akmal Khatami³, Wahyu Rifia Rizki⁴

^{1,2,3,4}Fakultas Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

Email : ^{1*}rangga.roris@gmail.com, ²farrel.m197@gmail.com, ³akmalkhatami2318@gmail.com,

⁴wahyurizki927@gmail.com

(* : coresponding author)

Abstrak - Penyakit arteri koroner merupakan salah satu penyebab kematian utama di seluruh dunia, sehingga sangat penting untuk mengembangkan metode klasifikasi yang akurat untuk deteksi dini penyakit ini. Berbagai pendekatan berbasis pembelajaran mesin telah digunakan untuk mengklasifikasikan penyakit jantung koroner, namun metode tradisional seringkali kurang akurat dan konsisten. Penelitian ini menggunakan pendekatan pembelajaran ansambel yang menggabungkan algoritma klasifikasi yang berbeda untuk meningkatkan kinerja dan akurasi pengenalan. Tujuan utama penelitian ini adalah untuk mengevaluasi efektivitas metode pembelajaran ensemble dalam mengatasi kekurangan metode tunggal khususnya pada klasifikasi penyakit jantung koroner. Solusi ini dievaluasi menggunakan beberapa artikel jurnal sebagai dasar studi tinjauan literatur yang melibatkan berbagai teknik ansambel seperti hutan acak dan peningkatan gradien. Hasil penelitian ini menunjukkan bahwa metode pembelajaran ansambel secara signifikan meningkatkan akurasi klasifikasi dibandingkan metode klasifikasi tradisional dan memberikan hasil yang lebih andal dalam mendeteksi pasien yang berisiko penyakit jantung koroner.

Kata Kunci : *Ensemble Learning*, Klasifikasi Penyakit Jantung Koroner, *Random Forest*, *Gradient Boosting*.

Abstract - Coronary artery disease is one of the leading causes of death worldwide, so it is very important to develop accurate classification methods for early detection of this disease. Various machine learning-based approaches have been used to classify coronary heart disease, but traditional methods are often less accurate and consistent. This study uses an Ensemble Learning approach that combines different classification algorithms to improve performance and recognition accuracy. The main objective of this study is to evaluate the effectiveness of Ensemble Learning methods in overcoming the shortcomings of single methods, especially in the classification of coronary heart disease. This solution is evaluated using several journal articles as the basis for a literature review study involving various ensemble techniques such as Random Forest and Gradient Boosting. The results of this study indicate that Ensemble Learning methods significantly improve classification accuracy compared to traditional classification methods and provide more reliable results in detecting patients at risk of coronary heart disease.

Keywords: *Ensemble Learning*, *Coronary Heart Disease Classification*, *Random Forest*, *Gradient Boosting*.

1. PENDAHULUAN

Penyakit jantung koroner menjadi salah satu ancaman kesehatan yang paling serius secara global, bertanggung jawab atas jutaan kematian setiap tahun. Deteksi dini penyakit ini sangat penting untuk memberikan penanganan yang tepat, sehingga mengurangi risiko komplikasi atau kematian. Namun, klasifikasi penyakit jantung koroner yang akurat merupakan tantangan tersendiri, mengingat kompleksitas faktor risiko yang terlibat, seperti riwayat keluarga, pola makan, gaya hidup, serta kondisi kesehatan lainnya. Di dunia medis, metode klasifikasi tradisional seperti Support Vector Machine (SVM), Decision Tree, dan Logistic Regression sering digunakan, namun metode-metode ini memiliki keterbatasan dalam hal akurasi dan generalisasi pada data baru.

Seiring perkembangan teknologi dalam bidang kecerdasan buatan dan machine learning, pendekatan *Ensemble Learning* mulai dilirik sebagai alternatif untuk meningkatkan akurasi dan keandalan sistem klasifikasi. *Ensemble Learning*, yang merupakan kombinasi dari beberapa algoritma, diharapkan mampu mengatasi kelemahan metode tunggal dengan cara mengintegrasikan keputusan dari berbagai model. Misalnya, algoritma seperti *Random Forest* dan *Gradient Boosting* sering digunakan dalam klasifikasi karena mampu menangkap pola yang lebih kompleks dalam data medis. Dengan demikian, penelitian ini bertujuan untuk menganalisis dan mengevaluasi efektivitas metode *Ensemble Learning* melalui studi literature review terhadap beberapa jurnal yang berkaitan



dengan klasifikasi penyakit jantung koroner. Harapannya, hasil dari penelitian ini dapat memberikan wawasan baru dalam penerapan metode klasifikasi yang lebih akurat dan andal bagi kalangan medis.

2. METODE

Penelitian ini menggunakan pendekatan Research Literatur Review (SLR) untuk mengumpulkan dan menganalisis informasi dari sumber-sumber terkait klasifikasi penyakit jantung koroner dengan menggunakan metode pembelajaran ensemble. Fokus utama penelitian ini adalah mengevaluasi kinerja metode ensemble seperti *Random Forest* dan peningkatan gradien dalam meningkatkan akurasi deteksi penyakit jantung. Langkah-langkah yang dilakukan dalam proses ini adalah:

1. Pengumpulan data: Teknik pembelajaran ensemble digunakan untuk mengumpulkan informasi dari lima jurnal besar terkait klasifikasi penyakit arteri koroner. Jurnal dipilih berdasarkan kriteria inklusi berikut: – Jurnal yang diterbitkan dalam lima tahun terakhir untuk memastikan data relevan dan terkini. – Berfokus pada teknik pembelajaran ansambel, khususnya penggunaan algoritma seperti hutan acak dan peningkatan gradien. – Memberikan informasi tentang performa model, termasuk akurasi, sensitivitas, dan spesifisitas dalam mengklasifikasikan penyakit jantung.Rephrase.
2. Analisis literatur : Setelah pengumpulan data dilakukan analisis lebih lanjut dengan cara sebagai berikut. - Mengelompokkan hasil setiap jurnal berdasarkan metode ansambel yang digunakan, tujuan penelitian, dan hasil akhir yang diperoleh. - Membuat tabel yang merangkum informasi utama setiap penelitian, seperti nama peneliti, metode ansambel yang dibahas, tujuan dan hasil yang dicapai. - Mengidentifikasi pola atau tren penerapan metode pembelajaran ensemble untuk mengklasifikasikan penyakit jantung, termasuk tantangan dalam penerapan metode tersebut.
3. Penilaian Kualitas Jurnal: Setiap jurnal yang digunakan dalam penelitian ini dievaluasi berdasarkan metodologi penelitian, jumlah data yang digunakan, dan validitas hasil. Langkah ini penting untuk memastikan bahwa data yang diolah dalam studi tinjauan pustaka ini berasal dari publikasi yang dapat dipercaya.

3. HASIL DAN PEMBAHASAN

Dalam penelitian ini, kami melakukan analisis terhadap lima jurnal yang menyelidiki klasifikasi penyakit arteri koroner dengan menggunakan pendekatan pembelajaran ensemble.

Tabel berikut merangkum pokok-pokok penelitian yang dianalisis, meliputi informasi mengenai peneliti, metode yang digunakan, tujuan penelitian, dan hasil yang diperoleh :

No	Nama Peneliti dan Tahun	Metode Yang Dibahas	Tujuan Penelitian	Hasil Yang Didapat
1	Mohan, S., Thirumalai, C., & Srivastava, G.	<i>Hybrid Ensemble</i> menggunakan <i>Random Forest</i> dan <i>Gradient Boosting</i>	Memprediksi penyakit jantung dengan akurasi tinggi	Akurasi 93.4% pada dataset UCI Heart Disease
2	Singh, A. & Kumar, R.	<i>Random Forest</i> dan <i>XGBoost</i>	Meningkatkan akurasi prediksi penyakit jantung dengan pendekatan ensemble	Akurasi 92%

3	Al Mehedi Hasan, M., Shin, J., & Das, U.	<i>Random Forest</i> dan <i>Support Vector Machine</i>	Menentukan fitur prognostik untuk prediksi penyakit jantung	Akurasi 89%
4	Nur Halizah Alfajr & Sofi Defiyanti	<i>Random Forest</i>	memprediksi penyakit jantung menggunakan metode <i>Random Forest</i> dan Principal Component Analysis (PCA)	akurasi 98,23%, dengan presisi 100%, perolehan 96%, dan skor F1 98%.
5	D. H. Depari, Y. Widiastiwi, and M. M. Santoni	<i>Random Forest</i>	Membandingkan Decision Tree, Naive Bayes, dan <i>Random Forest</i> untuk klasifikasi penyakit jantung	Akurasi 94%

Beberapa penelitian tentang prediksi penyakit jantung menggunakan metode pembelajaran ensemble berbeda dalam penerapan teknik untuk meningkatkan akurasi model. Sebuah studi oleh Mohan, S., Thirumarai, C., dan Srivastava, G. (2019) menemukan bahwa pendekatan ansambel hibrid yang menggabungkan hutan acak dan peningkatan gradien mencapai akurasi 93,4% ketika diterapkan pada kumpulan data Kardiologi UCI . Dengan menggunakan teknik ini, Anda dapat meningkatkan performa model dengan memanfaatkan kekuatan kedua metode untuk meningkatkan prediksi. Penelitian lain oleh Singh, A. dan Kumar, R. (2020) juga menggunakan *Random Forest* dan XGBoost dalam pendekatan ansambel. Penelitian ini mampu mencapai akurasi sebesar 92%, menunjukkan bahwa menggabungkan algoritma *Random Forest* dengan XGBoost dapat memanfaatkan kekuatan peningkatan gradien yang terkandung dalam XGBoost untuk meningkatkan akurasi prediksi penyakit jantung. Selain itu, Al Mehedi Hasan, M., Shin, J., Das, U. (2021) menggunakan *Random Forest* and support vector machine (SVM) dengan tujuan menentukan fitur prognostik yang digunakan untuk prediksi dan prediksi penyakit jantung. Model yang digunakan berhasil mencapai akurasi sebesar 89%. Hal ini menunjukkan bahwa meskipun menggunakan metode yang berbeda, diperoleh hasil yang sangat baik dalam memprediksi penyakit jantung. Nur Halizah Alfajr dan Sofi Defiyanti (2024) menggunakan *Random Forest* dengan analisis komponen utama (PCA) untuk memprediksi penyakit jantung. Hasil yang diperoleh sangat mengesankan: presisi 98,23%, presisi 100%, recall 96%, dan skor F1 98%. Penggunaan PCA untuk mengurangi dimensi data dan meningkatkan kinerja klasifikasi dengan hutan acak telah terbukti sangat efektif dalam meningkatkan akurasi prediksi. Terakhir ada penelitian Pak D. H. Depari, Y. Widiastiwi, dan M. M. Santoni (2022) membandingkan beberapa teknik pembelajaran mesin seperti *Random Forest*, pohon keputusan, dan naif Bayes untuk mengklasifikasikan penyakit jantung. Mereka menemukan bahwa hutan acak memberikan hasil yang lebih baik dengan akurasi hingga 94%, yang menegaskan efektivitas hutan acak dalam menangani kumpulan data kompleks dengan banyak fitur. Secara keseluruhan, pembelajaran ansambel telah terbukti memberikan hasil yang sangat baik dalam memprediksi penyakit jantung, terutama bila digunakan dalam kombinasi dengan metode seperti hutan acak dan peningkatan gradien. Pendekatan ini tidak hanya meningkatkan akurasi prediksi tetapi juga meningkatkan kemampuan model dalam menangani data yang tidak seimbang atau data dengan banyak variabel.



4. KESIMPULAN

Kesimpulan penelitian ini menunjukkan bahwa pendekatan pembelajaran ansambel, terutama menggunakan metode seperti *Random Forest* dan peningkatan gradien, terbukti sangat efektif dalam meningkatkan akurasi klasifikasi penyakit arteri koroner. Berbagai penelitian yang dianalisis menunjukkan bahwa penggabungan algoritma ansambel dapat mengatasi kekurangan metode klasifikasi tradisional seperti mesin vektor pendukung dan pohon keputusan, yang kurang optimal dalam hal akurasi dan generalisasi terhadap data baru. Sebuah studi oleh Mohan dkk. (2019), Singh dan Kumar (2020), Al Mehedi Hasan dkk. (2021) menunjukkan bahwa metode ensemble dapat meningkatkan hasil klasifikasi penyakit jantung dengan akurasi yang lebih tinggi, yakni masing-masing 93,4%, 92%, dan 89%.

Selain itu, penelitian yang dilakukan Nur Halizah Alfajr dan Sofi Defiyanti (2024) memberikan bukti lebih lanjut tentang efektivitas teknik *Random Forest* melalui analisis komponen utama (PCA). Akurasi metode ini sebesar 98,23%, presisi sebesar 100%, dan recall sebesar 100%. Skor F1 sebesar 96%. 98%. Secara keseluruhan, penggunaan metode ansambel untuk memprediksi penyakit jantung koroner tidak hanya meningkatkan akurasi namun juga meningkatkan kemampuan model untuk menangani data yang kompleks dan beragam. Hasil ini memberikan bukti kuat bahwa pembelajaran ensemble dapat digunakan untuk mengembangkan sistem deteksi dini penyakit jantung koroner yang lebih akurat.

REFERENCES

- Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques. 45-48, 2019
- Singh, A. & Kumar, R. (2020). Heart Disease Prediction Using Machine Learning Algorithms. 90-94, 2020
- Al Mehedi Hasan, M., Shin, J., & Das, U. (2021). Identifying Prognostic Features for Predicting Heart Failure by Using Machine Learning Algorithm. 120-124, 2021
- Nur Halizah Alfajr & Sofi Defiyanti (2024). PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN METODE *RANDOM FOREST* DAN PENERAPAN PRINCIPAL COMPONENT ANALYSIS (PCA). 50-54, 2024
- D. H. Depari, Y. Widiastiwi, and M. M. Santoni (2022). Prediksi Penyakit Jantung Menggunakan Metode-Metode Machine Learning Berbasis Ensemble – Weighted Vote. 132-136, 2022
- Perani Rosyani, saprudin saprudin, resti amalia, "Klasifikasi Citra Menggunakan Metode Random Forest dan Sequential Minimal Optimization (SMO)", JUSTIN (Jurnal Sistem dan Teknologi Informasi), vol 9, No 2 (2021)
- Anisa Maulida, Arisky Rahmatulloh, Irwan Ahussalim, Robby Alvian, Perani Rosyani, "Analisis Metode Forward Chaining pada Sistem Pakar: Systematic Literature Review", vol.1 No.04: Juni (2023)
- Zhang, S., Li, H., Luo, J., & Yao, X. (2023). Towards Inference Efficient Deep *Ensemble Learning*. Journal of Machine Learning Research, 45(3), 45-55.
- Zubair, M. D. M., & Srinivas, M. R. L. S. (2021). *Ensemble Learning* for Heart Disease Prediction. Journal of Applied Computational Biology, 112(2), 112-121.
- Jiao, L., & Liang, L. (2022). A Survey on *Ensemble Learning* under the Era of Deep Learning. Artificial Intelligence Review, 45(1), 1-15.
- Ganaie, M. A., & Yoon, H. (2021). *Ensemble Learning* Approaches for Predictive Modeling: A Review. IEEE Access, 9, 32035-32047.
- Gumelar, G., Diphan, R., Agustin, W., Christina, M., & Rosyani, P. (2023). Literatur Review Sistem Pakar Mengidentifikasi Penyakit Jantung dan Paru-paru Menggunakan Metode Forward Chaining. Jurnal Ilmu Komputer, Teknik, dan Multimedia, 1(01), 33-37.
- Pandey, M., & Singh, A. (2019). *Random Forest* and *Gradient Boosting* for Credit Scoring Prediction. International Journal of Machine Learning & Cybernetics, 10(6), 1059-1072.
- Angraini, Y., Indra, M., Khoirusofi, M., Azis, I. N., & Rosyani, P. (2023). Systematic Literature Review: Sistem Pakar Diagnosa Penyakit Gigi Menggunakan Metode Forward Chaining. BINER: Jurnal Ilmu Komputer, Teknik dan Multimedia, 1(1), 1-7.
- Liu, Y., & Wang, J. (2020). *Ensemble Learning* for Risk Assessment in Healthcare Systems. Journal of Medical Systems, 44(10), 177.