



ISSN : 2339 - 1871

JURNAL ILMIAH BETRIK

Besemah Teknologi Informasi dan Komputer

Editor Office : LPPM Sekolah Tinggi Teknologi Pagar Alam, Jln. Masik Siagim No. 75
Simpang Mbacang, Pagar Alam, SUM-SEL, Indonesia
Phone : +62 852-7901-1390.
Email : betrik@sttpagaralam.ac.id | admin.jurnal@sttpagaralam.ac.id
Website : <https://ejournal.sttpagaralam.ac.id/index.php/betrik/index>

MODEL PREDIKSI PENYAKIT JANTUNG MENGGUNKAN MACHINE LEARNING

Hamid Rahman¹, Ramdani Agusman², Tata Sutabri³

Mahasiswa Universitas Bina Darma^{1,2}, Dosen Universitas Bina Darma³

Jalan Jenderal Ahmad Yani No.3 Palembang

Sur-el : hamidrahman@live.com¹, ramdani.agusman14@gmail.com², tata.sutabri@gmail.com³

Abstract : Heart disease is one of the leading causes of death worldwide. Early detection is very important in order to help reduce the risk of death and prevent more severe complications by enabling faster and more appropriate treatment. In this context, the application of machine learning (ML) technology in the medical field provides great potential to improve the accuracy of heart disease diagnosis and prediction. In this study, we aim to derive a ML model that can predict heart disease with the best performance using the Cleveland UCI heart disease dataset. To improve the performance of the model, we also used particle swarm optimization (PSO) algorithm to reduce the number of features in the dataset to have a significant impact on the performance of the ML model. The ML model was trained with the dataset that had been feature selected with the PSO algorithm and then tested, and the performance was compared. The highest performance was obtained for the XGBoost classification model trained on the dataset with the PSO algorithm, with 86% accuracy, 86% precision, 86% recall, and 86% f1-score. The results show that the combination of PSO and XGBoost algorithms has the best performance for use in heart disease prediction.

Keywords: Heart Disease, Feature Selection, PSO, Machine Learning

Abstrak : Penyakit jantung merupakan salah satu penyakit yang menjadi penyebab utama kematian diseluruh dunia. Deteksi dini menjadi sangat penting supaya dapat membantu mengurangi risiko kematian serta mencegah komplikasi yang lebih parah dengan memungkinkan penanganan yang lebih cepat dan tepat. Dalam konteks ini, penerapan teknologi *machine learning* (ML) di bidang medis memberikan potensi besar untuk meningkatkan akurasi diagnosis dan prediksi penyakit jantung. Dalam penelitian ini, peneliti bertujuan untuk mendapatkan model ML yang dapat memprediksi penyakit jantung dengan performa terbaik menggunakan dataset penyakit jantung Cleveland UCI. Untuk meningkatkan performa dari model, peneliti juga menggunakan algoritma *particle swarm optimization* (PSO) untuk mengurangi jumlah fitur dalam dataset agar dapat mempunyai dampak yang signifikan terhadap kinerja model ML. Model ML dilatih dengan dataset yang telah dipilih fiturnya dengan algoritma PSO kemudian diuji, dan kinerjanya dibandingkan. Performa tertinggi diperoleh untuk model klasifikasi XGBoost yang dilatih pada dataset dengan algoritma PSO, dengan akurasi sebesar 86%, *precision* 86%, *recall* 86%, dan *f1-score* 86%. Hasil penelitian menunjukkan bahwa kombinasi algoritma PSO dan XGBoost memiliki kinerja yang paling baik untuk digunakan dalam prediksi penyakit jantung.

Kata kunci: Penyakit Jantung, Seleksi Fitur, PSO, Machine Learning

1. PENDAHULUAN

Penyakit jantung merupakan salah satu penyakit yang menjadi penyebab utama kematian di seluruh dunia. Menurut data dari Organisasi Kesehatan Dunia (WHO), penyakit

kardiovaskular, termasuk penyakit jantung, menyumbang sekitar 32% dari total kematian global pada tahun 2019 [1]. Di Indonesia, penyakit jantung juga menduduki peringkat tertinggi sebagai penyebab kematian,

mengalahkan penyakit infeksi. Data dari Kementerian Kesehatan Republik Indonesia

menunjukkan bahwa prevalensi penyakit jantung terus meningkat, yang mengindikasikan bahwa penyakit ini telah menjadi beban kesehatan masyarakat yang signifikan [2]. Mengingat dampaknya yang besar terhadap individu dan masyarakat, maka upaya untuk menurunkan angka kematian akibat penyakit jantung harus dilakukan dengan pendekatan yang lebih inovatif dan terintegrasi.

Penyakit jantung sering kali berkembang tanpa menunjukkan gejala yang jelas pada tahap awal sehingga banyak pasien yang baru menyadari bahwa mereka menderita penyakit jantung setelah mengalami gejala parah, seperti serangan jantung atau gagal jantung. Hal ini menjadikan deteksi dini sangat penting dalam upaya pengendalian penyakit jantung. Deteksi dini dapat membantu mengurangi risiko kematian serta mencegah komplikasi yang lebih parah dengan memungkinkan penanganan yang lebih cepat dan tepat. Oleh karena itu, penting bagi sistem kesehatan untuk mengembangkan metode yang dapat memprediksi risiko penyakit jantung lebih awal.

Dalam konteks ini, penerapan teknologi *machine learning* (ML) di bidang medis memberikan potensi besar untuk meningkatkan akurasi diagnosis dan prediksi penyakit jantung. *Machine learning* adalah cabang dari kecerdasan buatan yang memungkinkan komputer untuk "belajar" dari data dan menghasilkan prediksi atau keputusan tanpa pemrograman eksplisit. Dengan menggunakan algoritma tertentu, *machine learning* dapat menganalisis data yang sangat besar dan kompleks, seperti data medis historis,

faktor risiko, serta hasil pemeriksaan fisik, untuk menghasilkan prediksi yang lebih cepat dan lebih akurat mengenai kemungkinan seseorang mengidap penyakit jantung [3]. Selain itu, penerapan *machine learning* juga menawarkan keuntungan signifikan dalam hal efisiensi. Dengan kemampuannya untuk menganalisis data dalam jumlah besar secara cepat, teknologi ini dapat mengurangi ketergantungan pada tenaga medis yang terbatas serta mempercepat proses diagnosis. Penelitian oleh [4], [5], [6] menunjukkan bahwa penerapan *machine learning* dalam pendeteksian otomatis telah berkembang pesat sehingga dalam konteks mendeteksi penyakit jantung dapat mengurangi waktu diagnosis, tetapi juga memungkinkan untuk melakukan deteksi lebih awal, bahkan sebelum gejala klinis muncul. Hal ini memungkinkan perawatan yang lebih cepat dan lebih efisien.

Beberapa algoritma *machine learning* yang umum digunakan dalam model prediksi penyakit jantung meliputi regresi logistik, pohon keputusan (*decision tree*), *random forest*, dan jaringan saraf tiruan (*neural networks*). Algoritma-algoritma ini memiliki kemampuan untuk mengidentifikasi pola-pola dalam data yang mungkin tidak terlihat oleh manusia dan dapat memberikan informasi yang lebih spesifik mengenai kemungkinan risiko penyakit jantung pada individu. Banyak penelitian menggunakan *machine learning* menunjukkan bahwa teknik *machine learning* memiliki potensi yang besar dalam meningkatkan akurasi prediksi penyakit jantung. Penelitian [7] menggunakan dataset *Cleveland Heart Disease* dan tiga algoritma pembelajaran mesin: *decision*

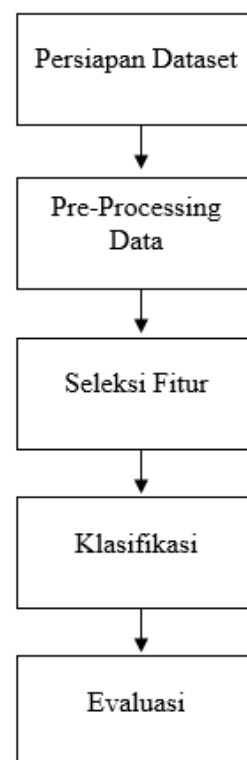
tree, random forest, dan model *hybrid* (*decision tree* dan *random forest*). Temuan eksperimental menunjukkan bahwa model prediksi penyakit jantung mendapatkan akurasi yang tinggi mencapai 88,7% dengan menggunakan model *hybrid*. Penelitian [8] membandingkan beberapa algoritma *machine learning* dengan tujuan mengembangkan model prediksi penyakit jantung. Algoritma yang digunakan pada penelitian ini diantaranya ET, LR, AB, CART, MNB, LDA, SVM, XGB, dan RF. Hasil penelitian ini menunjukkan algoritma SVM mempunyai nilai akurasi yang terbaik dengan mencapai nilai 96,72%. Penelitian [9] dilakukan dengan mengevaluasi beberapa teknik *machine learning* pada model prediksi penyakit jantung. Penelitian ini menggunakan algoritma RF, LR, KNN, DT, dan *Naïve Bayes* yang di optimalkan dengan *feature selection* menggunakan *chi-square*, ANOVA, dan *mutual information*. Penelitian ini mendapatkan nilai akurasi sebesar 94,51% pada algoritma *Random Forest*.

Penelitian ini bertujuan untuk mengembangkan dan mengkaji penerapan model prediksi penyakit jantung menggunakan *machine learning*, serta mengeksplorasi berbagai metode dan algoritma yang dapat diterapkan dalam menganalisis data medis pasien. Dengan pendekatan ini, diharapkan dapat tercipta model yang dapat meningkatkan akurasi diagnosis penyakit jantung dan memberikan kontribusi dalam pencegahan penyakit jantung melalui identifikasi faktor risiko lebih awal.

2. METODOLOGI PENELITIAN

Penelitian ini membahas tentang analisis performa dari berbagai algoritma

machine learning yang berbeda berdasarkan pemilihan fitur yang bermanfaat dari dataset dengan harapan dapat meningkatkan akurasi prediksi penyakit jantung. Dalam penelitian ini, kinerja model ML yang berbeda seperti XGBoost dan SVM dengan menggunakan *feature selection*, dengan tujuan untuk mendapatkan model dengan kinerja tertinggi. Dataset Cleveland yang digunakan dalam penelitian ini diperoleh dari repositori Kaggle Machine Learning.

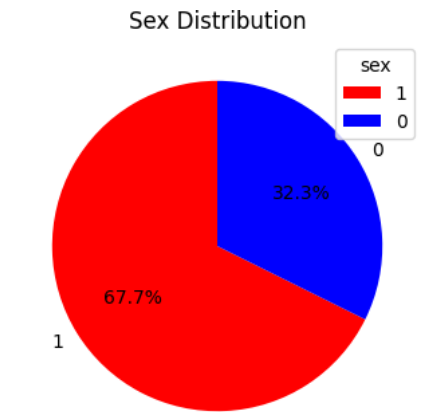


Gambar 1 Tahapan Penelitian

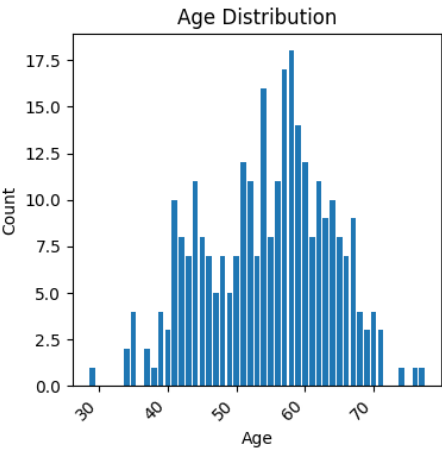
2.1 Persiapan Data

Dataset yang digunakan pada penelitian ini menggunakan *Cleveland Dataset University of California (UCI)*. Dataset *Cleveland UCI* menjadi salah satu dataset yang sering digunakan dalam mengembangkan model prediksi penyakit jantung. Dataset ini mempunyai total data sebanyak 297 dengan 14 kolom dimana 13 kolom

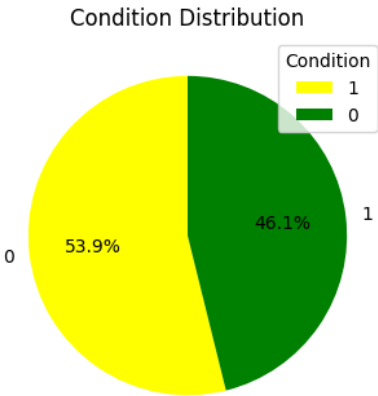
fitur dan 1 kolom label dimana terdapat 201 pasien pria dan 96 pasien wanita dan rentang umur pasien dari 29 tahun sampai 77 tahun. Kondisi pasien terdapat data 160 data bebas penyakit jantung dan 137 data mempunyai penyakit jantung. Gambar 2, 3, dan 4 menunjukkan persentase dari beberapa atribut seperti jenis kelamin, umur, dan kondisi pasien.



Gambar 2 Persentase Jenis Kelamin



Gambar 3 Persentase Umur Pasien



Gambar 4 Persentase Kondisi Pasien

Tabel 1 menunjukkan atribut dari dataset.

Tabel 1 Dataset

Name	Definition
Age	Age
Sex	Sex
Chest Pain	Cp
Resting Blood Pressure	Trestbps
Serum cholesterol	chol
Fasting sugar blood	Fbs
Rest electrocardiograph	Restecg
MaxHeart rate	Thalach
Exercise-Induced angina	exang
ST depression	oldpeak
Slope	Slop
No. of vessels	Ca
Thalassemia	Thal
condition	condition

2.2 Pre-Processing Data

Pre-Processing data memainkan peran penting dalam meningkatkan kinerja dan ketahanan algoritme pembelajaran mesin. Proses berulang ini sering kali melibatkan tugas-tugas

seperti menangani nilai yang hilang, menghapus duplikasi, menstandarkan fitur, dan mengkodekan variabel kategorikal. Pemrosesan ulang data yang tepat tidak hanya memastikan integritas data, tetapi juga berkontribusi secara signifikan terhadap keberhasilan upaya pembelajaran mesin, yang memungkinkan model untuk menggeneralisasi dengan lebih baik dan membuat prediksi yang lebih andal.

2.3 Feature Selection

Pemilihan fitur adalah komponen penting dari proses pembelajaran mesin karena ada beberapa kasus di mana dataset terdiri dari banyak fitur yang tidak penting, dan ini dapat berdampak buruk pada keakuratan algoritma.

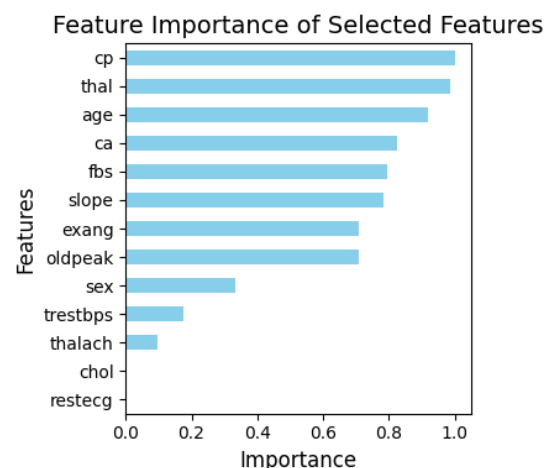
Atribut dalam set data pada dasarnya adalah properti atau karakteristik yang berfungsi sebagai dasar untuk menentukan apakah seseorang menderita penyakit tertentu. Atribut-atribut ini mencakup berbagai faktor, termasuk detak jantung, jenis nyeri dada, jenis kelamin, usia, keberadaan angina yang diakibatkan oleh olahraga, dan berbagai fitur lainnya. Tabel 1 memberikan penjelasan rinci tentang 14 fitur dari kumpulan data.

2.3.1 Particle Swarm Optimization

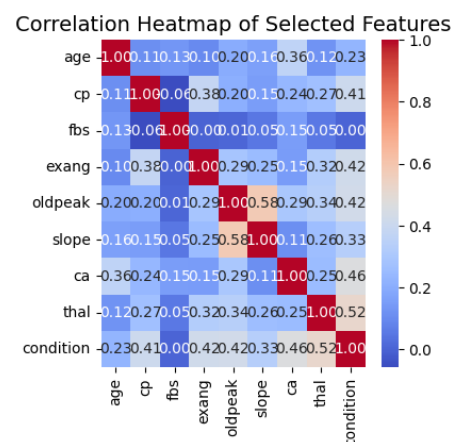
Algoritma *Particle Swarm Optimization* (PSO) merupakan teknik komputasi dalam melakukan optimasi masalah menggunakan pendekatan meta-heuristic [10]. PSO diperkenalkan pertama kali oleh J. Kennedy dan R. C. Eberhart pada tahun 1995 yang terinspirasi dari perilaku yang dilakukan sehari-hari dari sekelompok binatang yang bersifat alami [11]. Pada proses klasifikasi, algoritma ini biasanya dilakukan dalam proses seleksi fitur dari suatu dataset. Proses seleksi fitur menggunakan

algoritma PSO dinilai sangat efektif karena mampu mengeksplorasi berbagai kombinasi fitur tanpa melakukan pencarian acak.

Sehingga pada dataset Cleveland UCI, peneliti melakukan seleksi fitur untuk melihat nilai penting (*feature importance*) dari setiap fitur yang ada. Hasil menunjukkan pada gambar 2, terdapat 8 fitur yang mempunyai nilai penting yang lebih tinggi jika dibandingkan dengan fitur lainnya. Delapan fitur yang dipilih pada seleksi fitur ini meliputi 'age', 'cp', 'fbs', 'exang', 'oldpeak', 'slope', 'ca', 'thal'. Gambar 4 dan 5 menunjukkan korelasi antar setiap fitur dari dataset.



Gambar 5 Plot Bar Seleksi Fitur



Gambar 6 Korelasi Antar Fitur

2.4 Klasifikasi

Klasifikasi dilakukan dengan menggunakan model *machine learning* yaitu: *MultiLayer Perceptron Classifier* (MLPC), *Extreme Gradient Boosting* (XGBoost), *Support Vector Machine* (SVM), *Decision Tree* (DT), dan *Logistic Regression* (LR).

2.4.1 MultiLayer Perceptron Classifier

MPLC merupakan salah satu jenis dari jaringan saraf tiruan. Sebuah *Multilayer Perceptron* (MLP) terdiri dari satu *input layer*, satu atau lebih *hidden layer*, dan terakhir satu *output layer*. MLPC memiliki kemampuan untuk melakukan pembelajaran dari suatu pola yang kompleks dalam sebuah data, sehingga diharapkan efektif untuk melakukan tugas klasifikasi.

2.4.2 Extreme Gradient Boosting

Algoritma *Extreme Gradient Boosting* (XGBoost) adalah algoritma yang digunakan untuk regresi dan klasifikasi. Ini adalah metode ensemble di mana prediktor yang kuat dihasilkan dengan menggabungkan banyak prediktor yang buruk. Untuk meminimalkan *loss function*, algoritma ini beroperasi dengan menerapkan pohon keputusan pada residual dari model saat ini secara berulang-ulang. Dengan setiap iterasi, mengarahkan pengamatan pada kesalahan yang terjadi selama iterasi sebelumnya. Dengan menjumlahkan prediksi yang dihasilkan oleh setiap pohon dalam model, perkiraan akhir dapat dipastikan [12]. XGBoost mampu menangani data yang tidak seimbang, yang sering terjadi dalam dataset medis. Dengan teknik seperti pengambilan sampel adaptif, XGBoost dapat meningkatkan akurasi model.

2.4.3 Support Vector Machine

Algoritma SVM adalah algoritma untuk menganalisis data dalam bentuk klasifikasi dan regresi yang digunakan untuk memeriksa data dan mengidentifikasi pola [13], [14]. Metode ini berusaha untuk mengidentifikasi hyperplane optimal yang memaksimalkan margin antara kelas-kelas yang berbeda dalam ruang fitur berdimensi tinggi. SVM dapat menangani data non-linear dengan menggunakan berbagai fungsi kernel, seperti radial basis function (RBF) dan kernel polinomial. SVM dikenal karena kemampuannya untuk menggeneralisasi dengan baik dan bekerja secara efektif dalam ruang dimensi tinggi [15].

2.4.4 Decision Tree

Algoritma *Decision Tree* (DT) adalah model algoritma yang dibuat untuk tugas klasifikasi dan regresi [16]. Model ini berbentuk struktur seperti pohon di mana setiap simpul internal menandakan keputusan berdasarkan fitur, dan setiap simpul daun menandakan kelas atau nilai numerik [17]. Banyak peneliti telah menggunakannya sebagai alat untuk melakukan klasifikasi dalam domain kesehatan untuk mengevaluasi data dan mendapatkan kesimpulan. Algoritma DT menghasilkan model yang mengantisipasi nilai variabel target. Dengan menghasilkan aturan keputusan mendasar dari properti data dan melakukan segmentasi data ke dalam unit-unit seperti cabang. Nilai input dapat diklasifikasikan sebagai nilai kontinu atau tidak kontinu. Nilai probabilitas atau penunjukan kelas dikembalikan oleh simpul daun. Pohon tersebut dapat ditransformasikan menjadi kumpulan prinsip-prinsip keputusan [18].

2.4.5 Logistic Regression

Logistic Regression (LR) adalah jenis dari

supervised learning. *Logistic regression* digunakan untuk menghitung atau memprediksi probabilitas antara 0 atau 1 [19]. Contoh dari *logistic regression* adalah menerapkan pembelajaran mesin untuk menentukan apakah dalam suatu lalu lintas jaringan terdapat *malware* atau tidak.

Adapun dalam proses klasifikasi dibagi menjadi dua percobaan. Percobaan pertama dilakukan dengan melakukan *pre-processing* data terlebih dahulu kemudian membagi dataset menjadi data *training* dan *testing*. Setelah dataset dibagi menjadi data *training* dan *testing*, data tersebut langsung dilakukan proses pelatihan dan pengujian. Percobaan kedua dilakukan dengan melakukan *pre-processing* dataset dan membagi dataset menjadi data *training* dan *testing*. Setelah itu dilakukan *feature selection* menggunakan algoritma PSO untuk memilih fitur yang penting dari dataset. Kemudian setelah *feature selection* dilakukan, data kemudian dilakukan proses pelatihan dan pengujian.

2.5 Evaluasi

Evaluasi pada penelitian ini menggunakan *performance metric* dalam menilai keefektifan dan ketepatan pengklasifikasi. Penilaian dilakukan dengan menggunakan *confusion matrix* (CM) untuk mengevaluasi performa suatu model *machine learning* [20]. CM menggambarkan hasil dari proses klasifikasi terkait *false positive* (FN), *true negative* (TP), *true positive* (TN), dan *false positive* (FP). TN menunjukkan contoh yang secara akurat diprediksi sebagai bagian dari kelas negatif, sedangkan TP menandakan contoh yang secara akurat diprediksi sebagai bagian dari kelas positif. Prediksi yang tidak akurat pada kelas positif (FP) dan prediksi yang salah pada kelas

negatif (FN) masing-masing dilambangkan sebagai FP dan FN.

Tabel 2 Confusion Matrix

TP	FP
FN	TN

Selain itu beberapa metrik digunakan dalam analisis, seperti F1-score, recall, presisi, dan akurasi. Metrik-metrik tersebut memberikan wawasan yang signifikan mengenai berbagai aspek kemampuan pengklasifikasi.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Presisi = \frac{TP}{FP+TP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

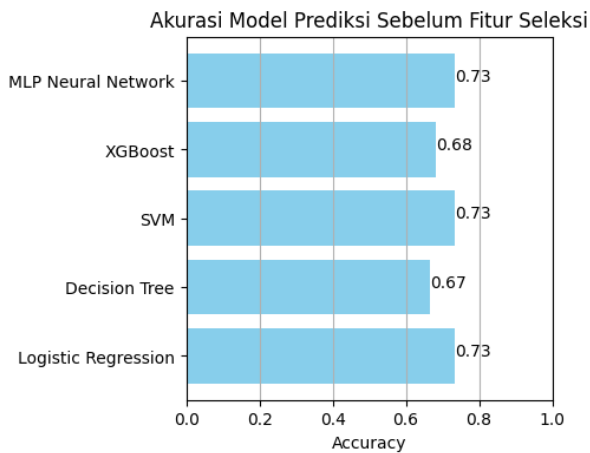
$$F - Measure = \frac{2 \times Presisi \times Recall}{Presisi+Recall} \quad (4)$$

3. HASIL DAN PEMBAHASAN

Pada penelitian ini dilakukan dua kali percobaan klasifikasi. Percobaan pertama klasifikasi dilakukan tanpa menggunakan seleksi fitur dan percobaan kedua dilakukan menggunakan seleksi fitur.

3.1 Percobaan Pertama

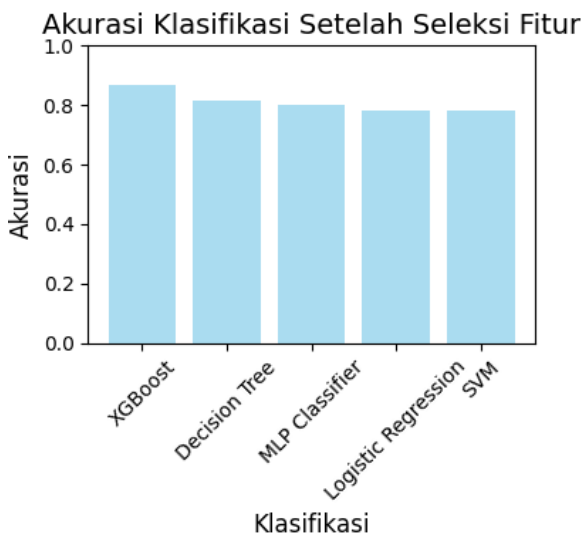
Percobaan pertama dilakukan dengan *pre-processing* data tanpa menggunakan seleksi fitur. Hasil akurasi yang didapatkan MLP 73%, XGBoost 68%, SVM 73%, DT 67%, dan LR 73%. Hasil yang di dapat pada percobaan pertama ini menunjukkan bahwa performa model secara keseluruhan masih kurang baik.



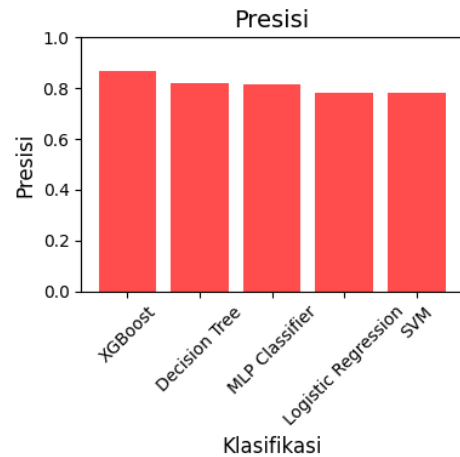
Gambar 7 Akurasi Percobaan Pertama

3.2 Percobaan Kedua

Percobaan kedua dilakukan dengan *pre-processing* data dan seleksi fitur menggunakan algoritma PSO. Percobaan kedua ini diharapkan dapat meningkatkan akurasi. Hasil dari percobaan kedua ini menunjukkan peningkatan akurasi dari seluruh model yang di ajukan. Hasil akurasi dari model percobaan kedua ini mencapai MLPC 80%, XGBoost 86%, SVM 78%, DT 81%, dan LR 78%.

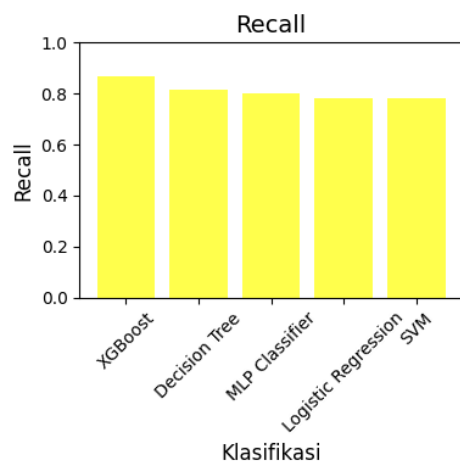


Gambar 8 Akurasi Percobaan Kedua

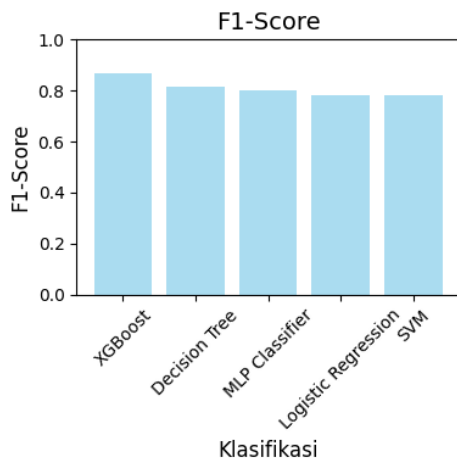


Gambar 9 Nilai Presisi dari setiap Model

Selain akurasi, terdapat metrik performa lain untuk melihat performa masing – masing model seperti presisi, *recall*, *F1-Score*. Model MLPC mendapatkan nilai presisi sebesar 81%, nilai *recall* sebesar 80%, dan nilai *F1-Score* 79%, XGBoost mendapatkan nilai presisi sebesar 86%, nilai *recall* sebesar 86%, dan nilai *F1-Score* 86%, SVM mendapatkan nilai presisi sebesar 78%, nilai *recall* sebesar 78%, dan nilai *F1-Score* 78%, DT mendapatkan nilai presisi sebesar 82%, nilai *recall* sebesar 81%, dan nilai *F1-Score* 81%, dan LR mendapatkan nilai presisi sebesar 78%, nilai *recall* sebesar 78%, dan nilai *F1-Score* 78%.

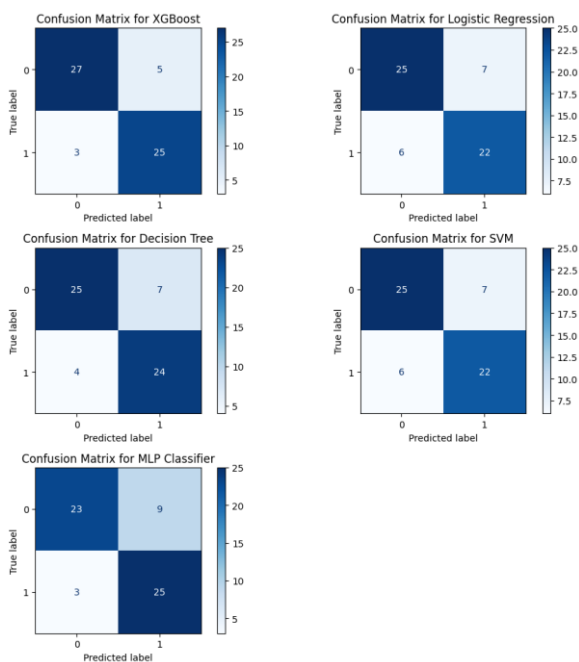


Gambar 10 Nilai Recall dari setiap Model



Gambar 11 Nilai F1-Score dari setiap Model

Gambar 11 menunjukkan jumlah data yang benar dan salah dari model MLPC, XGBoost, SVM, DT, dan LR.



Gambar 12 Confusion Matrix dari setiap Model

4. KESIMPULAN

Berdasarkan hasil kedua percobaan, dapat disimpulkan bahwa kombinasi antara pre-processing data dan seleksi fitur menggunakan algoritma PSO merupakan pendekatan yang efektif untuk meningkatkan kinerja model pembelajaran mesin. Seleksi fitur membantu

dalam mengurangi dimensi data dan meningkatkan relevansi fitur, sehingga model dapat menghasilkan prediksi yang lebih akurat. Di antara model-model yang diuji, XGBoost menunjukkan kinerja terbaik.

- Penambahan tahap seleksi fitur menggunakan algoritma PSO memberikan dampak positif yang signifikan terhadap peningkatan akurasi semua model.
- Seleksi fitur menggunakan algoritma PSO membantu dalam memilih fitur-fitur yang paling relevan, sehingga model dapat lebih fokus pada informasi yang penting.
- XGBoost menunjukkan kinerja terbaik dengan akurasi tertinggi (86%) dan nilai presisi, recall, serta F1-Score yang sangat baik. Ini mengindikasikan bahwa XGBoost sangat efektif dalam menangani dataset ini.
- MLPC dan DT juga menunjukkan peningkatan akurasi yang cukup signifikan setelah seleksi fitur.
- SVM dan LR mengalami peningkatan akurasi, namun tidak sebesar model-model lainnya.
- Eksplorasi Algoritma Seleksi Fitur Lainnya: Mencoba algoritma seleksi fitur lainnya seperti *filter method*, *wrapper method*, atau *embedded method* dapat memberikan hasil yang berbeda dan mungkin lebih baik.
- Penyetelan *Hyperparameter*: Melakukan penyetelan *hyperparameter* secara cermat pada masing-masing model dapat meningkatkan kinerja model lebih lanjut.

- Analisis Fitur: Melakukan analisis terhadap fitur-fitur yang terpilih dapat memberikan wawasan yang lebih dalam mengenai karakteristik data dan hubungan antara fitur dengan target variabel.

DAFTAR PUSTAKA

- [1] "Cardiovascular diseases (CVDs)." Accessed: Nov. 28, 2024. [Online]. Available: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))
- [2] "Profil Kesehatan Indonesia 2023." Accessed: Nov. 28, 2024. [Online]. Available: <https://www.kemkes.go.id/id/profil-kesehatan-indonesia-2023>
- [3] P. Rajpurkar, A. Y. Hannun, M. Haghighpanahi, C. Bourn, and A. Y. Ng, "Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks," Jul. 2017, Accessed: Nov. 30, 2024. [Online]. Available: <https://arxiv.org/abs/1707.01836v1>
- [4] A. A. A. Mohamed, A. Hançerlioğullari, J. Rahebi, M. K. Ray, and S. Roy, "Colon Disease Diagnosis with Convolutional Neural Network and Grasshopper Optimization Algorithm," *Diagnostics 2023, Vol. 13, Page 1728*, vol. 13, no. 10, p. 1728, May 2023, doi: 10.3390/DIAGNOSTICS13101728.
- [5] R. Liu, C. Ren, M. Fu, Z. Chu, and J. Guo, "Platelet Detection Based on Improved YOLO_v3," *Cyborg and Bionic Systems*, vol. 2022, Jan. 2022, doi: 10.34133/2022/9780569.
- [6] A. A. Ahmad and H. Polat, "Prediction of Heart Disease Based on Machine Learning Using Jellyfish Optimization Algorithm," *Diagnostics 2023, Vol. 13, Page 2392*, vol. 13, no. 14, p. 2392, Jul. 2023, doi: 10.3390/DIAGNOSTICS13142392.
- [7] M. Kavitha, G. Gnaneswar, R. Dinesh, Y. R. Sai, and R. S. Suraj, "Heart Disease Prediction using Hybrid machine Learning Model," *Proceedings of the 6th International Conference on Inventive Computation Technologies, ICICT 2021*, pp. 1329–1333, Jan. 2021, doi: 10.1109/ICICT50816.2021.9358597.
- [8] A. Saboor, M. Usman, S. Ali, A. Samad, M. F. Abrar, and N. Ullah, "A Method for Improving Prediction of Human Heart Disease Using Machine Learning Algorithms," *Mobile Information Systems*, vol. 2022, 2022, doi: 10.1155/2022/1410169.
- [9] N. Biswas *et al.*, "Machine Learning-Based Model to Predict Heart Disease in Early Stage Employing Different Feature Selection Techniques," *Biomed Res Int*, vol. 2023, 2023, doi: 10.1155/2023/6864343.
- [10] A. Sa'adah, A. Sasmito, and A. A. Pasaribu, "Comparison of Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) for Estimating the Susceptible-Exposed-Infected-Recovered (SEIR) Model Parameter Values," *Journal of Information Systems Engineering and Business Intelligence*, vol. 10, no. 2, pp. 290–301, Jun. 2024, doi: 10.20473/JISEBI.10.2.290-301.
- [11] A. M. Rizki and A. L. Nurlaili, "Algoritme Particle Swarm Optimization (PSO) untuk Optimasi Perencanaan Produksi Agregat Multi-Site pada Industri Tekstil Rumahan," *Journal of Computer, Electronic, and Telecommunication*, vol. 1, no. 2, Jan. 2021, doi: 10.52435/complete.v1i2.73.
- [12] L. V. Fulton, D. Dolezel, J. Harrop, Y. Yan, and C. P. Fulton, "Classification of Alzheimer's Disease with and without Imagery using Gradient Boosted Machines and ResNet-50," *Brain Sci*, vol. 9, no. 9, Sep. 2019, doi: 10.3390/BRAINSKI9090212.
- [13] M. Ahmed and I. Husien, "Heart Disease Prediction Using Hybrid Machine Learning: A Brief Review," 2024, *Department of Agribusiness, Universitas Muhammadiyah Yogyakarta*. doi: 10.18196/jrc.v5i3.21606.
- [14] D. Anggraini and T. Sutabri, "Pengembangan Aplikasi Penyaringan Spam e-Mail Menggunakan Teknik Machine Learning dengan Metode Support Vector Machines," *IJM: Indonesian Journal of Multidisciplinary*, vol. 2, no. 3, pp. 106–114, Apr. 2024, Accessed: Nov. 30, 2024. [Online].

Available:

<https://journal.csspublishing.com/index.php/ijm/article/view/720>

- [15] H. A. Taher and A. M. Abdulazeez, "Machine Learning Approaches for Heart Disease Detection: A Comprehensive Review," 2023.
- [16] J. R. Quinlan, "Induction of decision trees," *Machine Learning 1986 1:1*, vol. 1, no. 1, pp. 81–106, Mar. 1986, doi: 10.1007/BF00116251.
- [17] N. Khalili and M. A. Rastegar, "Optimal cost-sensitive credit scoring using a new hybrid performance metric," *Expert Syst Appl*, vol. 213, p. 119232, Mar. 2023, doi: 10.1016/J.ESWA.2022.119232.
- [18] M. A. Kadhim and A. M. Radhi, "Heart disease classification using optimized Machine learning algorithms," *Iraqi Journal For Computer Science and Mathematics*, vol. 4, no. 2, pp. 31–42, Feb. 2023, doi: 10.52866/IJCSM.2023.02.02.004.
- [19] G. N. Ahmad *et al.*, "Mixed Machine Learning Approach for Efficient Prediction of Human Heart Disease by Identifying the Numerical and Categorical Features," *Applied Sciences (Switzerland)*, vol. 12, no. 15, Aug. 2022, doi: 10.3390/app12157449.
- [20] H. RAHMAN and T. Sutabri, "Analysis DDoS Attack Using Machine Learning On Software-Defined Network Architectures," *JSAI (Journal Scientific and Applied Informatics)*, vol. 7, no. 3, pp. 531–536, Nov. 2024, doi: 10.36085/JSAI.V7I3.7301.