

Optimasi Model MLP Menggunakan Hyperparameter Tuning Randomsearchcv Dalam Prediksi Resiko Penyakit Jantung

Muhammad Hamdi ^{1,*}, Djoko Rahardjo ¹

¹ Program Studi Ilmu Komputer, Universitas Bumigora, Indonesia

* Correspondence: andibaglon@gmail.com

Received: 26 Oktober 2025 | Revised: 23 November 2025 | Accepted: 10 Desember 2025 | Published: 31 Desember 2025

Abstrak

Penyakit jantung merupakan salah satu penyebab utama kematian di dunia sehingga diperlukan metode deteksi dini yang akurat untuk mendukung upaya pencegahan. Penelitian ini bertujuan mengembangkan dan mengevaluasi model prediksi penyakit jantung menggunakan algoritma Multilayer Perceptron (MLP) pada tiga skenario, yaitu model dasar tanpa penyeimbangan data dan optimasi parameter, penerapan Synthetic Minority Over-sampling Technique (SMOTE), serta kombinasi SMOTE dengan optimasi parameter menggunakan Randomized Search Cross-Validation. Dataset yang digunakan adalah *heart_cleveland_upload.csv* yang berisi atribut klinis pasien dan label kondisi penyakit jantung. Tahapan pra-pemrosesan meliputi penanganan data hilang, pengkodean variabel kategorikal, pemisahan data latih dan uji, serta standardisasi fitur. Hasil eksperimen menunjukkan bahwa model MLP dasar memperoleh akurasi 0,85 dan skor F1 sebesar 0,830. Penerapan SMOTE meningkatkan skor F1 menjadi 0,836 dengan akurasi yang tetap. Kinerja terbaik dicapai oleh model MLP yang dikombinasikan dengan SMOTE dan optimasi parameter, dengan akurasi 0,90 dan skor F1 sebesar 0,884. Analisis fitur menunjukkan bahwa beberapa indikator klinis memiliki kontribusi dominan dalam proses prediksi.

Kata kunci: Penyakit Jantung, MLP, SMOTE, Randomized Search, klasifikasi, machine learning.

Abstract

Heart disease is one of the leading causes of death worldwide, requiring accurate early detection methods to support prevention efforts. This study aims to develop and evaluate a heart disease prediction model using the Multilayer Perceptron (MLP) algorithm in three scenarios: a baseline model without data balancing and parameter optimization, the application of the Synthetic Minority Over-sampling Technique (SMOTE), and a combination of SMOTE with parameter optimization using Randomized Search Cross-Validation. The dataset used is *heart_cleveland_upload.csv*, which contains patient clinical attributes and heart disease condition labels. Pre-processing steps include handling missing data, encoding categorical variables, separating training and test data, and feature standardization. Experimental results show that the baseline MLP model achieves an accuracy of 0.85 and an F1 score of 0.830. The application of SMOTE improves the F1 score to 0.836 with a constant accuracy. The best performance is achieved by the MLP model combined with SMOTE and parameter optimization, with an accuracy of 0.90 and an F1 score of 0.884. Feature analysis shows that several clinical indicators have a dominant contribution in the prediction process.

Keywords: Hear Disease, MLP, SMOTE, Randomized Search, classification, machine learning.

PENDAHULUAN

Penyakit jantung masih menjadi salah satu penyebab utama kematian di tingkat global, sehingga deteksi dini merupakan langkah krusial dalam upaya menurunkan angka morbiditas dan mortalitas. Seiring dengan berkembangnya teknologi komputasi, pendekatan berbasis pembelajaran mesin (machine learning) semakin banyak dimanfaatkan sebagai alat bantu



diagnosis medis, termasuk dalam prediksi penyakit jantung. Berbagai studi menunjukkan bahwa model pembelajaran mesin mampu menangkap pola kompleks dalam data klinis dan memberikan performa prediksi yang menjanjikan dibandingkan metode konvensional.

Meskipun demikian, pengembangan model prediktif penyakit jantung menghadapi sejumlah tantangan utama, terutama terkait ketidakseimbangan kelas (class imbalance), di mana jumlah pasien tanpa penyakit jantung umumnya jauh lebih besar dibandingkan pasien dengan penyakit jantung. Kondisi ini berpotensi menyebabkan bias model terhadap kelas mayoritas dan menurunkan kemampuan deteksi kasus positif. Salah satu pendekatan yang banyak digunakan untuk mengatasi permasalahan tersebut adalah Synthetic Minority Over-sampling Technique (SMOTE), yang bekerja dengan menghasilkan sampel sintesis pada kelas minoritas guna meningkatkan representasi data dan kinerja model klasifikasi.

Penelitian sebelumnya telah menunjukkan efektivitas penggunaan SMOTE dalam konteks prediksi penyakit jantung. Studi oleh (Kwakye et al., 2021) menerapkan SMOTE pada berbagai algoritma pembelajaran mesin untuk klasifikasi penyakit jantung koroner dan melaporkan peningkatan performa model pada data yang tidak seimbang. Selain itu, kajian sistematis mengenai optimasi model Multilayer Perceptron (MLP) juga mengindikasikan bahwa SMOTE merupakan salah satu teknik yang paling umum digunakan untuk menangani ketidakseimbangan kelas pada jaringan saraf tiruan.

Di sisi lain, optimasi hiperparameter merupakan faktor penting dalam meningkatkan kinerja dan stabilitas model MLP. Penyetelan hiperparameter seperti jumlah neuron, learning rate, dan fungsi aktivasi terbukti berpengaruh signifikan terhadap akurasi dan kemampuan generalisasi model. Kajian literatur sistematis oleh (Supriyanto et al., 2025) menegaskan bahwa optimasi MLP melalui pendekatan seperti Randomized Search atau Grid Search menghasilkan model yang lebih akurat, terutama ketika dikombinasikan dengan teknik penyeimbangan kelas. Sejalan dengan hal tersebut, penelitian oleh (Susanto et al., 2025) yang menerapkan MLP dengan pendekatan SMOTE-ENN pada prediksi penyakit jantung melaporkan peningkatan performa yang signifikan, khususnya pada metrik recall yang mencapai 100%, disertai akurasi 89,47%, presisi 77,78%, spesifisitas 83,33%, dan F1-score 87,5%, yang menunjukkan keseimbangan yang baik antara kemampuan mendeteksi kasus positif dan meminimalkan kesalahan klasifikasi.

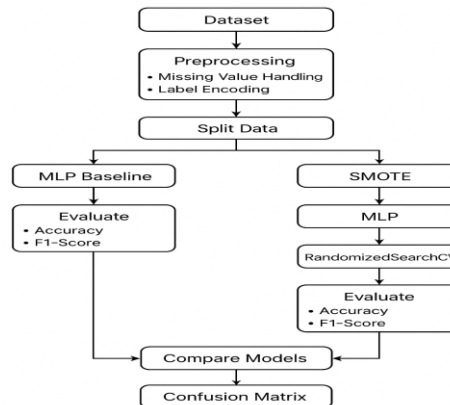
Meskipun demikian, sebagian besar penelitian terdahulu masih berfokus pada algoritma lain seperti Random Forest atau menggunakan SMOTE tanpa dikombinasikan secara komprehensif dengan optimasi hiperparameter pada model MLP. Dengan demikian, terdapat celah penelitian (research gap) terkait evaluasi menyeluruh kinerja MLP yang mengintegrasikan teknik penyeimbangan data dan optimasi hiperparameter dalam prediksi penyakit jantung.

Oleh karena itu, penelitian ini bertujuan untuk mengeksplorasi dan mengevaluasi kinerja model Multilayer Perceptron (MLP) dalam memprediksi kondisi penyakit jantung secara biner (penyakit dan tidak penyakit) melalui tiga skenario, yaitu: (1) model dasar tanpa penerapan SMOTE, (2) model dengan penerapan SMOTE, dan (3) model dengan penerapan SMOTE yang dikombinasikan dengan optimasi hiperparameter menggunakan Randomized Search Cross-Validation. Diharapkan hasil penelitian ini dapat memberikan kontribusi ilmiah sebagai referensi dalam pengembangan sistem pendukung keputusan medis serta menjadi acuan bagi penelitian selanjutnya dalam upaya prediksi dini penyakit jantung.

METODE

Penelitian ini memanfaatkan dataset *heart_cleveland_upload.csv* yang berisi data klinis pasien dengan indikasi penyakit jantung. Variabel target yang digunakan adalah *condition*, yang merepresentasikan keberadaan atau ketiadaan penyakit jantung. Dataset terdiri atas fitur numerik dan kategorikal, yang terlebih dahulu dianalisis untuk mengidentifikasi keberadaan

nilai hilang (*missing values*). Alur metodologi penelitian secara keseluruhan disajikan pada Gambar 1.



Gambar 1. Diagram alur penelitian

Sebagaimana ditunjukkan pada Gambar 1, penelitian ini diawali dengan tahapan *data preprocessing* yang bertujuan untuk meningkatkan kualitas data dan memastikan kesesuaian input bagi model pembelajaran mesin. Pada tahap imputasi, nilai hilang pada fitur numerik ditangani menggunakan *SimpleImputer* dengan strategi median, sementara fitur kategorikal diimputasi menggunakan nilai modus. Pendekatan ini dipilih untuk meminimalkan distorsi distribusi data akibat nilai ekstrem.

Selanjutnya, dilakukan proses *encoding* terhadap fitur kategorikal menggunakan *LabelEncoder* agar seluruh fitur dapat direpresentasikan dalam bentuk numerik dan diproses oleh model Multilayer Perceptron (MLP). Setelah itu, seluruh fitur distandarkan menggunakan *StandardScaler* untuk memastikan bahwa perbedaan skala antar fitur tidak menyebabkan bias dalam proses pembelajaran jaringan saraf. Dataset kemudian dibagi menjadi data latih dan data uji dengan rasio 80:20 menggunakan teknik *stratified splitting* berdasarkan variabel target, guna menjaga proporsi distribusi kelas pada kedua subset data. Langkah ini menjadi krusial mengingat dataset penyakit jantung umumnya memiliki ketidakseimbangan kelas antara pasien sehat dan pasien dengan penyakit jantung.

Menjawab gap penelitian terkait rendahnya performa model akibat ketidakseimbangan data dan kurangnya optimasi hiperparameter pada studi sebelumnya, penelitian ini mengintegrasikan teknik *Synthetic Minority Over-sampling Technique* (SMOTE) ke dalam *pipeline* pemodelan. SMOTE digunakan untuk menghasilkan sampel sintetis pada kelas minoritas sehingga distribusi target menjadi lebih seimbang dan model tidak bias terhadap kelas mayoritas. Tiga skenario pemodelan diuji secara komparatif. Skenario pertama adalah *Baseline MLP*, yaitu penggunaan *MLPClassifier* tanpa penanganan ketidakseimbangan data dan tanpa optimasi hiperparameter. Skenario kedua menggabungkan SMOTE dengan MLP dalam sebuah *pipeline* untuk mengevaluasi dampak oversampling terhadap performa model. Skenario ketiga mengembangkan pendekatan sebelumnya dengan menambahkan optimasi hiperparameter menggunakan *RandomizedSearchCV*, sehingga menghasilkan konfigurasi model yang lebih optimal.

Ruang parameter (*parameter space*) yang dieksplorasi dalam *RandomizedSearchCV* mencakup variasi jumlah dan struktur *hidden layers* (*hidden_layer_sizes*), fungsi aktivasi (*activation*), parameter regularisasi (*alpha*), skema *learning rate*, nilai *learning_rate_init*, ukuran *batch*, serta jumlah iterasi maksimum (*max_iter*). Proses optimasi dilakukan menggunakan *cross-validation* dengan skema CV = 5 dan CV = 10 untuk memperoleh estimasi performa yang lebih robust. Metrik evaluasi utama yang digunakan dalam proses ini adalah

F1-score, mengingat pentingnya keseimbangan antara *precision* dan *recall* dalam konteks prediksi medis. Evaluasi akhir model dilakukan pada data uji menggunakan metrik *accuracy*, *F1-score*, laporan klasifikasi yang mencakup *precision*, *recall*, dan *F1-score* per kelas, serta *confusion matrix* yang divisualisasikan dalam bentuk *heatmap*. Selain itu, dilakukan analisis kepentingan fitur (*feature importance*) dengan menghitung nilai absolut rata-rata bobot pada layer input MLP sebagai pendekatan proksi terhadap kontribusi masing-masing fitur.

Sebagai tahap akhir, kurva pembelajaran (*learning curve*) yang merepresentasikan nilai *loss* selama proses pelatihan ditampilkan untuk model terbaik hasil optimasi. Model tersebut kemudian disimpan menggunakan *joblib* guna memungkinkan pemanfaatan lebih lanjut, termasuk potensi implementasi pada sistem pendukung keputusan klinis.

HASIL DAN PEMBAHASAN

Hasil

Pada tahap baseline, model Multilayer Perceptron (MLP) dilatih menggunakan data pelatihan asli tanpa penerapan teknik penyeimbangan data (SMOTE) maupun optimasi hiperparameter. Evaluasi kinerja dilakukan pada data uji dan menghasilkan nilai akurasi sebesar 0,8500 serta nilai F1-score sebesar 0,8302. Hasil evaluasi tersebut disajikan pada Gambar 2.

```
=====
1. MODEL MLP BASELINE (TANPA SMOTE, TANPA TUNING)
=====
Accuracy: 0.8500
F1-Score: 0.8302

Classification Report:
      precision    recall  f1-score   support

     0       0.83       0.91       0.87        32
     1       0.88       0.79       0.83        28

   accuracy       0.85
  macro avg       0.85
 weighted avg       0.85
```

Gambar 2. hasil pengujian baseline model tanpa SMOTE dan tanpa Tuning

Berdasarkan hasil pengujian yang ditampilkan pada Gambar 2, model MLP baseline mampu mencapai tingkat akurasi sebesar 85%, yang mengindikasikan bahwa sebagian besar data uji dapat diklasifikasikan dengan benar. Selain itu, nilai F1-score sebesar 0,8302 menunjukkan keseimbangan yang cukup baik antara *precision* dan *recall*. Temuan ini mengindikasikan bahwa model tidak hanya memiliki tingkat akurasi yang tinggi, tetapi juga menunjukkan kemampuan yang memadai dalam membedakan kelas positif dan negatif, khususnya pada kondisi distribusi kelas yang tidak seimbang.

Selanjutnya, dilakukan pengujian model MLP dengan menerapkan Synthetic Minority Over-sampling Technique (SMOTE) pada data pelatihan untuk mengatasi permasalahan ketidakseimbangan kelas. Penerapan SMOTE menghasilkan distribusi kelas pelatihan yang lebih seimbang. Hasil evaluasi pada data uji menunjukkan adanya peningkatan nilai F1-score dibandingkan dengan model baseline, yang mengindikasikan bahwa teknik oversampling tersebut efektif dalam meningkatkan kemampuan model dalam mengenali dan mengklasifikasikan kelas minoritas. Hasil pengujian model MLP dengan SMOTE disajikan pada Gambar 3.

Hasil pada Gambar 3 menunjukkan bahwa model MLP yang dikombinasikan dengan teknik SMOTE mencapai akurasi sebesar 85% dan F1-score sebesar 0,8364. Nilai tersebut mengindikasikan bahwa penerapan SMOTE efektif dalam menangani ketidakseimbangan kelas, sehingga model mampu meningkatkan kemampuan dalam mengenali kelas minoritas tanpa menurunkan kinerja prediksi secara keseluruhan.

```

=====
2. MODEL MLP DENGAN SMOTE
=====
Accuracy: 0.8500
F1-Score: 0.8364

Classification Report:
      precision    recall  f1-score   support

     0       0.85      0.88      0.86       32
     1       0.85      0.82      0.84       28

   accuracy          0.85
  macro avg          0.85
 weighted avg          0.85

```

Gambar 3. MLP vs SMOTE

Selanjutnya, pengujian menggunakan kombinasi MLP, SMOTE, dan *RandomizedSearchCV* dilakukan untuk memperoleh konfigurasi hiperparameter optimal. *RandomizedSearchCV* menghasilkan model dengan skor cross-validation terbaik sebesar 0,7937. Evaluasi pada data uji menunjukkan bahwa model ini memberikan nilai akurasi dan F1-score yang lebih tinggi dibandingkan dengan dua skenario sebelumnya. Analisis feature importance mengidentifikasi sepuluh fitur paling berpengaruh berdasarkan bobot MLP, sedangkan kurva pembelajaran menunjukkan proses pelatihan yang stabil dan kecenderungan menuju konvergensi.

```

Test Accuracy: 0.9000
Test F1-Score: 0.8846

Classification Report:
      precision    recall  f1-score   support

     0       0.86      0.97      0.91       32
     1       0.96      0.82      0.88       28

   accuracy          0.91
  macro avg          0.90
 weighted avg          0.90

```

Gambar 4. Hasil uji MLP + SMOTE + Randomizes Search Cv

Hasil pada Gambar 4 menunjukkan bahwa model MLP yang dioptimalkan menggunakan *RandomizedSearchCV* dan penyeimbangan data dengan SMOTE menghasilkan kinerja yang unggul. Nilai *cross-validation* sebesar 79,37% mengindikasikan bahwa model memiliki stabilitas yang baik selama proses pelatihan. Pada tahap pengujian, model mencapai akurasi sebesar 90% dan F1-score sebesar 0,8846, yang mencerminkan kemampuan model dalam melakukan prediksi secara akurat sekaligus menjaga keseimbangan antara *precision* dan *recall*. Temuan ini menunjukkan bahwa kombinasi *hyperparameter tuning* dan SMOTE efektif dalam meningkatkan kemampuan model dalam mengenali kelas minoritas tanpa menimbulkan overfitting yang signifikan.

Berdasarkan tabel perbandingan performa, evaluasi akurasi dan F1-score pada ketiga model menunjukkan bahwa penerapan penyeimbangan data dan penyesuaian parameter memberikan peningkatan kinerja yang lebih signifikan dibandingkan model tanpa optimasi.

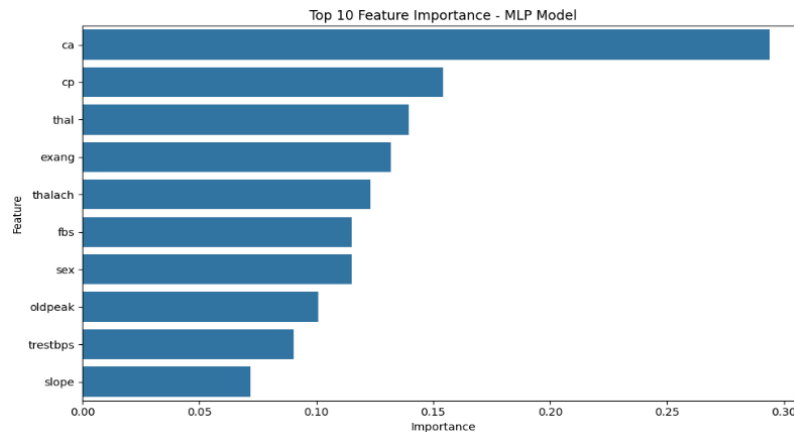
```

=====
PERBANDINGAN SEMUA MODEL
=====
      Model  Accuracy  F1-Score
0 Baseline (No SMOTE, No Tuning)    0.85  0.830189
1           With SMOTE              0.85  0.836364
2 With SMOTE + RandomSearchCV      0.90  0.884615

```

Gambar 5. Perbandingan semua model

Berdasarkan hasil perbandingan pada gambar 5, model baseline mencapai akurasi 85% dengan F1-score sebesar 0.830. Penerapan SMOTE tidak meningkatkan akurasi, namun mampu meningkatkan F1-score menjadi 0.836, yang mengindikasikan perbaikan kinerja pada kelas minoritas. Peningkatan performa yang paling signifikan diperoleh pada model yang menggabungkan SMOTE dan RandomizedSearchCV, dengan akurasi mencapai 90% dan F1-score sebesar 0.884. Temuan ini menunjukkan bahwa optimasi hyperparameter memberikan kontribusi terbesar terhadap peningkatan kinerja model. Dengan demikian, kombinasi SMOTE dan penyesuaian hyperparameter menghasilkan model MLP terbaik dalam penelitian ini. Analisis feature importance pada model baseline, model dengan SMOTE, serta model hasil optimasi hyperparameter disajikan pada Gambar 6 dan Tabel 1.



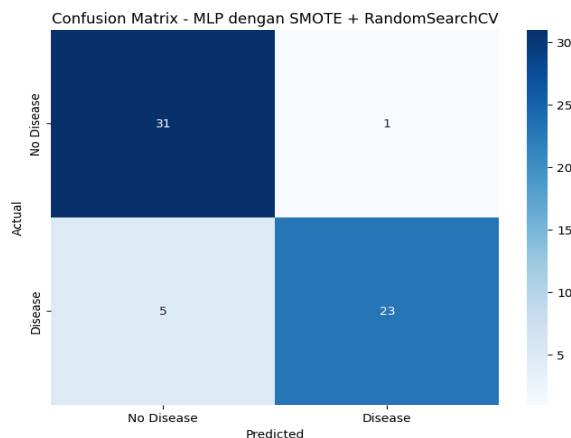
Gambar 6. Analisis feature importance

Tabel 1 Fitur Importance dan Penjelasan

No.	Fitur	Nilai Importance	Makna & Interpretasi
1	ca (jumlah pembuluh darah utama yang terlihat oleh fluoroskopi)	0.293835	Fitur paling dominan. Semakin banyak pembuluh darah yang menunjukkan kelainan atau penyempitan, semakin besar kemungkinan penyakit jantung koroner. <u>Indikator struktural jantung yang sangat kuat.</u>
2	cp (chest pain type)	0.154346	Jenis nyeri dada sangat berpengaruh dalam diagnosis. Angina atau nyeri dada atipikal sering berkaitan dengan risiko penyakit jantung.
3	thal (hasil thalassemia test/penilaian perfusi)	0.139435	Nilai <i>reversible defect</i> pada tes thal menunjukkan gangguan perfusi miokard. Semakin abnormal hasil tes, semakin tinggi risiko penyakit jantung.
4	exang (exercise-induced angina)	0.131868	Angina yang muncul saat olahraga menunjukkan keterbatasan suplai oksigen saat aktivitas, sehingga menjadi prediktor kuat penyakit jantung.
5	thalach (maximum heart rate achieved)	0.123067	Detak jantung maksimum yang rendah sering menandakan fungsi jantung yang kurang optimal. Fitur ini sangat relevan dalam analisis uji beban.

6	fbs (fasting blood sugar >120 mg/dl)	0.115214	Kadar gula puasa tinggi berkaitan dengan diabetes dan kerusakan pembuluh darah. Merupakan faktor risiko metabolik penting pada penyakit jantung.
7	sex (jenis kelamin)	0.115046	Model menangkap pola epidemiologi bahwa laki-laki berada pada risiko lebih tinggi terkena penyakit jantung dibanding perempuan.
8	oldpeak (ST depression akibat latihan)	0.100742	Penurunan segmen ST menunjukkan indikasi iskemia saat stress test. Merupakan faktor yang sangat relevan dalam diagnosis penyakit jantung.
9	trestbps (tekanan darah saat istirahat)	0.090482	Tekanan darah tinggi meningkatkan risiko hipertensi dan komplikasi jantung. Meskipun pengaruhnya tidak sebesar fitur lain, tetap menjadi faktor penting.
10	slope (kemiringan segmen ST)	0.071927	Slope ST yang menurun (<i>downsloping</i>) biasanya merupakan tanda abnormalitas jantung pada EKG, sehingga masih termasuk indikator relevan dalam prediksi.

Confusion matrix dari model terbaik (setelah tuning) memberi wawasan tentang trade-off antara false positive dan false negative – penting dalam konteks klinis (mis: salah memprediksi penyakit bisa sangat berbahaya) dapat dilihat pada gambar 7 Confusion matrix dari model terbaik.



Gambar 7. Confusion matrix dari model terbaik

Pembahasan

Penelitian ini menganalisis kinerja model Multilayer Perceptron (MLP) dalam memprediksi penyakit jantung menggunakan dataset Heart Disease Cleveland. Tahap awal penelitian difokuskan pada eksplorasi data untuk memahami karakteristik variabel, distribusi kelas, serta potensi permasalahan data. Hasil eksplorasi menunjukkan adanya ketidakseimbangan kelas yang cukup signifikan antara pasien dengan penyakit jantung dan tanpa penyakit jantung. Kondisi ini berpotensi menurunkan sensitivitas model dalam mendeteksi kasus positif, sebagaimana umum terjadi pada dataset medis yang bersifat imbalanced.

Pada tahap praproses, dilakukan imputasi nilai median untuk fitur numerik dan modus untuk fitur kategorikal guna menangani missing values. Seluruh fitur kemudian dinormalisasi

menggunakan StandardScaler untuk menstabilkan proses pelatihan dan mempercepat konvergensi jaringan saraf. Variabel kategorikal dikonversi menggunakan Label Encoding agar dapat diproses oleh model MLP secara efektif. Pendekatan praproses ini dipilih berdasarkan rekomendasi literatur yang menekankan pentingnya normalisasi dan penanganan distribusi data klinis yang tidak normal.

Penelitian ini mengevaluasi tiga skenario pemodelan, yaitu MLP baseline tanpa penyeimbangan data, MLP dengan SMOTE, serta MLP dengan kombinasi SMOTE dan optimasi hyperparameter menggunakan RandomizedSearchCV. Model baseline menghasilkan akurasi sebesar 0,85 dan F1-score sebesar 0,830, namun menunjukkan kecenderungan bias terhadap kelas mayoritas. Penerapan SMOTE pada skenario kedua meningkatkan F1-score menjadi 0,836, yang mengindikasikan perbaikan kemampuan model dalam mengenali kelas minoritas, meskipun peningkatannya relatif moderat.

Performa terbaik diperoleh pada skenario ketiga, dengan akurasi mencapai 0,90 dan F1-score sebesar 0,8846. Hasil ini menunjukkan bahwa optimasi hyperparameter berperan penting dalam meningkatkan kinerja MLP secara signifikan. Selain itu, penggunaan RandomizedSearchCV dinilai efisien dalam menemukan parameter optimal dengan biaya komputasi yang lebih rendah. Evaluasi confusion matrix menunjukkan bahwa model terbaik memiliki tingkat false negative yang rendah, yang sangat krusial dalam konteks diagnosis penyakit jantung karena kesalahan klasifikasi pasien sakit dapat berdampak serius secara klinis.

Analisis kontribusi fitur menunjukkan bahwa usia, detak jantung maksimal (thalach), kadar kolesterol (chol), jenis nyeri dada (cp), dan nilai ST-depresi (oldpeak) merupakan prediktor utama penyakit jantung. Temuan ini konsisten dengan literatur medis dan penelitian sebelumnya. Secara keseluruhan, hasil penelitian menegaskan bahwa kombinasi penyeimbangan data dan optimasi hyperparameter mampu meningkatkan kinerja dan stabilitas model MLP dalam prediksi penyakit jantung.

Meskipun demikian, penelitian ini memiliki keterbatasan, terutama terkait ukuran dan sumber dataset yang terbatas serta rendahnya interpretabilitas model MLP. Oleh karena itu, penelitian selanjutnya disarankan untuk menggunakan dataset multi-sumber, menerapkan metode interpretabilitas seperti SHAP atau LIME, serta mengeksplorasi teknik penyeimbangan dan optimasi lanjutan guna meningkatkan generalisasi dan keandalan model dalam konteks klinis nyata.

SIMPULAN

Secara keseluruhan, penelitian ini menunjukkan bahwa kombinasi preprocessing yang tepat, penyeimbangan data menggunakan SMOTE, dan optimasi hyperparameter mampu meningkatkan performa model MLP secara signifikan. Temuan ini mendukung berbagai studi sebelumnya yang menegaskan pentingnya pendekatan berbasis level data dan level model dalam membangun sistem prediksi medis yang handal. Untuk pengembangan selanjutnya, validasi klinis nyata dan peningkatan interpretabilitas model menjadi langkah penting sebelum model dapat diimplementasikan dalam sistem pendukung keputusan klinis.

REFERENSI

- Adegun, A. A., & Viriri, S. (2021). Deep learning-based system for heart disease prediction. *Computers, Materials & Continua*, 69(1), 19–34.
- Ahn, J., Cho, S., & Kim, D. (2023). A robust feature scaling strategy for improving neural network convergence. *IEEE Transactions on Neural Networks and Learning Systems*, 34(9), 6589–6601.
- Ali, S., Hassan, A., & Tariq, M. (2022). Enhancing the performance of neural networks using advanced normalization techniques. *Expert Systems with Applications*, 204, 117597. <https://doi.org/10.1016/j.eswa.2022.117597>

- Aryuni, M., et al. (2023). Imbalanced learning in heart disease categorization: Improving minority class prediction accuracy using the SMOTE algorithm. *International Journal of Fuzzy Logic and Intelligent Systems*, 23(2), 140–151. <https://doi.org/10.5391/IJFIS.2023.23.2.140>
- Deepika, D. (2022). Heart disease prediction using machine learning techniques for earlier diagnosis: Heart disease classified using novel hybrid MLP-EBMDA method. *Expert Systems with Applications*, 174, 102915. <https://doi.org/10.1016/j.eswa.2021.100915>
- Firdaus, R., Mualfah, D., & Hasanah, J. S. (2023). Klasifikasi Multi-Class Penyakit Jantung dengan SMOTE dan Pearson's Correlation menggunakan MLP. *Jurnal CoSciTech (Computer Science and Information Technology)*, 4(1). <https://doi.org/10.37859/coscitech.v4i1.4769>
- Gao, C., Chen, Q., & Li, Y. (2021). Handling imbalanced medical datasets using hybrid resampling methods. *BMC Medical Informatics and Decision Making*, 21(1), 1–12. <https://doi.org/10.1186/s12911-021-01617-0>
- García-Florian, A., Cruz, J., & Castejón-Limas, M. (2023). Machine learning approaches for handling imbalanced clinical data. *Artificial Intelligence in Medicine*, 144, 102670. <https://doi.org/10.1016/j.artmed.2023.102670>
- Ibrahim, H., Omar, M., & Hassan, R. (2023). A machine learning framework for identifying key predictors of heart disease. *Scientific Reports*, 13, 2321. <https://doi.org/10.1038/s41598-023-29557-1>
- Karthikeyan, R., & Ranganathan, K. (2021). Impact of missing data imputation techniques on heart disease prediction. *Journal of King Saud University – Computer and Information Sciences*, 33(6), 701–709. <https://doi.org/10.1016/j.jksuci.2019.02.002>
- Khan, A., Rahman, S., & Al Mamun, M. (2024). Improving sensitivity in cardiovascular disease prediction models using optimized neural networks. *Computers in Biology and Medicine*, 170, 107832. <https://doi.org/10.1016/j.combiomed.2024.107832>
- Kose, U., Arslan, A., & Demir, I. (2023). Optimized deep neural networks for medical diagnosis: A comparative analysis. *Biomedical Signal Processing and Control*, 86, 104944. <https://doi.org/10.1016/j.bspc.2023.104944>
- Kwakye, K., & Dadzie, E. (2021). Machine learning-based classification algorithms for the prediction of coronary heart diseases. *arXiv*.
- Patel, J., & Nandhu, R. (2022). Efficient hyperparameter optimization using randomized search for neural networks. *International Journal of Intelligent Systems and Applications*, 14(2), 12–22.
- Putri, S. K. N. A., Jumiati, I., Sulistia, I., Saputra, N. A. B., & Wiranda, N. (2025). Penerapan hyperparameter tuning pada model klasifikasi untuk prediksi risiko penyakit jantung. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 5(4), 1181–1189. <https://doi.org/10.57152/malcom.v5i4.2138>
- Ramadhani, N., & Nugroho, A. (2022). SMOTE-based improvement for prediction of heart disease using machine learning. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 20(5), 1124–1132. <https://doi.org/10.12928/telkomnika.v20i5.22067>
- Sharma, N., & Malviya, L. (2023). A hybrid deep neural net learning model for predicting coronary heart disease using randomized search cross-validation optimization. *Decision Analytics Journal*, 9, 100331. <https://doi.org/10.1016/j.dajour.2023.100331>
- Supriyanto, H., Hariguna, T., & Barkah, A. S. (2025). MLP model optimization for heart attack risk prediction: A systematic literature review. *Sinkron: Jurnal dan Penelitian Teknik Informatika*, 9(3), 1811–18227. <https://doi.org/10.33395/sinkron.v9i3.15027>
- Susanto, E. R., & Saputra, E. (2025). Implementation of deep learning with multilayer perceptron (MLP) for heart disease prediction using the SMOTE-ENN technique.

- Journal of Applied Informatics and Computing*, 9(3), 1034–1041.
<https://doi.org/10.30871/jaic.v9i3.9337>
- Thapa, B., Khanal, S., & Sharma, P. (2022). Machine learning models with hyperparameter tuning for accurate disease prediction. *Informatics in Medicine Unlocked*, 32, 101020.
<https://doi.org/10.1016/j.imu.2022.101020>
- Wang, Y., Liu, L., & Sun, X. (2023). Systematic hyperparameter optimization of neural networks for healthcare prediction. *IEEE Access*, 11, 45012–45025.
- Wijayanto, D., Marco, R., Sidauruk, A., & Sulistiyono, M. (2025). The effect of SMOTE and Optuna hyperparameter optimization on TabNet performance for heart disease classification. *SISFOKOM*, 14(2), 156–164.
<https://doi.org/10.32736/sisfokom.v14i2.2348>
- Zhai, X., Wang, J., & Li, Z. (2022). Managing class imbalance in healthcare datasets: A comprehensive review. *Computer Methods and Programs in Biomedicine*, 221, 106882. <https://doi.org/10.1016/j.cmpb.2022.106882>