Journal of Applied Informatics and Computing (JAIC)
Vol.9, No.3, June 2025, pp. 1034~1041
e-ISSN: 2548-6861

Implementation of Deep Learning with Multilayer Perceptron (MLP) for Heart Disease Prediction Using the SMOTE-ENN Technique

Erik Saputra [1], Erliyan Redy Susanto [2*]
[1] Sistem Informasi, Fakultas Teknik dan Ilmu Komputer, Universitas Teknokrat Indonesia, Indonesia
[2] Magister Ilmu Komputer, Fakultas Teknik dan Ilmu Komputer, Universitas Teknokrat Indonesia, Indonesia
erik_saputra@teknokrat.ac.id [1], erliyan.redy@teknokrat.ac.id [2*]

Article Info

Article history:

Received 2025-03-25
Revised 2025-05-06
Accepted 2025-05-07

Kata Kunci:

Heart Disease Prediction,
Deep Learning,
Multilayer Perceptron (MLP),
SMOTE-ENN,
Data Imbalance.

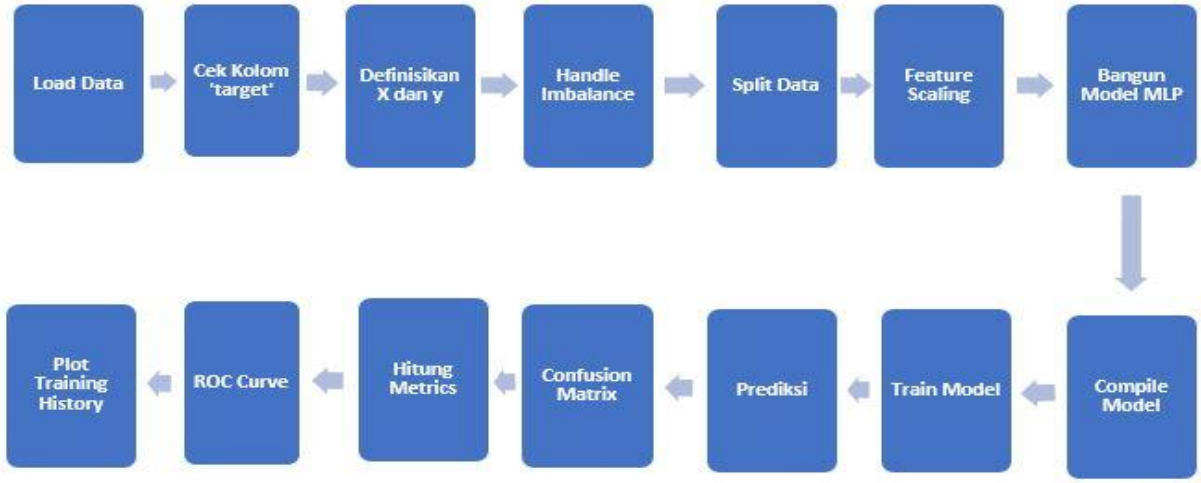
ABSTRAK

Heart disease is a leading cause of global mortality, with its prevalence increasing annually. This study aims to develop a heart disease prediction model using a Multilayer Perceptron (MLP) combined with the SMOTE-ENN resampling technique to address data imbalance issues. The dataset used was obtained from the UCI Machine Learning Repository and includes patients' clinical and demographic features. The initial dataset consisted of [number of data] records, with an imbalanced class distribution between patients with and without heart disease. After applying SMOTE-ENN, the class distribution became more balanced, allowing the model to learn patterns more effectively. The MLP model was designed with two hidden layers comprising 64 and 32 neurons, respectively, using the ReLU activation function in the hidden layers and a sigmoid function in the output layer. Evaluation results showed that the model achieved an accuracy of 89.47%, precision of 77.78%, recall of 100%, and an F1-score of 87.5%. To validate the effectiveness of SMOTE-ENN, comparisons were made with other methods such as SMOTE and undersampling, as well as baseline models like Logistic Regression and Decision Tree. The results demonstrate that SMOTE-ENN outperforms other techniques in handling class imbalance, leading to better overall model performance.

I. PENDAHULUAN

Penyakit jantung merupakan salah satu penyebab utama kematian secara global, dengan angka prevalensi yang terus meningkat setiap tahunnya[1]. Menurut data Organisasi Kesehatan Dunia (WHO)[2], penyakit kardiovaskular bertanggung jawab atas sekitar 17,9 juta kematian per tahun, atau setara dengan 31% dari total kematian di dunia[3]. Fenomena ini diperparah oleh gaya hidup modern yang kurang sehat, seperti pola makan tidak seimbang, kurangnya aktivitas fisik, serta kebiasaan merokok. Deteksi dini penyakit jantung menjadi krusial untuk mengurangi risiko kematian dan meningkatkan kualitas hidup pasien[4]. Namun, diagnosis manual sering kali menghadapi tantangan seperti ketidakakuratan dan ketergantungan pada pengalaman klinisi, sehingga diperlukan pendekatan berbasis kecerdasan buatan untuk meningkatkan akurasi prediksi[5].

Permasalahan utama dalam prediksi penyakit jantung adalah ketidakseimbangan data (imbalanced data) dan tingginya variasi fitur medis yang memengaruhi hasil diagnosis[6]. Beberapa penelitian sebelumnya telah mengimplementasikan metode machine learning seperti *Support Vector Machine* (SVM), Logistic Regression, dan *Artificial Neural Network* (ANN), namun hasilnya masih memiliki keterbatasan dalam hal akurasi dan kemampuan generalisasi. Pada studi sebelumnya metode SVM mencapai akurasi sekitar 85%, sedangkan Logistic Regression dan ANN masing-masing memiliki akurasi di kisaran 82-88%[7]. Oleh karena itu, diperlukan pengembangan model yang lebih robust untuk meningkatkan performa prediksi[8].



Gambar 1. Tahapan Penelitian

Studi ini mengusulkan penggunaan Multilayer Perceptron (MLP) yang dikombinasikan dengan teknik resampling SMOTEENN untuk mengatasi masalah ketidakseimbangan data[9]. MLP dipilih karena kemampuannya menangkap pola kompleks dalam data melalui arsitektur jaringan saraf terdalam[10]. Hasil penelitian menunjukkan bahwa model yang dibangun mencapai akurasi 0.89%, precision 0.77%, recall 100%, specificity 0.83%, dan F1-score 0.87%, yang lebih unggul dibandingkan dengan pendekatan sebelumnya. Tingginya nilai recall mengindikasikan bahwa model ini sangat efektif dalam mengidentifikasi pasien berisiko, sehingga dapat menjadi alat pendukung keputusan medis yang andal.

Kontribusi penelitian ini terletak pada peningkatan performa prediksi penyakit jantung melalui optimasi MLP dan penanganan data tidak seimbang, yang belum banyak dieksplorasi dalam studi-studi sebelumnya[11]. Selain itu, hasil penelitian ini dapat menjadi referensi bagi pengembangan sistem diagnosis berbasis AI di fasilitas kesehatan[12], sehingga membantu tenaga medis dalam mengambil keputusan yang lebih cepat dan akurat. Dengan demikian, implementasi model ini diharapkan dapat mengurangi angka kesalahan diagnosis dan meningkatkan efektivitas penanganan pasien penyakit jantung[13].

II. METODE

Penelitian ini dimulai dengan pengumpulan data dari dataset Heart Disease, yang berisi berbagai faktor risiko terkait penyakit jantung[14]. Setelah data dikumpulkan, langkah pertama adalah memeriksa kolom target untuk memastikan bahwa label yang menunjukkan adanya penyakit jantung sudah sesuai dan siap diproses[15]. Selanjutnya, variabel independen (X) dan dependen (y) didefinisikan, di mana X terdiri dari fitur-fitur prediktor dan y adalah kolom target yang telah diperiksa sebelumnya.

A. Load Data

Dataset yang digunakan dalam penelitian ini adalah dataset Heart Disease, yang merupakan kumpulan data kesehatan pasien yang berkaitan dengan indikator penyakit jantung[16]. Dataset ini mencakup berbagai fitur klinis dan demografis seperti usia, jenis kelamin, tekanan darah (*resting blood pressure*), kadar kolesterol serum (*serum cholesterol*), kadar gula darah puasa (*fasting blood sugar*), hasil elektrokardiografi (*resting electrocardiographic results*), detak jantung maksimum yang dicapai (*maximum heart rate achieved*), serta adanya angina yang diinduksi oleh olahraga (*exercise-induced angina*). Selain itu, dataset ini juga mencakup informasi tambahan seperti depresi ST yang diinduksi oleh olahraga relatif terhadap istirahat (*ST depression induced by exercise relative to rest*) dan kemiringan segmen ST puncak latihan (*slope of the peak exercise ST segment*).

Kolom target dalam dataset ini menunjukkan adanya penyakit jantung, di mana nilai 1 menunjukkan bahwa pasien memiliki penyakit jantung dan nilai 0 menunjukkan bahwa pasien tidak memiliki penyakit jantung. Dataset ini diperoleh dari repositori data publik seperti UCI *Machine Learning Repository* atau sumber lain yang relevan, dan telah digunakan secara luas dalam penelitian terkait prediksi penyakit jantung.

Sebelum digunakan, dataset ini melalui proses pra-pemrosesan awal seperti penanganan missing value, normalisasi, dan encoding variabel kategorikal untuk memastikan kualitas data yang baik sebelum dimasukkan ke dalam model. Penggunaan dataset ini diharapkan dapat memberikan gambaran yang komprehensif tentang faktor-faktor yang berkontribusi terhadap penyakit jantung, sehingga memungkinkan pengembangan model prediktif yang akurat.

B. Kolom Target

Sebelum memproses data lebih lanjut, langkah pertama yang dilakukan adalah memeriksa distribusi dari kolom

target. Kolom target dalam dataset ini berfungsi sebagai indikator utama yang menentukan apakah seorang pasien memiliki penyakit jantung (dengan nilai 1) atau tidak (dengan nilai 0). Pemeriksaan ini penting untuk mengidentifikasi apakah dataset mengalami masalah ketidakseimbangan kelas (*class imbalance*), di mana satu kelas mungkin memiliki jumlah sampel yang jauh lebih banyak dibandingkan kelas lainnya.

Ketidakseimbangan kelas dapat memengaruhi performa model machine learning, karena model cenderung bias terhadap kelas mayoritas dan kurang mampu memprediksi kelas minoritas dengan akurat. Untuk mengatasi hal ini, dilakukan analisis distribusi kelas dengan menghitung jumlah sampel untuk setiap kelas. Jika ditemukan ketidakseimbangan yang signifikan, langkah-langkah penanganan seperti teknik oversampling, undersampling, atau kombinasi keduanya (seperti SMOTEENN) akan dipertimbangkan untuk menyeimbangkan distribusi kelas[17].

Selain itu, pemeriksaan kolom target juga melibatkan analisis statistik deskriptif untuk memahami karakteristik data, seperti persentase pasien yang memiliki penyakit jantung dibandingkan dengan yang tidak. Informasi ini tidak hanya membantu dalam memahami struktur dataset, tetapi juga memberikan wawasan awal yang berguna untuk menentukan pendekatan pemodelan yang tepat. Dengan demikian, pemeriksaan kolom target merupakan langkah kritis yang memastikan bahwa dataset siap untuk diproses lebih lanjut tanpa bias yang signifikan.

C. *Variabel*

Dalam penelitian ini, variabel dibagi menjadi dua kategori utama: variabel independen (X) dan variabel dependen (y). Variabel independen mencakup seluruh fitur prediktor yang digunakan untuk memprediksi penyakit jantung, meliputi parameter klinis seperti tekanan darah sistolik (trestbps), kadar kolesterol serum (chol), detak jantung maksimum (thalach), serta hasil elektrokardiografi (restecg). Selain itu, termasuk pula variabel demografis seperti usia (age) dan jenis kelamin (sex), serta indikator terkait olahraga seperti angina yang diinduksi latihan (exang) dan depresi ST (oldpeak). Setiap variabel ini dipilih berdasarkan relevansi klinisnya terhadap diagnosis penyakit jantung dan kajian literatur sebelumnya. Sementara itu, variabel dependen (y) merupakan kolom target yang merepresentasikan diagnosis penyakit jantung dalam bentuk biner, dimana nilai 1 menunjukkan adanya penyakit jantung dan nilai 0 menunjukkan tidak adanya penyakit jantung. Definisi yang jelas terhadap kedua kelompok variabel ini sangat penting untuk memastikan konsistensi dalam proses pemodelan dan interpretasi hasil. Selain itu, pemisahan variabel ini juga memfasilitasi tahapan analisis selanjutnya seperti normalisasi data, pembagian dataset, dan evaluasi model secara tepat.

TABEL 1.
VARIABEL MENAMPILKAN FITUR-FITUR UTAMA YANG DIGUNAKAN DALAM ANALISIS INI.

Variabel	**Keterangan**
Age	Usia pasien (dalam tahun)
Sex	Jenis kelamin (1 = laki-laki, 0 = perempuan)
chest_pain_type	Tipe nyeri dada (0–3, dengan makna tertentu)
resting_bp	Tekanan darah saat istirahat (mmHg)
Cholestoral	Kadar kolesterol dalam darah (mg/dL)
fasting_blood_sugar	Gula darah setelah puasa (>120 mg/dL: 1, lainnya: 0)
Restecg	Hasil elektrokardiografi saat istirahat (0–2)
max_hr	Detak jantung maksimum yang tercapai
Exang	Angina akibat olahraga (1 = ya, 0 = tidak)
Oldpeak	Depresi segmen ST yang diinduksi oleh olahraga
Slope	Kemiringan segmen ST (0–2)
num_major_vessels	Jumlah pembuluh darah utama terdeteksi fluoroskopi
Thal	Hasil tes thalassemia (1 = normal, 2 = cacat tetap, 3 = cacat reversibel)
Target	Diagnosis penyakit jantung (1 = memiliki penyakit, 0 = tidak)

SMOTEENN (*Synthetic Minority Over-sampling Technique + Edited Nearest Neighbors*)

Metode SMOTEENN mengkombinasikan SMOTE untuk oversampling dan ENN untuk undersampling:

SMOTE:

Menambahkan data sintetis untuk kelas minoritas menggunakan interpolasi antara titik data minoritas yang ada.

$$\mathcal{X}_{new} = \mathcal{X}_i + \lambda (\mathcal{X}_j - \mathcal{X}_i), \lambda \sim U(0, 1)$$

di mana x_i dan x_j adalah dua titik minoritas yang dipilih secara acak, dan λ adalah bilangan acak antara 0 dan 1.

Edited Nearest Neighbors (ENN):

Menghapus sampel yang diklasifikasikan salah oleh k-NN (biasanya dengan $k = 3$).

Normalisasi (*StandardScaler*)

Fitur diskalakan menggunakan ***StandardScaler***, yang melakukan transformasi:

$$X^I = \frac{X - \mu}{\sigma}$$

di mana:
- μ adalah rata-rata dari fitur,
- σ adalah standar deviasi dari fitur.

Multi-Layer Perceptron (MLP)

Jaringan saraf buatan yang digunakan adalah MLP dengan *Dense Layers*.

Layer pertama:

$$h_1 = f(W_1 X + b_1)$$

Dropout**:**

Mengatur neuron dengan probabilitas ppp, untuk mencegah overfitting.

Hidden layer:

$$h_2 = f(W_2 h_1 + b_2)$$

Output layer:

$$\hat{y} = \sigma(W_3 h_2 + b_3)$$

dengan fungsi aktivasi sigmoid**:**

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

Loss Function (Binary Cross-Entropy)

Karena ini adalah klasifikasi biner, digunakan fungsi loss *binary cross-entropy***:**

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]$$

di mana:
y_i adalah nilai sebenarnya (0 atau 1),
y_i adalah prediksi model.

Evaluasi Model

Akurasi:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Presisi:

$$Precision = \frac{TP}{TP + FP}$$

Recall (Sensitivitas):

$$Recall = \frac{TP}{TP + FP}$$

Spesifisitas:

$$Specificity = \frac{TN}{TN + FP}$$

F1-score:

$$F1 = 2 \times \frac{Precicion \times Recall}{Precision + Recall}$$

ROC Curve dan AUC

ROC curve menggambarkan True Positive Rate (TPR) vs False Positive Rate (FPR)**:**

TPR (Recall):

$$TPR = \frac{TP}{TP + FN}$$

FPR:

$$FPR = \frac{FP}{FP + TN}$$

AUC (Area Under Curve) dihitung sebagai integral dari kurva ROC:

$$AUC = \int_0^1 TPR(FPR)\, d(FPR)$$

D. Handle Imbalance

Ketidakseimbangan data (*class imbalance*) merupakan masalah umum dalam dataset klasifikasi, di mana jumlah sampel pada satu kelas (biasanya kelas minoritas) jauh lebih sedikit dibandingkan kelas lainnya (kelas mayoritas)[18]. Dalam konteks prediksi penyakit jantung, ketidakseimbangan ini dapat terjadi jika jumlah pasien yang memiliki penyakit jantung (kelas minoritas) jauh lebih sedikit daripada pasien yang tidak memiliki penyakit jantung (kelas mayoritas). Jika tidak ditangani, model machine learning cenderung bias terhadap kelas mayoritas, sehingga performa prediksi untuk kelas minoritas menjadi kurang akurat.

Untuk mengatasi masalah ini, penelitian ini menerapkan teknik SMOTE-ENN (*Synthetic Minority Over-sampling Technique combined with Edited Nearest Neighbors*)[19]. SMOTEENN adalah metode hybrid yang menggabungkan dua pendekatan: oversampling pada kelas minoritas dan undersampling pada kelas mayoritas[20].

SMOTE (Synthetic Minority Over-sampling Technique)

Teknik ini menciptakan sampel sintetis untuk kelas minoritas dengan cara menginterpolasi antara sampel-sampel yang sudah ada. Dengan demikian, jumlah sampel pada kelas minoritas meningkat, sehingga mengurangi ketidakseimbangan[21].

ENN (Edited Nearest Neighbors)

Setelah oversampling, teknik ENN digunakan untuk membersihkan data dengan menghapus sampel yang dianggap sebagai noise atau outlier, terutama dari kelas mayoritas[22].

Dengan menggabungkan kedua teknik ini, SMOTEENN tidak hanya menyeimbangkan distribusi kelas tetapi juga meningkatkan kualitas data dengan mengurangi noise[23]. Hasilnya adalah dataset yang lebih seimbang dan bersih, yang siap digunakan untuk melatih model machine learning[24]. Pendekatan ini diharapkan dapat meningkatkan kemampuan model dalam memprediksi kelas minoritas (pasien dengan penyakit jantung) tanpa mengorbankan performa secara keseluruhan[25].

E. Split Data

Setelah melakukan penanganan ketidakseimbangan data, langkah selanjutnya adalah membagi dataset menjadi dua subset utama: data latih (training data) **dan** data uji (testing data). Pembagian ini dilakukan dengan proporsi yang umum

digunakan dalam penelitian *machine learning*, yaitu 80% untuk data latih dan 20% untuk data uji[26].

Data latih digunakan untuk melatih model machine learning agar dapat mempelajari pola antara fitur dan label target, sedangkan data uji digunakan untuk mengevaluasi kemampuan model dalam memprediksi data baru. Pembagian data dilakukan secara acak namun tetap menjaga keseimbangan distribusi kelas dengan menggunakan teknik stratifikasi, sehingga kedua subset memiliki representasi yang proporsional dari setiap kelas[27]. Hal ini mencegah terjadinya bias dan memastikan model tidak hanya baik pada data latih, tapi juga mampu menggeneralisasi dengan baik pada data yang belum pernah dilihat sebelumnya. Dengan demikian, pembagian data secara tepat sangat penting untuk menjamin akurasi dan keandalan model prediksi penyakit jantung.

F. Feature Scaling

Setelah dataset *dibagi* menjadi data latih dan data uji, langkah selanjutnya adalah melakukan penskalaan fitur *(feature scaling*)**.** Penskalaan fitur adalah proses menyesuaikan nilai-nilai fitur dalam dataset agar berada dalam rentang yang serupa[28]. Hal ini penting karena fitur-fitur dalam dataset sering kali memiliki skala yang berbeda-beda. Misalnya, usia pasien mungkin berkisar antara 0 hingga 100, sedangkan kadar kolesterol serum bisa mencapai ratusan atau bahkan ribuan. Perbedaan skala ini dapat menyebabkan ketidakstabilan dalam proses pelatihan model, terutama pada algoritma yang sensitif terhadap besaran nilai, seperti *Multilayer Perceptron* (MLP).

Dalam penelitian ini, teknik penskalaan yang digunakan adalah normalisasi atau standardisasi.

Normalisasi:

Teknik ini mengubah nilai fitur ke dalam rentang tertentu, biasanya antara 0 dan 1. Normalisasi cocok digunakan ketika distribusi data tidak mengikuti distribusi normal (Gaussian). Rumus yang umum digunakan untuk normalisasi adalah:

$$X_{normalized} = \frac{X - X_{min}}{X_{\max} - X_{min}}$$

di mana X_{min} dan X_{max} adalah nilai minimum dan maksimum dari fitur tersebut.

*Standardisasi**:***

Teknik ini mengubah nilai fitur sehingga memiliki mean (rata-rata) 0 dan standar deviasi 1. Standardisasi lebih cocok digunakan ketika data mengikuti distribusi normal atau ketika algoritma yang digunakan mengasumsikan distribusi Gaussian. Rumus standardisasi adalah:

$$X_{standardized} = \frac{X - \mu}{\sigma}$$

di mana μ adalah mean dan σ adalah standar deviasi dari fitur tersebut.

Penskalaan fitur dilakukan secara terpisah pada data latih dan data uji untuk menghindari data *leakage*, yaitu kebocoran informasi dari data uji ke data latih yang dapat menyebabkan overestimasi performa model. Dengan melakukan penskalaan fitur, model MLP dapat lebih cepat mencapai konvergensi selama pelatihan, dan performa prediksi dapat ditingkatkan karena semua fitur berkontribusi secara seimbang dalam proses pembelajaran. Dengan demikian, penskalaan fitur merupakan langkah penting yang memastikan stabilitas dan efisiensi dalam pelatihan model deep learning.

III. HASIL DAN PEMBAHASAN

Model Multilayer Perceptron (MLP) yang menggunakan teknik SMOTEENN untuk menangani ketidakseimbangan data menunjukkan kinerja yang baik dalam memprediksi penyakit jantung. Model mencapai akurasi 89.47%, presisi 77.78%, dan recall 100%, yang menunjukkan kemampuan luar biasa dalam mendeteksi semua pasien yang sakit tanpa melewatkan satu pun kasus (false negatives). Namun, spesifisitas sebesar 83.33% mengindikasikan masih adanya false positives, di mana beberapa pasien sehat diprediksi salah sebagai sakit. F1-score sebesar 87.5% mencerminkan keseimbangan yang baik antara presisi dan recall, meskipun ada ruang untuk perbaikan terutama dalam mengurangi false positives. Kurva ROC dengan AUC yang tinggi menunjukkan model mampu membedakan pasien sakit dan sehat dengan efektif. Untuk pengembangan lebih lanjut, penyesuaian threshold prediksi, tuning hyperparameter, atau regularisasi dapat diterapkan guna meningkatkan presisi dan spesifisitas sehingga model menjadi lebih optimal dalam aplikasi medis.

A. SMOTEENN (Synthetic Minority Over-sampling Technique + Edited Nearest Neighbors)

SMOTEENN adalah teknik yang sangat berguna untuk menangani ketidakseimbangan kelas dalam dataset medis, termasuk prediksi penyakit jantung.

TABEL 2.
TEKNIK SMOTE-ENN

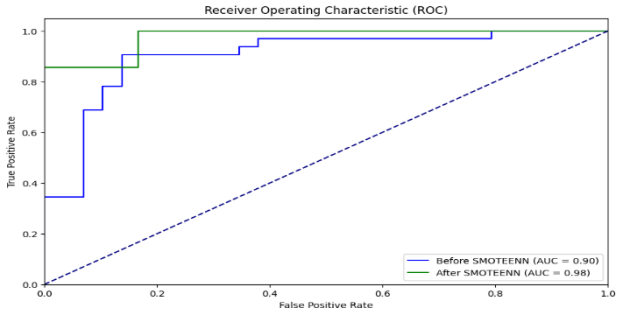
Matrix	Sebelum SMOTEENN	Sesudah SMOTEENN
Accuracy	0.86	0.89
Precision	0.87	0.77
Recall	0.87	1.00
F1-Score	0.87	0.87
ROC AUC	0.90	0.97

Tabel 2 menunjukkan bahwa setelah menerapkan teknik SMOTEENN meningkatkan performa model dalam mendeteksi kelas minoritas. Akurasi naik dari 0.86 menjadi 0.89, dan ROC AUC meningkat signifikan dari 0.90 menjadi 0.97, menunjukkan kemampuan model yang lebih baik dalam membedakan kelas. Recall mencapai 1.00, artinya semua contoh kelas minoritas terdeteksi, tetapi presisi turun dari 0.87 menjadi 0.77 akibat meningkatnya false positives. F1-Score stabil di 0.87. Secara keseluruhan, SMOTEENN efektif untuk

meningkatkan recall dengan trade-off pada presisi, sehingga cocok digunakan jika prioritas utama adalah deteksi penuh kelas minoritas.

B. Receiver Operating Characteristic (ROC)

ROC adalah alat penting dalam evaluasi diagnostik penyakit jantung karena membantu menilai akurasi tes secara visual dan kuantitatif melalui AUC. Dengan ROC, dokter dapat membuat keputusan yang lebih baik dalam diagnosis dan pengobatan pasien dapat dilihat pada gambar 2.

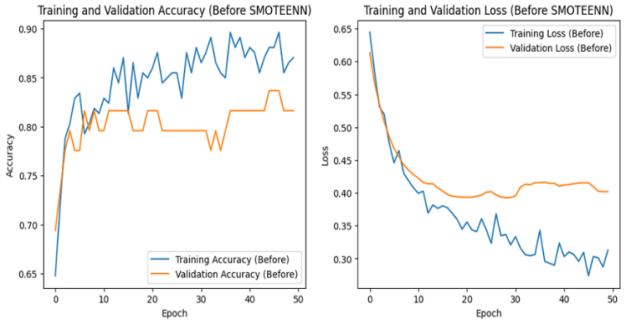


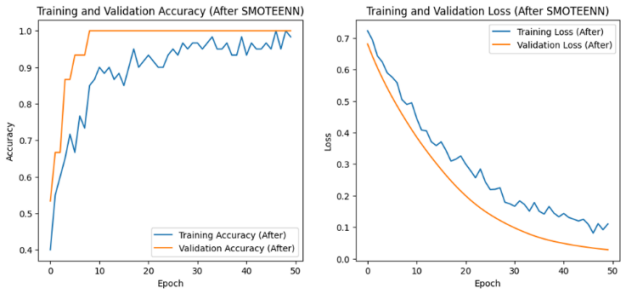
Gambar 2 Receiver operating characteristic (ROC)

Dengan demikian dapat disimpulkan hasil penelitian ini menunjukkan bahwa Multilayer Perceptron (MLP) dengan teknik SMOTEENN memberikan kinerja yang sangat baik untuk memprediksi penyakit jantung yang diperkuat dengan nilai ROC-AUC Sebelum SMOTEENN (AUC = 0.90), Sesudah SMOTEENN (AUS = 0.98)

C. Training And Validation Accuracy

Model mengalami overfitting sebelum penerapan SMOTEENN, seperti yang ditunjukkan oleh perbedaan signifikan antara akurasi dan loss pada data pelatihan dan validasi. Teknik SMOTEENN diharapkan dapat meningkatkan keseimbangan kelas dan kemampuan generalisasi model. Dapat dilihat perbedaan antara sebelum dan sesudah pada gambar 3.

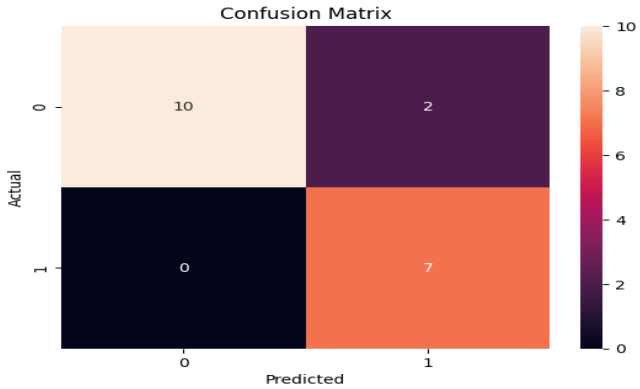




Gambar 3. Training And Validation Accuracy

D. Confusion Matrix

Berdasarkan confusion matrix yang diperoleh, model memiliki akurasi sebesar 89.47%, yang menunjukkan bahwa model mampu mengklasifikasikan sebagian besar sampel dengan benar, baik untuk pasien dengan maupun tanpa penyakit jantung. Dalam hal presisi, model mencatat nilai sebesar 77.78%, yang berarti dari semua pasien yang diprediksi memiliki penyakit jantung, sekitar 77.78% memang benar-benar sakit, sementara sisanya merupakan false positives. Recall atau sensitivitas mencapai 100%, menunjukkan bahwa model tidak melewatkan satu pun kasus penyakit jantung, yang sangat penting dalam aplikasi medis untuk memastikan semua pasien yang berisiko terdeteksi. Dapat dilihat pada gambar 4.



Gambar 4. Confusion Matrix Multilayer Perceptron (MLP) dengan teknik SMOTE-ENN

Selain itu, specificity model tercatat sebesar 83.33%, yang berarti model cukup baik dalam mengenali pasien sehat, meskipun masih terdapat beberapa kasus false positives. Terakhir, F1-score dihitung sebesar 87.5%**,** yang mencerminkan keseimbangan antara presisi dan recall, menjadikannya indikator yang baik terhadap performa keseluruhan model. Dengan recall yang sangat tinggi, model sangat andal dalam mendeteksi penyakit jantung, namun adanya false positives menunjukkan bahwa masih ada ruang untuk perbaikan dalam meningkatkan presisi dan specificity.

TABEL 3.
HASIL PENGUJIAN PEMBAGIAN DATA UJI

Confusion Matrix	Multilayer Perceptron (MLP)
Accuracy	0.89
Precision	0.77
Recall	1.00
Specificity	0.83
F1-score	0.87

Penelitian ini mengembangkan model prediksi penyakit jantung menggunakan dataset "heart.csv" dengan pendekatan deep learning berbasis Multi-Layer Perceptron (MLP). Setelah memvalidasi keberadaan kolom target, ketidakseimbangan data diatasi menggunakan SMOTEENN. Dataset dibagi menjadi data latih dan uji (80:20), lalu fitur dinormalisasi menggunakan StandardScaler. Model MLP terdiri dari tiga lapisan dengan aktivasi ReLU dan sigmoid, serta dropout untuk mengurangi overfitting. Optimasi dilakukan menggunakan Adam dengan learning rate 0.001 dan fungsi loss binary cross-entropy selama 50 epoch. Evaluasi mencakup akurasi, presisi, recall, F1-score, ROC, dan AUC, dengan hasil menunjukkan performa baik dan stabil. Penelitian ini membuktikan efektivitas MLP dalam memprediksi penyakit jantung, terutama dengan penanganan ketidakseimbangan data menggunakan SMOTEENN.

IV. KESIMPULAN

Penelitian ini mengimplementasikan model Multilayer Perceptron (MLP) dengan teknik SMOTEENN untuk memprediksi penyakit jantung menggunakan dataset yang tidak seimbang. Hasilnya menunjukkan peningkatan signifikan dalam performa model, terutama pada recall (sensitivitas) yang mencapai 100% , sehingga semua kasus penyakit jantung dapat terdeteksi. Model juga mencatatkan akurasi 89.47% , presisi 77.78% , spesifisitas 83.33% , dan F1-Score 87.5% , menunjukkan keseimbangan yang baik antara kemampuan mendeteksi kasus positif dan menghindari kesalahan prediksi. Meskipun demikian, masih terdapat beberapa false positives yang perlu diperbaiki melalui penyesuaian threshold, regularisasi, tuning hyperparameter, atau penambahan data. Dengan pengembangan lebih lanjut, model ini berpotensi menjadi alat bantu diagnosis dini yang efektif bagi tenaga medis, membantu mengurangi beban penyakit jantung secara global. Secara keseluruhan, kombinasi MLP dan SMOTEENN terbukti efektif untuk dataset tidak seimbang, dengan implikasi penting di bidang informatika medis dan aplikasi prediksi lainnya.

DAFTAR PUSTAKA

[1] O. Gaidai, Y. Cao, and S. Loginov, "Global Cardiovascular Diseases Death Rate Prediction," *Curr. Probl. Cardiol.*, vol. 48, no. 5, p. 101622, May 2023, doi: 10.1016/j.cpcardiol.2023.101622.

[2] S. Sidaria, E. Huriani, and S. D. Nasution, "Self Care dan Kualitas Hidup Pasien Penyakit Jantung Koroner," *JIK J. ILMU Kesehat.*, vol. 7, no. 1, p. 41, Apr. 2023, doi: 10.33757/jik.v7i1.631.

[3] A. A. Robert and M. A. Al Dawish, "Cardiovascular Disease among Patients with Diabetes: The Current Scenario in Saudi Arabia," *Curr. Diabetes Rev.*, vol. 17, no. 2, pp. 180–185, Feb. 2021, doi: 10.2174/1573399816666200527135512.

[4] V. Shorewala, "Early detection of coronary heart disease using ensemble techniques," *Informatics Med. Unlocked*, vol. 26, p. 100655, 2021, doi: 10.1016/j.imu.2021.100655.

[5] N. Ghaffar Nia, E. Kaplanoglu, and A. Nasab, "Evaluation of artificial intelligence techniques in disease diagnosis and prediction," *Discov. Artif. Intell.*, vol. 3, no. 1, p. 5, Jan. 2023, doi: 10.1007/s44163-023-00049-5.

[6] M. M. Ahsan and Z. Siddique, "Machine learning-based heart disease diagnosis: A systematic literature review," *Artif. Intell. Med.*, vol. 128, p. 102289, Jun. 2022, doi: 10.1016/j.artmed.2022.102289.

[7] F. Handayani, "Komparasi Support Vector Machine, Logistic Regression Dan Artificial Neural Network Dalam Prediksi Penyakit Jantung," *J. Edukasi dan Penelit. Inform.*, vol. 7, no. 3, p. 329, Dec. 2021, doi: 10.26418/jp.v7i3.48053.

[8] S. Chen, E. Xie, C. Ge, R. Chen, D. Liang, and P. Luo, "CycleMLP: A MLP-Like Architecture for Dense Visual Predictions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 12, pp. 14284–14300, Dec. 2023, doi: 10.1109/TPAMI.2023.3303397.

[9] H. Guan, Y. Zhang, M. Xian, H. D. Cheng, and X. Tang, "SMOTE-WENN: Solving class imbalance and small sample problems by oversampling and distance scaling," *Appl. Intell.*, vol. 51, no. 3, pp. 1394–1409, Mar. 2021, doi: 10.1007/s10489-020-01852-8.

[10] J. Naskath, G. Sivakamasundari, and A. A. S. Begum, "A Study on Different Deep Learning Algorithms Used in Deep Neural Nets: MLP SOM and DBN," *Wirel. Pers. Commun.*, vol. 128, no. 4, pp. 2913–2936, Feb. 2023, doi: 10.1007/s11277-022-10079-4.

[11] L. T. M and K. B, "HybMLP: Revolutionizing Cardiovascular Disease Prediction with Hybrid Multi-Layer Perceptron and Gradient Boosting," in *2025 International Conference on Multi-Agent Systems for Collaborative Intelligence (ICMSCI)*, Jan. 2025, pp. 1501–1507, doi: 10.1109/ICMSCI62561.2025.10894362.

[12] N. Kalra, P. Verma, and S. Verma, "Advancements in AI based healthcare techniques with FOCUS ON diagnostic techniques," *Comput. Biol. Med.*, vol. 179, p. 108917, Sep. 2024, doi: 10.1016/j.compbiomed.2024.108917.

[13] U. Nagavelli, D. Samanta, and P. Chakraborty, "Machine Learning Technology-Based Heart Disease Detection Models," *J. Healthc. Eng.*, vol. 2022, pp. 1–9, Feb. 2022, doi: 10.1155/2022/7351061.

[14] Y. Amelia, "PERBANDINGAN METODE MACHINE LEARNING UNTUK MENDETEKSI PENYAKIT JANTUNG," *IDEALIS Indones. J. Inf. Syst.*, vol. 6, no. 2, pp. 220–225, Jul. 2023, doi: 10.36080/idealis.v6i2.3043.

[15] D. Pardede, B. H. Hayadi, and Iskandar, "Kajian Literatur Multi Layer Perceptron Seberapa Baik Performa Algoritma Ini," *J. ICT Apl. Syst.*, vol. 1, no. 1, pp. 23–35, Jun. 2022, doi: 10.56313/.v1i1.127.

[16] 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," *Mob. Inf. Syst.*, vol. 2022, pp. 1–9, Mar. 2022, doi: 10.1155/2022/1410169.

[17] F. Gurcan and A. Soylu, "Learning from Imbalanced Data: Integration of Advanced Resampling Techniques and Machine Learning Models for Enhanced Cancer Diagnosis and Prognosis," *Cancers (Basel).*, vol. 16, no. 19, p. 3417, Oct. 2024, doi: 10.3390/cancers16193417.

[18] K. Ghosh, C. Bellinger, R. Corizzo, P. Branco, B. Krawczyk, and N. Japkowicz, "The class imbalance problem in deep learning," *Mach. Learn.*, vol. 113, no. 7, pp. 4845–4901, Jul. 2024, doi: 10.1007/s10994-022-06268-8.

[19] F. Yang, K. Wang, L. Sun, M. Zhai, J. Song, and H. Wang, "A hybrid sampling algorithm combining synthetic minority over-sampling technique and edited nearest neighbor for missed abortion diagnosis," *BMC Med. Inform. Decis. Mak.*, vol. 22, no. 1, p. 344, Dec. 2022, doi: 10.1186/s12911-022-02075-2.

[20] T.-T.-H. Le, Y. Shin, M. Kim, and H. Kim, "Towards unbalanced multiclass intrusion detection with hybrid sampling methods and ensemble classification," *Appl. Soft Comput.*, vol. 157, p. 111517,

May 2024, doi: 10.1016/j.asoc.2024.111517.
[21] D. Elreedy, A. F. Atiya, and F. Kamalov, "A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning," *Mach. Learn.*, vol. 113, no. 7, pp. 4903–4923, Jul. 2024, doi: 10.1007/s10994-022-06296-4.
[22] A. G. Putrada, M. Abdurohman, D. Perdana, and H. H. Nuha, "Shuffle Split-Edited Nearest Neighbor: A Novel Intelligent Control Model Compression for Smart Lighting in Edge Computing Environment," 2023, pp. 219–227.
[23] G. Husain *et al.*, "SMOTE vs. SMOTEENN: A Study on the Performance of Resampling Algorithms for Addressing Class Imbalance in Regression Models," *Algorithms*, vol. 18, no. 1, p. 37, Jan. 2025, doi: 10.3390/a18010037.
[24] F. Farahnakian *et al.*, "Addressing imbalanced data for machine learning based mineral prospectivity mapping," *Ore Geol. Rev.*, vol. 174, p. 106270, Nov. 2024, doi: 10.1016/j.oregeorev.2024.106270.
[25] F. Yazdi and S. Asadi, "Enhancing Cardiovascular Disease Diagnosis: The Power of Optimized Ensemble Learning," *IEEE Access*, vol. 13, pp. 46747–46762, 2025, doi: 10.1109/ACCESS.2025.3550015.
[26] Q. H. Nguyen *et al.*, "Influence of Data Splitting on Performance of Machine Learning Models in Prediction of Shear Strength of Soil," *Math. Probl. Eng.*, vol. 2021, pp. 1–15, Feb. 2021, doi: 10.1155/2021/4832864.
[27] A. Buabeng, A. Simons, N. K. Frempong, and Y. Y. Ziggah, "A novel hybrid predictive maintenance model based on clustering, smote and multi-layer perceptron neural network optimised with grey wolf algorithm," *SN Appl. Sci.*, vol. 3, no. 5, p. 593, May 2021, doi: 10.1007/s42452-021-04598-1.
[28] N. Tavakolizadeh and M. Bagheri, "Multi-attribute Selection for Salt Dome Detection Based on SVM and MLP Machine Learning Techniques," *Nat. Resour. Res.*, vol. 31, no. 1, pp. 353–370, Feb. 2022, doi: 10.1007/s11053-021-09973-8.
[29] M. Desai and M. Shah, "An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN)," *Clin. eHealth*, vol. 4, pp. 1–11, 2021, doi: 10.1016/j.ceh.2020.11.002.
[30] J. Zhang, C. Li, Y. Yin, J. Zhang, and M. Grzegorzek, "Applications of artificial neural networks in microorganism image analysis: a comprehensive review from conventional multilayer perceptron to popular convolutional neural network and potential visual transformer," *Artif. Intell. Rev.*, vol. 56, no. 2, pp. 1013–1070, Feb. 2023, doi: 10.1007/s10462-022-10192-7.