

Validasi Efektivitas Logistic Regression untuk Diagnosa Penyakit Jantung melalui Pendekatan Machine Learning

Dedik Fabiyanto^{1*}, Zico Pratama Putra²

Fakultas Teknologi Informasi, Universitas Nusa Mandiri^{1,2}
14220030@nusamandiri.ac.id^{1*}, zico.zpp@nusamandiri.ac.id²

*) Corresponding Author

(received: 24-06-24, revised: 09-08-24, accepted: 12-09-24)

Abstract

This study aims to evaluate the effectiveness of Logistic Regression in diagnosing heart disease using the UCI Heart Disease dataset. The dataset comprises 303 patient records with 14 features, including age, gender, blood pressure, and cholesterol, divided into 60% training data, 20% validation data, and 20% testing data. The research begins with problem formulation and literature review, followed by data preprocessing, which includes missing value imputation, feature normalization, and dataset splitting. Various machine learning models, including k-Nearest Neighbors (kNN), Support Vector Machine (SVM), Logistic Regression, Naïve Bayes, and Neural Network, are applied and compared. The models are evaluated using accuracy, precision, recall, and F1 score metrics. Results show that Logistic Regression achieves the highest accuracy of 0.89, with superior precision and recall. Naïve Bayes and kNN achieve accuracies of 0.87 and 0.85, respectively, while Neural Network shows the lowest accuracy of 0.77. Tuning hyperparameters of the Neural Network did not reveal a clear pattern for improving accuracy. This study concludes that Logistic Regression is the most effective model for heart disease prediction compared to others, particularly in medical applications that require interpretability and efficiency.

Keywords: Logistic Regression, Machine Learning, Neural Network, Heart Disease Prediction, UCI Heart Disease

Abstrak

Penelitian yang bertujuan untuk mengevaluasi efektivitas Logistic Regression dalam diagnosa penyakit jantung telah dilakukan dengan menggunakan dataset UCI Heart Disease. Dataset ini terdiri dari 303 data pasien dengan 14 fitur, termasuk usia, jenis kelamin, tekanan darah, dan kolesterol, yang dibagi menjadi 60% data pelatihan, 20% data validasi dan 20% data pengujian. Penelitian dimulai dengan perumusan masalah dan studi literatur, diikuti oleh proses preprocessing data yang mencakup imputasi nilai hilang, normalisasi fitur, dan pembagian dataset. Beberapa model machine learning, yaitu k-Nearest Neighbors (kNN), Support Vector Machine (SVM), Logistic Regression, Naïve Bayes, dan Neural Network, diterapkan dan dibandingkan. Model dievaluasi menggunakan metrik akurasi, presisi, recall, dan F1 score. Hasil penelitian menunjukkan bahwa Logistic Regression memperoleh akurasi tertinggi sebesar 0.89, dengan keunggulan dalam presisi dan recall. Naïve Bayes dan kNN menunjukkan akurasi masing-masing 0.87 dan 0.85, sedangkan Neural Network memperoleh akurasi terendah 0.77. Tuning hyperparameter pada Neural Network tidak menunjukkan pola yang jelas dalam meningkatkan akurasi. Pada penelitian ini Logistic Regression adalah model yang paling efektif untuk prediksi penyakit jantung dibandingkan dengan model lainnya, terutama dalam konteks aplikasi medis yang membutuhkan interpretabilitas dan efisiensi.

Kata Kunci: Logistic Regression, Machine Learning, Neural Network, Prediksi Penyakit Jantung, UCI Heart Disease

I. Pendahuluan

Penyakit jantung merupakan salah satu penyebab kematian utama di seluruh dunia, Menurut Organisasi Kesehatan Dunia (WHO), lebih dari 17,9 juta orang meninggal setiap tahun akibat penyakit kardiovaskular, yang mencakup penyakit jantung koroner, serangan jantung, dan stroke [1]. Menjadikannya fokus utama dalam

bidang kesehatan. Kondisi ini meliputi berbagai masalah seperti penyakit arteri koroner, gagal jantung, dan aritmia. Mengingat kompleksitas dan dampaknya, upaya untuk meningkatkan diagnosis dan prediksi penyakit jantung sangat penting. Teknologi machine learning menawarkan solusi potensial untuk tantangan ini dengan memungkinkan analisis data yang lebih mendalam dan akurat. Berbagai teknik machine learning telah dikembangkan dan diterapkan untuk meningkatkan akurasi diagnosis dan prediksi penyakit jantung. Beberapa metode yang sering digunakan meliputi k-Nearest Neighbors (k-NN), Support Vector Machine (SVM), Regresi Logistik, Naïve Bayes, dan Neural Network.

Pada tahun 2016 dilakukan penelitian penggunaan metode *Logistic Regression* dalam kombinasi dengan teknik seleksi fitur untuk meningkatkan akurasi diagnosa penyakit jantung. Dengan mengurangi jumlah fitur yang tidak relevan, peneliti berhasil meningkatkan performa model Logistic Regression secara signifikan. Hasil studi ini menunjukkan bahwa Logistic Regression dapat dioptimalkan lebih lanjut dengan teknik pra-pemrosesan data yang tepat, sehingga membuatnya lebih efektif dalam praktik klinis. Dengan teknik ini, model Logistic Regression berhasil mencapai akurasi sebesar 87%, menunjukkan peningkatan dibandingkan dengan model tanpa seleksi fitur [2].

Penelitian mengenai pengembangan model prediksi penyakit jantung berdasarkan data dari Framingham Heart Study, salah satu dataset paling terkenal dalam penelitian penyakit kardiovaskular. *Logistic Regression* digunakan sebagai model utama untuk memprediksi risiko penyakit jantung berdasarkan faktor-faktor seperti tekanan darah, kadar kolesterol, usia, jenis kelamin, dan riwayat merokok. Hasil penelitian menunjukkan bahwa model Logistic Regression digunakan untuk memprediksi risiko penyakit jantung berdasarkan berbagai faktor risiko seperti usia, tekanan darah, kolesterol, dan kebiasaan merokok. Model ini mencapai akurasi sekitar 77% dalam memprediksi kejadian penyakit jantung selama periode 10 tahun. Model Logistic Regression memiliki akurasi yang cukup tinggi dan memberikan interpretasi yang mudah dipahami, menjadikannya alat yang efektif dalam prediksi klinis [3].

Selanjutnya penelitian mengevaluasi performa berbagai model *machine learning*, termasuk *Logistic Regression*, dalam mendiagnosis penyakit jantung menggunakan dataset UCI Heart Disease. Dalam studi ini, peneliti membandingkan Logistic Regression dengan algoritma lain seperti Random Forest, SVM, dan Neural Networks. Hasil penelitian menunjukkan model Logistic Regression menghasilkan akurasi sebesar 85%, dengan precision dan recall masing-masing 84% dan 86%. Studi ini juga menunjukkan bahwa Logistic Regression memiliki performa yang sebanding dengan model lain seperti Random Forest dan SVM. Keunggulan metode *Logistic Regression* dalam studi ini adalah kesederhanaannya dan kemampuan untuk memberikan probabilitas prediktif yang dapat membantu dokter dalam pengambilan keputusan [4].

Penelitian lainnya memfokuskan pada validasi model Logistic Regression dalam mendeteksi penyakit jantung pada populasi pasien dengan gejala yang tidak spesifik. Dengan menggunakan data dari lebih dari 10.000 pasien, Penelitian ini menemukan bahwa Logistic Regression memiliki sensitivitas sebesar 83% dan spesifisitas sebesar 79% dalam mendeteksi penyakit jantung pada populasi pasien dengan gejala yang tidak spesifik. Model ini juga mampu memberikan nilai prediktif positif dan negatif yang cukup baik, sehingga berguna dalam pengambilan keputusan klinis. Logistic Regression berhasil menunjukkan sensitivitas dan spesifisitas yang cukup tinggi. Penelitian ini menekankan pentingnya validasi eksternal model untuk memastikan bahwa model dapat diterapkan secara efektif pada populasi yang lebih luas dan beragam [5].

Penelitian lain mengenai investigasi kinerja berbagai algoritma pembelajaran mesin dan analisis komparatif yang sesuai dalam memprediksi penyakit kardiovaskular. Kinerja algoritma diamati untuk parameter hiper default (DHP), validasi silang pencarian grid (GSCV), dan metode validasi silang pencarian acak (RSCV). Hasilnya akurasi sebesar 92% ditemukan pada kedua pengklasifikasi ensemble voting keras dan lunak (EVCH dan EVCS). Namun, diamati bahwa algoritma Adaboost mengungguli EVCH dan EVCS dalam hal presisi dan spesifisitas. Oleh karena itu, analisis komparatif keseluruhan di antara semua algoritma dilakukan secara ekstensif di mana akurasi, presisi, sensitivitas, spesifisitas, dan skor F1 diambil dalam tindakan serta menggunakan Jupyter notebook 6.0.3 untuk simulasi [6].

Penelitian mengenai Analisis Perbandingan Kinerja Algoritma Klasifikasi dalam Mendeteksi Penyakit Jantung telah dilakukan dengan membandingkan *Support Vector Machine (SVM)*, *Random Forest*, *Logistic Regression*, dan AdaBoost dengan tujuan untuk memprediksi penyakit jantung. Hasil yang diperoleh adalah algoritma random forest, SVM, dan AdaBoost merupakan algoritma yang mempunyai nilai terbaik dan nilai yang sama pada hasil pengujian. Untuk nilai akurasi 0.985366, presisi 0.985714, recall 0.985437, dan f1-score 0.985364 [6]. Penelitian lainnya Perbandingan Metode Machine Learning Untuk Mendeteksi Penyakit Jantung dilakukan oleh Yutri Amelia. Dalam penelitian ini menggunakan dataset dari UCI Machine Learning Repository. Pada

penelitian ini, metode yang digunakan meliputi Random Forest, Support Vector Machine (SVM), XGBoost, K-Nearest Neighbor (KNN), Decision Tree, Logistic Regression serta Multi-Layer Perceptron Classifier (MLP). Dari penelitian ini didapatkan akurasi terbaik menggunakan algoritma XGBoost dengan akurasi mencapai 95,08% [7]. Penelitian lain mengenai investigasi kinerja berbagai algoritma pembelajaran mesin dan analisis komparatif yang sesuai dalam memprediksi penyakit kardiovaskular. Kinerja algoritma diamati untuk parameter hiper default (DHP), validasi silang pencarian grid (GSCV), dan metode validasi silang pencarian acak (RSCV). Hasilnya akurasi sebesar 92% ditemukan pada kedua pengklasifikasi ensemble voting keras dan lunak (EVCH dan EVCS). Namun, diamati bahwa algoritma Adaboost mengungguli EVCH dan EVCS dalam hal presisi dan spesifisitas. Oleh karena itu, analisis komparatif keseluruhan di antara semua algoritma dilakukan secara ekstensif di mana akurasi, presisi, sensitivitas, spesifisitas, skor F1, dan ROC-AUC diambil dalam tindakan serta menggunakan Jupyter notebook 6.0.3 untuk simulasi [8]. Studi menggunakan SMOTE dan Random Forest untuk menangani dataset yang tidak seimbang dan meningkatkan akurasi prediksi penyakit jantung [9]. E Erlin, Y Desnelita, N Nasution (meneliti dampak SMOTE pada kinerja classifier Random Forest menggunakan data penyakit jantung yang tidak seimbang. Studi ini menunjukkan peningkatan signifikan dalam akurasi dan efektivitas model [10]. Pada perbandingan kinerja berbagai algoritma klasifikasi dalam mendeteksi penyakit jantung, AS Prabowo, FI Kurniadi menyoroti peningkatan akurasi dan efektivitas yang dicapai melalui metode machine learning [11].

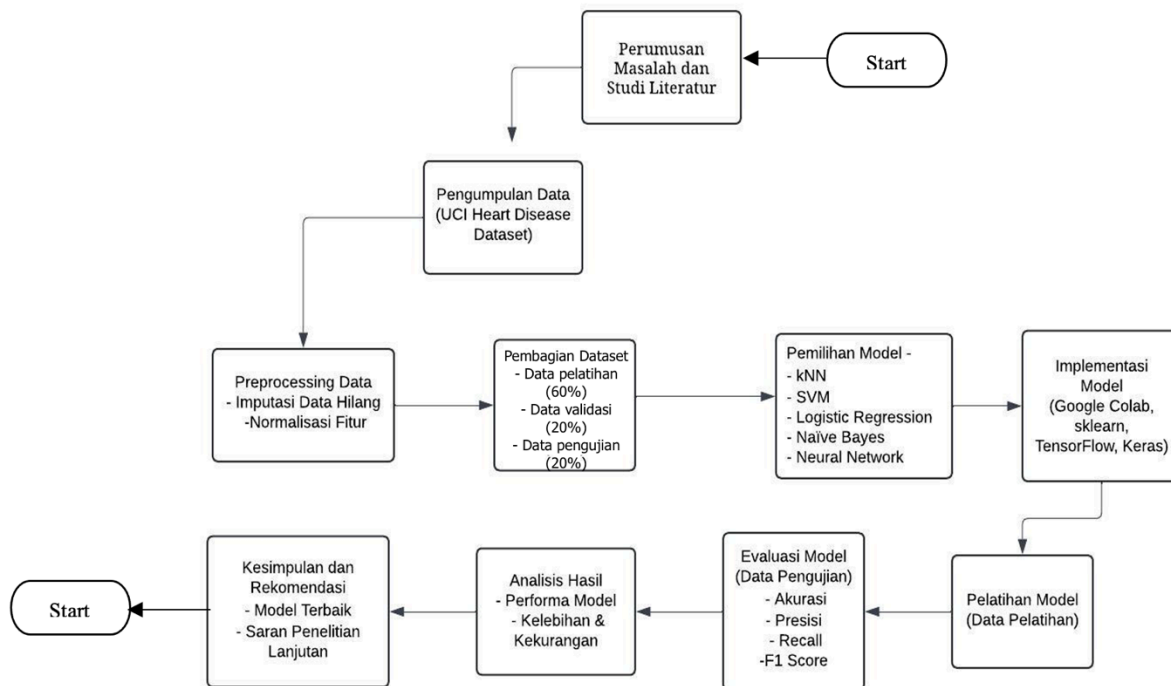
Pada tahun 2023, H Hidayat, A Sunyoto, H Al Fatta klasifikasi menggunakan algoritma Random Forest untuk klasifikasi penyakit jantung, menunjukkan akurasi tinggi dan ketangguhan dalam menangani dataset yang kompleks [12]. DK Saputro, MFR Ajie, S Azizah menerapkan regresi logistik untuk mendeteksi penyakit jantung pada pasien, menunjukkan kemampuan model dalam meningkatkan akurasi prediksi dan proses pengambilan keputusan [13]. Y Purbolingga, DM Putri, A Rahmawati pada studi lainnya juga ditemukan perbandingan algoritma CatBoost dan XGBoost untuk klasifikasi penyakit jantung, menyoroti kinerja superior mereka dalam hal akurasi prediksi dan efisiensi komputasi [14]. AP Ariyanti, MI Mazdadi, M Muliadi mengeksplorasi penggunaan Extreme Learning Machine (ELM) yang dikombinasikan dengan Particle Swarm Optimization (PSO) untuk klasifikasi penyakit jantung, menunjukkan peningkatan kinerja dan akurasi [15]. EE Fauziah, AF Zulfikar (menggunakan algoritma decision tree ID3 untuk klasifikasi risiko penyakit jantung, menunjukkan efektivitasnya dalam menangani dataset besar dan meningkatkan hasil klasifikasi [16]. R Sahobby menggunakan XGBoost dan Randomized Search Optimizer untuk memprediksi penyakit jantung, menunjukkan akurasi tinggi dan efektivitas algoritma dalam menangani dataset medis [17]. RD Muhammad machine Learning Autogluon untuk prediksi penyakit jantung, menyoroti fleksibilitas dan akurasi tinggi dalam analisis data medis [18].

Penelitian ini bertujuan untuk menganalisis efektivitas metode-metode machine learning tersebut dalam konteks prediksi dan diagnosis penyakit jantung. Dengan menggunakan data medis yang komprehensif, penelitian ini akan mengevaluasi performa masing-masing metode berdasarkan beberapa metrik evaluasi seperti akurasi, sensitivitas, spesifisitas, dan area under the curve (AUC). Hasil dari penelitian ini diharapkan dapat memberikan wawasan yang lebih mendalam tentang kekuatan dan kelemahan setiap metode, serta memberikan panduan bagi praktisi kesehatan dalam memilih dan menerapkan metode yang paling sesuai untuk meningkatkan hasil diagnosis dan prediksi penyakit jantung. Setiap metode ini menawarkan pendekatan unik dalam menganalisis data medis untuk mengidentifikasi pola dan faktor risiko yang berkaitan dengan penyakit jantung, dengan tujuan meningkatkan hasil diagnosis dan memberikan prediksi yang lebih tepat dan cepat. Pemilihan metode yang tepat bergantung pada karakteristik data yang tersedia serta tujuan spesifik dari analisis yang dilakukan. Misalnya, metode k-NN lebih cocok untuk dataset yang lebih kecil dan terstruktur dengan baik, sementara neural network dapat memberikan hasil yang lebih baik pada dataset yang besar dan kompleks.

Keterbaruan penelitian ini dibandingkan dengan penelitian-penelitian sebelumnya yaitu dalam pendekatan evaluasi yang lebih komprehensif terhadap efektivitas model *Logistic Regression* dalam mendiagnosis penyakit jantung. Penelitian-penelitian sebelumnya hanya berfokus pada peningkatan akurasi melalui teknik seleksi fitur atau validasi pada populasi tertentu, penelitian ini menggabungkan evaluasi terhadap berbagai model *machine learning* yang berbeda, yaitu k-Nearest Neighbors (kNN), Support Vector Machine (SVM), Naïve Bayes, dan Neural Network, untuk menilai keunggulan *Logistic Regression*. Dataset UCI Heart Disease yang diproses melalui beberapa tahap, seperti imputasi data hilang dan normalisasi fitur, akan membantu penelitian ini dalam menunjukkan bahwa *Logistic Regression* tetap menjadi model paling efektif. Ini adalah kontribusi baru untuk memahami kinerja model regresi logistik dalam berbagai kasus medis.

II. Metodologi Penelitian

Penelitian ini menggunakan dataset dari UCI Heart Disease yang berisi 303 data pasien dengan 14 fitur, termasuk usia, jenis kelamin, tekanan darah, kolesterol, dan lain-lain. Dataset dibagi menjadi data pelatihan (60%), data validasi(20%) dan data pengujian (20%). Tahapan-tahapan penelitian mulai dari perumusan masalah studi literatur hingga analisa dan rekomendasi metode machine learning yang digunakan pada gambar 1.



Gambar 1. Tahapan Penelitian

A. Perumusan Masalah dan Studi Literatur

Penelitian dimulai dengan merumuskan masalah utama yang ingin diselesaikan, yaitu bagaimana memvalidasi efektivitas Logistic Regression dalam diagnosa penyakit jantung. Pada tahap ini, dilakukan juga studi literatur untuk memahami penelitian-penelitian sebelumnya yang berkaitan dengan topik yang sama, guna mendapatkan landasan teori yang kuat.

B. Dataset dan Pre-processing

1. Pengumpulan Data

Data diambil dari UCI Heart Disease, yang merupakan sumber data yang telah banyak digunakan dalam penelitian sebelumnya. Dataset ini menyediakan informasi yang kaya dan beragam yang esensial untuk menganalisis faktor-faktor risiko penyakit jantung. Penelitian ini menggunakan dataset dari UCI Heart Disease yang berisi 303 data pasien dengan 14 fitur termasuk usia, jenis kelamin, tekanan darah, kolesterol, serta memiliki 14 atribut [19].

Tabel 1. Deskripsi Dataset

Variabel	Deskripsi
Age	Umur
Sex	Jenis kelamin
Cp	Nyeri Dada / Chest Pain
Trestbps	Hasil tekanan darah selama dirawat di rumah sakit
Chol	Kolestrol
Fbs	Gula darah puasa
Restecg	Hasil elektrokardiografi saat istirahat

Variabel	Deskripsi
Thalach	Detak jantung maksimum
Exang	Latihan induksi angina
Oldpeak	ST depression
Slope	ST segment slope
Ca	Jumlah pembuluh darah besar
Thal	Hasil tegangan thallium
Target/output	1= pasien sakit jantung, 0=pasien sehat

2. Persiapan Data

Berikut adalah langkah-langkah yang diambil untuk menyiapkan dataset: Pertama, dataset dimuat ke dalam lingkungan pemrograman seperti Python, di mana peneliti menggunakan pustaka seperti pandas untuk memuat dan mengelola data. Setelah dataset dimuat, langkah berikutnya adalah melakukan pemeriksaan data untuk memastikan tidak ada nilai yang hilang atau anomali yang dapat mempengaruhi analisis. Jika terdapat fitur kategorikal, fitur tersebut perlu dikonversi menjadi bentuk numerik agar dapat digunakan dalam algoritma pembelajaran mesin. Untuk memastikan semua fitur berada pada skala yang sama, data numerik biasanya dinormalisasi atau distandarisasi.

3.

Sebelum data digunakan untuk pelatihan model, dilakukan preprocessing atau pra-pemrosesan. Tahap ini meliputi imputasi data yang hilang (mengisi data yang hilang) dan normalisasi fitur (menyelaraskan skala berbagai fitur dalam dataset). Langkah ini penting untuk memastikan bahwa data bersih dan siap digunakan oleh algoritma Machine Learning. Selanjutnya data yang hilang diimputasi menggunakan metode median untuk fitur numerik dan modus untuk fitur kategorikal. Hal ini dilakukan untuk menangani kekurangan data yang dapat mempengaruhi hasil analisis. Selanjutnya, normalisasi dilakukan pada fitur numerik untuk memastikan bahwa semua fitur berada dalam skala yang sama, yang sangat penting untuk algoritma yang sensitif terhadap skala fitur.

4. Pembagian Dataset

Dataset yang telah diproses sebelumnya dibagi menjadi set pelatihan, validasi, dan pengujian dengan perbandingan 60:20:20. Set pelatihan digunakan untuk melatih model pembelajaran mesin, sehingga model dapat memahami data dengan baik. Set validasi digunakan untuk menyempurnakan hyperparameter dan mengurangi overfitting, yang bertujuan untuk mengoptimalkan kinerja model. Terakhir, set pengujian digunakan untuk mengevaluasi kinerja akhir model, memberikan ukuran efektivitas yang dapat diandalkan dan akurat. Proses ini memastikan bahwa model bekerja dengan baik pada data yang belum pernah dilihat sebelumnya.

C. Pemilihan Model Machine Learning

Pada tahap ini, model Machine Learning dipilih dan dibandingkan, termasuk k-Nearest Neighbors (kNN), Support Vector Machine (SVM), Logistic Regression, Naive Bayes, dan Neural Network. Tujuannya adalah untuk menentukan model mana yang paling efektif dalam memprediksi penyakit jantung.

1. k-Nearest Neighbors (kNN)

Algoritma berbasis instans yang mengklasifikasikan data berdasarkan mayoritas tetangga terdekat. Fungsi klasifikasi nya sebagai berikut

$$dis(x_1, x_2) = \sqrt{\sum_{i=0}^n (x_{1i} - x_{2i})^2} \quad (1)$$

2. Support Vector Machine (SVM)

Model yang menemukan hyperplane optimal yang memisahkan kelas data. Fungsi model SVM dapat di tuliskan sebagai berikut :

$$f(x) = w \cdot x - b \quad (2)$$

3. Logistic Regression (Logres)

Model probabilistik yang menggunakan fungsi logistik untuk melakukan klasifikasi. Fungsi klasifikasi nya sebagai berikut :

$$P(x) = \frac{1}{1 + e^{-(w \cdot x + b)}} \quad (3)$$

4. Naïve Bayes (Naïve)

Model probabilistik yang mengasumsikan independensi antar fitur. Dengan tampilan sebagai berikut :

$$p(C_k | \mathbf{x}) = \frac{p(C_k) \prod p(\mathbf{x} | C_k)}{p(\mathbf{x})} \quad (4)$$

5. Neural Network (NN)

Model yang meniru jaringan saraf biologis dengan beberapa lapisan (input, hidden, dan output). Ditampilkan dalam bentuk berikut :

$$\begin{aligned} z(l) &= W(l)x + b(l) \\ a(l) &= f(z(l)) \end{aligned} \quad (5)$$

D. Implementasi dan Pelatihan Model

Penelitian ini diimplementasikan menggunakan Google Colab dan pustaka Python seperti Scikit-learn (sklearn), TensorFlow, dan Keras untuk membangun dan mengevaluasi model machine learning. Implementasi dilakukan dengan langkah-langkah yang terstruktur, mulai dari pemilihan model, pelatihan, hingga evaluasi. Model-model yang digunakan dilatih menggunakan data pelatihan. Proses pelatihan ini bertujuan untuk mengoptimalkan model agar dapat membuat prediksi yang akurat saat diberi data baru.

E. Evaluasi Model

Model dievaluasi menggunakan metrik akurasi, presisi, recall, dan F1 score. Evaluasi ini penting untuk memahami sejauh mana model dapat memprediksi dengan benar kondisi kesehatan jantung pasien. Precision adalah metrik yang digunakan untuk mengevaluasi akurasi prediksi positif yang dibuat oleh model. Precision didefinisikan sebagai rasio antara jumlah True Positive (TP) dengan jumlah total prediksi positif (yang terdiri dari True Positive dan False Positive (FP)). Rumusnya adalah sebagai berikut:

$$Precision = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Positive\ (FP)} \quad (6)$$

Recall, juga dikenal sebagai Sensitivity atau True Positive Rate, adalah metrik yang digunakan untuk mengukur kemampuan model dalam mendeteksi semua kasus positif yang sebenarnya ada. Recall didefinisikan sebagai rasio antara jumlah True Positive (TP) dengan jumlah total kasus yang sebenarnya positif (yang terdiri dari True Positive dan False Negative (FN)). Rumusnya adalah sebagai berikut:

$$Recall = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Negative\ (FN)} \quad (7)$$

F1 Score adalah metrik yang digunakan untuk menyeimbangkan antara Precision dan Recall dalam satu nilai yang komprehensif. F1 Score adalah rata-rata harmonis dari Precision dan Recall, yang berarti mempertimbangkan kedua metrik ini secara seimbang. F1 Score memberikan gambaran tentang trade-off antara ketepatan prediksi positif dan kemampuan mendeteksi semua kasus positif. Rumusnya adalah sebagai berikut:

$$F1\ Score = 2 \times \frac{recall \times precision}{recall + precision} \quad (8)$$

Akurasi adalah ukuran seberapa dekat hasil prediksi model dengan nilai sebenarnya. Rumus tersebut adalah:

$$Accuracy = \frac{(TP + TN)}{(TP + FN + TN + FP)} \tag{9}$$

F. Analisis Hasil

Hasil dari evaluasi model dianalisis untuk menilai performa model. Analisis ini meliputi identifikasi kelebihan dan kekurangan dari masing-masing model yang telah diuji.

G. Kesimpulan dan Rekomendasi

Tahap terakhir adalah menarik kesimpulan berdasarkan hasil penelitian dan memberikan rekomendasi. Kesimpulan mencakup model mana yang terbaik untuk digunakan, dan rekomendasi untuk penelitian lebih lanjut atau peningkatan model di masa depan.

III. Hasil dan Diskusi

Data pelatihan yang diperiksa pada proses oversampling dilakukan untuk memastikan distribusi kelas yang seimbang. Hasil menunjukkan bahwa ada 97 sampel kelas positif (1) dan 97 sampel kelas negatif (0) dalam data pelatihan, memastikan bahwa oversampling berhasil menyeimbangkan kelas.

A. Akurasi Machine Learning

Model yang Digunakan pada penelitian ini adalah k-Nearest Neighbors (kNN), Support Vector Machine (SVM), Logistic Regression (Logres), Naïve Bayes (Naïve), dan Neural Network (NN). Evaluasi Model Model dievaluasi menggunakan metrik akurasi, presisi, recall dan F1 score. Tabel 2 berikut merupakan hasil evaluasi dari masing-masing model yang digunakan.

Tabel 2. Hasil Matrik Performansi

Metric	k-NN	Logistic Regression	Naive Bayes	SVM	Neural Network
Precision (0)	0.92	0.92	0.86	0.88	0.82
Precision (1)	0.81	0.86	0.88	0.83	0.74
Recall (0)	0.76	0.83	0.86	0.79	0.64
Recall (1)	0.94	0.94	0.88	0.91	0.88
F1-Score (0)	0.83	0.87	0.86	0.84	0.72
F1-Score (1)	0.87	0.90	0.88	0.87	0.81
Accuracy	0.85	0.89	0.87	0.85	0.77
Macro Avg (P)	0.86	0.89	0.87	0.86	0.78
Macro Avg (R)	0.85	0.88	0.87	0.85	0.76
Macro Avg (F1)	0.85	0.88	0.87	0.85	0.76
Weighted Avg (P)	0.86	0.89	0.87	0.86	0.78
Weighted Avg (R)	0.85	0.89	0.87	0.85	0.77
Weighted Avg (F1)	0.85	0.88	0.87	0.85	0.77
Support (0)	29	29	29	29	29
Support (1)	32	32	32	32	32

Penelitian ini menunjukkan performa model Logistic Regression (Logres) unggul dengan akurasi rata-rata sebesar 0.89, menunjukkan superioritasnya dalam tugas klasifikasi biner seperti prediksi penyakit jantung. Model ini menggunakan fungsi logistik yang memetakan hasil prediksi menjadi probabilitas, yang memungkinkan Logistic Regression untuk memaksimalkan log-likelihood dan meminimalkan error, terutama pada dataset dengan proporsi kelas yang seimbang.

Sebagai perbandingan, model Naïve Bayes, meskipun akurasinya mencapai 0.87, namun terpengaruh oleh asumsi independensi antar fitur yang tidak selalu valid. Model kNN dan SVM, dengan akurasi sebesar 0.85, memiliki kelebihan masing-masing namun lebih sensitif terhadap skala dan distribusi data, serta memerlukan

tuning parameter yang lebih hati-hati. Neural Network, dengan akurasi 0.77, menunjukkan kelemahan ketika jumlah data yang tersedia terbatas dan membutuhkan sumber daya komputasi yang besar untuk mencapai performa yang lebih baik.

Keunggulan model Logistic Regression dalam presisi dan recall yang juga mencapai 0.89, mengindikasikan kemampuan model ini dalam konsistensi identifikasi kasus positif tanpa mengorbankan tingkat false positives. Alasan utama mengapa Logistic Regression outperform model lainnya terletak pada beberapa faktor penting:

1. Linearity Assumption

Logistic Regression mengasumsikan hubungan linear antara fitur input dan log-odds dari output. Dalam prediksi penyakit jantung, hubungan linear ini memadai untuk menangkap pola antara faktor risiko seperti tekanan darah dan kolesterol dengan hasil diagnosis. Sebuah studi menunjukkan bahwa Logistic Regression efektif dalam memprediksi kondisi kesehatan karena kemampuannya untuk menangani hubungan linear secara langsung, membuatnya cocok untuk masalah klasifikasi seperti diagnosis penyakit jantung [20].

2. Simplicity and Interpretability

Logistic Regression menawarkan model yang lebih sederhana dan dapat diinterpretasikan dibandingkan model yang lebih kompleks seperti Neural Network. Sifat ini membuat Logistic Regression menjadi pilihan yang baik dalam kasus medis, di mana interpretasi yang jelas dari hasil model sangat penting. Model yang mudah dipahami ini memungkinkan praktisi kesehatan untuk mengimplementasikannya tanpa memerlukan pengetahuan teknis mendalam [21].

3. Efficient Training

Logistic Regression lebih efisien dalam pelatihan, terutama untuk dataset berukuran kecil hingga sedang. Ini karena Logistic Regression tidak memerlukan penyesuaian parameter yang rumit, seperti yang diperlukan oleh SVM atau Neural Network. Efisiensi ini juga memungkinkan model untuk cepat mencapai konvergensi dengan performa optimal [21].

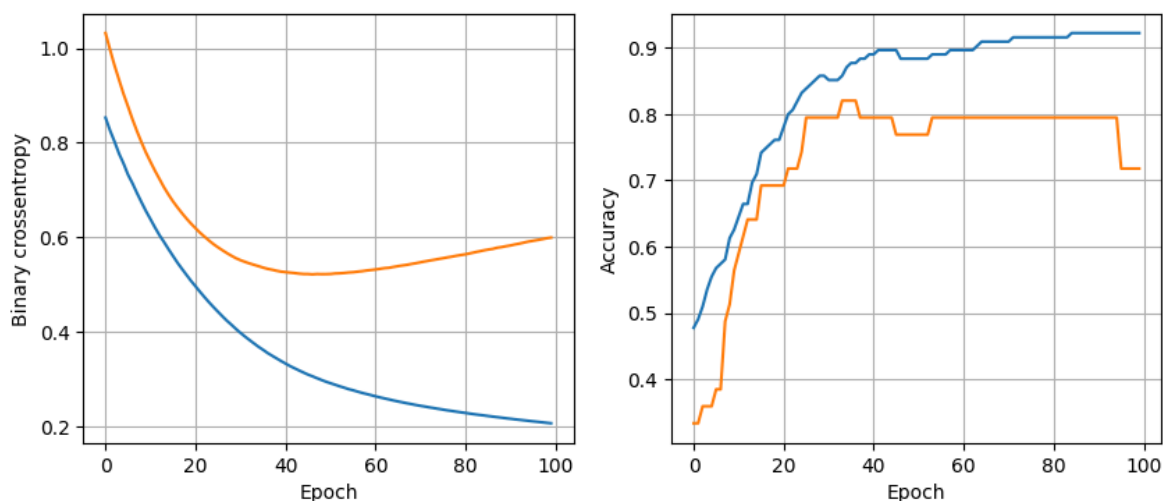
4. Handling of Overfitting

Dengan penggunaan regularisasi seperti L2, Logistic Regression dapat mengurangi risiko overfitting, yang merupakan tantangan umum pada model yang lebih kompleks seperti Neural Network, terutama ketika data yang tersedia terbatas. Regularisasi membantu dalam menjaga keseimbangan antara bias dan varians, yang esensial untuk kinerja model yang stabil pada data yang tidak terlihat [20] [21].

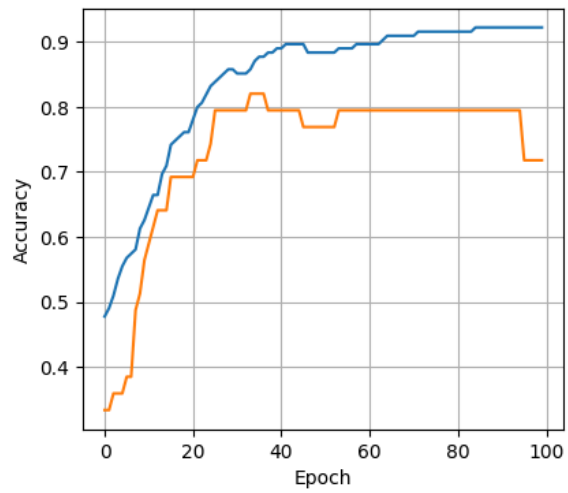
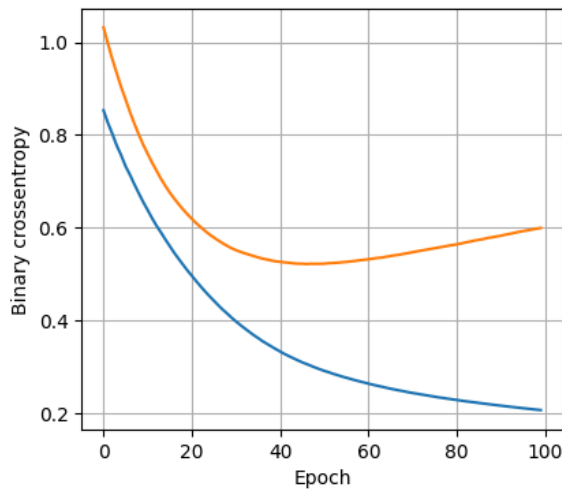
B. Hasil Tuning Hyperparameter Neural Network

Pada proses penyetelan hiperparameter Neural Network, berbagai konfigurasi diuji untuk menentukan kombinasi parameter yang optimal untuk kinerja terbaik. Konfigurasi didalamnya termasuk variasi dalam jumlah node, Dropout, Learning rate, dan ukuran batch. Hasil dari eksperimen ini diilustrasikan pada gambar 2 dan hasilnya dirangkum dalam Tabel 3.

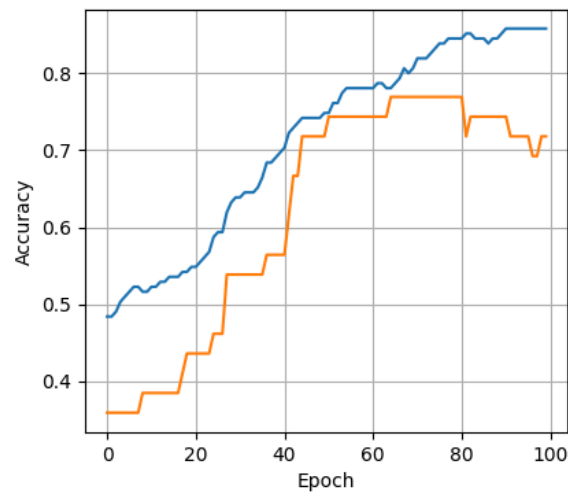
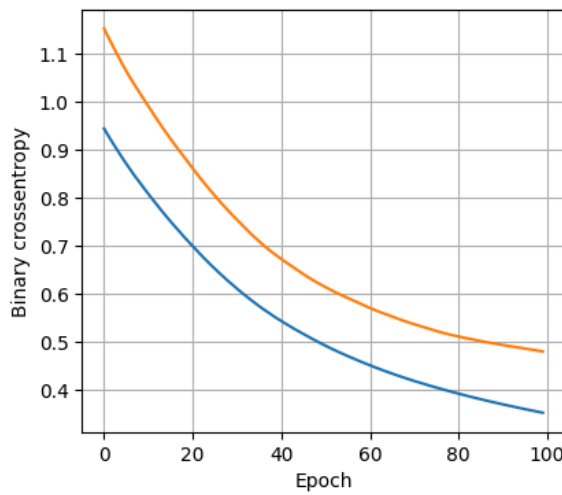
(a)



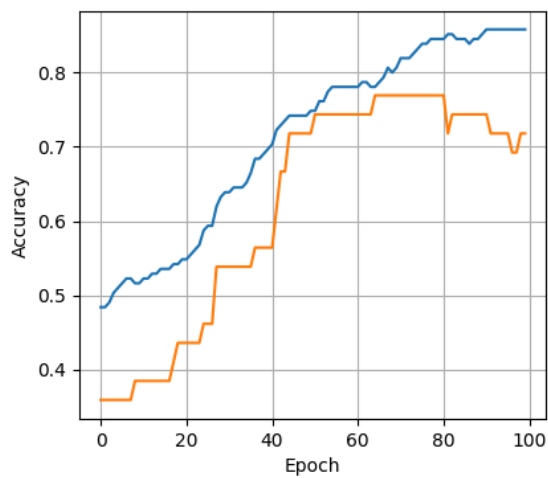
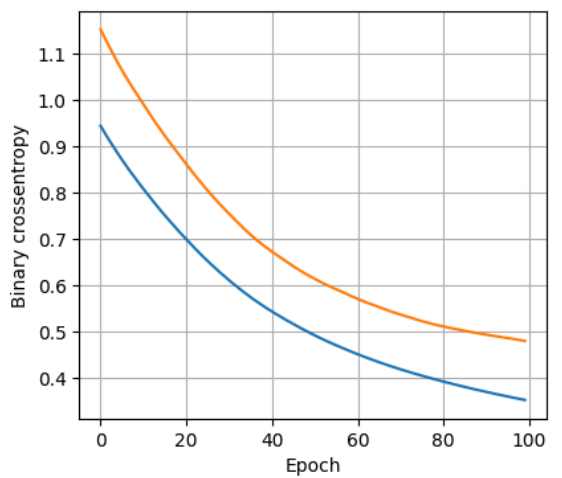
(b)



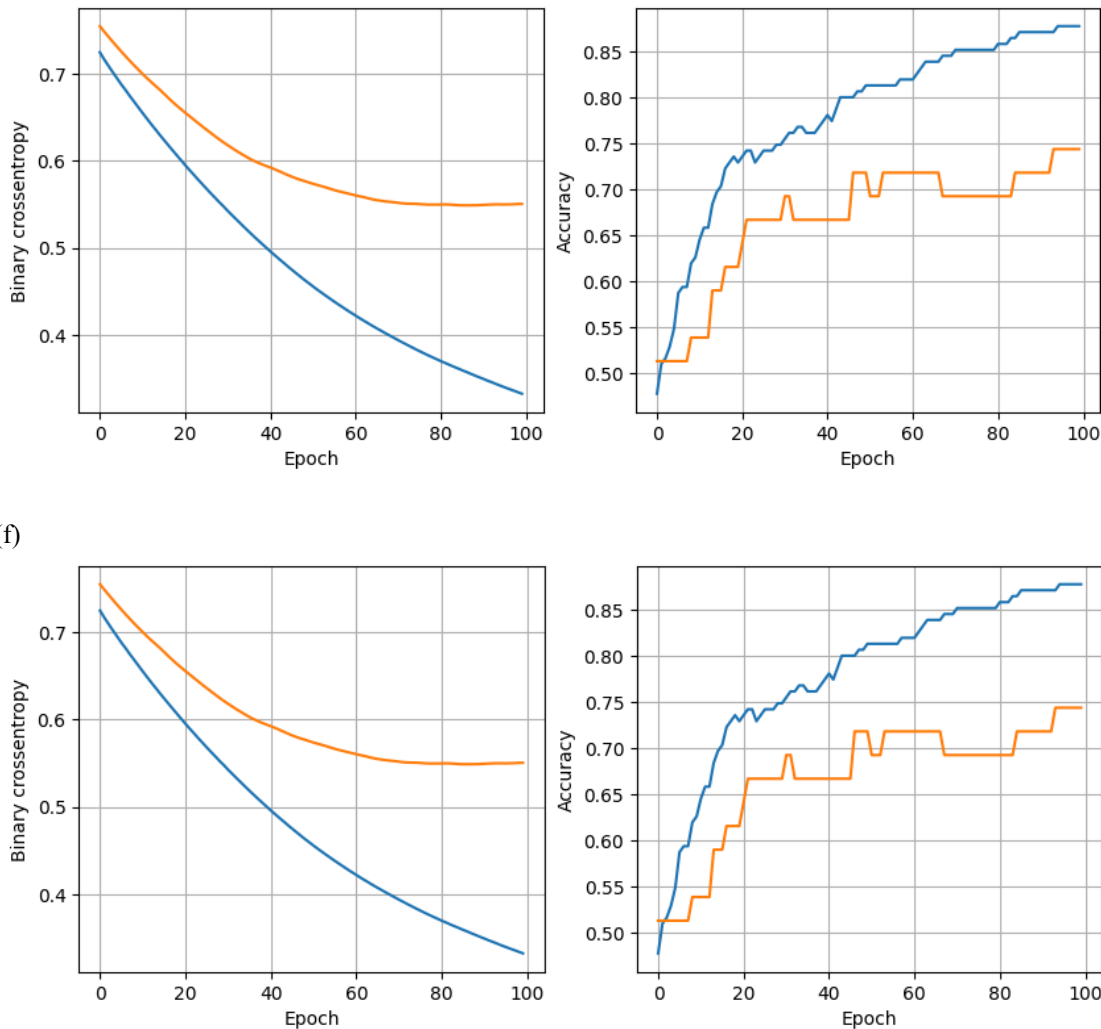
(c)



(d)



(e)



Gambar 2. Hasil Tuning Hyperparameter Nueral Network

Tabel 3. Hasil Konfigurasi Hyperparameter Tuning (a)-(f)

Configuration	Nodes	Dropout	Learning Rate	Batch Size	Loss	Accuracy
a	16	0	0.01	32	0.4166	0.8361
b	16	0	0.01	64	0.4243	0.8525
c	16	0	0.01	128	0.4156	0.8033
d	16	0	0.005	32	0.4488	0.8525
e	16	0	0.005	64	0.4348	0.8197
f	16	0	0.005	128	0.3825	0.8197

Terdapat dua kelompok konfigurasi yang diuji, yaitu dengan learning rate 0.01 dan 0.005. Pada konfigurasi dengan learning rate 0.01, akurasi model tertinggi tercapai pada batch size 64 dengan nilai 0.8525. Sedangkan pada konfigurasi dengan learning rate 0.005, akurasi model tertinggi tercapai pada batch size 32 dengan nilai 0.8525. Pada konfigurasi dengan learning rate 0.01, akurasi model tertingginya dicapai pada batch size 64. Sedangkan pada konfigurasi dengan learning rate 0.005, akurasi model tertingginya dicapai pada batch size 32.

Secara umum, tidak terdapat pola yang jelas mengenai pengaruh batch size terhadap akurasi model. Dalam beberapa kasus, batch size yang lebih besar dapat meningkatkan akurasi model, namun dalam kasus lain, batch size yang lebih kecil justru lebih optimal. Hal ini menunjukkan bahwa performa model tidak hanya dipengaruhi oleh ukuran batch, tetapi juga oleh kombinasi faktor lain seperti learning rate dan parameter model lainnya. Oleh karena itu, eksperimen tambahan dengan berbagai konfigurasi diperlukan untuk memahami secara mendalam bagaimana batch size dan learning rate mempengaruhi performa model.

Dalam pengujian lebih lanjut, berbagai kombinasi learning rate dan batch size diuji untuk menemukan kombinasi optimal yang menghasilkan akurasi tertinggi. Misalnya, dengan menguji learning rate 0.001 dan 0.0005 dengan batch size 16, 32, 64, dan 128, ditemukan bahwa setiap kombinasi memiliki hasil yang bervariasi tergantung pada kompleksitas data dan arsitektur model yang digunakan.

Dalam penelitian ini, pada model Neural Network (NN) dilakukan tuning meskipun performanya lebih rendah dibandingkan dengan model Logistic Regression (Logres) dengan alasan model Neural Network memiliki potensi untuk menemukan pola-pola kompleks dalam data yang tidak dapat ditangkap oleh model yang lebih sederhana seperti Logistic Regression. Hal ini dilakukan untuk mengeksplorasi potensinya yang lebih besar dalam menangani data yang kompleks dan untuk menyediakan wawasan yang lebih mendalam bagi pengembangan model pada penelitian-penelitian selanjutnya.

Penelitian sebelumnya menunjukkan akurasi sekitar 85% dengan menggunakan Logistic Regression, tetapi penelitian ini berhasil mencapai akurasi yang lebih tinggi yaitu 89% dengan dataset dan metodologi yang serupa, hasil penelitian ini menunjukkan konsistensi yang lebih tinggi dalam akurasi dan validitas model. Penelitian ini juga menunjukkan bahwa meskipun model Neural Network memiliki potensi dalam pengolahan data besar, model Logistic Regression tetap unggul dalam hal kesederhanaan dan interpretabilitas, yang merupakan aspek penting dalam dunia medis. Hal ini menegaskan validitas hasil penelitian ini, sekaligus memperkuat posisi Logistic Regression sebagai model yang andal dan efisien dalam diagnosis penyakit jantung.

IV. Kesimpulan

Berdasarkan hasil penelitian dan analisis yang dilakukan peneliti menyimpulkan bahwa model Logistic Regression menunjukkan performa terbaik dalam klasifikasi penyakit jantung dengan akurasi 0.89. Model ini unggul dalam metrik presisi, recall, dan F1 score, menunjukkan kemampuannya dalam memprediksi kondisi kesehatan jantung secara efektif. Model Naïve Bayes menunjukkan akurasi 0.87 dan dapat memprediksi dengan baik, namun terbatas oleh asumsi independensi fitur yang mungkin tidak selalu berlaku. Model k-Nearest Neighbors (kNN) dan Support Vector Machine (SVM) masing-masing memiliki akurasi 0.85. Keduanya menawarkan kelebihan tertentu tetapi lebih sensitif terhadap skala dan distribusi data serta memerlukan parameter tuning yang lebih hati-hati. Model Neural Network (NN), meskipun potensial untuk menangkap pola-pola kompleks, menunjukkan akurasi terendah (0.77) dan memerlukan sumber daya komputasi yang signifikan serta tuning hyperparameter yang mendalam untuk meningkatkan performanya.

Pengujian berbagai konfigurasi hyperparameter pada model Neural Network menunjukkan bahwa tidak ada pola yang konsisten mengenai pengaruh ukuran batch terhadap akurasi model. Akurasi model tertinggi dicapai pada konfigurasi yang berbeda tergantung pada learning rate dan ukuran batch. Konfigurasi dengan learning rate 0.01 dan ukuran batch 64 serta learning rate 0.005 dan ukuran batch 32 menunjukkan akurasi tertinggi yang serupa. Ini menandakan bahwa eksperimen lebih lanjut dengan berbagai kombinasi parameter diperlukan untuk mengoptimalkan performa Neural Network.

Berdasarkan analisis dan pembahasan, model yang direkomendasikan untuk mendiagnosa penyakit jantung adalah Logistic Regression yang merupakan pilihan terbaik untuk tugas klasifikasi penyakit jantung dalam penelitian ini karena kesederhanaannya, kemampuannya untuk menangani hubungan linear, dan efektivitasnya dalam memprediksi dengan akurat. Model ini juga lebih mudah diinterpretasikan dalam konteks medis. Pentingnya interpretasi hasil dalam konteks medis, model Logistic Regression dapat disarankan untuk diterapkan dalam praktek klinis untuk diagnosis penyakit jantung. Kejelasan dan kesederhanaan model ini membuatnya lebih mudah dipahami oleh praktisi kesehatan dibandingkan dengan model yang lebih kompleks. Sedangkan model Neural Network dapat digunakan untuk eksplorasi lebih lanjut pada dataset yang lebih besar atau kompleks. Tuning hyperparameter tambahan dan penggunaan dataset yang lebih banyak dapat membantu meningkatkan performa model ini.

Perlu dilakukan eksperimen lebih lanjut dengan Neural Network untuk menemukan konfigurasi hyperparameter yang optimal, termasuk variasi learning rate dan ukuran batch. Pengujian tambahan dengan data yang lebih besar atau kompleks dapat memberikan wawasan lebih dalam mengenai performa model ini. Pertimbangan

teknik regularisasi tambahan pada model seperti SVM dan Neural Network untuk mengatasi overfitting dan meningkatkan generalisasi model. Selain itu perlu juga dilakukan studi lebih lanjut untuk mengeksplorasi dan membandingkan berbagai model machine learning lainnya dengan dataset yang lebih besar dan lebih bervariasi. Penelitian ini juga dapat diperluas untuk mencakup teknik ensembling atau meta-learning yang mungkin meningkatkan performa klasifikasi.

Daftar Pustaka

- [1] World Health Organization, "Cardiovascular diseases," World Health Organization, 22 Maret 2023. [Online]. Available: https://www-who-int.translate.google/?_x_tr_sl=en&_x_tr_tl=id&_x_tr_hl=id&_x_tr_pto=sc&_x_tr_hist=true#tab=tab_3. [Accessed 11 Juni 2024].
- [2] M. A. Jabbar, B. Deekshatulu and P. Chandra, "Heart Disease Prediction System using Associative Classification and Genetic Algorithm," in *International Conference on Computational Intelligence and Computing Research (ICCCIC)*, 2016.
- [3] W. B. Kannel and D. L. McGee, "Diabetes and Cardiovascular Disease: The Framingham Study," *Circulation*, vol. 61, pp. 1194-1200, 2017.
- [4] M. Amin, K. Agarwal and R. Beg, "Comparative Study of Machine Learning Algorithms for Heart Disease Prediction," *International Journal of Engineering and Advanced Technology (IJEAT)*, vol. 8, no. 6, pp. 156-160, 2019.
- [5] M. Al-Mallah, M. B. Elshazly, H. Tanaka, C. A. Brawner and M. Blaha, "Outcomes of Preoperative Percutaneous Coronary Intervention in Heart Transplant Patients," *American Heart Journal*, vol. 230, pp. 100-106, 2020.
- [6] L. Yahaya, N. D. Oye and J. E. Garba, "A Comprehensive Review on Heart Disease Prediction Using Data Mining and Machine Learning Techniques," *American Journal of Artificial Intelligence*, vol. 4 (1), no. doi: 10.11648/j.ajai.20200401.12, pp. 20-29, 23 April 2020.
- [7] Y. Amelia, "Perbandingan Metode Machine Learning Untuk Mendeteksi Penyakit Jantung," *Idealis: Indonesia Journal Information System*, vol. 6 (2), no. ISSN 2684-7280, pp. 220-225, Juli 2023.
- [8] M. A. A.-R. Asif, M. M. Nishat, F. Faisal, R. R. Dip, M. H. Uday, M. F. Shikder and R. Ahsan, "Performance Evaluation and Comparative Analysis of Different Machine Learning Algorithms in Predicting Cardiovascular Disease," *Engineering Letters*, vol. 29, no. 2, pp. 1-11, Juni 2021.
- [9] A. Rahim, I. Pratiwi and M. Fikri, "Klasifikasi Penyakit Jantung Menggunakan Metode Synthetic Minority Over-Sampling Technique Dan Random Forest Classifier," *Indonesian Journal of Computer Science (IJCS)*, vol. 12 (5), 2023.
- [10] E. Erlin, Y. Desnelita, N. Nasution and L. Suryati, "Dampak SMOTE terhadap Kinerja Random Forest Classifier berdasarkan Data Tidak seimbang," *MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, vol. 21 (3), no. <https://doi.org/https://doi.org/10.30812/matrik.v21i3.1726>, pp. 677-690, 2022.
- [11] A. Prabowo and F. Kurniadi, "Analisis Perbandingan Kinerja Algoritma Klasifikasi dalam Mendeteksi Penyakit Jantung," *Jurnal SISKOM-KB (Sistem Komputer dan Kecerdasan Buatan)*, vol. 7, no. Doi: 10.47970/siskom-kb.v7i1.468, pp. 56-61, 2023.
- [12] H. Hidayat, A. Sunyoto, and H. Al Fatta, "Klasifikasi Penyakit Jantung Menggunakan Random Forest Classifier," *Jurnal SISKOM-KB (Sistem Komputer dan Kecerdasan Buatan)*, vol. 7 (1), no. <https://doi.org/10.47970/siskom-kb.v7i1.464>, pp. 35-42, 2023.
- [13] D. Saputro, M. Ajie, S. Azizah and Hart, "Penerapan Logistic Regression untuk Mendeteksi Penyakit Jantung pada Pasien," *Prosiding Seminar Nasional Teknologi Informasi dan Bisnis (SENATIB) 2023*, 2023.
- [14] Y. Purbolinggo, D. M. Putri, F. Fahrizal, A. Rahmawati and B. W. Akramunnas, "Perbandingan Algoritma CatBoost dan XGBoost dalam Klasifikasi Penyakit Jantung," *Aptek*, vol. 15 (2), no. <https://doi.org/10.30606/aptek.v15i2.1930>, pp. 126-133, 2023.
- [15] A. Ariyanti, M. Mazdadi, A. Farmadi, M. Muliadi and R. Herteno, "Application of Extreme Learning Machine Method With Particle Swarm Optimization to Classify of Heart Disease," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 17, no. Doi: 10.22146/ijccs.86291, p. 281, 2023.
- [16] E. Fauziah and A. F. Zulfikar, "Penerapan Metode Decision Tree Menggunakan Algoritma Iterative Dichotomiser 3 (ID3) Untuk Klasifikasi Resiko Penyakit Jantung," *OKTAL*, vol. 2 (04), p. 1207-1219, 2023.

- [17] R. Sohabby, "Prediksi Penyakit Jantung dengan Menggunakan Algoritma XgBoost dan Randomized Search Optimizer," *UPN "Veteran" Yogyakarta Repositori*, no. Doi: <http://eprints.upnyk.ac.id/27337/>, 2023.
- [18] M. R. D, "Prediksi Penyakit Jantung dengan menggunakan Machine Learning Autogluon," *Universitas Islam Indonesia Repositori*, no. Doi: <https://dspace.uii.ac.id/handle/123456789/48775>.
- [19] F. Nekouei, "Heart Disease Prediction," Kaggle, [Online]. Available: <https://www.kaggle.com/code/farzadnekouei/heart-disease-prediction>. [Accessed 23 Juni 2024].
- [20] R. Prasad, P. Anjali, S. Adil and N. Deepa, "Heart Disease Prediction using Logistic Regression Algorithm using Machine Learning," *International Journal of Engineering and Advanced Technology (IJEAT)*, vol. 8, no. 38, 2019.
- [21] F. Hrvat, L. Spahic and A. Aleta, "Heart Disease Prediction Using Logistic Regression Machine Learning Model," in *In: Badnjević, A., Gurbeta Pokvić, L. (eds) MEDICON'23 and CMBEBIH'23. MEDICON CMBEBIH 2023 IFMBE Proceedings*, 2024.