

## Klasifikasi Penyakit Jantung Menggunakan Algoritma ANN dengan Persentase diatas 70%

Mohammad Soharto<sup>1</sup>, Mohammad Aldy Fermansyah Hadi<sup>2</sup>,  
Maulana Firdaus Al-Ayyubi<sup>3</sup>

Program Studi Teknik Informatika, Universitas Nurul Jadid, Indonesia

Email: [Rifakchelsea23@gmail.com](mailto:Rifakchelsea23@gmail.com)<sup>1</sup>, [mohaldyfermansyahhadi@gmail.com](mailto:mohaldyfermansyahhadi@gmail.com)<sup>2</sup>, [maulanafa278@gmail.com](mailto:maulanafa278@gmail.com)<sup>3</sup>

**Abstract.** Heart disease is one of the biggest causes of death in the world. This research examines the use of the Artificial Neural Network (ANN) algorithm to classify heart disease based on 303 medical data of heart disease patients obtained from the Kaggle dataset center. The data used includes medical parameters such as age, gender, blood pressure, cholesterol levels and other examination results. Various ANN architectures were tested to find the optimal configuration in terms of the number of hidden layers and neurons in each layer. Model performance is evaluated using metrics such as accuracy, precision, recall, and F1-score. From the results of the performance measurement research, an accuracy rate of 97.06%, precision of 92.30%, recall of 92.30%, and F1-Score of 92.30% were obtained.

**Keywords:** Artificial Neural Network, Heart Disease Classification, Performance Analysis

**Abstrak.** Penyakit jantung merupakan salah satu penyebab kematian terbesar didunia. Penelitian ini mengkaji penggunaan algoritma Jaringan Saraf Tiruan (Artificial Neural Network/ANN) untuk mengklasifikasikan penyakit jantung berdasarkan 303 data medis pasien penyakit jantung yang diperoleh dari pusat dataset Kaggle. Data yang digunakan mencakup parameter medis seperti usia, jenis kelamin, tekanan darah, kadar kolesterol, dan hasil pemeriksaan lainnya. Berbagai arsitektur ANN diuji untuk menemukan konfigurasi optimal dalam hal jumlah lapisan tersembunyi dan neuron pada setiap lapisan. Kinerja model dievaluasi menggunakan metrik seperti akurasi, presisi, recall, dan F1-score. Dari hasil penelitian ukuran performa diperoleh nilai tingkat akurasi 97.06%, presisi 92.30%, recall 92.30%, dan F1-Score sebesar 92.30%.

**Kata kunci:** Artificial Neural Network, Klasifikasi Penyakit Jantung, Analisis Performa

### 1. PENDAHULUAN

Penyakit jantung masih menjadi salah satu masalah kesehatan utama di dunia, dengan tingkat kematian yang tinggi di berbagai negara. Deteksi dini dan diagnosis yang akurat sangat penting untuk mencegah komplikasi serius dan meningkatkan peluang kesembuhan pasien. Serangan jantung adalah salah satu penyakit yang paling mematikan di dunia dan salah satu penyakit yang banyak penderitanya adalah penyakit jantung dengan angka kematian mencapai 17,90% dari semua penyakit jantung. Kurangnya akses untuk mencari informasi tentang penyakit jantung ini menyebabkan peningkatan angka kematian setiap tahunnya. Karena itu dibutuhkan sebuah sistem klasifikasi yang dapat memberikan informasi tentang penyakit jantung serta dapat melakukan pengecekan klasifikasi secara dini tentang penyakit jantung yang dialami oleh seseorang.

Kami menggunakan berbagai parameter medis yang umum digunakan dalam diagnosis penyakit jantung, seperti usia, jenis kelamin, tekanan darah, kadar kolesterol, dan hasil

pemeriksaan lainnya. Data ini kemudian diproses dan dimasukkan ke dalam model ANN untuk pelatihan dan pengujian. Dengan menggunakan ANN, Kami harap dapat mencapai tingkat akurasi yang tinggi dalam klasifikasi penyakit jantung. Maka hal ini akan membantu dalam pengembangan sistem pendukung keputusan medis yang lebih efektif, yang dapat memberikan rekomendasi yang lebih akurat dan mempercepat proses diagnosis.

Penelitian ini juga diharapkan tidak hanya memberikan kontribusi akademis dalam bidang teknologi dan kesehatan, tetapi juga memiliki dampak praktis dalam meningkatkan kualitas pelayanan kesehatan.

## 2. METODOLOGI PENELITIAN

Proses pengumpulan dan pengolahan data dalam riset ini bertujuan untuk mengekstrak informasi penting pada data. Proses pengumpulan dan ekstraksi informasi tersebut dapat dilakukan menggunakan perangkat lunak dengan bantuan perhitungan statistika, matematika, ataupun teknologi Artificial Intelligence (AI). Jaringan Saraf Tiruan, atau Artificial Neural Network (ANN), adalah salah satu metode dalam bidang kecerdasan buatan yang terinspirasi oleh cara kerja otak manusia. ANN terdiri dari sekumpulan unit pemrosesan yang dikenal sebagai neuron, yang terorganisir dalam beberapa lapisan. Setiap neuron dalam jaringan ini berfungsi untuk menerima input, memprosesnya, dan menghasilkan output berdasarkan bobot dan fungsi aktivasi tertentu.

Jaringan neuron buatan terdiri atas kumpulan grup neuron yang tersusun dalam 3 lapisan yaitu 1) Lapisan Input (Input Layer) berfungsi sebagai penghubung jaringan ke dunia luar (sumber data). 2) Lapisan Tersembunyi (Hidden Layer). Suatu jaringan dapat memiliki lebih dari satu hidden layer atau bahkan tidak bisa punya sama sekali. 3) Lapisan Output (Output Layer). Prinsip kerja pada lapisan ini sama dengan prinsip kerja pada hidden layer dan ini juga digunakan fungsi sigmoid.

Hasil klasifikasi ANN dengan parameter kemudian akan diuji validitasnya menggunakan Confusion Matrix berupa Precision, Recall, dan Accuracy.

- a) Akurasi : didefinisikan sebagai tingkat kedekatan antara nilai prediksi dengan nilai aktual.
- b) Presisi: didefinisikan sebagai rasio item relevan yang dipilih terhadap semua item yang terpilih.
- c) *Recall*: didefinisikan sebagai rasio dari item relevan yang dipilih terhadap total jumlah item relevan yang tersedia.

- d) *F-Measure*: harmonic mean antara nilai presisi dan recall, *F-measure* juga kadang disebut dengan nama *F1-Score*.

$$\begin{aligned}
 AKURASI &= ((TP + TN))/((TP + TN + FP + FN)) \\
 PRESISI &= TP/(TP + FP) \\
 RECALL &= TP/(TP + FN) \\
 F - measure &= 2 (Presisi \times Recall)/(Presisi + Recall)
 \end{aligned}$$

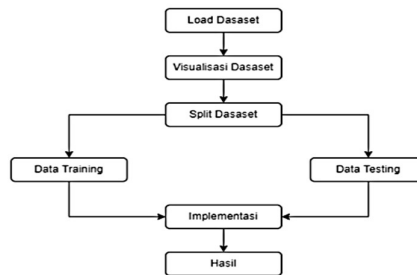
Keterangan Variabel :

- TP : True Positive
- TN : True Negative
- FP : False Positive
- FN : False Negative

Receiver Operating Characteristic (ROC) digunakan untuk mengetahui tingkat sensitivity dan specificity dari klasifikasi. Sensitivity adalah proporsi jumlah positif benar yang diidentifikasi dengan benar. Sedangkan specificity adalah proporsi negatif benar yang diidentifikasi dengan benar.

$$\begin{aligned}
 \text{Sensitivity (\%)} &= \frac{\text{jumlah positif benar} \times 100\%}{\text{jumlah positif benar} + \text{jumlah negatif salah}} \\
 \text{Specificity (\%)} &= \frac{\text{jumlah negatif benar} \times 100\%}{\text{jumlah negatif benar} + \text{jumlah positif salah}}
 \end{aligned}$$

Rancangan atau model pada penelitian ini seiring dengan alur proses metode ANN dimana dimulai dari mengumpulkan dataset, visualisasi dataset, split atau membagi dataset dataset, implementasi metode ANN dan menghitung performa metode ANN pada dataset tersebut.



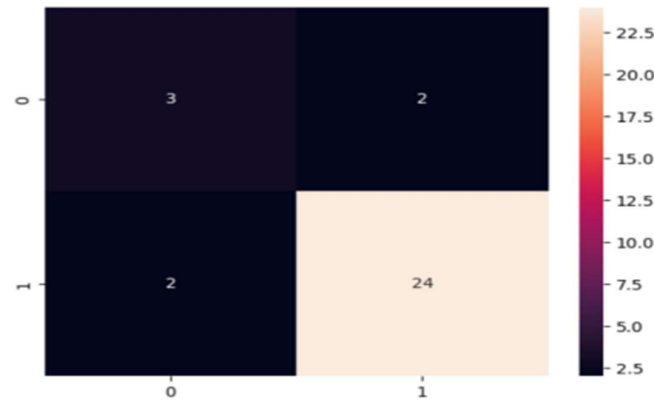
Data yang diolah pada penelitian ini memiliki beberapa atribut seperti CP (jenis nyeri dada), trestbps (tekanan darah istirahat (dalam mmHg saat masuk ke rumah sakit)), chol (kolesterol serum dalam mg/dl), fbs (gula darah puasa > 120 mg/dl) (1 = benar; 0 = salah), restecg (hasil elektrokardiografi istirahat), thalach (detak jantung maksimum tercapai), exang (angina akibat olahraga (1 = ya; 0 = tidak)), oldpeak (depresi ST yang diinduksi oleh olahraga relatif terhadap istirahat).

Teknik analisis dilakukan dengan cara menghitung performa metode ANN dimana performa yang diukur yaitu akurasi, recall, presisi dan f-measure menggunakan scikit-learn library sebagai machine learning tools, Tabel dibawah ini akan menampilkan perintah yang dijalankan.

Load	<pre>dataset = pd.read_csv('/content/drives/MyDrive/heart.csv')</pre>
Split	<pre>x = dataset.iloc[:, 2:].values y = dataset.iloc[:, 1].values x_train, x_test, y_train, y_test = train_test_split(x, y, test_size = 0.10, random_state = 0)</pre>
Test	<pre>y_pred = classifier.predict(x_test) y_pred = (y_pred &gt; 0.5) plt.plot(range(100),history.history['accuracy']) plt.title('history accuracy') plt.ylabel('accuracy') plt.xlabel('epoch') plt.show</pre>
Train	<pre>classifier = Sequential classifier.add(Dense(units=16,activation='relu')) classifier.add(Dense(units=16,activation='relu')) classifier.add(Dense(units=1,activation='sigmoid'))</pre>
Result	<pre>cm = confusion_matrix(y_test, y_pred) sns.heatmap(cm, annot=True) plt.savefig('a.png') score01 = precision_score(y_test, y_pred) score02 = recall_score(y_test, y_pred) score = f1_score(y_test, y_pred)</pre>

### 3. HASIL DAN PEMBAHASAN

Kesimpulan pengujian dilakukan setelah proses metode ANN terdapat 27 data benar 31 data yang ditampilkan dengan heatmap. Dan juga akan ditampilkan hasil uji serta validasi.



Hasil pelatihan model menunjukkan bahwa ANN mampu mencapai akurasi sebesar 97.06%, Metrik evaluasi lainnya juga diperhitungkan, dengan hasil sebagai berikut :

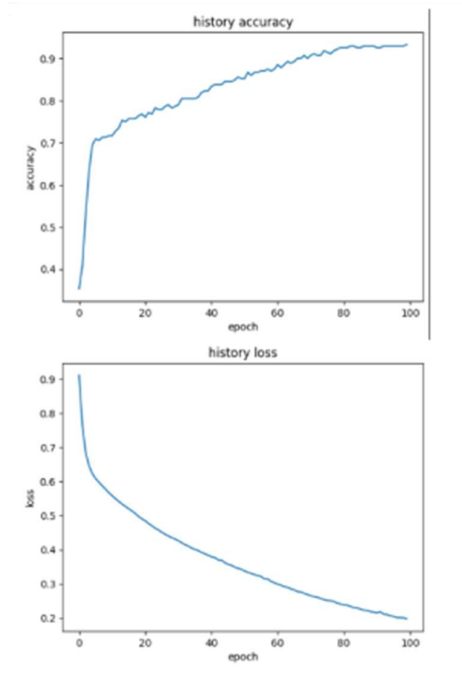
1	Presisi	92.30%
2	Recall	92.30%
3	F1-Score	92.30%

Hasil penelitian menunjukkan bahwa algoritma ANN dapat mengklasifikasikan penyakit jantung dengan tingkat akurasi yang tinggi. Akurasi sebesar 97.06% pada data pengujian menunjukkan bahwa model ANN memiliki kemampuan generalisasi yang baik dan dapat diaplikasikan pada data yang tidak terlihat sebelumnya.

Presisi yang mencapai 92.30% menunjukkan bahwa sebagian besar prediksi positif yang dibuat oleh model adalah benar, yang sangat penting dalam konteks medis untuk menghindari overdiagnosis. Recall sebesar 92.30% menandakan bahwa model ini juga sangat sensitif dalam mendeteksi kasus penyakit jantung, sehingga risiko underdiagnosis dapat diminimalkan. F1-score yang berada pada angka 92.30% mencerminkan keseimbangan yang baik antara presisi dan recall, menunjukkan bahwa model ini cukup andal dalam klasifikasi penyakit jantung.

Dalam penelitian ini, preprocessing data seperti ROC (Receiver Operating Characteristic) yang hanya bisa mencapai spesifikasi 40% dan Sensitivitas hanya 42.30%. Data akan ditampilkan sebagai berikut beserta dengan diagram akurasi dan loss.

Diagram Accurasi Dan Loss :



ANN memerlukan sejumlah besar data untuk dilatih dengan baik. Jika dataset yang tersedia kecil, model mungkin tidak dapat mencapai performa yang diharapkan. Overfitting dapat menjadi masalah jika model terlalu kompleks dibandingkan dengan data yang ada. Ini bisa diatasi dengan teknik seperti dropout, regularisasi, dan validasi silang. Meskipun hasil yang dicapai cukup memuaskan, penelitian ini juga memiliki beberapa keterbatasan. Salah satunya adalah ketergantungan pada dataset yang digunakan; hasil mungkin berbeda jika menggunakan dataset yang lebih beragam atau lebih besar. Selain itu, model ANN yang digunakan masih bisa dioptimalkan lebih lanjut dengan teknik-teknik seperti fine-tuning parameter, penggunaan arsitektur yang lebih kompleks, atau dengan menggabungkan metode machine learning lainnya.

#### 4. KESIMPULAN

Berdasarkan hasil penelitian sebelumnya yang hanya dapat nilai akurasi sebesar 73,77% dan nilai presisi 80,43% recall 84,09% dan F-measure 82,22%. maka kami dapat memperbaiki beberapa kesalahan seperti text\_size yang sebelumnya menggunakan 0.2 maka kami menggunakan 0.10 serta memperkecil epoch atau perulangan yang sebelumnya range 150, kami menggunakan range epoch sebesar 100 yang mungkin jadi penyebab tidak tingginya

suatu penelitian maka kami menarik beberapa kesimpulan, nilai akurasi paling yaitu sebesar 97.06% dan nilai presisi 92.30% recall 92.30% dan F-measure 92.30% Pada penelitian kami.

## **5. DAFTAR PUSTAKA**

- [1]. Do, M.S., Park, Y.J., and Lee, J.Y., "The Effect of Spreading Gain Control on a CDMA slotted ALOHA System", IEEE Trans, Computer Commun, Vol, 26, pp 996-1006, July 2005.
- [2]. Idris, N.R.N., Toh, C.L., and Elbuluk, E., "A New Torque and Flux Controller for DTC of Induction Machine", IEEE Transactions on Industry Application, Vol. 42, No.6, pp. 1358-1366, Nov/Dec 2006.
- [3]. M. Hosoz and H. M. Ertunc, "Artificial Neural Network Analysis of an Automobile Air Conditioning System" Energy Conv. Man., Vol. 47, pp. 1574-1587, July 2006.
- [4]. Ahmad, T., & Alghamdi, A. S. (2019). "Memprediksi Penyakit Jantung Menggunakan Jaringan Syaraf Tiruan." *Jurnal Universitas Indonesia - Ilmu Komputer dan Informasi*, 31(4), 450-458.
- [5]. Anderson, R., & Vasilakos, A. V. (2020). "Tinjauan Jaringan Syaraf Tiruan dalam Aplikasi Medis: Peluang dan Tantangan." *Jurnal Informatika Medis*, 141, 104232.
- [6]. D. Cahyanti, A. Rahmayani, and S. Ainy, "Analisis performa metode Knn pada Dataset pasien pengidap Kanker Payudara," *Indones. J. Data Sci.*, vol. 1, no. 2, pp. 39–43, 2020.
- [7]. A. Maulida, "Penerapan Metode Klasifikasi K-Nearest Neighbor pada Dataset Penderita Penyakit Diabetes," *Indones. J. Data Sci.*, vol. 1, no. 2, pp. 29–33, 2020.
- [8]. Hasran, "Klasifikasi Penyakit Jantung Menggunakan Metode K-Nearest Neighbor," *Indones. J. Data Sci.*, vol. 1, no. 1, pp. 1–4, 2020.
- [9]. Chaurasia, V., & Pal, S. (2014). "Pendekatan Data Mining untuk Mendeteksi Penyakit Jantung." *Jurnal Ilmu Komputer dan Teknologi Informasi*, 2, 56-66.