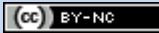


# Klasifikasi Penyakit Jantung Menggunakan Metode Artificial Neural Network

David Galih Pradana<sup>a,1</sup>, Muhammad Luthfi Alghifari<sup>a,2</sup>, Muhammad Farhan Juna<sup>a,3</sup>, Shulun Dwisiwi Palaguna<sup>a,4</sup>

<sup>a</sup> Universitas Amikom Yogyakarta, Jalan Ring Road Utara Condong Catur Depok-Sleman, Yogyakarta dan 55283, Indonesia

<sup>1</sup> david.12@students.amikom.ac.id; <sup>2</sup> muhammad.alghifari@students.amikom.ac.id; <sup>3</sup> muhammad.juna@students.amikom.ac.id; <sup>4</sup> shulun.palaguna@students.amikom.ac.id;

INFORMASI ARTIKEL	ABSTRAK
Diterima : 30 – 05 – 2022 Direvisi : 11 – 06 – 2022 Diterbitkan : 31 – 07 – 2022	Secara global, penyebab kematian nomor satu setiap tahunnya adalah penyakit jantung. Penyakit jantung adalah sebuah kondisi yang menyebabkan jantung tidak dapat melaksanakan tugasnya dengan baik. Klasifikasi merupakan salah satu bidang ilmu pada datamining dimana proses klasifikasi dilakukan dengan melakukan training terlebih dahulu pada sebagai data training dilanjutkan dengan proses testing untuk data yang baru. Artificial Neural Network (ANN) merupakan model penalaran yang didasarkan pada otak manusia. Penelitian ini mencakup pengukuran performa (akurasi, presisi, recall dan f-score) metode ANN dengan 304 data pasien penyakit jantung yang diperoleh dari pusat dataset Kaggle. Hasil dari pengukuran performa diperoleh nilai akurasi 73,77%, presisi 80,43%, recall 84,09% dan f1-score sebesar 82,22%.
<b>Kata Kunci:</b> ANN Analisis performa Data Mining Jantung Validasi	 

## I. Pendahuluan

Jantung adalah organ vital yang berfungsi sebagai pemompa darah untuk memenuhi kebutuhan oksigen dan nutrisi ke seluruh tubuh. Apabila jantung mengalami gangguan, peredaran darah dalam tubuh dapat terganggu sehingga menjaga kesehatan jantung sangatlah penting agar terhindar dari berbagai jenis penyakit jantung. Penyakit jantung adalah sebuah kondisi yang menyebabkan jantung tidak dapat melaksanakan tugasnya dengan baik. Hal ini disebabkan matinya sebagian otot jantung yang disebabkan karena penyempitan arteri koroner. Penyebab penyakit jantung pada umumnya terdapat dua faktor resiko yaitu faktor resiko yang tidak dapat diubah dan dapat diubah. Faktor resiko yang tidak dapat diubah antara lain usia, jenis kelamin, serta genetik atau keturunan. Sedangkan faktor resiko yang dapat diubah adalah hipertensi, kolesterol tinggi, obesitas, diabetes, kurang aktivitas fisik, dan konsumsi alkohol berlebih. 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 12,90% 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 [6]. Untuk melakukan sebuah klasifikasi sistem membutuhkan metode yang tepat dalam mengelola pengetahuan yang diadopsi dari pakar sehingga diperoleh hasil yang akurat. Salah satu metode yang dapat digunakan dalam klasifikasi ini adalah metode ANN (Artificial Neural Network). Penelitian ini mencoba untuk menghitung akurasi metode ANN jika diimplementasikan pada kasus klasifikasi penyakit jantung.

## II. Metode

### A. Data Mining

Data Mining adalah proses pengumpulan dan pengolahan data yang bertujuan untuk mengekstrak informasi penting pada data[1]–[10]. Proses pengumpulan dan ekstraksi informasi tersebut dapat dilakukan menggunakan perangkat lunak dengan bantuan perhitungan statistika, matematika, ataupun teknologi Artificial Intelligence (AI).

## B. Artificial Neural Network (ANN)

Algoritma Artificial Neural Network (ANN) merupakan model penalaran yang didasarkan pada otak manusia [11]–[13]. ANN terdiri dari sejumlah prosesor sangat sederhana dan saling berhubungan yang disebut neuron. Neuron yang terhubung dengan pembobotan (weight) melewati sinyal dari neuron satu ke neuron yang lain. ANN merupakan model yang meniru cara kerja jaringan neural biologis. Dengan melakukan proses belajar jaringan syaraf tiruan dapat mengatur dirinya untuk menghasilkan suatu respon yang konsisten terhadap rangkaian masukan. Jaringan saraf tiruan dirancang dan dilatih untuk memiliki kemampuan seperti manusia. Setiap neuron dapat memiliki beberapa masukan dan mempunyai satu keluaran. Jalur masukan pada suatu neuron bisa berisi data mentah atau data hasil olahan neuron sebelumnya. Sedangkan hasil keluaran suatu neuron dapat berupa hasil akhir atau berupa bahan masukan bagi neuron berikutnya.

Jaringan neuron buatan terdiri atas kumpulan grup neuron yang tersusun dalam 3 lapisan yaitu:

- Lapisan Input (Input Layer) berfungsi sebagai penghubung jaringan ke dunia luar (sumber data). Neuron-neuron ini tidak melakukan apapun pada data, hanya meneruskan data ini ke lapisan berikutnya.
- Lapisan Tersembunyi (Hidden Layer). Suatu jaringan dapat memiliki lebih dari satu hidden layer atau bahkan tidak bisa punya sama sekali. Jika jaringan memiliki beberapa lapisan tersembunyi, maka lapisan tersembunyi paling bawah yang menerima dari input dari input layer.
- Lapisan Output (Output Layer). Prinsip kerja pada lapisan ini sama dengan prinsip kerja pada hidden layer dan ini juga digunakan fungsi sigmoid. Tetapi keluaran dari lapisan ini sudah dianggap dari keluaran proses.

Hasil klasifikasi ANN dengan parameter kemudian akan diuji validitasnya menggunakan Confusion Matrix berupa Precision, Recall, dan Accuracy[14]–[23].

### 1) Akurasi

Akurasi didefinisikan sebagai tingkat kedekatan antara nilai prediksi dengan nilai actual.

### 2) Presisi

Presisi didefinisikan sebagai rasio item relevan yang dipilih terhadap semua item yang terpilih. Presisi dapat diartikan sebagai kecocokan antara permintaan informasi dengan jawaban terhadap permintaan tersebut.

### 3) Recall

Recall didefinisikan sebagai rasio dari item relevan yang dipilih terhadap total jumlah item relevan yang tersedia.

### 4) F-Measure

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

$$AKURASI = ((TP + TN))/((TP + TN + FP + FN))$$

$$PRESISI = TP/(TP + FP)$$

$$RECALL = TP/(TP + FN)$$

$$F - measure = 2 (Presisi \times Recall)/(Presisi + Recall)$$

Keterangan Variabel :

TP : True Positive

TN : True Negative

FP : False Positive

FN : False Negative

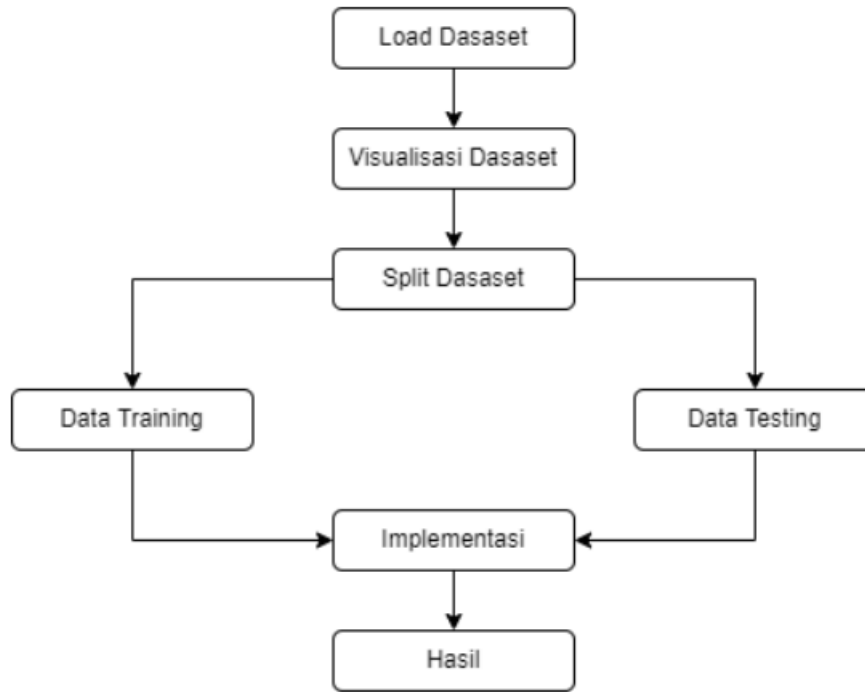
Receiver Operating Characteristic (ROC) digunakan untuk mengetahui tingkat sensitivity dan specificity dari klasifikasi. Sensitivity adalah proporsi jumlah positif benar yang diidentifikasi dengan benar. Sedangkan specificity adalah proporsi negatif benar yang diidentifikasi dengan benar. Perhitungan sensitivity dan specificity dapat dihitung dengan menggunakan rumus di bawah ini:

$$\text{Sensitivity (\%)} = \frac{\text{jumlah positif benar} \times 100\%}{\text{jumlah positif benar} + \text{jumlah negatif salah}}$$

$$\text{Specificity (\%)} = \frac{\text{jumlah negatif benar} \times 100\%}{\text{jumlah negatif benar} + \text{jumlah positif salah}}$$

### C. Rancangan Model

Rancangan atau model pada penelitian ini seiring dengan alur proses metode ANN dimana dimulai dari mengumpulkan dataset, visualisasi dataset, split atau membagi dataset, implementasi metode ANN dan menghitung performa metode ANN pada dataset tersebut. Flowchart model ditunjukkan pada [Gambar 1](#) berikut.



Gambar 1. Flowchart Model

### D. Sample dan Data

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). Data tersebut diperoleh dari website kaggle yang dipublish pada tahun 2018 dan dikelola oleh Bi Developer at Vision B.I.

### E. Teknik Analisis

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. Tahap awal yang dilakukan adalah mengumpulkan data, data yang kami gunakan pada penelitian ini menggunakan data dari website kaggle. Selanjutnya adalah melakukan split data dimana 10% digunakan untuk data testing dan 90% digunakan sebagai data training. Tahap selanjutnya adalah menerapkan algoritma ANN menggunakan data testing dan data training. Pada tahap terakhir dilakukan perhitungan performa dari seluruh data testing. Tabel dibawah menunjukkan beberapa perintah yang akan dijalankan dengan library scikit-learn library.

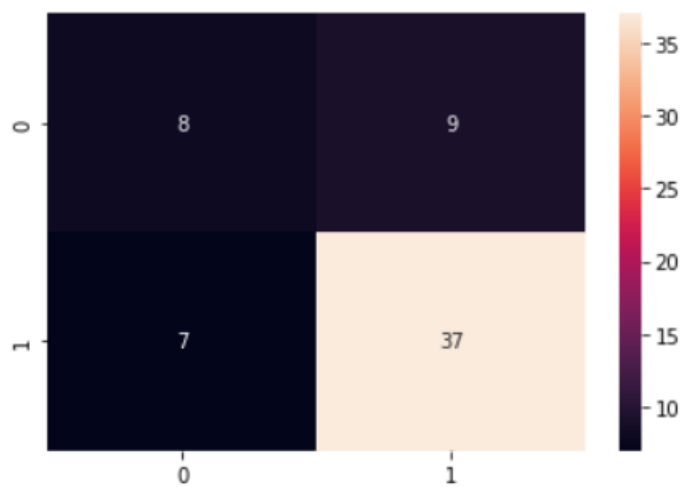
Tabel 1.

Load	<code>dataset = pd.read_csv('drive/MyDrive/DATASET/heart.csv')</code>
Split	<code>x = dataset.iloc[:, 2:].values</code> <code>y = dataset.iloc[:, 1].values</code> <code>x_train,x_test,y_train,y_test=train_test_split(x,y,test_size = 0.2,</code> <code>random_state = 0)</code>
Test	<code>y_pred = classifier.predict(x_test)</code> <code>y_pred = (y_pred &gt; 0.5)</code> <code>Result : plt.plot (range(150),history.history['accuracy'])</code> <code>plt.title('History Accuracy')</code> <code>plt.ylabel('accuracy')</code>

	plt.xlabel('epoch') plt.show()
Train	classifier = Sequential( classifier.add(Dense(units=16,activation='relu')) classifier.add(Dense(units=16,activation='relu')) classifier.add(Dense(units=1,activation='sigmoid'))
Result	cm = confusion_matrix(y_test, y_pred) sns.heatmap(cm, annot=True) plt.savefig('a.png') precision_score(y_test, y_pred) recall_score(y_test, y_pred) f1_score(y_test, y_pred)

### III. Hasil dan Pembahasan

Kesimpulan pengujian dilakukan setelah proses metode ANN terdapat 45 data benar 61 data yang ditampilkan dengan heatmap seperti pada Gambar 2. Hasil uji performa dan validasi ditampilkan pada Tabel 2 dan Tabel 3.



Gambar 2. Heatmap

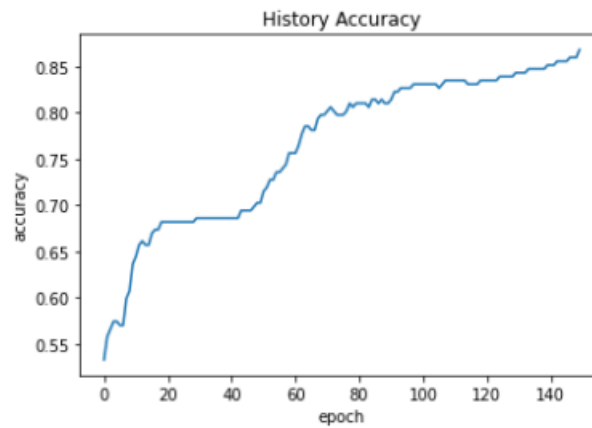
Tabel 2. Table uji Validasi

<i>Accuracy</i>	73,77%
<i>Recall</i>	84,09%
<i>Precision</i>	80,43%
<i>F1-Score</i>	82,22%

Tabel 3. ROC

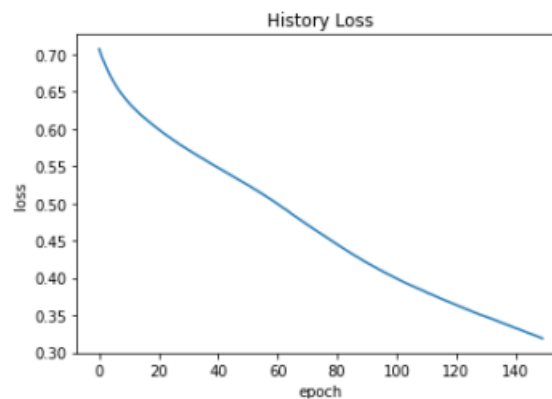
<i>Sensitivity(%)</i>	80,43%
<i>Specificity(%)</i>	53,33%

*History accuracy* ditunjukkan pada [Gambar 3](#).



Gambar 3. *History Accuracy*

*History loss* ditunjukkan pada [Gambar 4](#).



Gambar 4. *History loss*

#### IV. Kesimpulan

Berdasarkan hasil penelitian ini maka penulis dapat menarik beberapa kesimpulan, nilai akurasi paling yaitu sebesar 73,77% dan nilai presisi 80,43% recall 84,09% dan F-measure 82,22%. Pada penelitian selanjutnya dapat mengimplementasikan penelitian ke tahap selanjutnya berupa decision support system serta dapat menghitung usability aplikasi.

#### Daftar Pustaka

- [1] M. M. Baharuddin, T. Hasanuddin, and H. Azis, "Analisis Performa Metode K-Nearest Neighbor untuk Identifikasi Jenis Kaca," *Ilk. J. Ilm.*, vol. 11, no. 28, pp. 269–274, 2019.
- [2] Rizky Ade Putranto, Triastiti Wuryandari, and Sudarno, "Perbandingan Analisis Klasifikasi Antara Decision Tree Dan Support Vector Machine Multiclass Untuk Penentuan Jurusan Pada Siswa Sma," *J. Gaussian*, vol. 4, no. 4, pp. 1007–1016, 2015.
- [3] W. Safira Azis and dan Dedy Atmajaya, "Pengelompokan Minat Baca Mahasiswa Menggunakan Metode K-Means," *Ilk. J. Ilm.*, vol. 8, no. 2, pp. 89–94, 2016.
- [4] S. Chugh, K. Arivu Selvan, and R. K. Nadesh, "Prediction of heart disease using apache spark analysing decision trees and gradient boosting algorithm," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 263, no. 4, pp. 0–10, 2017, doi: 10.1088/1757-899X/263/4/042078.
- [5] H. Zhang, "The optimality of Naive Bayes," *Proc. Seventeenth Int. Florida Artif. Intell. Res. Soc. Conf. FLAIRS 2004*, vol. 2, pp. 562–567, 2004.

- [6] Y. Salim, Y. Puspitasari, H. Azis, and R. Anas, "The use of augmented reality to educate preschoolers on preventing dental malocclusion," *Bull. Soc. Informatics Theory Appl.*, vol. 3, no. 2, pp. 56–60, 2019.
- [7] L. Nurhayati and H. Azis, "Perancangan Sistem Pendukung Keputusan Untuk Proses Kenaikan Jabatan Struktural Pada Biro Kepegawaian," *Semin. Nas. Teknol. Inf. dan Multimed.*, pp. 6–7, 2016.
- [8] H. Azis and R. Wardoyo, "Penerapan Network Steganography Menggunakan Metode Modifikasi LACK Dan Layanan Message Authentication Code Pada Voip Network Steganography System with modification of LACK and Message Authentication Code on VoIP," *Semin. Nas. Komun. dan Inform.*, pp. 13–19, 2015.
- [9] H. Azis, F. T. Admojo, and E. Susanti, "Analisis Perbandingan Performa Metode Klasifikasi pada Dataset Multiclass Citra Busur Panah," *Techno.Com*, vol. 19, no. 3, 2020.
- [10] F. Muharram, H. Azis, and A. R. Manga, "Analisis Algoritma pada Proses Enkripsi dan Dekripsi File Menggunakan Advanced Encryption Standard (AES)," *Pros. Semin. Nas. Ilmu Komput. dan Teknol. Inf.*, vol. 3, no. 2, pp. 112–115, 2018.
- [11] D. D. Novita, A. B. Sesunan, M. Telaumbanua, S. Triyono, and T. W. Saputra, "Identifikasi Jenis Kopi Menggunakan Sensor E-Nose Dengan Metode Pembelajaran Jaringan Syaraf Tiruan Backpropagation," *J. Ilm. Rekayasa Pertan. dan Biosist.*, vol. 9, no. 2, pp. 205–217, 2021, doi: 10.29303/jrpb.v9i2.241.
- [12] Herman *et al.*, "Comparison of Artificial Neural Network and Gaussian Naïve Bayes in Recognition of Hand-Writing Number," *Proc. - 2nd East Indones. Conf. Comput. Inf. Technol. Internet Things Ind. EIConCIT 2018*, no. 1, pp. 276–279, 2018, doi: 10.1109/EIConCIT.2018.8878651.
- [13] I. Fauzan, "Artificial Intelligence (AI) Pada Proses Pengawasan dan Pengendalian Kepegawaian - Sebuah Eksplorasi Konsep Setelah Masa Pandemi Berakhir," *J. Civ. Serv.*, vol. 14, no. 1, pp. 31–42, 2020.
- [14] F. Tangguh and Y. Islami, "Analisis performa algoritma Stochastic Gradient Descent (SGD) dalam mengklasifikasi tahu berformalin," *Indones. J. Data Sci.*, vol. 3, no. 1, pp. 1–8, 2022.
- [15] H. Nursan and Muslim, "Penerapan Metode Digital Watermarking dan Privilege pada Dokumen Skripsi," *Indones. J. Data Sci.*, vol. 1, no. 1, pp. 19–22, 2020.
- [16] N. Litha and T. Hasanuddin, "Analisis Performa Metode Moving Average Model untuk Prediksi Jumlah Penderita Covid-19," *Indones. J. Data Sci.*, vol. 1, no. 3, pp. 87–95, 2020.
- [17] Sugiarti and Mirnawati, "Implementasi Algoritma Government Standard (GOST) dalam Pengamanan File Dokumen," *Indones. J. Data Sci.*, vol. 1, no. 2, pp. 52–56, 2020.
- [18] Hasran, "Klasifikasi Penyakit Jantung Menggunakan Metode K-Nearest Neighbor," *Indones. J. Data Sci.*, vol. 1, no. 1, pp. 1–4, 2020.
- [19] K. Ilunga, T. Lumbala, and K. Mulenda, "Application of Big data to configuration management in a PLM context," *Indones. J. Data Sci.*, vol. 3, no. 1, pp. 9–16, 2022.
- [20] 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.
- [21] A. Prasetya Wibawa, W. Lestar, A. Bella Putra Utama, I. Tri Saputra, and Z. Nabila Izdihar, "Multilayer Perceptron untuk Prediksi Sessions pada Sebuah Website Journal Elektronik," *Indones. J. Data Sci.*, vol. 1, no. 3, pp. 57–67, 2020, doi: 10.33096/ijodas.v1i3.15.
- [22] I. P. Putri, "Analisis Performa Metode K- Nearest Neighbor (KNN) dan Crossvalidation pada Data Penyakit Cardiovascular," *Indones. J. Data Sci.*, vol. 2, no. 1, pp. 21–28, 2021, doi: 10.33096/ijodas.v2i1.25.
- [23] 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.