

## PERBANDINGAN ALGORITMA *SUPPORT VECTOR MACHINES (SVM)* DAN *NEURAL NETWORK* UNTUK KLASIFIKASI PENYAKIT JANTUNG

Muhammad Dion Febrian Tino<sup>1</sup>, Herliyani Hasanah<sup>2</sup>, Tri Djoko Santosa<sup>3</sup>

<sup>1,2,3</sup>Teknik Informatika, Fakultas Ilmu Komputer, Universitas Duta Bangsa Surakarta

Email: [190103111@fikom.udb.ac.id](mailto:190103111@fikom.udb.ac.id)<sup>1</sup>, [herliyani\\_hasanah@udb.ac.id](mailto:herliyani_hasanah@udb.ac.id)<sup>2</sup>, [tri\\_djoko@udb.ac.id](mailto:tri_djoko@udb.ac.id)<sup>3</sup>

### ABSTRACT

*One of the most important organs of the body is the heart. Blood can be distributed well throughout the body if there is a heart. requirements for sample characterization of coronary disease. The algorithms used are Support Vector Machines (SVM) and Neural Net algorithms. This dataset has a total of 416 instances. The SVM calculation has 14 incorrect predictions and 70 correct predictions. There are 74 incorrect predictions and 344 correct predictions in the Confusion Matrix Neural Network algorithm. The accuracy of the SVM algorithm is higher than the Neural Net can be seen from the results of AUC 0.833, CA 0.83, F1 0.835, Accuracy 0.835 and Review 0.835. Based on experiments on the two algorithms, the SVM algorithm is better at classifying heart disease with an accuracy of 83%, while the Neural Net algorithm is better at solving problems related to heart disease with an accuracy value of 82%.*

### ABSTRAK

Salah satu organ tubuh yang paling penting adalah jantung. Darah dapat didistribusikan dengan baik ke seluruh tubuh jika ada jantung. persyaratan untuk contoh karakterisasi penyakit koroner. algoritma yang digunakan adalah *Support Vector Machines (SVM)* dan algoritma *Neural Net*. Kumpulan dataset ini memiliki jumlah 416 instance. Perhitungan SVM memiliki 14 prediksi salah dan 70 prediksi benar. Ada 74 prediksi salah dan 344 prediksi benar pada algoritma Confusion Matrix Neural Network. Keakuratan algoritma SVM lebih tinggi dari pada *Neural Net* dapat dilihat dari hasil AUC 0.833, CA 0.83, F1 0.835, Accuracy 0.835 serta Review 0.835. Dari percobaan terhadap kedua model tersebut, model SVM dapat memiliki hasil lebih baik dengan melakukan klasifikasi yang menghasilkan akurasi 83%, sedangkan algoritma *Neural Net* lebih baik dalam menyelesaikan masalah terkait penyakit jantung dengan nilai akurasi 82% .

Kata Kunci: Jantung, Confusion Matrix, Support Vector Machines, Nerual Network

---

### Riwayat Artikel :

Tanggal diterima : 25-05-2023

Tanggal revisi : 28-05-2023

Tanggal terbit : 29-05-2023

### DOI :

<https://doi.org/10.31949/infotech.vxix.xxx>

**INFOTECH journal** by Informatika UNMA is licensed under CC BY-SA 4.0 Copyright

© 20xx By Author



## 1. PENDAHULUAN

### 1.1. Latar Belakang

Salah satu organ tubuh yang paling penting adalah jantung. Darah dapat didistribusikan dengan baik ke seluruh tubuh jika terdapat jantung. Organ lain akan berhenti bekerja dan orang tersebut akan mati jika jantung di dalam tubuh tidak berfungsi dengan baik. (Hasanah, 2023).

Menurut data WHO kematian akibat penyakit jantung mencapai angka 26,6 juta kematian dari populasi orang dewasa (Utomo & Mesran, 2020). Serta berdasarkan Riskesdas tahun 2018 penyakit jantung meningkat setiap tahunnya, peningkatan ini terjadi pada pria (1,3%) dan wanita (1,6%). Jadi kurang lebih 2.784.064 jiwa di Indonesia memiliki riwayat heart disease. (Bahar & Moordingsih, 2021). Penyakit jantung susah di deteksi sejak dini, sehingga membuat banyak masyarakat yang mengidap penyakit tersebut tidak dapat diselamatkan.

Kami memerlukan pola klasifikasi penyakit jantung untuk membantu diagnosis awal. Klasifikasi yang digunakan merupakan Algoritma SVM dan *Neural Net* yang merupakan perhitungan dasar dengan eksekusi yang sangat baik.

Untuk menemukan solusi untuk masalah ini, akan dilakukan perbandingan antara algoritma *Support Vector Machines* (SVM) dengan algoritma *Neural Net* yang akan dilakukan dengan bahasa *python* untuk pengolahan data penyakit jantung. Dengan cara ini akan memungkinkan untuk mendapatkan hasil yang tepat dari kedua algoritma untuk penyakit jantung.

### 1.2. Tinjauan Pustaka

#### 1. Penyakit Jantung

Salah satu organ tubuh yang paling penting adalah jantung. Darah dapat didistribusikan dengan baik ke seluruh tubuh jika terdapat jantung. Organ lain akan berhenti bekerja dan orang tersebut akan mati jika jantung di dalam tubuh tidak berfungsi dengan baik. (Hasanah, 2023).

#### 2. Data Mining

*Data mining* adalah proses menentukan sebuah pola yang dapat diartikan memisahkan banyak data yang besar lalu di simpan dalam repository, untuk pengolahan data tersebut menggunakan teknik statistik, teknologi pengenalan pola dan matematika (Nabila et al., 2021).

#### 3. Algoritma *Support Vector Machines* (SVM)

*Support Vector Machines* (SVM) merupakan suatu metode *machine learning* yang bertujuan untuk memisahkan dua buah kelas (*hyperplane*) agar mendapatkan *hyperplane* yang terbaik. Untuk menggunakan algoritma SVM dibutuhkan data training, yang dimana data training tersebut akan diolah dan diubah menjadi prediksi data baru atau

biasa disebut dengan testing data (Suryati et al., 2023).

#### 4. Algoritma *Neural Network*

Algoritma *Neural Network* adalah sebuah syaraf yang ditiru dalam model non liner yang dibangun menggunakan individu dari model regresi (Dairoh et al., 2021).

#### 5. *Confusion Matrix*

*Confusion matrix* biasa disebut dengan error matrix. *Confusion matrix* biasanya dapat menemukan perbandingan klasifikasi dari data asli dengan klasifikasi yang di olah pada sistem (Daffa Setiawan Suparno, 2021).

### 1.3. Metodologi Penelitian

#### 1. Sumber Data

Penelitian ini menggunakan dataset yang diperoleh dari website *UCI Machine Learning Repository* yaitu kumpulan data penyakit jantung.

#### 2. Variable Data

Dataset ini berisi kumpulan dari dua dataset jantung yang di gabungkan dengan 76 atribut, akan tetapi untuk eksperimen hanya mengacu pada 12 atribut. Pada dataset ini memiliki total 416 *instance*, yang dimana dataset dari Hungaria memiliki 294 observasi dan dataset dari Switzerland memiliki 122 observasi.

#### 3. Tahapan Penelitian

##### 1) Identifikasi Masalah

Identifikasi masalah pada penelitian ini adalah bagaimana kita dapat memprediksi menggunakan kondisi medis bahwan seseorang terkena penyakit jantung.

##### 2) Pengumpulan dan Praprosesing Data

Data yang diolah ialah data dari website *UCI Machine Learning Repository*. Jumlah dari dataset keseluruhan yaitu 416 *instance* yang siap untuk di olah.

##### 3) Model Mining

Pada tahapan ini penelitian menggunakan model *Support Vector Machines* (SVM) dan algoritma *Neural Network*. Semua algoritma diimplementasikan menggunakan bahasa *python* dan pustaka *scikit*.

##### 4) *Learner*

Pada tahapan ini dilakukan penelitian dari dataset.

##### 5) Evaluasi

Tahapan evaluasi prediksi menggunakan AUC, CA, F1, *Precision*, *Recall*, *Confusion Matrix* dan ROC analysis.

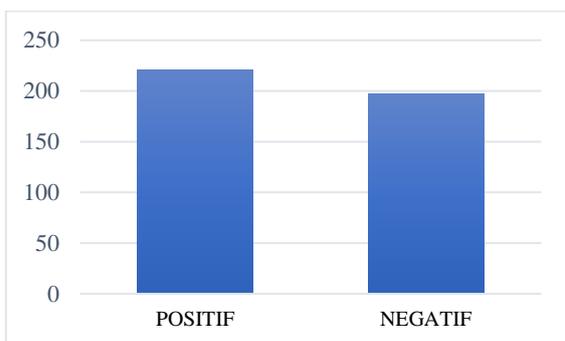
2. PEMBAHASAN

2.1 Implementasi Algoritma *Support Vector Machines* (SVM) dan Algoritma *Neural Network*.

Data yang digunakan pada penelitian ini merupakan dataset pasien jantung dan kondisi klinis yang diambil dari website UCI Machine Learning. Dataset ini, memiliki 416 *instance* dengan 12 atribut. Pada dataset pasien jantung dari UCI Machine Learning, terdapat 11 atribut yang menjadi kondisi klinis pasien agar dapat digunakan untuk atribut prediksi. Pada atribut terakhir memiliki 2 nilai, yang pertama nilai 1 untuk para pasien penyakit jantung dan yang kedua nilai 0 untuk para pasien tidak terkena penyakit jantung. Berikut merupakan gambaran atribut dalam bentuk tabulasi :

Tabel 1. Atribut Prediksi

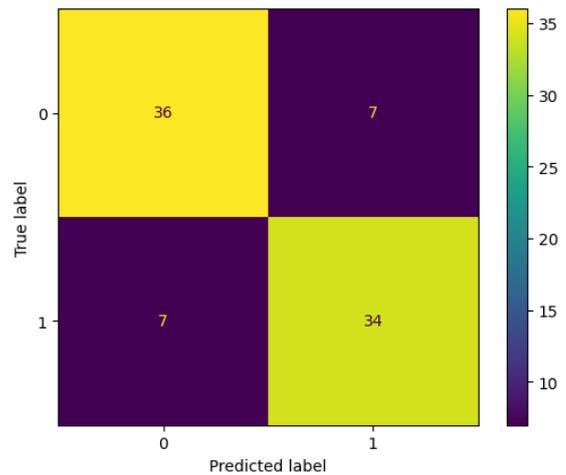
No	Atrbut	Tipe Data	Keterangan
X <sub>1</sub>	Age	numerik	Usia pasien
X <sub>2</sub>	Sex	numerik	Jenis kelamin
X <sub>3</sub>	ChestPainType	numerik	Jenis sakit dada
X <sub>4</sub>	RestingBP	numerik	Tekanan darah
X <sub>5</sub>	Cholesterol	numerik	Jumlah koletrol
X <sub>6</sub>	FastingBS	numerik	Tekanan gula darah
X <sub>7</sub>	RestingECG	numerik	Hasil tes tekanan electrodiogram
X <sub>8</sub>	MaxHR	numerik	Detak jantung maksimum
X <sub>9</sub>	ExerciseAngina	numerik	Nyeri dada ketika olahraga
X <sub>10</sub>	Oldpeak	numerik	Oldpeak
X <sub>11</sub>	ST_Slope	numerik	Kemiringan detak jantung setelah berolahraga
Y	HeartDisease	text	1: positif 0: negatif



Gambar 1. Distribusi Data

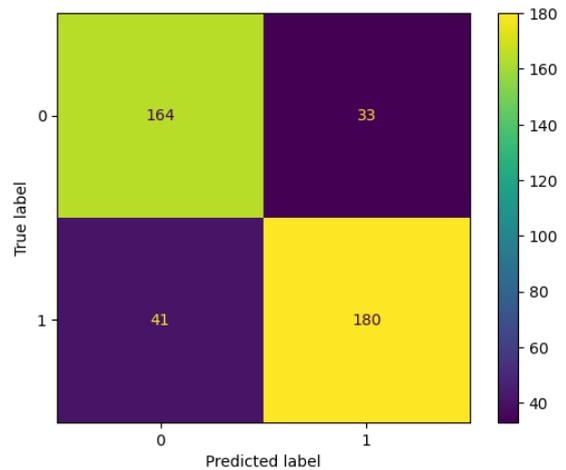
2.2 Evaluasi

Pada tahapan evaluasi prediksi, meliputi AUC,CA,F1, *Precision*, *Recall*, *Confussion Matrix* dan *ROC analysis*. CA merupakan sebuah akurasi yang diambil dari dataset. *Precision* merupakan keadaan akurasi data yang dimana terjadi dua kemungkinan yaitu 1 dan 0. *Recall* berfungsi sebagai rasio. F1 sebagai pebanding antara *recall* dan *precision*. Berikut merupakan hasil prediksi menggunakan *confusion matrix* pada gambar di bawah .



Gambar 2. *Confusion Matrix* SVM

*Confusion Matrix* algoritma *Support Vector Machines* (SVM) pada gambar 2 memiliki total data testing 84, terdapat 70 prediksi yang benar dan 14 prediksi yang salah. Jadi tingkat akurasi pada dataset dapat dihitung  $\frac{70}{84} 100\% = 83\%$ .



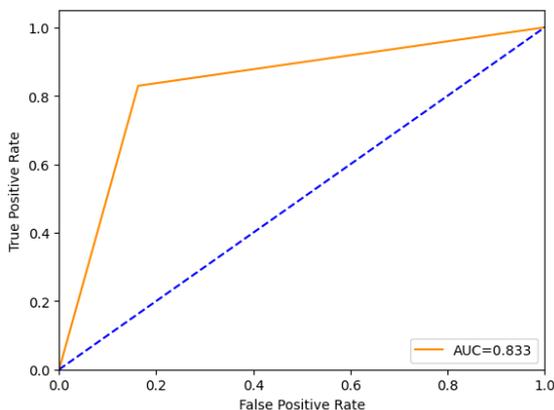
Gambar 3. *Confusion Matrix* Neural Net

*Confusion Matrix* algoritma *Neural Network* ada gambar 3 memiliki total data testing 418, terdapat 344 prediksi yang benar dan 74 prediksi yang salah. Jadi tingkat akurasi pada dataset dapat dihitung  $\frac{344}{418} 100\% = 82\%$ .

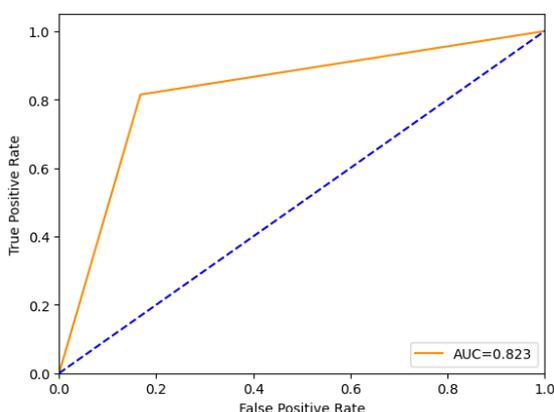
Tabel 2. Kinerja Model

Model	AUC	CA	F1	Precision	Recall
SVM	0.833	0.83	0.835	0.835	0.835
NN	0.823	0.82	0.825	0.825	0.82

Dapat dilihat pada tabel 2 Algoritma *Support Vector Machine* (SVM) mengungguli algoritma NN dalam akurasi. Nilai yang dihasilkan ialah AUC 0,833, CA 0,83, F1 0,835, *Precision* 0,835 serta *Recall* 0,835.



Gambar 4. ROC SVM



Gambar 5. ROC Neural Net

Dari analisis kedua kurva ROC diatas dapat dilihat nilai *TP Rate* dan *FP Rate* kurang dari 1. Kedua kurva ROC diatas memiliki nilai kualitas *classifier good* karena kedua ROC memiliki rentan akurasi 0,80-0,90. Jadi dapat diketahui bahwa model SVM memiliki tingkat akurasi pada *classifite* lebih baik dibandingkan algoritma *Neural Net*, dikarenakan nilai kruva SVM lebih besar dibandingkan kurva *Neural Net*.

#### 4. KESIMPULAN

Berdasarkan hasil penelitian diatas, algoritma *Support Vector Machines* (SVM) lebih akurat dibandingkan dengan algoritma *Neural Net* dalam menyelesaikan klasifikasi penyakit jantung, algritma SVM memiliki akurasi 83%. Diketahui dari hasil ROC bahwa classifier algoritma SVM memiliki nilai akurasi yang lebih tinggi dengan nilai AUC sebesar 0,833, sedangkan classifier algoritma *Neural Net* memiliki nilai akurasi dengan nilai AUC sebesar 0,823.

#### PUSTAKA

- Bahar, R. N. A., & Moordiningsih, M. (2021). Meningkatkan Kesejahteraan Subjektif Penderita Penyakit Jantung Melalui Pelatihan Kebersyukuran. *Jurnal Studia Insania*, 8(2), 119. <https://doi.org/10.18592/jsi.v8i2.3881>
- Daffa Setiawan Suparno. (2021). Pengenalan Pola Untuk Mengetahui Jumlah Target Pengunjung Mall Berdasarkan Usia, Gender, Pendapatan, Pertahun, Pengeluaran, Tujuannya Untuk Mempermudah Mengetahui Target Pasar Menggunakan Metode EDA, K-Means, Hierarchical Clustering, Confusion Matrix. *SAKTI – Sains, Aplikasi, Komputasi Dan Teknologi Informasi.*, 3, 61–69.
- Dairoh, D., Bakti, V. K., & Naufal, M. (2021). Neural Network dan Particle Swarm Optimization untuk Penunjang Keputusan Antipasi Mahasiswa Pra Lulus Bekerja Sesuai Bidang. *MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 21(1), 151–158. <https://doi.org/10.30812/matrik.v21i1.1164>
- Hasanah, H. (2023). Perbandingan Tingkat Akurasi Algoritma Support Vector Machines (SVM) dan C45 dalam Prediksi Penyakit Jantung. In *Universitas Nusantara PGRI Kediri. Kediri* (Vol. 2).
- Nabila, Z., Rahman Isnain, A., & Abidin, Z. (2021). ANALISIS DATA MINING UNTUK CLUSTERING KASUS COVID-19 DI PROVINSI LAMPUNG DENGAN ALGORITMA K-MEANS. *Jurnal Teknologi Dan Sistem Informasi (JTSI)*, 2(2), 100. <http://jim.teknokrat.ac.id/index.php/JTSI>
- Suryati, E., Ari Aldino, A., Penulis Korespondensi, N., & Suryati Submitted, E. (2023). Analisis Sentimen Transportasi Online Menggunakan Ekstraksi Fitur Model Word2vec Text Embedding Dan Algoritma Support Vector Machine (SVM). 4(1), 96–106. <https://doi.org/10.33365/jtsi.v4i1.2445>
- Utomo, D. P., & Mesran, M. (2020). Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 4(2), 437. <https://doi.org/10.30865/mib.v4i2.2080>