

IMPLEMENTASI DATA MINING UNTUK MENDIAGNOSA PENYAKIT JANTUNG DENGAN ALGORITMA NAÏVE BAYES DI RS TK II MOH. RIDWAN MAURAKSA

¹Sultan Rahma Fernanda, ²Putri Dina Mardika

Univesitas Indrprasta PGRI
Jalan Raya Tengah No 80, Kelurahan Gedong, Pasar Rebo, Jakarta Timur
[1Sultanrahma280@gmail.com](mailto:Sultanrahma280@gmail.com), [2putridinamar@gmail.com](mailto:putridinamar@gmail.com)

ABSTRAK

Kesehatan jantung memainkan peran penting dalam kehidupan manusia karena jantung mendistribusikan darah yang kaya oksigen ke seluruh tubuh. Penyakit kardiovaskular, termasuk serangan jantung, merupakan salah satu penyebab utama kematian di dunia. Menurut laporan WHO, lebih dari 17,8 juta orang meninggal setiap tahun karena penyakit jantung. Faktor risiko termasuk gaya hidup, usia, dan kondisi sosial ekonomi memiliki pengaruh signifikan terhadap prevalensi penyakit ini. Deteksi dini melalui algoritma pembelajaran mesin, seperti Naïve Bayes, menunjukkan akurasi tinggi dalam menganalisis data medis. Penelitian ini menggunakan algoritma Naïve Bayes untuk memprediksi risiko penyakit jantung pada pasien di RS TK II Ridwan Mauraksa, dan menunjukkan akurasi sebesar 80% pada 25 data uji. Hasil ini selaras dengan temuan sebelumnya yang menunjukkan efektivitas algoritma pembelajaran mesin dalam prediksi medis. Dengan pengaplikasian yang lebih luas, algoritma ini dapat membantu penanganan dini dan mengurangi angka kematian akibat penyakit jantung. Studi ini berpotensi diterapkan di rumah sakit lain untuk mendukung upaya deteksi dini penyakit jantung.

Kata Kunci: Penyakit Jantung, Naïve Bayes, Data Mining

ABSTRACT

Heart health Heart health plays an important role in human life as it distributes oxygen-rich blood throughout the body. Cardiovascular diseases, including heart attacks, are one of the leading causes of death in the world. According to a WHO report, more than 17.8 million people die each year from heart disease. Risk factors including lifestyle, age, and socioeconomic conditions have a significant influence on the prevalence of these diseases. Early detection through machine learning algorithms, such as Naïve Bayes, shows high accuracy in analyzing medical data. This study used the Naïve Bayes algorithm to predict the risk of heart disease in patients at Ridwan Mauraksa TK II Hospital, and showed an accuracy of 80% on 25 test data. This result is in line with previous findings showing the effectiveness of machine learning algorithms in medical prediction. With wider application, this algorithm can help early treatment and reduce mortality from heart disease. This study could potentially be applied in other hospitals to support early detection of heart disease.

Keywords: Heart Disease, Naïve Bayes, Data Mining

PENDAHULUAN

Berbagai penelitian telah dilakukan untuk mendukung prediksi dan diagnosis penyakit jantung menggunakan metode data mining. (Riani et al., 2019) mengimplementasikan metode Naïve Bayes untuk memprediksi penyakit jantung dengan membandingkan hasil probabilitas posterior yang menghasilkan diagnosis bahwa pasien sehat. (Sabransyah et al., 2017) menggunakan metode Naïve Bayes untuk menganalisis karakteristik pasien rawat inap di RSUD AWS, menggambarkan variabel seperti jenis kelamin, umur, kadar kolesterol, dan riwayat penyakit jantung keluarga. (Sabransyah et al., 2017) menggunakan seleksi fitur dan

algoritma Support Vector Machine (SVM) untuk klasifikasi penyakit jantung koroner, dengan hasil bahwa seleksi fitur terbaik meningkatkan akurasi klasifikasi. (Yudistira & Putra, 2021), menemukan bahwa Naïve Bayes memiliki akurasi tertinggi namun masih memerlukan Studi sebelumnya oleh (Alhamad et al., 2019) menunjukkan bahwa metode ensemble berbasis Weighted Vote lebih unggul dalam meningkatkan akurasi prediksi penyakit jantung dibandingkan algoritma individual (Alhamad et al., 2019) Penelitian serupa juga menyoroti pentingnya pengumpulan dan pengelolaan data yang akurat untuk meningkatkan kualitas prediksi diagnosis (Giacomini et al., 2023) Dengan meningkatnya penerapan machine learning di bidang kesehatan, penting untuk

mengembangkan model diagnosis yang lebih efisien melalui implementasi algoritma seperti Naïve Bayes (Gagliano, 2017)

Tujuan Penelitian Tujuan utama penelitian ini adalah untuk membangun sistem pendukung keputusan dalam mendiagnosa penyakit jantung di RS TK II Moh. Ridwan Meuraksa. Sistem ini bertujuan untuk membantu dokter dalam meningkatkan akurasi diagnosa, meningkatkan efisiensi proses diagnosis, dan mengurangi keterlambatan penanganan pasien penyakit jantung. Studi oleh menunjukkan bahwa sistem pendukung keputusan berbasis machine learning dapat mempercepat proses diagnosis dan memberikan hasil yang lebih akurat dibandingkan metode konvensional (Allah et al., 2022)

METODE PENELITIAN

Penelitian ini menggunakan algoritma Naive Bayes sebagai metode utama untuk mendiagnosa penyakit jantung, yang didasarkan pada Teorema Bayes dan prinsip probabilitas kondisional. Proses klasifikasi dimulai dengan memasukkan data dari dataset yang dipilih, seperti yang tersedia di Kaggle. Setelah melalui tahap pra-pemrosesan data, dataset dibagi untuk melatih model Naive Bayes. Algoritma ini kemudian mengevaluasi probabilitas prior dari masing-masing kelas, seperti "Penyakit Jantung" (C=1) dan "Tidak Ada Penyakit Jantung" (C=0). Pada langkah berikutnya, probabilitas likelihood dihitung untuk setiap fitur yang ada dalam dataset. Misalnya, jika terdapat 50 pasien dengan penyakit jantung dan 253 pasien tanpa penyakit jantung, maka probabilitas priori masing-masing kelas dihitung terlebih dahulu. Selanjutnya, algoritma menghitung probabilitas posterior untuk menentukan kelas mana yang paling mungkin berdasarkan fitur-fitur yang ada. Kelas dengan probabilitas posterior tertinggi dipilih sebagai prediksi akhir. Keseluruhan proses ini memberikan hasil klasifikasi yang cukup akurat dan dapat diandalkan dalam mendeteksi risiko penyakit jantung. Algoritma Naive Bayes, dengan asumsi independensi antar fitur, membuktikan efektivitasnya dalam mendiagnosa penyakit jantung dengan cepat dan efisien, serta menunjukkan potensi besar untuk diterapkan

dalam sistem diagnostik medis berbasis data mining.

HASIL DAN PEMBAHASAN

Berisi Hasil dari Rumus dasar dari Naïve Bayes untuk menghitung probabilitas kelas C (misalnya, pasien memiliki penyakit jantung atau tidak) berdasarkan fitur X adalah:

$$P(C|X) = \frac{P(C) \cdot P(X)}{P(X)}$$

Melakukan perhitungan manual dengan menggunakan data sederhana:

Age	Sex	Cp	Condition
63	1	1	0
67	1	4	1
67	1	4	1
37	1	3	0
41	0	2	0
56	1	2	0
62	0	4	1
57	0	4	0
63	1	4	1
53	1	4	1

Gambar 1. Contoh Data

1. Hitung Probabilitas prior P(C) P(C=1): Jumlah Kasus Positif (6/10) =0.6 P(C=0): Jumlah Kasus Negatif (4/10) =0.4
2. Hitung Likelihood (PX|C) Age: P (Age = 63|C=1):2 dari 6 (0.3330) P (Age = 63|C=0):1 dari 4 (0.2)
3. Sex: P (Sex = 1|C=1):5 dari 6 (0.833) P (Sex = 1|C=0):3 dari 4 (0.75) Cp: P (Cp = 4|C=1):4 dari 6 (0.667) 56 P (Cp = 4|C=0):3 dari 4 (0.25)
4. Hitung Evidence P(X) P(X)=P(Age=63) · P(Sex=1) · P(CP=4) P(Age=63): 3 dari 10 (0.3) P(Sex=1): 8 dari 10 (0.8) P(CP=4): 6 dari 10 (0.6) P(X)=0.3×0.8×0.6=0.144
5. Hitung Probabilitas P(C|X) Untuk menentukan apakah pasien dengan fitur tersebut memiliki penyakit jantung atau tidak:

$$P(1|X) = \frac{P(X | C = 1) \cdot P(C = 1)}{P(X)}$$

$$P(0|X) = \frac{P(X | C = 0) \cdot P(C = 0)}{P(X)}$$

$$P(1|X) = \frac{(0.333 \times 0.833 \times 0.667) \times 0.6}{0.144} \approx 0.64$$

$$P(0|X) = \frac{(0.25 \times 0.75 \times 0.25) \times 0.4}{0.144} \approx 0.14$$

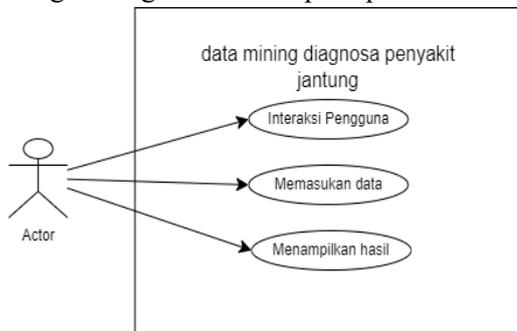
Kesimpulan dari hasil perhitungan manual Karena $P(C=1|X) > P(C=0|X)P(C = 1|X) > P(C = 0|X)P(C=1|X) > P(C=0|X)$, pasien dengan fitur tersebut lebih mungkin memiliki penyakit jantung. Dengan penjelasan sebagai berikut:

1. Probabilitas Prior (Prior Probability): Ini adalah probabilitas awal sebelum mempertimbangkan fitur XXX.
2. Likelihood: Menghitung probabilitas fitur XXX diberikan kelas CCC.
3. Evidence: Probabilitas keseluruhan dari fitur XXX.
4. Posterior Probability: Probabilitas bahwa pasien memiliki penyakit jantung, diberikan fitur yang diketahui.

Permodelan Perangkat Lunak

Use case diagram

Use case diagram menggambarkan fungsionalitas yang diharapkan dari sebuah sistem. Sebuah use case diagram merepresentasikan interaksi antara aktor dengan sistem. Berikut merupakan Use Case Diagram digambarkan seperti pada Gambar 1



Gambar 1. Use Case Diagram

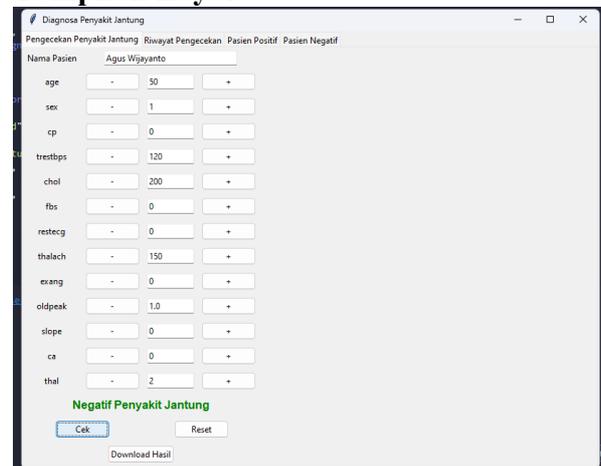
Activity Diagram

Activity diagram menunjukkan alur kerja atau proses dalam sistem. Berikut adalah aktivitas untuk aplikasi diagnosis penyakit jantung:



Gambar 2. Activity Diagram

Tampilan Layar



Gambar 3. Tampilan Layar

Pada menu ini akan menampilkan inputan data yang akan di mining, data yang memiliki hasil yang serupa dari data-data yang di kumpulkan akan menghasilkan output yang sama, contohnya seperti inputan di atas ini data yang dimiliki adalah data yang Negatif Penyakit Jantung.

Nilai Akurasi

```
Accuracy: 0.7666666666666667
Confusion Matrix:
[[22 10]
 [ 4 24]]
Classification Report:

```

	precision	recall	f1-score	support
0	0.85	0.69	0.76	32
1	0.71	0.86	0.77	28
accuracy			0.77	60
macro avg	0.78	0.77	0.77	60
weighted avg	0.78	0.77	0.77	60

```
Precision: 0.7058823529411765
Recall: 0.8571428571428571
F1-Score: 0.7741935483870968
```

Gambar 2. Nilai Akurasi

Dari hasil akurasi yang sudah di hasilkan maka akurasi yang di dapat adalah 76% , prcisionya 70% dan recall memiliki nilai 85% dari hasil data mining untuk mendiagnosa penyakit jantung dengan metode naïve bayes.

SIMPULAN DAN SARAN

Penelitian ini berhasil Berdasarkan hasil penelitian ini, terdapat beberapa saran yang dapat meningkatkan implementasi algoritma Naive Bayes dalam mendiagnosis penyakit jantung melalui pendekatan data mining. Pertama, meskipun akurasi yang dicapai sebesar 74% menunjukkan kemampuan algoritma dalam mengolah data medis, peningkatan akurasi masih dapat dicapai dengan pengujian lebih lanjut menggunakan dataset yang lebih besar dan beragam. Hal ini akan membantu meningkatkan generalisasi model sehingga lebih dapat diandalkan dalam berbagai kondisi klinis.

Kedua, pengembangan lebih lanjut aplikasi berbasis GUI sangat dianjurkan, dengan menambahkan fitur analisis data secara lebih mendalam, seperti visualisasi statistik dari hasil diagnosis atau peringatan dini jika ada anomali dalam data pasien. Ini akan memudahkan tenaga medis dalam menginterpretasi hasil dan mempercepat proses pengambilan keputusan.

Terakhir, meskipun aplikasi ini sudah mendukung penyimpanan hasil diagnosis, fitur integrasi dengan sistem rekam medis elektronik (EHR) dapat meningkatkan efisiensi dan aksesibilitas data pasien untuk tujuan pemantauan jangka panjang. Penelitian

lebih lanjut juga bisa dilakukan untuk menggabungkan algoritma lain, seperti Support Vector Machine (SVM) atau pendekatan ensemble, untuk mengeksplorasi potensi peningkatan akurasi dan keandalan model.

UCAPAN TERIMA KASIH

Saya ingin menyampaikan terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan dukungan dan kontribusi dalam penulisan artikel ini. Terima kasih kepada rekan-rekan yang telah berbagi pengetahuan dan wawasan, serta kepada para pembaca yang telah meluangkan waktu untuk membaca dan memberikan umpan balik. Semoga artikel ini bermanfaat dan dapat memberikan wawasan baru bagi semua yang membacanya.

DAFTAR PUSTAKA

- Ahmed, M., & Husien, I. (2024). Heart Disease Prediction Using Hybrid Machine Learning: A Brief Review. *Journal of Robotics and Control (JRC)*, 5(3), 884–892. <https://doi.org/10.18196/jrc.v5i3.21606>
- Alhamad, A., Azis, A. I. S., Santoso, B., & Taliki, S. (2019). Prediksi Penyakit Jantung Menggunakan Metode-Metode Machine Learning Berbasis Ensemble – Weighted Vote. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 5(3), 352. <https://doi.org/10.26418/jp.v5i3.37188>
- Allah, E. M. A., El-Matary, D. E., Eid, E. M., & Dien, A. S. T. El. (2022). Performance Comparison of Various Machine Learning Approaches to Identify the Best One in Predicting Heart Disease. *Journal of Computer and Communications*, 10(02), 1–18. <https://doi.org/10.4236/jcc.2022.102001>
- Furizal, Ma'arif, A., & Rifaldi, D. (2023). Application of Machine Learning in Healthcare and Medicine: A Review. *Journal of Robotics and Control (JRC)*, 4(5), 621–631. <https://doi.org/10.18196/jrc.v4i5.19640>
- Gagliano, M. (2017). Applications of Machine Learning in Medical Diagnosis. *The International Conference on Computational Science, November*.

- Giacomini, M., Mora, S., Lazarova, E., Maggi, N., & Ruggiero, C. (2023). Improving data quality in healthcare: medical informatics tools and standards. *Convegno Nazionale Di Bioingegneria, June*.
- Maral, P. V. (2017). Heart Disease Prediction Using Naive Bayes and K-Means Techniques. *NOVATEUR PUBLICATIONS International Journal of Research Publications in Engineering and Technology, 3(6)*, 2454–7875.
- Naser, M. A., Majeed, A. A., Alsabab, M., Al-Shaikhli, T. R., & Kaky, K. M. (2024). A Review of Machine Learning's Role in Cardiovascular Disease Prediction: Recent Advances and Future Challenges. *Algorithms, 17(2)*. <https://doi.org/10.3390/a17020078>
- Riani, A., Susianto, Y., & Rahman, N. (2019). Implementasi Data Mining Untuk Memprediksi Penyakit Jantung Menggunakan Metode Naive Bayes. *Journal of Innovation Information Technology and Application (JINITA), 1(01)*, 25–34. <https://doi.org/10.35970/jinita.v1i01.64>
- Sabransyah, M., Novia Nasution, Y., Deny, F., & Amijaya, T. (2017). Aplikasi Metode Naive Bayes dalam Prediksi Risiko Penyakit Jantung Naive Bayes Method for a Heart Risk Disease Prediction Application. *Jurnal*

EKSPONENSIAL, 8(2).

- Yudistira, N., & Putra, A. F. (2021). Algoritma Decision Tree Dan Smote Untuk Klasifikasi Serangan Jantung Miokarditis Yang Imbalance. *Jurnal Litbang Edusaintech, 2(2)*, 112–122. <https://doi.org/10.51402/jle.v2i2.48>

Biografi Penulis

Sultan Rahma Fernanda,

Riwayat Pendidikan SDN 04 kramat jati dari tahun 2007 s.d. 2013. Pendidikan menengah di SMPN 20 Jakarta dari tahun 2013 – 2016. SMA di Widya Manggala dari tahun 2016 – 2018. Pendidikan tinggi di Universitas Indraprasta PGRI dari tahun 2020 – 2024



Putri Dina Mardika,

Universitas Indraprasta PGRI Jakarta, Fakultas Teknik dan Ilmu Komputer, Program Studi Teknik Informatika sejak tahun 2014 hingga saat ini. Jabatan fungsional saat ini sedang berproses Lektor dan telah tersertifikasi dosen pada tahun 2021. Research Interest Artificial Intelligence

