

**KLASIFIKASI PENYAKIT JANTUNG MENGGUNAKAN
METODE *NAIVE BAYES CLASSIFIER***

SKRIPSI



Oleh:

MUHAMMAD HERIYAN

2355201128

TEKNIK INFORMATIKA

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS LANCANG KUNING
PEKANBARU
2026**

HALAMAN PERSETUJUAN

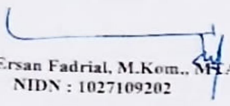
HALAMAN PERSETUJUAN

Nama : MUHAMMAD HERIYAN
Nim : 2355201128
Program Studi : Teknik Informatika
Judul : Klasifikasi Penyakit Jantung Menggunakan Metode *Naive Bayes*

Pekanbaru, 26 Januari 2026

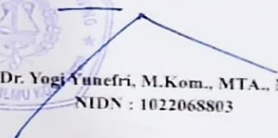
Disetujui oleh

Ketua

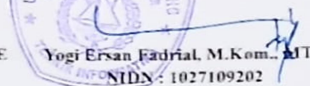

Yogi Ersan Fadrial, M.Kom., MTA
NIDN : 1027109202

Mengetahui.

Dekan Fakultas Ilmu Komputer


Dr. Yogi Yusefri, M.Kom., MTA., MCE
NIDN : 1022068803

Ketua Prodi Teknik Informatika


Yogi Ersan Fadrial, M.Kom., MTA
NIDN : 1027109202

HALAMAN PENGESAHAN

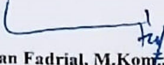
HALAMAN PENGESAHAN

Nama : MUHAMMAD HERIYAN
Nim : 2355201128
Program Studi : Teknik Informatika
Judul Skripsi : Klasifikasi Penyakit Jantung Menggunakan Metode *Naive Bayes*

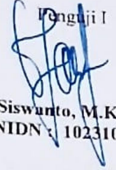
Skripsi ini telah dipertahankan di depan Dewan Penguji pada tanggal 13 Januari 2026 ujian komprehensif sebagai persyaratan untuk mendapatkan gelar Sarjana Komputer (S.Kom).

Dewan Penguji,

Ketua


Yogi Ersan Fadrial, M.Kom., MTA
NIDN : 1027109202

Penguji I


Didik Siswanto, M.Kom., MTA
NIDN : 1023108501

Penguji II


Assoc. Prof. Dr. Ahmad Zamsuri, M.Kom.,
MTA., MCF
NIDN : 1007078602

Mengetahui,

Dekan
Fakultas Ilmu Komputer


Dr. Yogi Yusefri, M.Kom., MTA., MCE
NIDN : 1022068803

Ketua Program Studi
Teknik Informatika


Yogi Ersan Fadrial, M.Kom., MTA
NIDN : 1027109202

KARYA ILMIAH

PERNYATAAN KEASLIAN KARYA ILMIAH

Dengan ini saya menyatakan bahwa dalam tugas akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan disuatu perguruan tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis oleh orang lain, kecuali yang secara acuan ditulis dalam naskah ini dan disebutkan dalam Daftar Pustaka.

Apabila ditemukan suatu jiplakan atau plagiat maka saya bersedia menerima akibat berupa sanksi akademis dan sanksi lain yang diberikan oleh yang berwenang sesuai ketentuan peraturan dan perundang-undangan yang berlaku.

Pekanbaru, 09 Januari 2026

Yang Membuat Pernyataan



Muhammad Heriyan

MOTTO DAN PERSEMBAHAN

Kedua Orang Tua (Herril St. Mudo(Alm) dan Khairiyah)

Penulis mengucapkan rasa terima kasih dan penghargaan yang sebesar-besarnya kepada kedua orang tua tercinta atas doa, kasih sayang, serta dukungan moril dan materil yang terus mengiringi perjalanan penulis, sehingga studi dapat diselesaikan dengan baik hingga berhasil meraih gelar Sarjana Komputer (S.Kom.).

Adik-adik (Fitria Herdiyani dan Vinna Herizkha)

Penulis mengucapkan terima kasih kepada adik-adik tercinta atas doa dan dukungan yang diberikan sehingga skripsi ini dapat diselesaikan dengan baik dan tepat waktu.

Dosen Pembimbing

Terima kasih kepada Bapak Yogi Ersan Fadrial, S.Kom., M.Kom., MTA selaku dosen pembimbing skripsi atas bimbingan, arahan, masukan, serta motivasi yang diberikan selama proses penyusunan skripsi ini.

Teman-teman

Terima kasih kepada teman seperjuangan Muhammad Ibnu Sabil, S.Kom, M. Farel Asyrofi, S.Kom dan Rizal Triwardana, S.Kom. yang turut saling mendukung dan memberikan bantuan, kritik, saran, serta motivasi untuk berjuang.

“Cinta menumbuhkan ketulusan, perasaan melahirkan kesabaran, dan ilmu memberi arah dalam setiap langkah. Dengan ketiganya, setiap proses perjuangan menjadi bermakna dan mengantarkan pada keberhasilan yang hakiki.”

KATA PENGANTAR

Alhamdulillah Robbil 'Alamin, puji dan syukur ke hadirat Allah *Subhanahu wa Ta'ala*. Karena atas limpahan segala rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi ini yang berjudul **KLASIFIKASI PENYAKIT JANTUNG MENGGUNAKAN METODE *NAIVE BAYES CLASSIFIER*** dengan baik dan tepat waktu. Tak lupa, shalawat serta salam penulis panjatkan kepada junjungan besar, Nabi Muhammad *Shallallahu 'Alaihi wa Sallam*.

Penyusunan skripsi ini tidak terlepas dari bantuan berbagai pihak baik berupa bimbingan, kritik, saran, bantuan, dukungan, motivasi maupun doa. Oleh karena itu, ucapan terima kasih penulis sampaikan kepada :

1. Bapak Assoc.prof. Dr. Yogi Yunefri, S.Kom., M.Kom., MTA., MCE.
Selaku Dekan Fakultas Ilmu Komputer Universitas Lancang Kuning.
2. Bapak Afriansyah, S.Kom., M.Kom., MTA. Selaku Wakil Dekan I
Fakultas Ilmu Komputer Universitas Lancang Kuning.
3. Ibu Assoc. prof. Dr. Lucky Lhaura Van FC, S.Kom., M.Kom., MTA.
Selaku Wakil Dekan II Fakultas Ilmu Komputer Universitas Lancang
Kuning.
4. Bapak Sutejo, M.Kom, Wakil Dekan III Fakultas Ilmu Komputer
5. Bapak Yogi Ersan Fadrial, S.Kom., M.Kom., MTA. Selaku Dosen
Pembimbing Skripsi yang telah memberikan bimbingan, arahan, dan
motivasi kepada penulis dalam penyusunan skripsi ini.

6. Bapak Assoc. Prof. Didik Siswanto, M.Kom., MTA selaku dosen penguji 1 yang telah memberikan arahan,dan motivasi kepada penulis dalam menyusun skripsi ini
7. Bapak Assoc. Prof. Dr. Ahmad Zamsuri, M.Kom, MTA., MCF selaku dosen penguji 2 yang telah memberikan arahan,dan motivasi kepada penulis dalam menyusun skripsi ini
8. Bapak dan Ibu Dosen Fakultas Ilmu Komputer Universitas Lancang Kuning, yang telah memberikan ilmu dan pengetahuan yang sangat bermanfaat selama masa studi.
9. Kedua Orang Tua dan Keluarga tercinta, yang selalu memberikan doa, dukungan, dan semangat yang tiada henti kepada penulis.
10. Teman-teman seperjuangan, yang telah memberikan dukungan moral, berbagi pengalaman, dan semangat.
11. Semua pihak yang tidak bisa disebutkan satu per satu, terima kasih atas segala dukungan serta bantuannya.

Semoga penulisan skripsi ini bermanfaat bagi pembaca dan pengembang selanjutnya serta dapat memberikan kontribusi yang berarti bagi perkembangan ilmu pengetahuan, khususnya di bidang teknik informatika. Penulis menyadari bahwa penulisan skripsi ini masih jauh dari kata sempurna serta masih terdapat kekurangan yang disebabkan oleh keterbatasan kemampuan dan pengalaman.

Oleh karena itu, dengan segala kerendahan hati penulis mengharapkan adanya kritik dan saran yang bersifat membangun guna untuk menyempurnakan penulisan kedepannya, terima kasih.

Pekanbaru, 27 Januari 2026

Muhammad Heriyan

KLASIFIKASI PENYAKIT JANTUNG MENGGUNAKAN METODE *NAIVE BAYES CLASSIFIER*

Muhammad Heriyan¹, Yogi Ersan Fadrial², Didik Siswanto³, Ahmad Zamsuri⁴
heriyanahmad17@gmail.com¹, yogiersanfadrial@unilak.ac.id²,
didik@gmail.com³, ahmadzamsuri@unilak.ac.id⁴

ABSTRAK

Penyakit jantung merupakan salah satu penyebab utama kematian di dunia dan sering kali sulit terdeteksi pada tahap awal karena gejala yang tidak spesifik. Proses diagnosis secara konvensional membutuhkan biaya tinggi serta keterbatasan tenaga medis, sehingga diperlukan pendekatan alternatif berbasis teknologi. Penelitian ini bertujuan untuk membangun model klasifikasi penyakit jantung menggunakan metode *Naïve Bayes Classifier* dengan memanfaatkan dataset publik dari Kaggle yang telah divalidasi oleh dokter. Data yang digunakan terdiri dari 916 data pasien dengan atribut medis seperti usia, jenis kelamin, tekanan darah, kolesterol, detak jantung maksimum, *exercise angina*, dan *ST slope*. Hasil penelitian menunjukkan bahwa metode *Naïve Bayes Classifier* mampu mengklasifikasikan penyakit jantung dengan tingkat akurasi yang baik serta memberikan hasil yang efisien dan mudah diinterpretasikan. Penelitian ini diharapkan dapat menjadi referensi dalam pengembangan sistem pendukung keputusan di bidang kesehatan, khususnya untuk deteksi dan klasifikasi penyakit jantung.

Kata kunci: Penyakit Jantung, Klasifikasi, *Naïve Bayes Classifier*, Data Mining.

CLASSIFICATION OF HEART DISEASE USING THE NAIVE BAYES CLASSIFIER METHOD

Muhammad Heriyan¹, Yogi Ersan Fadrial², Didik Siswanto³, Ahmad Zamsuri⁴
heriyanahmad17@gmail.com¹, yogiersanfadrial@unilak.ac.id²,
didik@gmail.com³, ahmadzamsuri@unilak.ac.id⁴

ABSTRACT

Heart disease is one of the leading causes of death worldwide and is often difficult to detect at an early stage due to non-specific symptoms. Conventional diagnostic processes require high costs and are limited by the availability of medical professionals, making technology-based approaches necessary. This study aims to develop a heart disease classification model using the Naïve Bayes Classifier method by utilizing a public dataset from Kaggle that has been validated by medical experts. The dataset consists of 916 patient records with medical attributes such as age, gender, resting blood pressure, cholesterol level, maximum heart rate, exercise-induced angina, and ST slope. The results show that the Naïve Bayes Classifier method is able to classify heart disease effectively with satisfactory accuracy and efficient performance. This research is expected to contribute as a reference for the development of decision support systems in the healthcare sector, particularly for the detection and classification of heart disease.

Keywords: Heart Disease, Classification, Naïve Bayes Classifier, Data Mining.

DAFTAR ISI

HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iii
KATA PENGANTAR.....	vii
DAFTAR ISI	xii
DAFTAR GAMBAR.....	xiv
DAFTAR TABEL.....	xv
Lampiran.....	xvi
BAB I PENDAHULUAN	1
A. Latar Belakang.....	1
B. Rumusan Masalah	4
C. Batasan Masalah	4
D. Tujuan Penelitian.....	4
E. Manfaat Penelitian	5
F. Sistematika Penulisan.....	6
BAB II TINJAUAN PUSTAKA.....	7
A. Tinjauan Penelitian Terdahulu	7
B. Landasan Teori	16
1. Klasifikasi.....	16
2. Penyakit Jantung	17
3. Jenis-Jenis Penyakit Jantung.....	17
4. <i>Naïve Bayes Classifier</i>	19
BAB III METODOLOGI PENELITIAN.....	23
A. Tahapan Penelitian	23
B. Lokasi dan Waktu Penelitian	30
C. Data Yang Digunakan.....	30
D. Teknik Pengumpulan Data.....	32
E. Metode/Pemodelan Yang Digunakan	33

BAB IV GAMBARAN UMUM OBJEK PENELITIAN	34
A. Gambaran Umum Klinik Narita.....	34
BAB V HASIL DAN PEMBAHASAN	40
A. Pengumpulan Data	40
B. <i>Data Preparation</i>	44
C. Pemodelan Naïve Bayes	49
D. Evaluasi dan Analisa	55
E. Evaluasi Akurasi	56
F. Evaluasi menggunakan <i>Confusion Matrix</i>	57
G. Pengujian Sistem.....	63
BAB VI KESIMPULAN DAN SARAN.....	68
A. Kesimpulan	68
B. Saran.....	69
DAFTAR PUSTAKA	71
Lampiran.....	76

DAFTAR GAMBAR

Gambar 2. 1 Teorema Bayes	20
Gambar 2. 2 Confusion Matrix.....	21
Gambar 2. 3 Metrik Evaluasi	22
Gambar 3. 1 Tahapan Penelitian.....	24
Gambar 3. 2 Visualisasi Metode Naïve Bayes	33
Gambar 4. 1 Struktur Keanggotaan	37
Gambar 4. 2 Logo Klinik	38
Gambar 5. 1 Hasil Data Cleaning Dataset	45
Gambar 5. 2 Hasil Data Transformation Dataset.....	46
Gambar 5. 3 Hasil Normalisasi Data Numerik.....	47
Gambar 5. 4 Hasil Pembagian Dataset	49
Gambar 5. 5 Diagram Alur Pemodelan Naive Bayes	50
Gambar 5. 6 Kurva Model Klasifikasi Jenis Penyakit Jantung.....	57
Gambar 5. 7 Confusion Matrix Pada Jenis Penyakit Jantung	59
Gambar 5. 8 Tampilan Utama Pengujian Sistem.....	64
Gambar 5. 9 Tampilan Input Data Pasien	65
Gambar 5. 10 Pilihan Model Prediksi.....	65
Gambar 5. 11 Hasil Prediksi Model yang dipakai	66

DAFTAR TABEL

Tabel 5. 1 Deskripsi Atribut Dataset.....	41
Tabel 5. 2 Deskripsi Atribut Dataset Terpilih.....	42

Lampiran

Lampiran 1 Validasi Dataset Foto dengan dr. Riane Parlin, Sp. KKLP, MKM	76
Lampiran 2 Hasil Akhir Foto dengan dr. Riane Parlin, Sp. KKLP, MKM	78
Lampiran 3 Lembar Wawancara.....	79
Lampiran 4 Dataset.....	80
Lampiran 5 Surat Penelitian.....	83
Lampiran 6 Surat Bimbingan	84

BAB I

PENDAHULUAN

A. Latar Belakang

Penyakit jantung adalah kondisi fungsi jantung tidak dapat berfungsi secara optimal. Kondisi ini terjadi akibat matinya sebagian otot jantung yang dipicu oleh penyempitan arteri koroner. Faktor penyebab penyakit jantung umumnya terbagi menjadi dua kategori, yaitu faktor risiko yang tidak dapat diubah dan faktor risiko yang dapat diubah. Faktor risiko yang tidak dapat diubah meliputi usia, jenis kelamin, dan faktor genetik atau keturunan. Sementara itu, faktor risiko yang dapat diubah meliputi tekanan darah tinggi (hipertensi), kadar kolesterol yang tinggi, obesitas, diabetes, kurangnya aktivitas fisik, serta konsumsi alkohol secara berlebihan. Serangan jantung merupakan salah satu penyakit paling mematikan di dunia, dengan tingkat kematian mencapai 12,90% dari total kasus penyakit jantung. Minimnya akses informasi terkait penyakit jantung menjadi salah satu penyebab meningkatnya angka kematian setiap tahunnya. (Pradana et al., 2022)

Banyak pasien penderita penyakit jantung tidak menyadari gejala-gejala awal yang muncul, dan tidak sedikit penderita penyakit jantung meninggal akibat serangan jantung. Rendahnya kesadaran akan pola hidup sehat dan kurangnya informasi mengenai penyakit jantung membuat seseorang sulit mengenali tanda-tanda awal penyakit ini. Proses deteksi penyakit jantung biasanya dilakukan secara manual, yaitu dengan berkonsultasi langsung ke dokter spesialis jantung dan menjalani serangkaian pemeriksaan laboratorium, yang hasilnya harus

dikonsultasikan kembali dengan dokter spesialis. Proses ini tentunya memerlukan biaya yang cukup besar. Mengingat risiko kematian yang sangat tinggi, diperlukan sebuah sistem yang mampu mendeteksi penyakit jantung secara akurat dan dengan biaya yang lebih terjangkau. (Wibisono & Fahrurozi, 2021) Banyak dari mereka enggan memeriksakan kondisi jantungnya karena terkendala oleh biaya pengobatan yang tinggi, terbatasnya pelayanan pasien, kurangnya tenaga medis, terutama dokter spesialis jantung, serta jam kerja dokter yang terbatas. Penyakit jantung sering kali tidak disadari oleh penderitanya, sehingga ketika terdeteksi, kondisinya sudah parah dan bahkan bisa berujung pada kematian. Oleh karena itu, dibutuhkan sebuah teknologi yang dapat meniru cara berpikir manusia, seperti teknologi kecerdasan buatan. (Dona et al., 2021)

Beberapa metode klasifikasi yang sering digunakan meliputi *Naïve Bayes*, *K-Nearest Neighbor*, *Decision Tree*, *Random Forest*, dan *Support Vector Machines*. Pada penelitian yang dilakukan oleh (Retnasari & Rahmawati, 2017) yang berjudul “Diagnosa prediksi penyakit jantung dengan model algoritma *Naïve Bayes* dan algoritma C4.5” hasil penelitian tersebut menunjukkan bahwa algoritma *Naïve Bayes* untuk prediksi penyakit jantung mendapatkan nilai akurasi 86,67% dan algoritma C4.5 mendapatkan nilai 83,70%. Kemudian penelitian yang dilakukan oleh (Hardiani & Putri, 2024) yang berjudul “Implementasi Metode *Naïve Bayes Classifier* Untuk Klasifikasi Stunting Pada Balita” hasil penelitian tersebut menunjukkan bahwa algoritma *Naïve Bayes* untuk prediksi penyakit jantung mendapatkan nilai presisi sebesar 85%, nilai recall sebesar 91%, dan nilai akurasi sebesar 80%. Dari beberapa penelitian terdahulu memiliki kemampuan

klasifikasi yang baik dan algoritma ini telah terbukti memberikan hasil akurasi yang baik dalam berbagai studi sebelumnya.

Naïve Bayes Classifier adalah metode klasifikasi berbasis probabilistik yang sederhana, yang menghitung probabilitas dengan mempertimbangkan frekuensi dan kombinasi nilai dari dataset yang tersedia. Algoritma ini didasarkan pada teorema Bayes dan mengasumsikan bahwa setiap atribut bersifat independen atau tidak saling bergantung terhadap nilai pada variabel kelas (Sahar, 2020). Penelitian ini menggunakan *Framework* CRISP-DM yang merupakan suatu pendekatan terstruktur untuk pengelolaan dan pemodelan data, yang mencakup tahapan-tahapan yang sistematis mulai dari pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, hingga penerapan model. Kerangka kerja ini dirancang untuk membantu dalam proses eksplorasi dan analisis data secara efisien, terutama dalam proyek data mining dan *machine learning* (Hardiani & Putri, 2024).

Oleh karena itu, penelitian ini bertujuan untuk mengembangkan model klasifikasi penyakit jantung berbasis *Naïve Bayes Classifier* yang berjudul “Klasifikasi Penyakit Jantung Menggunakan Metode *Naïve Bayes Classifier*”. Dataset yang digunakan dalam penelitian ini berasal dataset publik dari Kaggle terdiri dari 919 data yang akan divalidasi dengan dokter. Pendekatan penelitian ini diharapkan dapat memberikan pemahaman yang lebih baik mengenai efektivitas algoritma dalam menganalisis data penyakit jantung, serta menentukan metode yang paling optimal berdasarkan metrik evaluasi tertentu. Dengan demikian, hasil penelitian ini dapat berkontribusi dalam pengembangan metode klasifikasi yang

lebih akurat dan efisien, serta menjadi referensi bagi penelitian lanjutan di bidang kesehatan berbasis teknologi.

B. Rumusan Masalah

Berdasarkan permasalahan yang telah diuraikan dalam latar belakang, penulis merumuskan masalah penelitian ini sebagai berikut:

1. Bagaimana membangun model klasifikasi penyakit jantung menggunakan algoritma *Naïve Bayes Classifier*?
2. Bagaimana mengukur akurasi hasil kinerja model *Naïve Bayes Classifier* dalam mengklasifikasikan risiko penyakit jantung ?

C. Batasan Masalah

Adapun batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Dataset publik yang diperoleh dari *Kaggle* akan melalui proses validasi / verifikasi oleh dokter guna memastikan bahwa atribut yang digunakan relevan dengan kondisi klinis penyakit jantung.
2. Penelitian difokuskan pada klasifikasi jenis penyakit jantung, bukan prediksi tingkat keparahan.
3. Penelitian menggunakan algoritma *Naïve Bayes Classifier* dan tidak membandingkan dengan algoritma lain.
4. Penelitian ini tidak mencakup pengembangan aplikasi atau implementasi sistem berbasis model yang dihasilkan.

D. Tujuan Penelitian

Tujuan penelitian terkait klasifikasi penyakit jantung menggunakan metode *Naïve Bayes Classifier* dirumuskan sebagai berikut:

1. Membuat model klasifikasi guna mengenali jenis penyakit jantung dengan memanfaatkan dataset publik dari *Kaggle* yang telah melalui proses validasi dokter.
2. Menggunakan algoritma *Naïve Bayes Classifier* sebagai metode utama dalam proses klasifikasi penyakit jantung.
3. Melakukan evaluasi terhadap performa model melalui metrik pengukuran seperti akurasi, presisi, *recall*, dan *F1-score*.
4. Memberikan sumbangsih ilmiah dalam pengembangan pemanfaatan data mining dan kecerdasan buatan untuk analisis di bidang kesehatan, khususnya penyakit jantung.

E. Manfaat Penelitian

Manfaat penelitian mengenai klasifikasi penyakit jantung menggunakan metode *Naïve Bayes Classifier* dapat dirumuskan sebagai berikut:

1. Menyediakan sistem berbasis data yang dapat membantu tenaga medis dalam mengidentifikasi jenis penyakit jantung berdasarkan gejala.
2. Memberikan kontribusi dalam pemanfaatan data wawancara sebagai sumber informasi dalam pengembangan sistem klasifikasi di bidang kesehatan.
3. Menjadi referensi akademik terkait penerapan algoritma *Naïve Bayes* pada data berbasis teks dalam bidang medis.

F. Sistematika Penulisan

Agar jelas dan mudah dipahami, maka penulis memberikan rangkuman dari pembuatan proposal ini untuk memberikan gambaran pokok, sebagai berikut :

BAB I PENDAHULUAN

Berisi latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Berisi teori-teori yang mendukung penelitian ini, seperti algoritma klasifikasi, *Naïve Bayes*, dan machine learning.

BAB III METODE PENELITIAN

Berisi penjelasan mengenai metode yang digunakan dalam penelitian ini, seperti metode pengumpulan data, analisis data, dan pengembangan sistem.

BAB IV JADWAL PELAKSANAAN

Berisi waktu jadwal pelaksanaan penelitian yang akan dilaksanakan.

BAB VI

KESIMPULAN DAN SARAN

A. Kesimpulan

1. Dalam penelitian ini, sistem klasifikasi penyakit jantung berhasil dirancang dengan memanfaatkan metode *Naïve Bayes* Multi-Class yang dipadukan dengan teknik *stacking* (*GaussianNB*, *ComplementNB*, dan *BernoulliNB*). Hasil pengujian menunjukkan bahwa model mampu menangkap pola-pola relevan dari data klinis pasien secara efektif serta menghasilkan prediksi yang konsisten dan andal.
2. Model *Stacked Naïve Bayes* v6 pada penelitian ini berhasil memperoleh akurasi sekitar 85,26%, menunjukkan bahwa model mampu mengidentifikasi sebagian besar kategori penyakit jantung dengan baik. Peningkatan kinerja tersebut tidak lepas dari proses *feature engineering*, normalisasi, serta penyeimbangan data yang memberikan kontribusi signifikan terhadap kualitas prediksi.
3. Hasil pengujian memperlihatkan bahwa model mampu mengenali beberapa kategori penyakit jantung dengan akurasi tinggi, seperti *Angina*, Arteri Koroner, dan *Cardiomyopathy*. Meski demikian, beberapa kelas dengan jumlah data yang terbatas—seperti *Aritmia*—masih menunjukkan kesalahan klasifikasi. Kondisi ini umum terjadi

pada model multi-kelas dan dapat ditingkatkan melalui penambahan atau pemerataan jumlah data pada kelas terkait.

4. Sistem yang dibangun tidak hanya mampu mengklasifikasikan jenis penyakit, tetapi juga menyediakan fitur prediksi interaktif, pencarian berdasarkan nama penyakit, serta evaluasi model melalui confusion matrix dan grafik metrik performa. Dengan fitur tersebut, sistem berfungsi tidak hanya sebagai alat prediksi, tetapi juga sebagai sarana edukasi dan analisis bagi pengguna.

B. Saran

1. Menambah Jumlah Data Pasien

Untuk meningkatkan akurasi dan kestabilan prediksi pada seluruh kelas penyakit jantung, jumlah data yang digunakan sebaiknya diperbanyak. Variasi data seperti usia, tekanan darah, kadar kolesterol, detak jantung, serta karakteristik klinis lainnya juga perlu diperluas. Dengan dataset yang lebih besar dan beragam, model dapat mempelajari lebih banyak pola nyata sehingga risiko kesalahan klasifikasi dapat diminimalkan.

2. Menguji Model Klasifikasi yang Lebih Unggul

Pada tahap pengembangan berikutnya, disarankan untuk mencoba algoritma yang lebih canggih seperti *Random Forest*, *XGBoost*, *SVM*, atau *Deep Neural Network*. Model-model tersebut biasanya lebih mampu menangkap pola non-linear dan berpotensi meningkatkan akurasi, terutama pada kelas penyakit dengan jumlah data yang terbatas.

3. Mengoptimalkan Teknik Balancing dan Feature Engineering

Selain *SMOTE* dan *SMOTETomek*, metode penyeimbangan lain seperti *ADASYN* atau *Borderline-SMOTE* dapat dipertimbangkan. Pada bagian rekayasa fitur, penambahan fitur baru seperti interaksi antar variabel, rasio fisiologis, atau fitur statistik lainnya dapat memperkuat kemampuan model dalam mengenali pola penyakit jantung.

4. Pengembangan Sistem Menjadi Aplikasi Sebenarnya

Sistem ini berpeluang besar untuk dikembangkan menjadi aplikasi web atau mobile sehingga dapat digunakan oleh masyarakat maupun tenaga medis sebagai alat analisis awal kondisi jantung. Namun demikian, prediksi dari sistem tetap harus dianggap sebagai pendukung, bukan alat diagnosis utama.

5. Bekerja Sama dengan Dokter Spesialis Jantung

Untuk meningkatkan validitas klinis, pengembangan lanjutan sebaiknya dilakukan dengan keterlibatan dokter spesialis jantung. Masukan dari ahli dapat membantu menyempurnakan logika klasifikasi, meningkatkan akurasi prediksi, dan membuat sistem lebih siap digunakan dalam konteks medis sebagai alat bantu pengambilan keputusan.

DAFTAR PUSTAKA

- An, Q., Rahman, S., Zhou, J., & Kang, J. J. (2023). A Comprehensive Review on Machine Learning in Healthcare Industry: Classification, Restrictions, Opportunities and Challenges. *Sensors*, 23(9).
<https://doi.org/10.3390/s23094178>
- Arumi, E. R., Sumarno Adi Subrata, & Anisa Rahmawati. (2023). Implementation of Naïve bayes Method for Predictor Prevalence Level for Malnutrition Toddlers in Magelang City. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 7(2), 201–207. <https://doi.org/10.29207/resti.v7i2.4438>
- Dona, D., Maradona, H., & Masdewi, M. (2021). Sistem Pakar Diagnosa Penyakit Jantung Dengan Metode Case Based Reasoning (Cbr). *ZONasi: Jurnal Sistem Informasi*, 3(1), 1–12. <https://doi.org/10.31849/zn.v3i1.6442>
- Fadrial, Y. E., Yunefri, Y., & Sutejo. (2021). CHATBOT PADA *SMART COOPERATIVE ORIENTED PROBLEM* MENGGUNAKAN *NATURAL LANGUAGE PROCESSING DAN NAIVE BAYES CLASSIFIER*. 4, 131–141.
- Ghani, L., & Novriani, H. (2016). Dominant Risk Factors for Coronary Heart Disease in Indonesia. *Buletin Penelitian Kesehatan*, 44(3), 153–164.
- Golpour, P., Ghayour-Mobarhan, M., Saki, A., Esmaily, H., Taghipour, A., Tajfard, M., Ghazizadeh, H., Moohebaty, M., & Ferns, G. A. (2020). Comparison of

support vector machine, naïve bayes and logistic regression for assessing the necessity for coronary angiography. *International Journal of Environmental Research and Public Health*, 17(18), 1–9.

<https://doi.org/10.3390/ijerph17186449>

Hardiani, T., & Putri, R. N. (2024). Implementasi Metode Naïve Bayes Classifier Untuk Klasifikasi Stunting Pada Balita. *Digital Transformation Technology*, 4(1), 621–627. <https://doi.org/10.47709/digitech.v4i1.4481>

Ismail, M., Hassan, N., & Bafjaish, S. S. (2020). Comparative Analysis of Naive Bayesian Techniques in Health-Related for Classification Task. *Journal of Soft Computing and Data Mining*, 1(2), 1–10.

<https://doi.org/10.30880/jscdm.2020.01.02.001>

Jackins, V., Vimal, S., Kaliappan, M., & Lee, M. Y. (2021). AI-based smart prediction of clinical disease using random forest classifier and Naive Bayes. *Journal of Supercomputing*, 77(5), 5198–5219.

<https://doi.org/10.1007/s11227-020-03481-x>

Muharni, S., & Andriyanto, S. (2021). Sistem Diagnosa Penyakit Jantung Berbasis Case Based Reasoning (CBR). *Prosiding Seminar Nasional Darmajaya*, 1, 1–11. <https://jurnal.darmajaya.ac.id/index.php/PSND/article/view/2910>

Mustafa, M. S., Ramadhan, M. R., & Thenata, A. P. (2018). Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier. *Creative Information Technology Journal*, 4(2), 151. <https://doi.org/10.24076/citec.2017v4i2.106>

Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier

- Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 5(2), 697–711.
- Pradana, D., Luthfi Alghifari, M., Farhan Juna, M., & Palaguna, D. (2022). Klasifikasi Penyakit Jantung Menggunakan Metode Artificial Neural Network. *Indonesian Journal of Data and Science*, 3(2), 55–60.
<https://doi.org/10.56705/ijodas.v3i2.35>
- Rayuwati, Husna Gemasih, & Irma Nizar. (2022). IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK MEMPREDIKSI TINGKAT PENYEBARAN COVID. *Jural Riset Rumpun Ilmu Teknik*, 1(1), 38–46.
<https://doi.org/10.55606/jurritek.v1i1.127>
- Retnasari, T., & Rahmawati, E. (2017). Diagnosa Prediksi Penyakit Jantung Dengan Model Algoritma Naïve Bayes Dan Algoritma C4.5. *Konferensi Nasional Ilmu Sosial & Teknologi (KNiST)*, 7-12Retnasari, T., Rahmawati, E. (2017). Diagnos.
- Sahar, S. (2020). Analisis Perbandingan Metode K-Nearest Neighbor dan Naïve Bayes Clasiffier Pada Dataset Penyakit Jantung. *Indonesian Journal of Data and Science*, 1(3), 79–86. <https://doi.org/10.33096/ijodas.v1i3.20>
- Sepharni, A., Hendrawan, I. E., & Rozikin, C. (2022). Klasifikasi Penyakit Jantung dengan Menggunakan Algoritma C4.5. *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, 7(2), 117.
<https://doi.org/10.30998/string.v7i2.12012>
- Sulaksono, J., & Darsono. (2021). Sistem Pakar Penentuan Penyakit Gagal Jantung Menggunakan Metode Naive Bayes Classifier. *Seminar nasional*

teknologi informasi dan multimedia 2015, 19–24.

- Surbakti, N. M., Angelyca Angelyca, Anita Talia, Cecilia Br Perangin-Angin, Dina Olivia Nainggolan, Nia Devi Friskaully, & Sikap Ruth Br Tumorang. (2024). Penggunaan Bahasa Pemrograman Python dalam Pembelajaran Kalkulus Fungsi Dua Variabel. *Algoritma : Jurnal Matematika, Ilmu pengetahuan Alam, Kebumihan dan Angkasa*, 2(3), 98–107.
<https://doi.org/10.62383/algoritma.v2i3.67>
- Suryadewiansyah, M. K., Endra, T., & Tju, E. (2020). Naïve Bayes dan Confusion Matrix untuk Efisiensi Analisa Intrusion Detection System Alert. *Jurnal Nasional Teknologi dan Sistem Informasi*, 8(2), 81–88.
- Tasia, E., Zaid, R., Zarier, I., Kenia, S., & Loka, P. (2023). Klasifikasi Penyakit Gagal Jantung Menggunakan Supervised Learning. *Sentimas*, 1–7.
<https://journal.irpi.or.id/index.php/sentimas/article/view/535>
- Titimeidara, M. Y., & Hadikurniawati, W. (2021). Implementasi Metode Naïve Bayes Classifier Untuk Klasifikasi Status Gizi Stunting Pada Balita. *Jurnal Ilmiah Informatika*, 9(01), 54–59. <https://doi.org/10.33884/jif.v9i01.3741>
- Wibisono, A. B., & Fahrurrozi, A. (2021). Perbandingan Algoritma Klasifikasi Dalam Pengklasifikasian Data Penyakit Jantung Koroner. *Jurnal Ilmiah Teknologi dan Rekayasa*, 24(3), 161–170.
<https://doi.org/10.35760/tr.2019.v24i3.2393>