

SKRIPSI

**PREDIKSI PENYAKIT JANTUNG KORONER
MENGUNAKAN ALGORITMA KLASIFIKASI
*NAIVE BAYES***



Disusun oleh :

Azaroby Dwi Anggoro 1519620011

PROGRAM STUDI

SISTEM DAN TEKNOLOGI INFORMASI

FAKULTAS TEKNIK

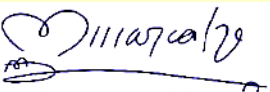

UNIVERSITAS NEGERI JAKARTA

2025

LEMBAR PENGESAHAN

Judul : Prediksi Penyakit Jantung Koroner Menggunakan Algoritma
Klasifikasi *Naive Bayes*
Penyusun : Azaroby Dwi Anggoro
NIM : 1519620011

Disetujui oleh:

NAMA DOSEN	TANDA TANGAN	TANGGAL
Murien Nugraheni, S.T., M.Cs (Dosen Pembimbing 1)		7 Februari 2025
Fuad Mumtas, M.T.I. (Dosen Pembimbing 2)		10 Januari 2025

PENGESAHAN PANITIA UJIAN SKRIPSI

NAMA DOSEN	TANDA TANGAN	TANGGAL
Lipur Sugiyanta, Ph.D (Ketua Penguji)		7/2-2025
Irma Permata Sari, M.Eng (Dosen Penguji 1)		30 Januari 2025

ABSTRAK

Azaroby Dwi Anggoro, Prediksi Penyakit Jantung Koroner Menggunakan Algoritma Klasifikasi *Naive Bayes*. Skripsi. Program Studi Sistem dan Teknologi Informasi, Fakultas Teknik Universitas Negeri Jakarta 2025. Dosen Pembimbing: Murien Nugraheni, S.T., M.Cs. dan Fuad Mumtas, M.T.I.

Penyakit Jantung Koroner (PJK) merupakan salah satu penyebab utama kematian di dunia, termasuk di Indonesia. Deteksi dini dan akurasi prediksi penyakit ini menjadi faktor penting dalam upaya pencegahan dan pengobatan. Penelitian ini bertujuan untuk membangun model prediksi penyakit jantung koroner menggunakan algoritma klasifikasi Naive Bayes. Dataset yang digunakan dalam penelitian ini diperoleh dari sumber terbuka dan telah melalui tahapan preprocessing data seperti normalisasi, penanganan outlier, serta pembagian data untuk pelatihan dan pengujian. Hasil penelitian menunjukkan bahwa model Naive Bayes dapat memprediksi risiko PJK dengan tingkat akurasi sebesar 89,06%. Evaluasi model menggunakan metrik precision dan recall mengindikasikan bahwa metode ini cukup andal dalam mengidentifikasi pasien dengan risiko PJK. Selain itu, parameter *var_smoothing* optimal ditemukan pada nilai $1e-09$ yang membantu meningkatkan performa model dengan mengurangi overfitting. Dari hasil penelitian ini, dapat disimpulkan bahwa algoritma Naive Bayes merupakan salah satu metode yang efektif dalam prediksi penyakit jantung koroner. Untuk pengembangan lebih lanjut, disarankan agar model diuji dengan dataset yang lebih beragam serta mengombinasikan algoritma lain guna meningkatkan akurasi dan generalisasi model dalam dunia medis.

Kata kunci: Penyakit Jantung Koroner, Prediksi, Naive Bayes, Machine Learning, Data Mining.

ABSTRACT

Azaroby Dwi Anggoro, *Coronary Heart Disease Prediction Using the Naive Bayes Classification Algorithm. Thesis. Information Systems and Technology Study Program, Faculty of Engineering, Jakarta State University 2025. Supervisor: Murien Nugraheni, S.T., M.Cs. and Fuad Mumtas, M.T.I.*

Coronary Heart Disease (CHD) is one of the main causes of death in the world, including in Indonesia. Early detection and accurate prediction of this disease are important factors in prevention and treatment efforts. This research aims to build a prediction model for coronary heart disease using the Naive Bayes classification algorithm. The dataset used in this research was obtained from open sources and has gone through data preprocessing stages such as normalization, handling outliers, and sharing data for training and testing. The research results show that the Naive Bayes model can predict CHD risk with an accuracy rate of 89.06%. Evaluation of the model using precision and recall metrics indicates that this method is quite reliable in identifying patients at risk of CHD. Additionally, the optimal var_smoothing parameter is found at the value $1e-09$ which helps improve the model performance by reducing overfitting. From the results of this research, it can be concluded that the Naive Bayes algorithm is an effective method for predicting coronary heart disease. For further development, it is recommended that the model be tested with a more diverse dataset and combine other algorithms to increase the accuracy and generalization of the model in the medical world.

Keywords: Coronary Heart Disease, Prediction, Naive Bayes, Machine Learning, Data Mining.

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa:

1. Skripsi ini yang merupakan Karya asli dan belum pernah diajukan untuk mendapatkan gelar akademik sarjana, baik di Universitas Negeri Jakarta maupun di Perguruan Tinggi lain.
2. Skripsi ini belum dipublikasikan, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka.
3. Pernyataan ini saya buat dengan sesungguhnya dan apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Universitas Negeri Jakarta.

Jakarta, 27 November 24

Yang Membuat Pernyataan



Azaroby Dwi Anggoro
NIM 1519620011



KEMENTERIAN PENDIDIKAN TINGGI, SAINS DAN TEKNOLOGI
UNIVERSITAS NEGERI JAKARTA
UPT PERPUSTAKAAN

Jalan Rawamangun Muka Jakarta 13220
Telepon/Faksimili: 021-4894221
Laman: lib.unj.ac.id

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademika Universitas Negeri Jakarta, yang bertanda tangan di bawah ini, saya:

Nama : Azaroby Dwi Anggoro
NIM : 1519620011
Fakultas/Prodi : Fakultas Teknik / Sistem dan Teknologi Informasi
Alamat email : azaroby.dwianggoro30@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada UPT Perpustakaan Universitas Negeri Jakarta, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah:

Skripsi Tesis Disertasi Lain-lain (... ..)

yang berjudul :

Prediksi Penyakit Jantung Koroner Menggunakan Algoritma Klasifikasi Naive Bayes

Dengan Hak Bebas Royalti Non-Eksklusif ini UPT Perpustakaan Universitas Negeri Jakarta berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya, dan menampilkan/mempublikasikannya di internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan Universitas Negeri Jakarta, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 11 Februari 2025

Penulis

(Azaroby Dwi Anggoro)

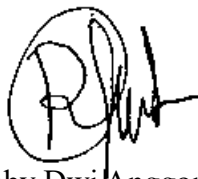
KATA PENGANTAR

Puji dan Syukur penulis ucapkan ke hadirat Allah SWT yang telah melimpahkan Rahmat serta karunia-Nya sehingga penulis dapat menyelesaikan Skripsi yang berjudul “ Prediksi Penyakit Jantung Koroner Menggunakan Algoritma Klasifikasi *Naive Bayes* “ Skripsi ini disusun sebagai salah satu persyaratan meraih gelar Sarjana Pendidikan pada Program Studi S1 Sistem dan Teknologi Informasi Universitas Negeri Jakarta. Penyelesaian skripsi ini tidak lepas dari bantuan berbagai pihak, oleh karena itu penulis menyampaikan ucapan terima kasih kepada :

1. Lipur Sugiyanta, Ph.D, selaku Koordinator Program Studi Sistem dan Teknologi Informasi, Fakultas Teknik, Universitas Negeri Jakarta,
2. Murien Nugraheni, S.T., M.Cs, selaku Dosen Pembimbing I Proposal Skripsi.
3. Fuad Mumtas, M.T.I., selaku Dosen Pembimbing II Proposal Skripsi.
4. Seluruh Dosen Sistem dan Teknologi Informasi Universitas Negeri Jakarta yang telah memberi bekal pengetahuan yang berharga.
5. Teman-teman seperjuangan yang setia saling membantu dalam proses Skripsi.

Penulis hanya dapat memanjatkan doa semoga semua pihak yang telah membantu penulis dalam penyusunan Skripsi ini mendapatkan pahala dari Allah SWT. Semoga Skripsi ini dapat bermanfaat Dan memberikan sumbangan yang berarti lagi pihak yang membutuhkan.

Jakarta, 27 November 2024



Azaroby Dwi Anggoro

DAFTAR ISI

PREDIKSI PENYAKIT JANTUNG KORONER MENGGUNAKAN ALGORITMA KLASIFIKASI <i>NAIVE BAYES</i>	I
LEMBAR PENGESAHAN	II
ABSTRAK.....	III
ABSTRACT	IV
LEMBAR PERNYATAAN.....	II
KATA PENGANTAR	III
DAFTAR ISI.....	II
DAFTAR GAMBAR	V
DAFTAR TABEL	VII
DAFTAR LAMPIRAN.....	VIII
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Identifikasi Masalah	4
1.3 Pembatasan Masalah	4
1.4 Rumusan Masalah	5
1.5 Tujuan Penelitian.....	5
1.6 Manfaat Penelitian	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Kerangka Teoritik.....	6
2.1.1 Penyakit Jantung Koroner.....	6
2.1.2 <i>Machine Learning</i>	10
2.1.3 Algoritma <i>Navie Bayes</i>	11
2.1.4 Metrik Evaluasi untuk Prediksi Penyakit Jantung Koroner.....	12
2.2 Penelitian relevan	14

2.3 Kerangka Berpikir	17
2.4 Hipotesis Penelitian.....	18
2.4.1 Hipotesis Nol (H0):.....	18
2.4.2 Hipotesis Alternatif (H1):	18
BAB III METODE PENELITIAN	19
3.1 Alat dan Bahan Tugas Akhir	19
3.1.1 Alat.....	19
3.1.2 Bahan	20
3.1.3 Data.....	21
3.2 Alur Penelitian.....	23
3.2.1 Flowchart	23
3.3 Penjabaran Langkah Penelitian	24
3.3.1 Pengumpulan Data	24
3.3.2 <i>Preprocessing</i> Data	28
3.3.3 Pembagian Data	30
3.3.4 Pembuatan Model	32
3.3.5 Evaluasi Model	34
BAB IV HASIL DAN PEMBAHASAN	36
4.1 HASIL PENELITIAN	36
4.1.1 Data <i>Preprocessing</i> Dan Cleaning	39
4.1.2 Hasil Pembagian Dataset	45
4.1.3 Skenario Penujian Naive Bayes	45
4.1.4 Hasil Pengujian Naive Bayes.....	50
4.1.5 Evaluasi Model menggunakan Metrik Klasifikasi.....	52
4.2 PEMBAHASAN	53

4.2.1 Mencakup Hasil Algoritma Naive Bayes dan Evaluasi.....	53
BAB V KESIMPULAN DAN SARAN	57
5.1 Kesimpulan	57
5.2 Saran.....	57
DAFTAR PUSTAKA.....	58
LAMPIRAN	61
DAFTAR RIWAYAT HIDUP	228



DAFTAR GAMBAR

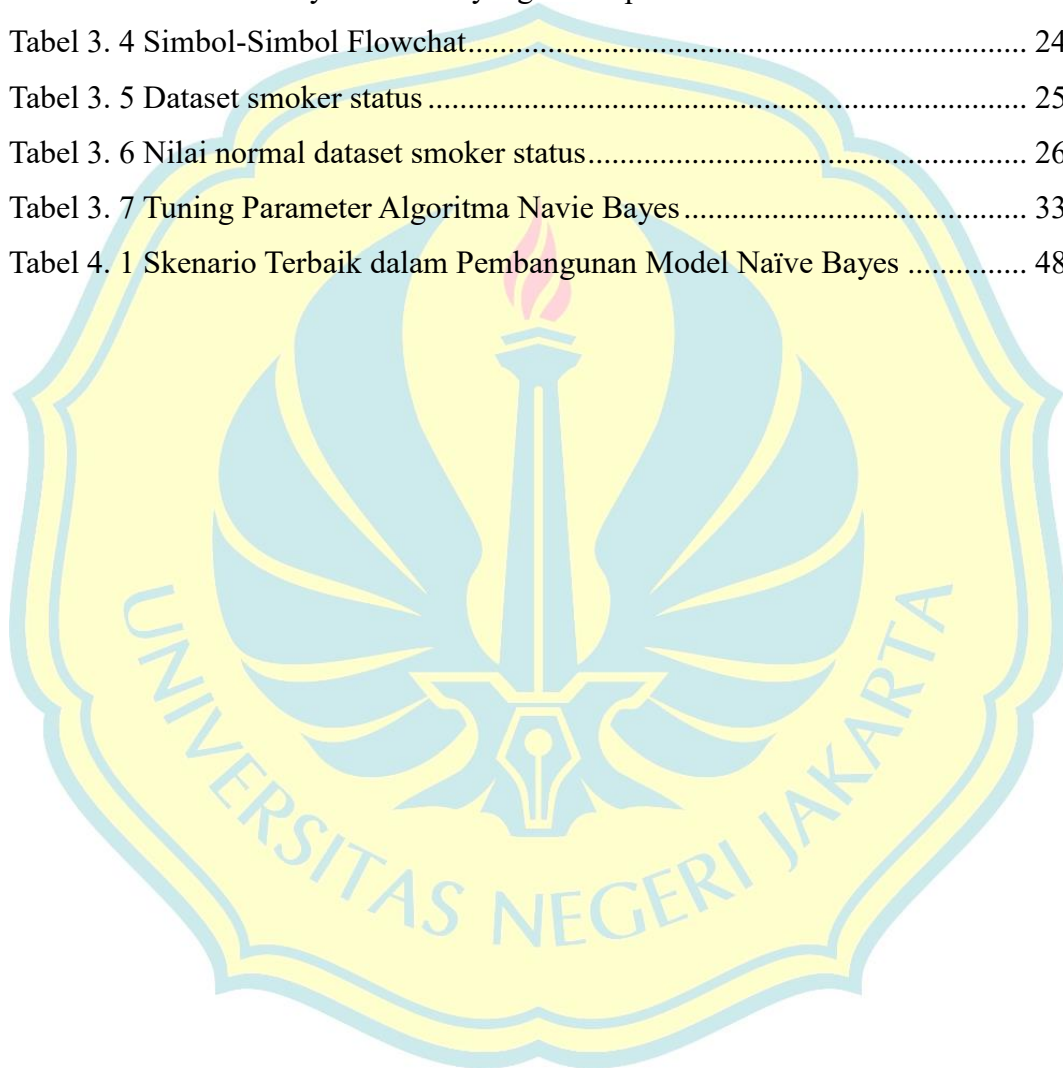
Gambar 2. 1 Kerangka Berpikir	18
Gambar 3. 1 Alur Penelitian.....	23
Gambar 4.1 Load Dataset.....	36
Gambar 4.2 Output Load Data	37
Gambar 4.3 Aturan Untuk Membuat Kolom Target.....	38
Gambar 4. 4 Menambah Kolam Target.....	38
Gambar 4.5 Menyimpan Dataset Baru	38
Gambar 4.6 Menghasilkan Target	39
Gambar 4.7 Output Penghapusan Data Duplikat	40
Gambar 4.8 Output Pembersihan Missing Value	41
Gambar 4.9 Cek Outlier	42
Gambar 4.10 Output Hasil Deteksi Outlier.....	42
Gambar 4.11 Memisahkan Fitur dan Target.....	43
Gambar 4.12 Inisialisasi MinMaxScaler.....	43
Gambar 4.13 Pencocokan Dan Transformasi.....	43
Gambar 4.14 Membuat DataFrame yang dinormalisasi	43
Gambar 4.15 Menggabungkan Target.....	44
Gambar 4.16 Output Normalisasi Data	44
Gambar 4.17 hasil pembagian dataset.....	45
Gambar 4.18 Inisiasi Naive Bayes	46
Gambar 4.19 Melatih Model.....	46
Gambar 4.20 Definisi parameter grid naïve bayes.....	46
Gambar 4.21 Pencarian Grid (Grid Search).....	47
Gambar 4. 22 Proses Pembangunan Model	47
Gambar 4. 23 Proses Pembangunan Model	47
Gambar 4.24 Mencari model terbaik naïve bayes.....	49
Gambar 4.25 Output hasil parameter terbaik naïve bayes	50
Gambar 4. 26 Proses Pengujian Model dengan Data Latih.....	50
Gambar 4. 27 Hasil Validasi Model dengan Data Latih.....	51
Gambar 4. 28 Proses Pengujian Model dengan Data Uji.....	51

Gambar 4. 29 Hasil Validasi Model dengan Data Uji	51
Gambar 4.30 Output Hasil Akurasi	52
Gambar 4.31 Output Laporan Klasifikasi	52
Gambar 4.32 Output Confusion Matrix	53
Gambar 4.33 Naïve Bayes Confusion Matrix	55



DAFTAR TABEL

Tabel 2. 1 Penelitian Yang Relevan.....	14
Tabel 3. 1 Alat Yang Digunakan Pada Penelitian.....	19
Tabel 3. 2 Bahan Yang Digunakan Pada Penelitian	20
Tabel 3. 3 Artibut Riwayat Perokok yang Description.....	22
Tabel 3. 4 Simbol-Simbol Flowchat.....	24
Tabel 3. 5 Dataset smoker status	25
Tabel 3. 6 Nilai normal dataset smoker status.....	26
Tabel 3. 7 Tuning Parameter Algoritma Navie Bayes.....	33
Tabel 4. 1 Skenario Terbaik dalam Pembangunan Model Naïve Bayes	48



DAFTAR LAMPIRAN

Lampiran 1. Surat Tugas Dosen Pembimbing.....	61
Lampiran 2. Lembar Konsultasi Dosen Pembimbing	62
Lampiran 3. Surat Pernyataan Dosen Pembimbing 1	65
Lampiran 4. Surat Pernyataan Dosen Pembimbing 2	66
Lampiran 5. Lembar persetujuan dosen pembimbing.....	67
Lampiran 6 Hasil Penerapan Model KNN dengan GridSearchCV	68
Lampiran 7. Hasil Validasi Data Prediksi dengan Data Latih.....	69
Lampiran 8. Kesalahan Prediksi (False Positive) pada Proses Validasi dengan Data Latih	74
Lampiran 9. Hasil Validasi Data Prediksi dengan Data Uji	85
Lampiran 10. Kesalahan Prediksi (False Positive) pada Proses Validasi dengan Data Uji.....	186
Lampiran 11. Dataset Awal Penelitian	194

