

**PREDIKSI RESIKO KOLESTEROL TINGGI
BERDASARKAN DATA KESEHATAN MENGGUNAKAN
IMPROVED NAÏVE BAYES**

Skripsi



Disusun Oleh :

Ardian Majid

200602002

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH GRESIK 2024**

KATA PENGANTAR

Alhamdulillah, puji syukur kehadirat Allah SWT yang telah melimpahkan rahmat, taufik serta hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul "Prediksi Resiko Kolesterol Tinggi Berdasarkan Data Kesehatan Menggunakan *Improved Naïve Bayes*".

Laporan skripsi ini digunakan sebagai persyaratan menjadi Sarjana Komputer pada Program Studi Teknik Informatika Universitas Muhammadiyah Gresik.

Terselesainya skripsi ini dengan baik berkat dukungan, motivasi, petunjuk, bimbingan dan do'a dari berbagai pihak. Oleh karena itu, ucapan terima kasih yang sebesar-besarnya penulis sampaikan kepada:

1. Allah SWT atas segala nikmat, ridho dan kuasanya.
2. Keluarga terutama kedua orang tua yang selalu memberikan semangat, do'a serta dukungan penuh.
3. Ibu Henny Dwi Bhakti, S.Si, M.Si dan Ibu Farhanna Mar'i, S.Kom, M.Kom. selaku dosen pembimbing yang senantiasa meluangkan waktunya untuk memberikan bimbingan, arahan serta masukan bagi penulis dalam penyelesaian skripsi ini.
4. Teman seperjuangan, Ubed, Alif, Dani, Rizal, Taqul, Bangkit, Oki dan Ilham yang tak pernah bosan untuk saling memberi semangat.
5. Game mobile legend yang selalu menemani saat penulis bosan.
6. Film anime Naruto, Jojutsu Kaisen, One Piece, dan Attack On Titan yang selalu memberikan motivasi yang membuat penulis terus berjuang.
7. Seluruh mahasiswa teknik informatika, khususnya angkatan 2020.

Permohonan maaf yang sebesar-besarnya jika dalam pengembangan program serta penulisan skripsi ini terdapat kekurangan dan keterlambatan baik yang penulis sengaja maupun tidak penulis sengaja. Sehingga saran dan kritik yang sifatnya membangun dari semua pihak berguna bagi penulis untuk memperbaiki skripsi ini. Sejatinya tidak ada yang sempurna di muka Bumi ini,

sehingga dengan belajarnya melalui kesalahan-kesalahan yang didahulu semoga menjadikannya esok yang lebih baik lagi. Aamiin Yarabbalamin.

Gresik, 15 Januari 2024

Penulis

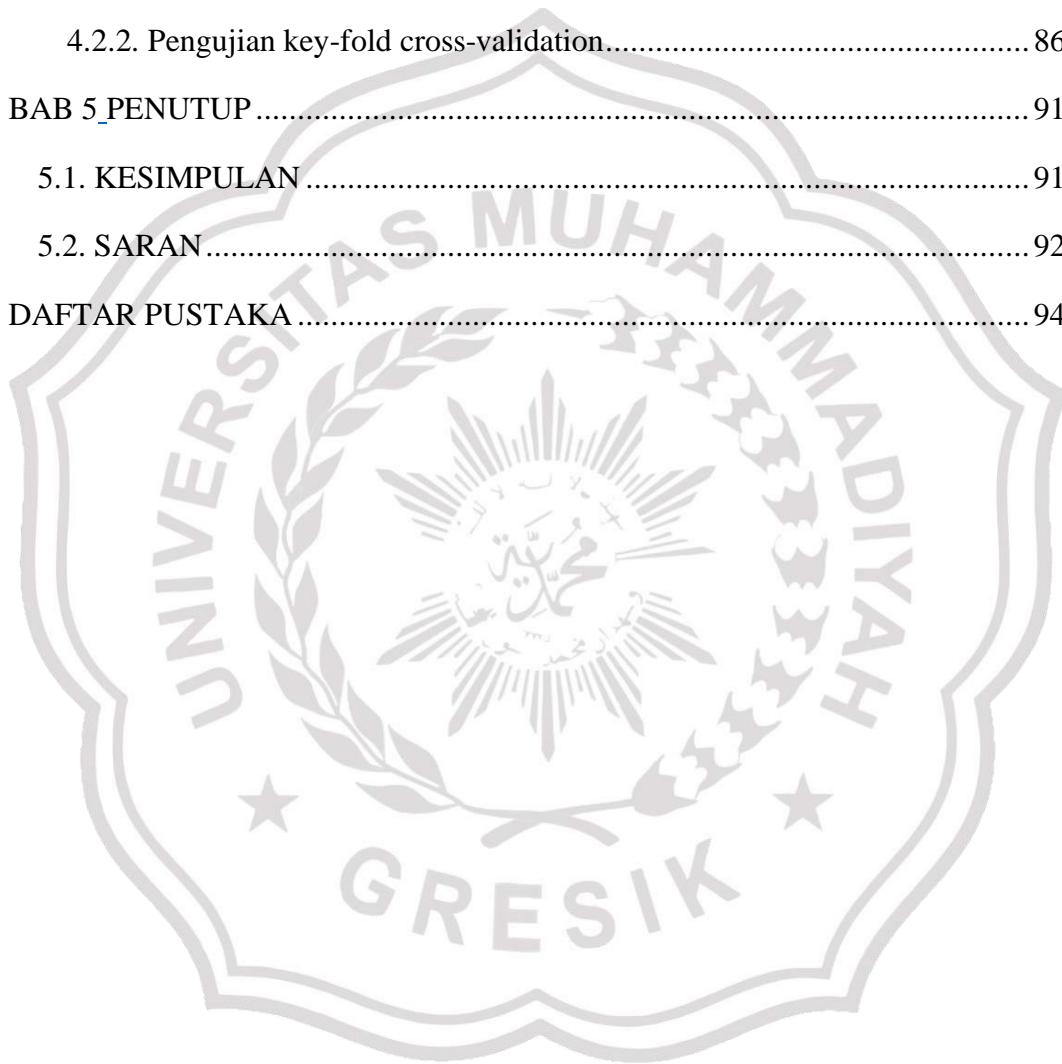


DAFTAR ISI

KATA PENGANTAR	i
LEMBAR PERSETUJUAN SKRIPSI	iii
LEMBAR PENGESAHAN SKRIPSI	iv
LEMBAR PERNYATAAN KEASLIAN SKRIPSI.....	v
DAFTAR ISI.....	vi
DAFTAR GAMBAR	ix
DAFTAR TABEL	x
ABSTRAK	xi
ABSTRACK.....	xiii
BAB 1 PENDAHULUAN	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	6
1.3. Tujuan Penelitian.....	7
1.4. Manfaat Penelitian.....	7
1.6. Metodologi Penelitian.....	8
1.7. Sistematika Penulisan	10
BAB 2 LANDASAN TEORI.....	11
2.1 Kolesterol Tinggi.....	11
2.2 <i>Machine Learning</i>	13
2.3 Data Mining.....	13
2.4 Prediksi.....	14
2.5 <i>Naïve Bayes</i>	15
2.6 <i>Improved Naïve Bayes</i>	16

2.7	Laplacian Correction.....	17
2.8	Penelitian Terkait.....	19
BAB 3 ANALISIS DAN PERANCANGAN SISTEM		32
3.1.	Analisis Sistem	32
3.2	Hasil Analisis	35
3.3	Perancangan Sistem	36
3.3.1	Sumber Data.....	38
3.3.2	Diagram Konteks.....	40
3.3.3	Diagram Berjenjang.....	40
3.3.4	Data Flow Diagram Level 1	42
3.3.5	Data Flow Diagram Level 2	43
3.4	Representasi Model.....	44
3.4.1	<i>Laplacian Correction</i>	55
3.5	Struktur Tabel.....	64
3.6	Desain Antarmuka	66
3.7	Spesifikasi Kebutuhan Sistem	69
3.7.1	Kebutuhan Perangkat Keras	69
3.7.2	Kebutuhan perangkat Lunak	70
3.8	Skenario Pengujian	70
3.8.1	K-fold Cross-Validation.....	70
3.8.2	Pengujian Sistem	71
BAB 4 IMPLEMENTASI DAN PENGUJIAN SISTEM.....		74
4.1.	Implementasi	74
4.1.1.	Halaman Perkenalan.....	74
4.1.2.	Halaman <i>Home</i>	75

4.1.3. Halaman <i>About</i>	81
4.1.4. Halaman <i>Login Admin</i>	82
4.1.3. Halaman Admin.....	83
4.2. Hasil Pengujian Sistem.....	84
4.2.1. Pengujian <i>Black Box</i>	84
4.2.2. Pengujian key-fold cross-validation.....	86
BAB 5_PENUTUP.....	91
5.1. KESIMPULAN.....	91
5.2. SARAN.....	92
DAFTAR PUSTAKA.....	94



DAFTAR GAMBAR

Gambar 3.1 Flowchart Evaluasi Kinerja Metode <i>Naive Bayes</i> dan <i>Improved Naive Bayes</i>	33
Gambar 3.2 <i>Block Diagram</i> sistem prediksi resiko kolesterol tinggi menggunakan <i>Improved Naive Bayes</i>	37
Gambar 3.3 Diagram Konteks Sistem Prediksi Resiko Kolesterol Tinggi.....	40
Gambar 3.4 Diagram Berjenjang pada sistem prediksi kolesterol tinggi.....	41
Gambar 3.5 DFD level 1.....	42
Gambar 3.6 DFD Level 2	43
Gambar 3.7 halaman perkenalan.....	66
Gambar 3.8 halaman <i>home</i>	67
Gambar 3.9 halaman <i>About</i>	68
Gambar 3.10 halaman <i>Login Admin</i>	68
Gambar 3.11 halaman <i>Admin</i>	69
Gambar 4.1 Halaman <i>Landing</i>	75
Gambar 4.2 Kode Program Perhitungan BMI.....	78
Gambar 4.3 Halaman Pertanyaan <i>slide</i> pertama	78
Gambar 4.4 Halaman Pertanyaan <i>slide</i> kedua	78
Gambar 4.5 Halaman Pertanyaan <i>slide</i> terakhir.....	79
Gambar 4.6 Halaman hasil perhitungan	80
Gambar 4.7 Halaman <i>About</i>	81
Gambar 4.8 Halaman <i>Login Admin</i>	82
Gambar 4.9 Halaman <i>Admin</i>	83
Gambar 4.10 Hasil akurasi <i>5-fold cross-validation</i>	88

DAFTAR TABEL

Tabel 2.1 Daftar Penelitian Terkait.....	19
Tabel 3.1 Data Atribut.....	39
Tabel 3.2 Data Latih.....	45
Tabel 3.3 Data Uji.....	46
Tabel 3.4 Hasil prediksi perhitungan <i>Naïve Bayes</i> Biasa.....	54
Tabel 3.5 Hasil prediksi perhitungan <i>Naïve Bayes</i> setelah ditingkatkan	63
Tabel 3.6 Perbandingan hasil antara <i>Naive Bayes</i> Biasa dengan <i>Improved Naive Bayes</i>	63
Tabel 3.7 admin.....	64
Tabel 3.8 data_pengguna.....	64
Tabel 3.9 data_training	65
Tabel 3.10 Data Atribut Akurasi.....	71
Tabel 3.11 Tabel <i>Blackbox</i>	72
Tabel 4.1 Deskripsi dan Pilihan Jawaban untuk Variabel Data	84
Tabel 4.2 Tabel Hasil <i>Blackbox</i>	84
Tabel 4.3 data dari <i>confusion matrix</i>	83
Tabel 4.4 Hasil <i>5-fold cross-validation</i>	85

**PREDIKSI RESIKO KOLESTEROL TINGGI
BERDASARKAN DATA KESEHATAN MENGGUNAKAN
IMPROVE NAÏVE BAYES**

Oleh

ARDIAN MAJID

200602002

Diajukan kepada Program Studi Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Gresik pada tanggal 15 Januari 2024 untuk memenuhi sebagian persyaratan memperoleh derajat sarjana S-1 Program Studi Teknik Informatika

ABSTRAK

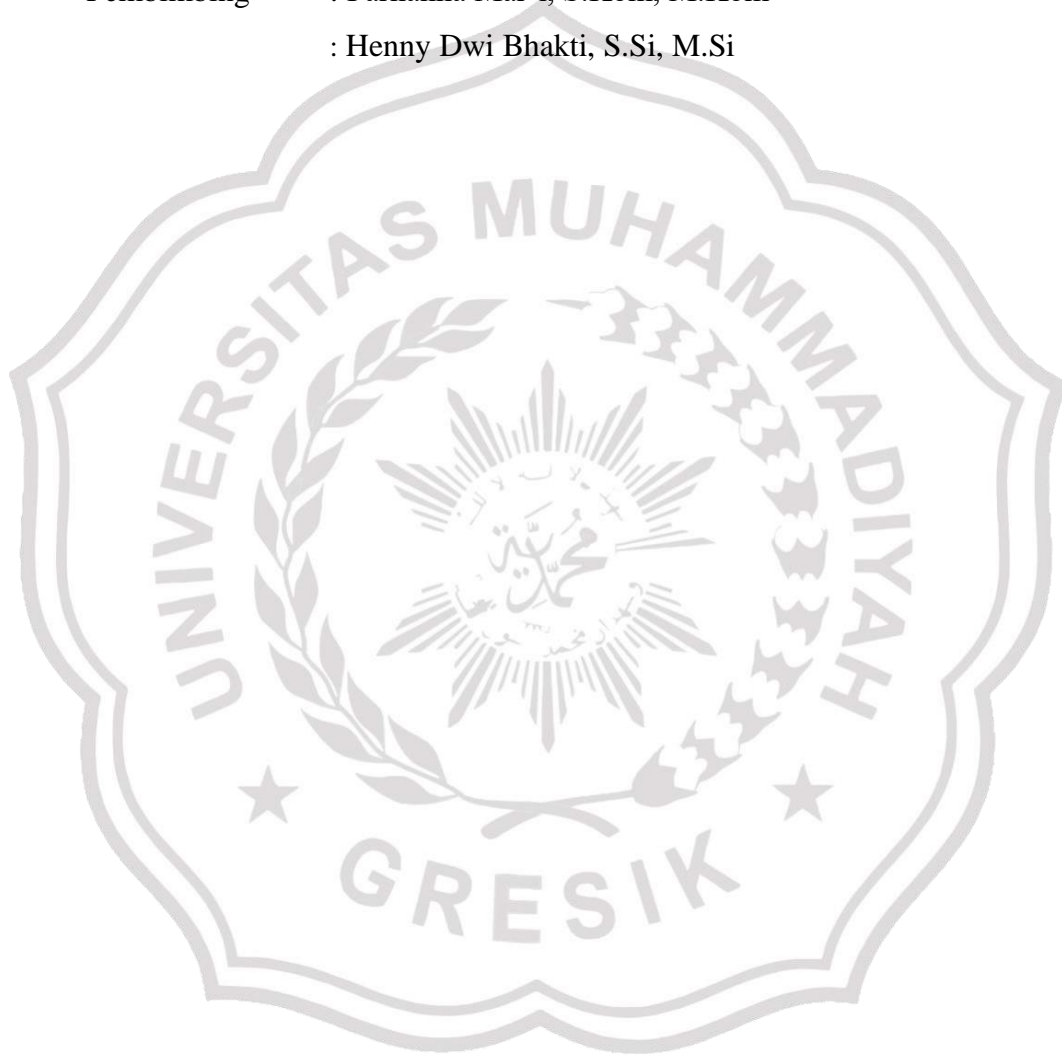
Kolesterol tinggi, yang secara medis dikenal sebagai hiperkolesterolemia, adalah salah satu faktor risiko utama untuk penyakit *kardiovaskular*, yang termasuk penyakit jantung dan stroke, yang merupakan penyebab utama kematian di seluruh dunia. Tingkat kolestrol yang tinggi dalam darah dapat mengakibatkan penumpukan plak pada dinding arteri, menyempitkannya dan meningkatkan risiko terjadinya penyumbatan pembuluh darah yang vital. Oleh karena itu, pengelolaan risiko kolestrol tinggi menjadi krusial dalam upaya pencegahan penyakit *kardiovaskular*. Dalam konteks ini, penelitian ini menggunakan metode *Naive Bayes* untuk memprediksi risiko kolestrol tinggi berdasarkan data kesehatan yang diperoleh dari Kaggle, sebuah platform sumber data terbuka. Hasil dari penelitian ini adalah model prediksi yang dapat mengestimasi risiko kolestrol tinggi pada individu berdasarkan karakteristik kesehatan mereka, memberikan kontribusi penting dalam upaya pencegahan penyakit *kardiovaskular* dengan memungkinkan prediksi risiko kolestrol tinggi berdasarkan data kesehatan individu. Selain itu, penelitian ini juga mengilustrasikan pentingnya pemanfaatan sumber data terbuka seperti

Kaggle dalam riset kesehatan untuk memahami dan mengatasi masalah kesehatan masyarakat.

Kata Kunci Kolesterol Tinggi, Penyakit *Kardiovaskular*, Prediksi Risiko, *Naive Bayes*, Data Kesehatan, Kaggle, Pencegahan, Riset Kesehatan.

Pembimbing : Farhanna Mar'i, S.Kom, M.Kom

: Henny Dwi Bhakti, S.Si, M.Si



**PREDICTION OF HIGH CHOLESTEROL RISK
BASED ON HEALTH DATA USING IMPROVED
NAÏVE BAYES**

BY

ARDIAN MAJID

200602002

Submitted to the Informatics Engineering Study Program, Faculty of Engineering, University of Muhammadiyah Gresik on January 15, 2024 to fulfill some of the requirements for obtaining a bachelor's degree in the Informatics Engineering Study Program.

ABSTRACT

High cholesterol, medically known as hypercholesterolemia, is a significant risk factor for cardiovascular diseases, including heart disease and stroke, which are leading causes of death worldwide. Elevated levels of cholesterol in the blood can result in the accumulation of plaque on the artery walls, narrowing them and increasing the risk of vital blood vessel blockages. Therefore, the management of high cholesterol risk becomes crucial in cardiovascular disease prevention efforts. In this context, this research employs the Naïve Bayes method to predict the risk of high cholesterol based on health data obtained from Kaggle, an open data platform. The outcome of this study is a prediction model capable of estimating the risk of high cholesterol in individuals based on their health characteristics, offering a significant contribution to cardiovascular disease prevention by enabling the prediction of high cholesterol risk based on individual health data. Additionally, this research underscores the importance of utilizing open data sources like Kaggle in health research to understand and address public health issues.

Keywords High Cholesterol, Cardiovascular Disease, Risk Prediction, Naïve Bayes, Health Data, Kaggle, Prevention, Health Research

Leader : Farhanna Mar'i, S.Kom, M.Kom

: Henny Dwi Bhakti, S.Si, M.Si

