

**ANALISIS IMPLEMENTASI METODE NAIVE BAYES, K-
NEAREST NEIGHBOR, DAN *LOGISTIC REGRESSION*
UNTUK MENGLASIFIKASI PENYAKIT JANTUNG**

SKRIPSI



Oleh :

AUDHY BRILLIANT PRATAMA

19081010123

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
2023**

**ANALISIS IMPLEMENTASI METODE NAIVE BAYES, K-
NEAREST NEIGHBOR, DAN *LOGISTIC REGRESSION*
UNTUK MENGLASIFIKASI PENYAKIT JANTUNG**

SKRIPSI

Diajukan Untuk Memenuhi Sebagai Persyaratan Dalam Menempuh Gelar Sarjana
Komputer Program Studi Informatika



Oleh :

AUDHY BRILLIANT PRATAMA

19081010123

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
2023**

LEMBAR PENGESAHAN SKRIPSI

Judul : ANALISIS IMPLEMENTASI METODE NAIVE BAYES,
K-NEAREST NEIGHBOR, DAN LOGISTIC REGRESSION
UNTUK MENGGKLASIFIKASI PENYAKIT JANTUNG
Oleh : AUDHY BRILLIANT PRATAMA
NPM : 19081010123

Telah Diseminarkan Dalam Ujian Skripsi Pada :
Hari Rabu, Tanggal 24 Mei 2023

Mengetahui

Dosen Pembimbing

Dosen Penguji

1.

Dr. Basuki Rahmat, S.Si., MT.

NIP : 19690723 2021211 002

1.

Dr. Rr. Ani Dajah Rahajoe, ST., M.Cs.

NIP : 19730512 200501 2003

2.

Eva Yuka Puspawati, S.Kom., M.Kom.

NIP : 19890705 2021212 002

2.

Afina Lina Nurlaili, S.Kom., M.Kom.

NIP : 1993121 3202203 2010

Menyetujui

Dekan Fakultas Ilmu Komputer

Koordinator Program Studi
Teknik Informatika

Dr. Ir. Novirina Hendrasarie, MT.

NPT : 19681126 199403 2 001

Fetty Tri Anggraeny, S.Kom., M.Kom.

NPT : 19820211 2021212 005

SURAT PERNYATAAN ANTI PLAGIAT

Saya mahasiswa program studi Informatika UPN “Veteran” Jawa Timur,
yang bertandatangan berikut ini:

Nama : Audhy Brilliant Pratama

NPM : 19081010123

Menyatakan bahwa Judul Skripsi/Tugas Akhir yang Saya ajukan dan akan
dikerjakan, yang berjudul:

**“ANALISIS IMPLEMENTASI METODE NAIVE BAYES, K-
NEAREST NEIGHBOR, DAN LOGISTIC REGRESSION
UNTUK MENGLASIFIKASI PENYAKIT JANTUNG”**

Bukan merupakan plagiat dari Skripsi/Tugas Akhir/Penelitian orang lain
juga bukan merupakan produk dan atau perangkat lunak yang beli dari pihak lain.
Saya juga menyatakan bahwa Skripsi/Tugas Akhir ini adalah pekerjaan Saya
sendiri, kecuali yang dinyatakan dalam Daftar Pustaka dan tidak pernah diajukan
untuk syarat memperoleh gelar di UPN “Veteran” Jawa Timur maupun di institusi
Pendidikan lain.

Jika ternyata di kemudian hari pernyataan ini terbukti tidak benar, maka
Saya siap menerima segala konsekuensinya.

Surabaya, 9 Mei 2023

Hormat Saya,



AUDHY BRILLIANT PRATAMA

NPM. 19081010123

ANALISIS IMPLEMENTASI METODE NAIVE BAYES, K-NEAREST NEIGHBOR, DAN LOGISTIC REGRESSION UNTUK MENGKLASIFIKASI PENYAKIT JANTUNG

Nama Mahasiswa : AUDHY BRILLIANT PRATAMA
NPM : 19081010123
Program Studi : Teknik Informatika
Dosen Pembimbing : Dr. Basuki Rahmat, S.Si, MT
Eva Yulia Puspaningrum, S.Kom, M.Kom

Abstrak

Organ jantung sangat penting bagi kelangsungan hidup semua makhluk hidup karena berperan dalam memompa darah, nutrisi, dan oksigen ke seluruh organisme. Apabila pasokan darah yang disalurkan oleh jantung terganggu akibat adanya sumbatan pada pembuluh darah utama atau kegagalan pada jantung dalam memompa darah ke seluruh tubuh, kondisi ini disebut sebagai penyakit jantung koroner yang bisa menimbulkan masalah kesehatan yang serius bahkan bisa berujung pada kematian. Penyakit jantung koroner tetap menjadi salah satu penyebab utama kematian di Indonesia. Gejala-gejala yang menunjukkan faktor risiko sangat membantu dalam mendiagnosis apakah seseorang mengalami penyakit jantung atau tidak. Selain itu, keterbatasan informasi dan media serta keterlambatan pemeriksaan awal untuk penyakit jantung juga menjadi masalah.

Metode yang akan digunakan pada penelitian ini adalah Naïve Bayes, KNN, Regresi Logistik, karena ketiga metode ini memiliki kecocokan untuk mengkaji permasalahan yang ada. Hal ini diperkuat dengan adanya penelitian terdahulu yang memiliki similaritas terkait permasalahan yang terjadi dengan metode yang digunakan, sehingga metode ini sudah ter-uji dan sesuai dengan penelitian ini.

Pada skenario pengujian yang dilakukan terdapat beberapa perubahan yang dilakukan pada arsitektur sistem seperti perubahan rasio antara penggunaan data training dan data testing pada semua metode masing-masing berjumlah 70:30, 80:20, dan 90:10. Lalu perubahan nilai K pada metode KNN mulai dari K=1 hingga K=20. Lalu pada hasil kesimpulan metode regresi logistik memiliki hasil performansi yang terbaik dengan nilai akurasi sebesar 86.159%.

Kata kunci: *Naïve Bayes, KNN, Regresi Logistik, Penyakit Jantung*

KATA PENGANTAR

Segala puji hanya milik Allah Subhanahu Wa Ta'ala, Ya Rabb semesta alam yang telah memberikan rahmat dan hidayah-Nya karena atas izin dan keridhoan-Nya penulis mampu untuk menyelesaikan penelitian skripsi beserta dengan laporan hasil yang berjudul “Analisis Implementasi Metode Naive Bayes, K-Nearest Neighbor, Dan Logistic Regression Untuk Mengklasifikasi Penyakit Jantung”. Sholawat serta salam senantiasa tercurahkan kepada Nabi Muhammad shallallahu ‘alaihi wa sallam yang telah meneladankan sikap ulet, ketelitian dan sabar sehingga penulis dapat menuntaskan skripsi ini dengan baik.

Banyak dukungan serta bantuan yang didapatkan selama melakukan penelitian hingga menyelesaikan laporan ini. Dengan rasa hormat serta banyak terima kasih penulis ucapkan kepada seluruh pihak yang telah membantu proses penyelesaian penilitan dari awal hingga akhir.

Penulis menyadari masih terdapat banyaknya kekurangan pada penelitian ini dikarenakan keterbatasan pengetahuan dan pemikiran. Oleh karena itu, penulis mengharapkan dan menerima kritik dan saran yang membangun dari semua pihak dalam penyempurnaan laporan ini. Semoga skripsi ini barokah, bermanfaat, dan berguna serta mampu memenuhi harapan berbagai pihak.

Surabaya, 9 Mei 2023

Penulis,

Audhy Brilliant Pratama

UCAPAN TERIMA KASIH

Penyelesaian skripsi ini melalui berbagai tahapan dan rintangan yang memerlukan pengorbanan yang sangat berat baik harta, jiwa, dan waktu. Banyak sekali hambatan yang dihadapi, namun berkat dukungan dan kehendak Allah subhanahu wa ta'ala lah sehingga penulis berhasil menyelesaikan penyusunan skripsi ini. Penulis menyadari bahwa penyelesaian skripsi ini tidak terlepas dari dukungan oleh semua pihak yang terlibat didalamnya. Oleh karena itu, pada kesempatan ini dengan segala kerendahan hati dan penuh rasa hormat penulis mengucapkan terima kasih yang tidak terhingga bagi semua pihak yang telah memberi bantuan baik moril maupun materil, baik secara langsung maupun tidak langsung dalam penyusunan skripsi ini hingga selesai, khususnya kepada:

1. Allah subhanahu wa ta'ala lah. Karena telah membawa penulis mulai dari awal penelitian hingga pada titik penyelesaian penelitian skripsi ini dengan lancar tanpa adanya hambatan, dimana hal ini sangat mustahil untuk dicapai tanpa kehendak-Nya.
2. Bapak Prof. Dr. Ir. Akhmad Fauzi, MMT.,IPU selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Ibu Dr. Novirina Hendrasarie, S.T., M.T., selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. Bapak Dr. Basuki Rahmat, S.Si, MT., selaku Pembimbing I dan sekaligus Wakil Dekan yang telah bersedia membimbing dan mengarahkan penulis selama menyusun skripsi dan memberikan banyak ilmu serta solusi pada setiap permasalahan atas kesulitan dalam penulisan skripsi ini.
5. Ibu Fetty Tri Anggraeny, S.Kom, M.Kom, selaku koordinator program studi jurusan Informatika Fakultas Ilmu Komputer UPN “Veteran” Jawa Timur.
6. Ibu Eva Yulia Puspaningrum, S.Kom., M.Kom. selaku Dosen Pembimbing II yang telah bersedia meluangkan waktu, tenaga dan pikiran untuk memberikan penulis masukan pada proses bimbingan

selama penyusunan skripsi ini

7. Seluruh Bapak dan Ibu dosen program studi Informatika UPN “Veteran” Jawa Timur yang telah mendidik dan memberikan ilmunya sehingga Penulis memiliki bekal untuk dapat melakukan penelitian ini.
8. Kedua Orang Tua, Kakak, dan Kerabat yang telah memberikan doa, kasih sayang, serta semangat pada saat melaksanakan penelitian dan juga dalam pengerjaan laporan, walaupun terkadang agak sering membuat penulis sedikit panik dan tertekan dalam mengerjakan penelitian skripsi ini.
9. Teman-teman Angkatan 2019 khususnya yang senantiasa menemani dan memberi motivasi dan bantuan ketika penulis sedang kesusahan dalam mengerjakan penelitian.
10. Para “NPC” dan member pada server “discord” yang selalu memberikan hiburan pada saat penulis merasa tertekan, dan memberi informasi baik seputar akademik maupun non-akademik sehingga penulis dapat mengetahui informasi yang beredar secara aktual.
11. Provider “Indihome” yang selalu menemani penulis dengan layanan yang diberikan untuk mengerjakan penelitian ini mulai dari proses awal pengerjaan hingga akhir penelitian dengan sepenuh hati walaupun terkadang agak sering mengalami gangguan jaringan.
12. Seseorang diluar sana yang entah darimana datangnya, yang masuk ke dunia penulis dan mengubah hidup si penulis?.
13. Semua pihak yang penulis tidak dapat sebutkan satu persatu atas segala bentuk kepedulian dan bantuan dalam pengerjaan penelitian ini.

DAFTAR ISI

SKRIPSI	1
SKRIPSI	2
LEMBAR PENGESAHAN SKRIPSI	2
SURAT PERNYATAAN ANTI PLAGIAT	3
KATA PENGANTAR	7
UCAPAN TERIMA KASIH	8
DAFTAR ISI.....	10
DAFTAR TABEL.....	14
DAFTAR GAMBAR	15
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Tujuan.....	4
1.4. Manfaat	4
1.4.1. Bagi Penulis.....	4
1.4.2. Bagi Pengguna.....	4
1.5. Batasan Masalah.....	5
BAB II TINJAUAN PUSTAKA	6
2.1. Penelitian Terdahulu.....	6
2.2. Penyakit Jantung	9
2.3. Pengertian <i>Data Mining</i>	11
2.4. Python.....	13
2.5. Jupyter Notebook	14

2.6. Pandas	14
2.7. Numpy	14
2.8. Matplotlib.....	14
2.9. Seaborn	15
2.10. Metode <i>Naive bayes</i>	15
2.11. Metode <i>K-Nearest Neighbor</i>	17
2.12. Metode <i>Logistic Regression</i>	18
2.13. <i>Confussion Matrix</i>	21
BAB III METODOLOGI.....	23
3.1. Gambaran Penelitian	23
3.2. Analisis Kebutuhan.....	23
3.2.1. Lingkungan Penelitian	23
3.2.2. Spesifikasi Perangkat Keras	24
3.2.3. Spesifikasi Perangkat Lunak	24
3.3. Tahapan Penelitian.....	25
3.4. Akuisisi Data	28
3.5. Pemrosesan Data.....	28
3.6. Pengujian Data.....	30
3.6.1. <i>Naïve Bayes</i>	31
3.6.2. KNN	35
3.6.3. <i>Logistic Regression</i>	39
3.7. Membuat Skema Pengujian Dengan <i>Confusion Matrix</i>	42
3.8. Skenario Uji Coba.....	42
BAB IV HASIL DAN PEMBAHASAN	44
4.1. Implementasi Sistem.....	44

4.1.1.	Proses Perhitungan <i>Naïve Bayes</i>	44
4.1.2.	Praproses Data <i>Naïve Bayes</i>	45
4.1.3.	Proses Perhitungan Data <i>Naïve Bayes</i>	47
4.1.4.	Penentuan diagnosis Penyakit Jantung <i>Naïve Bayes</i>	48
4.1.5.	Perhitungan Hasil Input User <i>Naïve Bayes</i>	51
4.1.6.	Program Utama <i>Naïve Bayes</i>	52
4.1.7.	User Input <i>Naïve Bayes</i>	54
4.1.8.	Proses Perhitungan KNN.....	57
4.1.9.	Praproses Data KNN	57
4.1.10.	Proses Perhitungan Data KNN	58
4.1.11.	Penentuan diagnosis Penyakit Jantung KNN	60
4.1.12.	Perhitungan Hasil Input User KNN	61
4.1.13.	Program Utama KNN	62
4.1.14.	User Input KNN.....	65
4.1.15.	Proses Perhitungan <i>Logistic Regression</i>	67
4.1.16.	Praproses Data <i>Logistic Regression</i>	68
4.1.17.	Proses Perhitungan Data <i>Logistic Regression</i>	69
4.1.18.	Penentuan diagnosis Dengan <i>Logistic Regression</i>	72
4.1.19.	Program Utama <i>Logistic Regression</i>	72
4.2.	Skenario Pengujian	74
4.2.1.	Perubahan rasio data <i>Naïve Bayes</i> 70:30	74
4.2.2.	Perubahan rasio data <i>Naïve Bayes</i> 80:20	76
4.2.3.	Perubahan rasio data <i>Naïve Bayes</i> 90:10	78
4.2.4.	Perubahan Nilai K Pada KNN Dengan Rasio 70:30	81
4.2.5.	Perubahan Nilai K Pada KNN Dengan Rasio 80:20	83

4.2.6. Perubahan Nilai K Pada KNN Dengan Rasio 90:10	84
4.2.7. Skenario Pengujian Pada Metode <i>Logistic Regression</i>	86
BAB V KESIMPULAN DAN SARAN	88
5.1 . Kesimpulan	88
5.2 . Saran.....	88
Daftar Pustaka	90
BIODATA PENULIS	93

DAFTAR TABEL

Tabel 2.1 Contoh tabel confusion matrix	21
Tabel 3.1 Spesifikasi Perangkat Keras.....	24
Tabel 3.2 Spesifikasi Perangkat Lunak.....	25
Tabel 3.3 Dataset Heart Disease	28
Tabel 3.4 Sample Data Uji	40
Tabel 4.1 Hasil Skenario Pengujian Pertama Naïve Bayes.....	76
Tabel 4.2 Hasil Skenario Pengujian Kedua Naïve Bayes	78
Tabel 4.3 Hasil Skenario Pengujian Ketiga Naïve Bayes	80
Tabel 4.4 Hasil Seluruh Skenario Pengujian Naïve Bayes	81
Tabel 4.5 Hasil Skenario Pengujian Pertama KNN.....	82
Tabel 4.6 Hasil Skenario Pengujian Kedua KNN	83
Tabel 4.7 Hasil Skenario Pengujian Ketiga KNN	84
Tabel 4.8 Hasil Seluruh Skenario Pengujian KNN	85
Tabel 4.9 Hasil Skenario Pengujian Logistic Regression	86
Tabel 4.10 Hasil Performansi Ketiga Metode	87

DAFTAR GAMBAR

Gambar 3.1 Kerangka Pemikiran	25
Gambar 3.2 Diagram Alur Sistem	27
Gambar 3.3 Diagram Alur Metode Naive Bayes	31
Gambar 3.4 Diagram Alur Metode KNN Dengan Euclidean Distance	35
Gambar 3.5 Diagram Alur Metode Logistic Regression	39
Gambar 4.1 Grafik Perbandingan Jumlah Cost dan Iterasi.....	74
Gambar 4.2 Hasil Perhitungan Prediksi pada Data testing Dengan rasio 70:30 ...	75
Gambar 4.3 Hasil Prediksi pada Data Testing Dengan Rasio 70:30	75
Gambar 4.4 Hasil Kinerja Metode Naïve Bayes Dengan Rasio 70:30	76
Gambar 4.5 Hasil Perhitungan Prediksi pada Data testing Dengan rasio 80:20 ...	77
Gambar 4.6 Hasil Prediksi pada Data Testing Dengan Rasio 80:20	77
Gambar 4.7 Hasil Kinerja Metode Naïve Bayes Dengan Rasio 80:20	78
Gambar 4.8 Hasil Perhitungan Prediksi pada Data testing Dengan rasio 90:10 ...	79
Gambar 4.9 Hasil Prediksi pada Data Testing Dengan Rasio 90:10	79
Gambar 4.10 Hasil Kinerja Metode Naïve Bayes Dengan Rasio 90:10	80