

**RANCANG BANGUN APLIKASI PREDIKSI PENYAKIT STROKE
MENGUNAKAN ALGORITMA GAUSSIAN NAÏVE BAYES**



Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

Michelle Shiela Haryanto

0000031135

UMN

UNIVERSITAS

MULTIMEDIA

NUSANTARA

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA**

TANGERANG

2024

**RANCANG BANGUN APLIKASI PREDIKSI PENYAKIT STROKE
MENGUNAKAN ALGORITMA GAUSSIAN NAÏVE BAYES**



Michelle Shiela Haryanto

0000031135

UMMN

**UNIVERSITAS
MULTIMEDIA
NUSANTARA**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA**

TANGERANG

2024

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Michelle Shiela Haryanto
NIM : 00000031135
Program Studi : Informatika

Menyatakan dengan sesungguhnya bahwa Tesis/Skripsi/Tugas Akhir/Laporan Magang/MBKM saya yang berjudul:

Rancang Bangun Aplikasi Prediksi Penyakit Stroke menggunakan Algoritma Gaussian Naïve Bayes

merupakan hasil karya saya sendiri, bukan merupakan hasil plagiat, dan tidak pula dituliskan oleh orang lain; Semua sumber, baik yang dikutip maupun dirujuk, telah saya cantumkan dan nyatakan dengan benar pada bagian Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan skripsi maupun dalam penulisan laporan karya ilmiah, saya bersedia menerima konsekuensi untuk dinyatakan TIDAK LULUS. Saya juga bersedia menanggung segala konsekuensi hukum yang berkaitan dengan tindak plagiarisme ini sebagai kesalahan saya pribadi dan bukan tanggung jawab Universitas Multimedia Nusantara.

Tangerang, 22 Mei 2024



(Michelle Shiela Haryanto)

HALAMAN PENGESAHAN

Skripsi dengan judul

RANCANG BANGUN APLIKASI PREDIKSI PENYAKIT STROKE MENGUNAKAN ALGORITMA GAUSSIAN NAÏVE BAYES

oleh

Nama : Michelle Shiela Haryanto
NIM : 00000031135
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika


Telah diujikan pada hari Senin, 03 Juni 2024

Pukul 09.00 s/s 10.30 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang



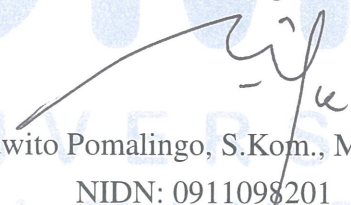
(Wirawan Istiono, S.Kom., M.Kom.)
NIDN: 0313048304

Penguji




(Adhi Kusnadi, S.T, M.Si.)
NIDN: 0303037304

Pembimbing



(Suwito Pomalingo, S.Kom., M.Kom.)
NIDN: 0911098201

Pjs. Ketua Program Studi Informatika,



(Dr. Eng. Niki Prastomo, ST., M.Sc.)
NIDN: 0419128203

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini:

Nama : Michelle Shiela Haryanto
NIM : 00000031135
Program Studi : Informatika
Jenjang : S1
Jenis Karya : Skripsi

Menyatakan dengan sesungguhnya bahwa:

- Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya di repositori Knowledge Center, sehingga dapat diakses oleh Civitas Akademika/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial dan saya juga tidak akan mencabut kembali izin yang telah saya berikan dengan alasan apapun.
- Saya tidak bersedia karena dalam proses pengajuan untuk diterbitkan ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*)**.

Tangerang, 22 Mei 2024

Yang menyatakan



Michelle Shiela Haryanto

** Jika tidak bisa membuktikan LoA jurnal/HKI selama enam bulan ke depan, saya bersedia mengizinkan penuh karya ilmiah saya untuk diunggah ke KC UMN dan menjadi hak institusi UMN.

Halaman Persembahan / Motto

"17. Janganlah hatimu iri kepada orang-orang yang berdosa, tetapi takutlah akan TUHAN senantiasa. 18. Karena masa depan sungguh ada, dan harapanmu tidak akan hilang."

Amsal 23:17-18 (TB)

UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan Skripsi ini dengan judul: Rancang Bangun Aplikasi Prediksi Penyakit Stroke menggunakan Algoritma Gaussian Naïve Bayes dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

Mengucapkan terima kasih

1. Bapak Dr. Ninok Leksono, selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, ST., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Dr. Eng. Niki Prastomo, ST., M.Sc., selaku Pjs. Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Bapak Suwito Pomalingo, S.Kom., M.Kom., sebagai Pembimbing pertama yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya tesis ini.
5. Orang tua, dan kedua adik saya yang telah memberikan motivasi dan dukungan, sehingga penulis dapat menyelesaikan tesis ini.

Semoga karya ilmiah ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 22 Mei 2024



Michelle Shiela Haryanto

RANCANG BANGUN APLIKASI PREDIKSI PENYAKIT STROKE MENGUNAKAN ALGORITMA GAUSSIAN NAÏVE BAYES

Michelle Shiela Haryanto

ABSTRAK

Stroke merupakan salah satu penyakit paling berbahaya yang menyebabkan kecacatan dan kematian di seluruh dunia. Penelitian menunjukkan peningkatan signifikan dalam insiden stroke, termasuk stroke iskemik dan hemoragik, dengan prevalensi tinggi terutama pada individu usia lanjut. Disparitas ekonomi mempengaruhi distribusi beban stroke, dengan negara berpenghasilan rendah melaporkan sebagian besar kasus kematian dan kecacatan. Di Indonesia, stroke adalah penyebab utama kematian, mempengaruhi pasien serta sistem kesehatan secara keseluruhan. Lima faktor risiko utama stroke adalah tekanan darah tinggi, indeks massa tubuh tinggi, glukosa darah tinggi, polusi udara, dan merokok. Deteksi dini stroke sangat penting untuk meningkatkan pemulihan dan mengurangi komplikasi. Berbagai metode deteksi telah dikembangkan, termasuk MRI dan CT, namun biayanya tinggi dan sifatnya invasif. Pembelajaran mesin menawarkan solusi lebih cepat dan murah, dengan algoritma Gaussian Naïve Bayes menunjukkan kinerja baik dalam prediksi stroke. Penelitian ini bertujuan merancang aplikasi web prediksi stroke menggunakan algoritma Gaussian Naïve Bayes. Proses penelitian meliputi perencanaan, analisis kebutuhan, perancangan sistem dan aplikasi, serta pengembangan model prediksi menggunakan dataset dari Kaggle. Evaluasi model menunjukkan akurasi 72%, presisi 72%, recall 75%, F1-score untuk kelas prediksi pertama 71%, dan F1-score untuk kelas prediksi kedua 73%. Aplikasi ini berhasil diimplementasikan dan menunjukkan kinerja yang memuaskan dalam memprediksi risiko stroke berdasarkan data yang dikumpulkan.

Kata kunci: *Gaussian Naïve Bayes, Naïve Bayes, Pembelajaran Mesin, Prediksi, Stroke*

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

Design and Development of a Stroke Prediction Application Using the Gaussian Naïve Bayes Algorithm

Michelle Shiela Haryanto

ABSTRACT

Stroke is one of the most dangerous diseases that causes disability and death throughout the world. Studies show a significant increase in the incidence of stroke, including ischemic and hemorrhagic stroke, with a particularly high prevalence in elderly individuals. Economic disparities influence the distribution of stroke burden, with low-income countries reporting the majority of deaths and disability. In Indonesia, stroke is the leading cause of death, affecting patients as well as the health system as a whole. The five main risk factors for stroke are high blood pressure, high body mass index, high blood glucose, air pollution, and smoking. Early detection of stroke is critical to improving recovery and reducing complications. Various detection methods have been developed, including MRI and CT, but they are expensive and invasive. Machine learning offers a faster and cheaper solution, with the Gaussian Naïve Bayes algorithm showing good performance in stroke prediction. This research aims to design a stroke prediction web application using the Gaussian Naïve Bayes algorithm. The research process includes planning, needs analysis, system and application design, and development of prediction models using datasets from Kaggle. Model evaluation shows accuracy 72%, precision 72%, recall 75%, F1-score for the first prediction class 71%, and F1-score for the second prediction class 73%. This application was successfully implemented and showed satisfactory performance in predicting stroke risk based on the collected data.

Keywords: *Gaussian Naive Bayes, Machine Learning, Naive Bayes, Prediction, Stroke*

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI ILMIAH	iv
HALAMAN PERSEMBAHAN/MOTO	v
KATA PENGANTAR	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
DAFTAR KODE	xiii
DAFTAR LAMPIRAN	xiv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Permasalahan	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	5
BAB 2 LANDASAN TEORI	6
2.1 Stroke	6
2.2 Naïve Bayes	6
2.3 Gaussian Naïve Bayes	8
2.4 Confusion Matrix	8
2.5 Evaluasi Peforma	9
2.6 Python	10
2.7 Penelitian Terdahulu	10
BAB 3 METODOLOGI PENELITIAN	24
3.1 Perencanaan	24
3.2 Analisa kebutuhan	24
3.2.1 Studi Literatur	24
3.2.2 Teknik Pengumpulan Data	24
3.3 Perancangan Sistem	25
3.3.1 Perancangan Model	25
3.3.2 Perancangan Aplikasi	26
3.4 Pengembangan	37
3.4.1 Data Preprocessing	37
3.4.2 Seleksi Model	38
3.4.3 Cross Validation	38
3.4.4 Penyimpanan Model	38
3.4.5 Prediksi dan Evaluasi	38
3.4.6 Deployment	39
3.5 Pengujian	39
3.6 Dokumentasi	39
3.7 Penulisan Laporan	39
3.8 Spesifikasi Sistem	39
BAB 4 HASIL DAN DISKUSI	41

4.1	Pembuatan Model	41
4.1.1	Library yang digunakan	41
4.1.2	Pemuatan Data	41
4.1.3	Data Preprocessing	42
4.1.4	Seleksi Model dan Data Train	44
4.1.5	Cross Validation	45
4.1.6	Penyimpanan Model Gaussian Naïve Bayes Setelah Pelatihan	45
4.1.7	Data Test	45
4.1.8	Prediksi dan Evaluasi	46
4.2	Pembuatan Aplikasi	48
4.2.1	Fungsi Index	48
4.2.2	Fungsi Prediksi	49
4.2.3	Fungsi Create Form	49
4.2.4	Fungsi Save Form	50
4.2.5	Fungsi untuk eksekusi python script	51
4.2.6	Fungsi Result	52
4.2.7	Fungsi Detail	52
4.2.8	Dataset stroke_dataset	53
4.2.9	Halaman Home	53
4.2.10	Halaman Prediksi	54
4.2.11	Halaman Form Prediksi	54
4.2.12	Halaman Result	55
4.2.13	Halaman Detail Prediksi	56
4.3	Diskusi	57
BAB 5	SIMPULAN DAN SARAN	59
5.1	Simpulan	59
5.2	Saran	59
	DAFTAR PUSTAKA	60



DAFTAR GAMBAR

Gambar 3.1	<i>Flowchart</i> Pemodelan	26
Gambar 3.2	<i>Flowchart</i> Halaman Home	27
Gambar 3.3	<i>Flowchart</i> Halaman Prediksi	28
Gambar 3.4	<i>Flowchart</i> Halaman Form Prediksi	29
Gambar 3.5	<i>Flowchart</i> Submit	30
Gambar 3.6	<i>Flowchart</i> Halaman Hasil Prediksi	31
Gambar 3.7	<i>Flowchart</i> Halaman Detail Prediksi	32
Gambar 3.8	Halaman Home	33
Gambar 3.9	Halaman Prediksi	33
Gambar 3.10	Halaman Form Prediksi	34
Gambar 3.11	Halaman Hasil Prediksi	35
Gambar 3.12	Halaman Detail Prediksi	36
Gambar 3.13	Entity Relationship Diagram	37
Gambar 4.1	Hasil perintah df.head	42
Gambar 4.2	Hasil pengecekan nilai null dalam dataset	42
Gambar 4.3	Hasil Cross Validation	45
Gambar 4.4	Hasil Ekstrasi Joblib	45
Gambar 4.5	Visualisasi Heatmap Confussion Matrix	47
Gambar 4.6	Menghitung Akurasi	47
Gambar 4.7	Menghitung Presisi	48
Gambar 4.8	Menghitung Recall	48
Gambar 4.9	Classification Report	48
Gambar 4.10	stroke_dataset	53
Gambar 4.11	Halaman Home	54
Gambar 4.12	Halaman Prediksi	54
Gambar 4.13	Halaman Form	55
Gambar 4.14	Lanjutan Halaman Form	55
Gambar 4.15	Halaman Result	56
Gambar 4.16	Tombol Detail	56
Gambar 4.17	Halaman Detail Prediksi	57
Gambar 4.18	Lanjutan 1 Halaman Detail Prediksi	57

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

DAFTAR TABEL

Tabel 2.1	Tabel Confussion Matrix	8
Tabel 2.2	Tabel Penelitian Terdahulu	11
Tabel 2.3	Tabel Lanjutan 1 Penelitian Terdahulu	12
Tabel 2.4	Tabel Lanjutan 2 Penelitian Terdahulu	13
Tabel 2.5	Tabel Lanjutan 3 Penelitian Terdahulu	14
Tabel 2.6	Tabel Lanjutan 4 Penelitian Terdahulu	15
Tabel 2.7	Tabel Lanjutan 5 Penelitian Terdahulu	16
Tabel 2.8	Tabel Lanjutan 6 Penelitian Terdahulu	17
Tabel 2.9	Tabel Lanjutan 7 Penelitian Terdahulu	18
Tabel 2.10	Tabel Lanjutan 8 Penelitian Terdahulu	19
Tabel 2.11	Tabel Lanjutan 9 Penelitian Terdahulu	20
Tabel 2.12	Tabel Lanjutan 10 Penelitian Terdahulu	21
Tabel 2.13	Tabel Lanjutan 11 Penelitian Terdahulu	22



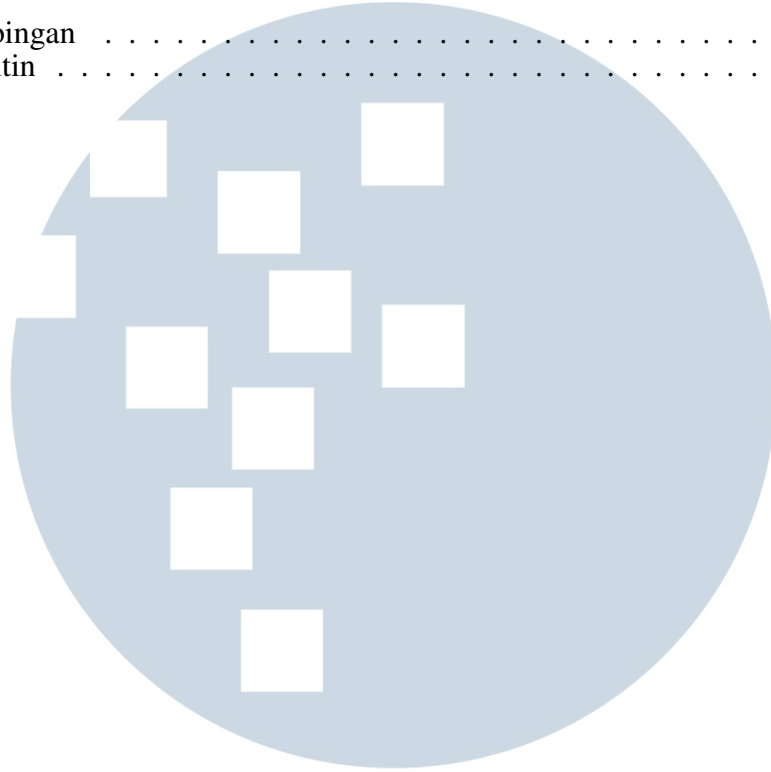
DAFTAR KODE

4.1	Read dataset and data	41
4.2	Pengecekan nilai null	42
4.3	Penghilangan nilai null	43
4.4	Label Encoder	43
4.5	Memisahkan data berdasarkan variabel target	43
4.6	Menghilangkan variabel independen dan menentukan variabel dependen	44
4.7	Membagi dataset menjadi data train dan data test	44
4.8	Standarisasi nilai fitur	44
4.9	Seleksi Model dan Data Train	44
4.10	Cross Validation	45
4.11	Penyimpanan model dengan ekstrasi joblib	45
4.12	Data Test	45
4.13	Confussion Matrix	46
4.14	Visualisasi Heatmap Confussion Matrix	46
4.15	Menghitung Akurasi	47
4.16	Menghitung Presisi	47
4.17	Menghitung Recall	48
4.18	Classification Report	48
4.19	Fungsi Index	49
4.20	Fungsi Prediksi	49
4.21	Fungsi Create Form	49
4.22	Fungsi Save	50
4.23	Fungsi untuk eksekusi python script	51
4.24	Fungsi Result	52
4.25	Fungsi Detail	53



DAFTAR LAMPIRAN

Form Bimbingan	63
Hasil Turnitin	64



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA