

**SKRIPSI**  
**PENERAPAN METODE NAIVE BAYES UNTUK KLASIFIKASI**  
**PENYAKIT DIABETES MELITUS**



**AINUN ANNISA**

**NIM: 205410074**

**PROGRAM STUDI INFORMATIKA**  
**FAKULTAS TEKNOLOGI INFORMASI**  
**UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA**

**2024**

**SKRIPSI**

**PENERAPAN METODE NAIVE BAYES UNTUK KLASIFIKASI  
PENYAKIT DIABETES MELITUS**

**Diajukan sebagai salah satu syarat untuk menyelesaikan studi**

**Program Sarjana**

**Program Studi Informatika**

**Fakultas Teknologi Infoemasi**

**Universitas Teknologi Digital Indonesia**

**Yogyakarta**

**Disusun Oleh**

**AINUN ANNISA**

**NIM : 205410074**

**PROGRAM STUDI INFORMATIKA**

**PROGRAM SARJANA**

**FAKULTAS TEKNOLOGI INFORMASI**

**UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA**

**YOGYAKARTA**

**2024**

**HALAMAN PERSETUJUAN**

**UJIAN SKRIPSI**

**Judul** : Penerapan Metode Naive Bayes Untuk  
Klasifikasi Penyakit Diabetes Melitus

**Nama** : Ainun Annisa

**NIM** : 205410074

**Jurusan** : Informatika

**Semester** : Ganjil

**Tahun Akademik** : 2023/2024



Telah diperiksa dan disetujui untuk diujikan di hadapan Dewan Penguji Skripsi

Yogyakarta, \_\_\_\_\_

Dosen pembimbing,

  
M. Agung Nugroho, M.Kom

NIDN. 0507078501

**HALAMAN PENGESAHAN**

**SKRIPSI**

**PENERAPAN METODE NAIVE BAYES UNTUK KLASIFIKASI**

**PENYAKIT DIABETES MELITUS**

Telah dipertahankan di depan Dewan Penguji Skripsi dan dinyatakan diterima untuk memenuhi sebagian persyaratan guna memperoleh Gelar

**Sarjana Komputer**


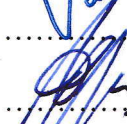
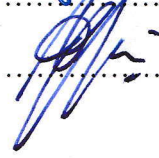
**Program Studi Informatika**

**Fakultas Teknologi Informasi**

**Universitas Teknologi Digital Indonesia**

**Yogyakarta**

Yogyakarta, \_\_\_\_\_

Dewan Penguji	NIDN	Tanda Tangan
Ilham Rais Arvianto, S.Pd., M.Pd.	0510048901	
Femi Dwi Astuti, S.Kom., M.Cs.	0516088701	
M. Agung Nugroho, S.Kom., M.Kom.	0507075501	

Mengetahui

Ketua Program Studi Informatika

  
Dini Fakta Sari, S.T., M.T.

NPP : 121172

## PERNYATAAN KEASLIAN SKRIPSI

Dengan ini penulis menyatakan bahwa naskah skripsi ini belum pernah diajukan untuk memperoleh gelar Sarjana Komputer di suatu Perguruan Tinggi, dan sepanjang pengetahuan penulis tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara sah diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta,.....2024



Ainun Annisa

NIM:205410074

## **HALAMAN PERSEMBAHAN**

Puji syukur penulis panjatkan kepada Allah SWT, yang telah memberikan kesehatan, rahmat dan hidayah, sehingga penulis masih diberikan kesempatan untuk menyelesaikan skripsi ini, sebagai salah satu syarat untuk mendapatkan gelar kesarjanaan. Walaupun jauh dari kata sempurna, namun penulis bangga telah mencapai pada titik ini, yang akhirnya skripsi ini bisa terselesaikan. Skripsi ini saya persembahkan untuk :

1. Allah SWT yang telah memberikan kemudahan dalam menyelesaikan skripsi ini. Puji syukur yang tak terhingga yang telah meridhoi dan mengabulkan segala do'a karena hanya atas Karunia-Nyalah skripsi ini dapat selesai.
2. Untuk (Alm) Bapak H.Sangkala selaku orang tua saya yang sudah meninggal sejak bulan Agustus 2023, yang belum sempat saya berikan kebahagiaan dan rasa bangga. Skripsi ini sebagai tanda bahwa perjuangan orang tua saya untuk memberikan pendidikan tinggi untuk anaknya tidak sia-sia. Semoga Beliau bangga disana.
3. Untuk Ibu saya, Ibu Nuryani yang selama ini selalu memberikan dukungan dan do'a yang tiada hentinya secara lahir dan batin.
4. Untuk keluarga besar saya yang selalu memberikan motivasi.
5. Untuk teman-teman Informatika 2020, IF2 UTDI, ML Skuy, Support System yang telah membantu dan menjadi teman selama di perkuliahan

serta Kos Kusuma KarangJambe yang sudah menjadi tempat kos terbaik dan ternyaman di Jogja.

6. Untuk orang-orang terdekatku Debo, Atri, Nova, Ellena, Tikah terima kasih untuk empat tahun ini yang selalu ada disaat susah dan senang dan saya sangat bersyukur memiliki kalian di hidup saya.
7. Untuk sahabat-sahabatku yang di Sorong, Eka, Intan yang selalu memberikan doa, semangat, dan menjadi pendengar yang baik, terima kasih kalian orang baik.
8. Untuk Nim 220511034 terimakasih atas dukungan, kebaikan, perhatian, dan kebijaksanaan, terima kasih karena memberitahu saya hidup dengan bahagia.
9. Untuk Kantor Disperindag tempat PKL, terima kasih atas kebaikan Ibu dan Bapak selama saya magang. Saya mendapatkan banyak pengalaman dan ilmu yang bermanfaat.
10. Untuk diri sendiri yang mau dan mampu bertahan, berjuang dan berusaha untuk menyelesaikan skripsi ini, terimakasih karena sudah bertahan untuk tetap kuat sampai detik ini.

## **MOTTO**

*“Barang siapa keluar untuk mencari sebuah ilmu, maka ia akan berada di jalan  
Allah hingga ia kembali”*

- HR Tirmidzi”

“Usaha membuat kamu menjadi dirimu, kamu akan menyesalinya suatu saat nanti  
jika kamu tidak melakukan yang terbaik sekarang”

- Jungkook BTS -



## **KATA PENGANTAR**

Segala puji dan syukur penulis panjatkan kepada Allah SWT Sang Maha Segalanya, atas seluruh curahan rahmat dan hidayahNya sehingga penulis mampu menyelesaikan skripsi yang berjudul “Penerapan Metode Naive Bayes Untuk Klasifikasi Penyakit Diabetes Melitus” yang masih jauh dari kata sempurna.

Dalam penyelesaian studi dan penulisan skripsi ini, penulis banyak memperoleh bantuan baik pengajaran, bimbingan dan arahan dari berbagai pihak baik secara langsung maupun tidak langsung. Untuk itu penulis menyampaikan penghargaan dan terima kasih yang tak terhingga kepada :

1. Ibu Sri Redjeki, S.Si., M.Kom., Ph.D. selaku Rektor Universitas Teknologi Digital Indonesia.
2. Dr. Bambang Purnomosidi DP, S.E. Akt., S.Kom., MMSI. selaku Dekan Fakultas Teknologi Informasi Universitas Teknologi Digital Indonesia.
3. Ibu Dini Fakta Sari, S.T., M.T. dan Ibu Femi Dwi Astuti, S.Kom., M.Cs. , Ketua dan Sekretaris Program Studi Informatika Universitas Teknologi Digital Indonesia.
4. Ibu Maria Mediatrix Sebatubun S.Kom., M.Eng selaku wali dosen selama perkuliahan.
5. Bapak M. Agung Nugroho, S.Kom., M.Kom selaku dosen pembimbing yang telah memberikan arahan dan bimbingan selama pengerjaan skripsi.
6. Bapak Ilham Rais Arvianto, S.Pd.,M.Pd. dan Ibu Femi Dwi Astuti, S.Kom., M.Cs. selaku dosen penguji pada sidang skripsi.

7. Teman-teman mahasiswa Universitas Teknologi Digital Indonesia yang telah terlibat dalam membantu terselesaikannya skripsi ini.

Sebagai manusia biasa penulis menyadari penyusunan skripsi ini jauh dari kata sempurna, oleh karena itu atas kesalahan dan kekurangan dalam penulisan skripsi ini, penulis memohon maaf dan bersedia menerima kritik yang membangun

Harapan penulis, semoga skripsi ini dapat memberikan manfaat bagi siapa saja yang membacanya.

Yogyakarta,.....

Ainun Annisa

NIM:205410074

## DAFTAR ISI

SKRIPSI .....	i
HALAMAN PERSETUJUAN .....	ii
HALAMAN PENGESAHAN .....	iii
PERNYATAAN KEASLIAN SKRIPSI .....	iv
HALAMAN PERSEMBAHAN .....	v
MOTTO .....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI .....	x
DAFTAR GAMBAR .....	xiii
DAFTAR TABEL .....	xvi
INTISARI .....	xvii
ABSTRACT .....	xviii
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang Masalah .....	1
1.2 Rumusan Masalah .....	3
1.3 Ruang Lingkup .....	3
1.4 Tujuan Penelitian .....	4
1.5 Manfaat Penelitian .....	4
1.6 Sistematika Penulisan .....	4

BAB II TINJAUAN PUSTAKA DAN DASAR TEORI.....	5
BAB III METODE PENELITIAN.....	5
BAB IV IMPLEMENTASI DAN PEMBAHASAN.....	5
BAB V PENUTUP.....	6
BAB II.....	7
TINJAUAN PUSTAKA DAN DASAR TEORI.....	7
2.1 Tinjauan Pustaka.....	7
2.2 Dasar Teori.....	12
2.2.1 Diabetes Melitus.....	12
2.2.2 Naive Bayes Classifier.....	12
2.2.3 Data Mining.....	14
2.2.4 Confusion Matrix.....	15
2.2.5 Python.....	17
BAB III METODE PENELITIAN.....	19
3.1 Bahan/Data.....	19
3.2 Peralatan.....	19
3.2.1 Kebutuhan Perangkat Keras.....	19
3.2.2 Kebutuhan Perangkat Lunak.....	19
3.2.3 Kebutuhan Input.....	20
3.2.4 Kebutuhan Proses.....	20
3.2.5 Kebutuhan Output.....	28
3.3 Analisis dan Rancangan Sistem.....	28
3.3.1 Use Case Diagram.....	28

3.3.2 Perancangan Antarmuka .....	29
BAB IV .....	32
<b>IMPLEMENTASI DAN PEMBAHASAN .....</b>	<b>32</b>
4.1 Hasil .....	32
4.1.1 Pengumpulan Data.....	32
4.1.4 Preprocessing Data .....	33
4.1.5 Membagi Data Training dan Data Testing .....	37
4.1.6 Model Naive Bayes .....	41
4.1.7 Confusion Matrix .....	44
4.2 Uji Coba Sistem.....	50
4.2.1 Antarmuka Aplikasi Web .....	51
BAB V.....	53
PENUTUP.....	53
3.4 Kesimpulan.....	53
3.5 Saran.....	54
DAFTAR PUSTAKA.....	55
LAMPIRAN.....	56
LAMPIRAN.....	57

## DAFTAR GAMBAR

Tabel 2. 1 Perbandingan Penelitian .....	9
Tabel 2. 2 Confusion Matrix .....	15
Gambar 3. 1 Alur Penelitian .....	21
Tabel 3. 1 Perhitungan Manual .....	23
Gambar 3. 2 Use Case Diagram .....	28
Gambar 3. 3 Halaman <i>Login</i> .....	30
Gambar 3. 4 Halaman Diagnosa Penyakit.....	30
Gambar 3. 5 Halaman Riwayat Penyakit .....	31
Gambar 4. 1 Kode Menampilkan Data.....	32
Gambar 4. 2 Hasil Menampilkan Data.....	33
Gambar 4. 3 Kode Missing Value .....	33
Gambar 4. 4 Hasil Missing Value .....	34
Gambar 4. 5 Kode Duplikat Data .....	34
Gambar 4. 6 Hasil Duplikat Data .....	34
Gambar 4. 7 Kode Memeriksa Data .....	34
Gambar 4. 8 Hasil Memeriksa Data .....	35
Gambar 4. 9 Kode Encoding Kategorikal .....	35
Gambar 4. 10 Hasil Encoding Kategorikal.....	35
Gambar 4. 11 Kode Menghitung Jumlah Kemunculan .....	35
Gambar 4. 12 Hasil Menghitung Jumlah Kemunculan .....	36
Gambar 4. 13 Visualisasi Data .....	36
Gambar 4. 14 Kode Pembagian Data .....	37

Gambar 4. 15 Kode Menampilkan Data Train dan Data Test .....	37
Gambar 4. 16 Hasil Data Training 90% .....	38
Gambar 4. 17 Hasil Data Testing 10% .....	38
Gambar 4. 18 Hasil Data Training 80% .....	39
Gambar 4. 19 Hasil Data Testing 20% .....	39
Gambar 4. 20 Hasil Data Training 50% .....	40
Gambar 4. 21 Hasil Data Testing 50% .....	40
Gambar 4. 22 Kode Model Naïve Bayes.....	42
Gambar 4. 23 Kode Melatih Data Training.....	42
Gambar 4. 24 Kode Mengakses Atribut .....	42
Gambar 4. 25 Hasil Mengakses Atribut .....	42
Gambar 4. 26 Kode Prediksi Data Testing.....	43
Gambar 4. 27 Hasil Prediksi.....	43
Gambar 4. 28 Kode Menghitung Probabilitas .....	43
Gambar 4. 29 Hasil Menghitung Probabilitas .....	44
Gambar 4. 30 Kode Confusion Matrix.....	45
Gambar 4. 31 Hasil Perbandingan 90:10.....	46
Gambar 4. 32 Hasil Perbandingan 80:20.....	47
Gambar 4. 33 Hasil Perbandingan 50:50.....	48
Gambar 4. 34 Visualisasi Hasil Confusion Matrix.....	50
Gambar 4. 35 Tampilan Command Prompt .....	51
Gambar 4. 36 Tampilan Interface Halaman Login.....	51
Gambar 4. 37 Tampilan Interface Halaman Diagnosa Penyakit .....	52

Gambar 4. 38 Halaman Interface Riwayat Penyakit .....52



## DAFTAR TABEL

Tabel 2. 1 Perbandingan Penelitian .....	9
Tabel 2. 2 Confusion Matrix .....	15
Tabel 3. 1 Perhitungan Manual .....	23
Tabel 4.1 Perbandingan Akurasi.....	50

## INTISARI

Diabetes Melitus adalah penyakit metabolik kronis yang ditandai dengan peningkatan kadar glukosa darah, yang mengarah dari waktu ke waktu ke kerusakan serius pada jantung, pembuluh darah, mata, ginjal, dan saraf. Diabetes melitus terdaftar sebagai penyakit penyumbang kematian terbesar didunia. Diabetes melitus dapat diklasifikasikan berdasarkan kemungkinan terkenanya dari atribut gejala diawal fasenya. Penyakit ini bisa dideteksi karena banyak gejala yang terdeteksi. Data yang digunakan pada analisis ini merupakan data dari kaggle yaitu Diabetes Prediction Dataset yang terdiri 8 variabel input dan 1 variabel kelas.

Analisis yang digunakan meliputi data preprocessing, model, dan evaluasi. Pengujian metode klasifikasi pada riset adalah Naive Bayes, data yang digunakan tidak seimbang yang berjumlah 100.000 data tetapi setelah preprocessing data didapatkan 96146 dengan kelas yang tidak terkena diabetes lebih banyak dibandingkan kelas yang terkena diabetes. Hasil klasifikasi pada penelitian ini didapatkan akurasi terbaik pada perbandingan 80:20 yaitu dengan akurasi = 90.38%, presisi = 45.84%, recall = 65.73% dan F1 score = 54.01%.

**Kata kunci :** *Data Mining, Klasifikasi, Naive Bayes, Diabetes Melituus*

## ABSTRACT

Diabetes Mellitus is a chronic metabolic disease characterized by elevated blood glucose levels, leading over time to serious damage to the heart, blood vessels, eyes, kidneys, and nerves. Diabetes mellitus is listed as the biggest cause of death in the world. Diabetes mellitus can be classified based on the possibility of being affected by the symptom attributes at the beginning of the phase. This disease can be detected because many symptoms are detected. The data used in this analysis is data from Kaggle, namely the Diabetes Prediction Dataset which consists of 8 input variables and 1 class variabel. The analysis used includes data preprocessing, models and evaluation.

The analysis used includes data preprocessing, models and evaluation. The classification method tested in the research was Naive Bayes, the data used was unbalanced, amounting to 100,000 data, but after preprocessing the data was obtained 96,146 with more classes that were not affected by diabetes, namely compared to classes that were affected. The classification results in this study obtained the best accuracy in a ratio of 80:20, namely with accuracy = 90.38%, precision = 45.84%, recall = 65.73% and F1 score = 54.01%.

*Keywords: Data Mining, Classification, Naive Bayes, Diabetes Mellitus*