



**PENERAPAN METODE *NAIVE BAYES* DALAM
KLASIFIKASI DATA PENYAKIT DIABETES**

SKRIPSI

**Karya tulis sebagai salah satu syarat untuk memperoleh gelar Sarjana
Komputer (S. Kom) dari Universitas Indo Global Mandiri.**

Oleh
M. ILHAM JAYA KUSUMA
NIM. 2018310008
(Program Studi Sarjana Sistem Komputer)

FAKULTAS ILMU KOMPUTER
UNIVERSITAS INDO GLOBAL MANDIRI
JANUARI 2024



**PENERAPAN METODE *NAIVE BAYES* DALAM
KLASIFIKASI DATA PENYAKIT DIABETES**

SKRIPSI

**Karya tulis sebagai salah satu syarat untuk memperoleh gelar Sarjana
Komputer (S. Kom) dari Universitas Indo Global Mandiri.**

Oleh
M. ILHAM JAYA KUSUMA
NIM. 2018310008
(Program Studi Sarjana Sistem Komputer)

FAKULTAS ILMU KOMPUTER
UNIVERSITAS INDO GLOBAL MANDIRI
JANUARI 2024

HALAMAN PENGESAHAN

Oleh
M. Ilham Jaya Kusuma
NIM: 2018310008
(Program Studi Sarjana Sistem Komputer)

Universitas Indo Global Mandiri

Menyetujui
Tim Pembimbing

Tanggal 24 Januari 2024

Pembimbing 1



Tasmi, S. Si., M. Kom
NIDN. 2017.01.02.30

Pembimbing 2



Ricky Maulana Fajri, S. Kom., M. Sc
NIDN. 2016.01.02.20

Mengetahui
Dekan



Rudi hermansyah, S.T., M. Eng., Ph. D
NIDN.2022.01.03.15

LEMBAR PERSETUJUAN DEWAN PENGUJI

Pada hari ini Senin Tanggal 15 Januari 2024 telah dilaksanakan Ujian Skripsi oleh Program Studi Sistem Komputer Fakultas Ilmu Komputer Universitas Indo Global Mandiri Palembang.

Menyetujui
Tim Penguji


Palembang 15 Januari 2024

Ketua Penguji




Tasmi, S. Si., M. Kom
NIDN. 2017.01.02.30

Penguji 1



Ir. Zulkifli., M. Sc
NIDN. 2011.01.01.11

Penguji 2



Fery Antony., S.T., M. Kom
NIDN. 2003.01.00.67

Mengetahui
Ketua Program Studi Sistem Komputer



Tasmi., S. Si., M. Kom
NIDN. 2017.01.02.30

SURAT KETERANGAN REVISI SKRIPSI

Kami yang bertanda tangan dibawah ini, menerangkan bahwa:

Nama : M. Ilham Jaya Kusuma
NPM : 2018310008
Judul Skripsi : Penerapan Metode *Naïve Bayes* dalam Klasifikasi Data Penyakit Diabetes.

Mahasiswa yang namanya tercantum diatas, telah selesai merevisi penulisan skripsi.

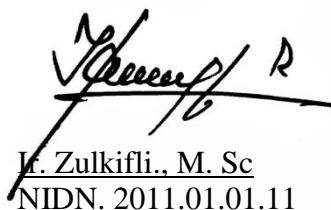
Menyetujui
Tim Penguji
Tanggal 25 Januari 2024

Ketua Penguji



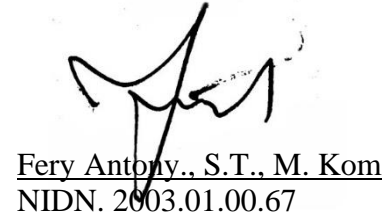
Tasmi, S. Si., M. Kom
NIDN. 2017.01.02.30

Penguji 1



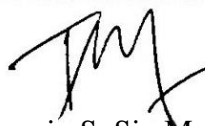
J. Zulkifli., M. Sc
NIDN. 2011.01.01.11

Penguji 2



Fery Antony., S.T., M. Kom
NIDN. 2003.01.00.67

Mengetahui
Ketua Program Studi Sistem Komputer



Tasmi., S. Si., M. Kom
NIDN. 2017.01.02.30

PENERAPAN METODE *NAIVE BAYES* DALAM KLASIFIKASI DATA PENYAKIT DIABETES

ABSTRAK

PENERAPAN METODE *NAIVE BAYES* DALAM KLASIFIKASI DATA PENYAKIT DIABETES

Penelitian ini fokus pada penerapan metode *Naive Bayes* untuk mengklasifikasikan data penyakit diabetes menggunakan bahasa pemrograman Python. Langkah-langkah implementasi mencakup pengumpulan data, pemrosesan data, pelatihan model *Naive Bayes*, dan evaluasi klasifikasi. Atribut data yang digunakan termasuk ketebalan kulit, insulin, tekanan darah, kadar gula darah, dan riwayat keluarga. Evaluasi model menunjukkan akurasi sebesar 0.7813, dengan perbedaan signifikan antara metrik pada kelas non-diabetes dan diabetes. Model memiliki recall negatif 87%, menemukan data non-diabetes dengan baik, namun recall positif hanya 59%, menemukan kasus diabetes dengan keterbatasan. Precision negatif sebesar 82%, sementara precision positif 68%. Hasil ini memberikan wawasan tentang kinerja *Naive Bayes* pada klasifikasi penyakit diabetes, dengan implikasi untuk meningkatkan identifikasi kasus sebenarnya dan mengatasi ketidakseimbangan data. Penelitian ini memberikan panduan praktis bagi peneliti atau praktisi yang tertarik dalam menerapkan metode ini pada analisis data medis.

Kata Kunci: Diabetes, Klasifikasi, *Naive Bayes*, *Python*

ABSTRACT

APPLICATION OF THE NAIVE BAYES METHOD IN CLASSIFICATION OF DIABETES DISEASE DATA

This research focuses on applying the Naive Bayes method to classify diabetes data using the Python programming language. Implementation steps include data collection, data processing, Naive Bayes model training, and classification evaluation. Data attributes used include skin thickness, insulin, blood pressure, blood sugar levels, and family history. Model evaluation showed an accuracy of 0.7813, with significant differences between metrics in the non-diabetic and diabetes classes. The model had a negative recall of 87%, finding non-diabetic data well, but a positive recall of only 59%, finding diabetes cases with limitations. Negative precision is 82%, while positive precision is 68%. These results provide insight into the performance of Naive Bayes in diabetes disease classification, with implications for improving identification of true cases and addressing data imbalances. This research provides practical guidance for researchers or practitioners interested in applying this method to medical data analysis.

Keyword: Diabetes, Classification, Naive Bayes, Python.

PEDOMAN PENGGUNAAN SKRIPSI

Skripsi S1 yang tidak dipublikasikan terdaftar dan tersedia di Perpustakaan Universitas Indo Global Mandiri, dan terbuka untuk umum dengan ketentuan bahwa hak cipta ada pada penulis dengan mengikuti aturan HaKI yang berlaku di Universitas Indo Global Mandiri. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau peringkasan hanya dapat dilakukan seizin penulis dan harus disertai dengan kaidah ilmiah untuk menyebutkan sumbernya.

Sitasi hasil penelitian Skripsi ini dapat di tulis dalam bahasa Indonesia sebagai berikut:

Kusuma, M.I.J. (Tahun): Penerapan Metode *Naive Bayes* Dalam Klasifikasi Data Penyakit Diabetes, Skripsi Program Sarjana, Universitas Indo Global Mandiri.

dan dalam bahasa Inggris sebagai berikut:

Kusuma, M.I.J. (Tahun): *Application of The Naive Bayes Method In Classification Of Diabetes Disease Data*, Postgraduate Thesis, Universitas Indo Global Mandiri.

Memperbanyak atau menerbitkan sebagian atau seluruh Skripsi haruslah seizin Dekan Fakultas Ilmu Komputer, Universitas Indo Global Mandiri.

KATA PENGANTAR

Puji dan syukur Kami panjatkan atas kehadiran Tuhan Yang Maha Esa karena akhirnya Skripsi ini bisa terselesaikan dengan baik tepat pada waktunya.

Skripsi yang Penulis buat dengan judul “**Penerapan Metode Naive Bayes dalam Klasifikasi Data Penyakit Diabetes**”. dibuat sebagai salah satu syarat untuk menyelesaikan studi pada Program Studi Sistem Komputer. Tidak Lupa Penulis mengucapkan terima kasih atas bantuan yang diberikan selama penyusunan Skripsi ini kepada :

1. Kedua orang tua dan saudara-saudara terima kasih atas kehadiran dan dukungan yang selalu ada di samping penulis, memberikan semangat yang tak henti kepada penulis.
2. Bapak Rudi Heriansyah, S.T., M. Eng, Ph. D sebagai Dekan Fakultas Ilmu Komputer.
3. Bapak Tasmi, S.Si., M.Kom sebagai ketua program studi Sistem Komputer dan sebagai dosen pembimbing I menjadi pembimbing dan memberikan berbagai saran yang berharga untuk mencapai keberhasilan skripsi ini.
4. Bapak Ricky Maulana Fajri, S.Kom., M.Sc sebagai dosen pembimbing II yang telah menyisihkan waktu untuk memberikan nasehat dan bimbingan yang sangat berarti demi berhasilnya penyelesaian skripsi ini.
5. Bapak Ir. Hastha Sunardi., M.T sebagai dosen pembimbing akademik yang telah memberikan arahan dan masukan berharga kepada saya dalam menyusun rencana mata kuliah selama proses pembelajaran.
6. Bapak dan ibu dosen yang selama ini telah melimpahkan ilmunya kepada penulis selama proses belajar mengajar di Fakultas Ilmu Komputer Universitas Indo Global Mandiri Palembang.
7. Teman-Teman Angkatan 2018 Program Studi Sistem Komputer.
8. Diah Wulandari yang telah menemani selama pengerjaan skripsi.
9. Old Men, Rizki Qisthinano dan Muhammad Arigoh Waluyo yang telah membantu dan memberikan saran masukan selama pengerjaan skripsi.

Penulis menyadari bahwa dalam penulisan Skripsi ini masih jauh dari kata sempurna, oleh karena itu Penulis mengharapkan saran dan kritik yang sifatnya membangun agar dapat digunakan demi perbaikan ini nantinya. Penulis juga berharap agar Skripsi ini akan memberikan banyak manfaat bagi yang membacanya.

Penulis

M. Ilham Jaya Kusuma

DAFTAR ISI

HALAMAN JUDUL LUAR.....	1
HALAMAN JUDUL DALAM.....	2
HALAMAN PENGESAHAN.....	iii
LEMBAR PERSETUJUAN DEWAN PENGUJI	iv
SURAT KETERANGAN REVISI SKRIPSI	v
SURAT PERNYATAAN BEBAS PLAGIAT.....	vi
ABSTRAK.....	vii
<i>ABSTRACT</i>	viii
PEDOMAN PENGGUNAAN SKRIPSI	ix
HALAMAN PERUNTUKAN	xi
DAFTAR ISI.....	xiv
DAFTAR LAMPIRAN.....	xviii
DAFTAR GAMBAR DAN ILUSTRASI.....	xix
DAFTAR TABEL.....	xx
BAB I PENDAHULUAN.....	2
1.1 Latar Belakang.....	2
1.2 Rumusan Masalah	6
1.3 Batasan Masalah.....	7
1.4 Tujuan dan Manfaat Penelitian.....	7

1.4.1	Tujuan Penelitian	7
1.4.2	Manfaat Penelitian	7
1.5	Sistematika penulisan	8
BAB I PENDAHULUAN.....		8
BAB II LANDASAN TEORI		8
BAB III METODOLOGI PENELITIAN.....		9
BAB IV HASIL DAN PEMBAHASAN		9
BAB V KESIMPULAN DAN SARAN.....		10
BAB II.....		11
TINJAUAN PUSTAKA.....		11
2.1	Landasan Teori	11
2.1.1	Analisa.....	11
2.1.2	Klasifikasi	12
2.1.3	<i>Data Mining</i>	13
2.1.4	<i>Naive Bayes</i>	14
2.1.5	<i>Knowledge Discovery in Database</i>	16
2.1.6	<i>Python</i>	19
2.2	Penelitian Sebelumnya	21
BAB III		32
METODOLOGI PENELITIAN		32

3.1	Identifikasi Masalah	32
3.1.1	Tahapan Penelitian	32
3.1.2	Deskripsi Penelitian	34
3.1.3	Studi Pustaka.....	35
3.2	Model Yang Digunakan	37
3.3	Metode Pemrosesan Data	38
3.4	Data Penelitian	40
3.5	Data Sampel	41
3.6	Alat dan Bahan	42
3.5.1.	Alat Penelitian.....	43
3.5.2.	Bahan Penelitian.....	44
BAB IV		47
HASIL DAN PEMBAHASAN.....		47
4.1	Pengumpulan Data	47
4.2	Persebaran Distribusi Data	48
4.2.1	Distirbusi Data Berdasarkan Kadar Gula.....	48
4.2.2	Distirbusi Data Berdasarkan Tekanan Darah.....	49
4.2.3	Distirbusi Data Berdasarkan Berat Badan.....	50
4.2.4	Distirbusi Data Berdasarkan Usia Pasien.....	52
4.3	Klasifikasi <i>Naïve Bayes</i>	53

4.5	Uji Data dan Model	59
4.6	Hasil Pembahasan.....	66
BAB V.....		69
PENUTUP.....		69
5.1	Kesimpulan.....	69
5.2	Saran.....	70
DAFTAR PUSTAKA		72
LAMPIRAN.....		76

DAFTAR LAMPIRAN

Lampiran 1. Surat Persetujuan Ujian Skripsi.

Lampiran 2. Surat Keterangan Siap Sidang.

Lampiran 3. Pernyataan Tidak Plagiat.

Lampiran 4. Lembar Revisi Ujian Skripsi.

Lampiran 5. Kartu Bimbingan.

Lampiran 6. Daftar Riwayat Hidup

DAFTAR GAMBAR DAN ILUSTRASI

Gambar II. 1 Tahapan Knowledge Discovery in Database	19
Gambar III. 1 Tahapan Penelitian	32
Gambar IV. 1 Perbandingan Data	45
2 Grafik Data Berdasarkan Kadar Gula	46
3 Grafik Data Berdasarkan Tekanan Darah	47
4 Grafik Distribusi Berdasarkan Berat Badan	49
5 Grafik Distribusi Berdasarkan Usia Pasien	51
6 <i>Source Code Library</i> dan Pemanggilan File Data	52
7 <i>Output Data</i> Penderita Penyakit Diabetes Format CSV	53
8 Label Data	55
9 Pemanggilan Library <i>Naive Bayes</i>	56
10 Ekstraksi Data untuk Klasifikasi dengan <i>Naive Bayes</i>	57
11 <i>Source Code</i> Pengujian Model	59
12 Hasil Pengujian Model <i>Naive Bayes</i>	60
13 Hasil F1-Score	62

DAFTAR TABEL

Tabel II. 1 Penelitian Terdahulu.....	25
Gambar III. 1 Sampel Data Penelitian	39
Gambar IV. 1 Hasil Persebaran Confission Matrix NB.....	60
2 Hasil Pengujian <i>Naive Bayes</i>	63
3 Hasil Recall dan Precision <i>Naive Bayes</i>	65