

**ANALISIS PENGARUH GAYA HIDUP TERHADAP RISIKO
TERKENA PENYAKIT DIABETES MELITUS TIPE 2
MENGUNAKAN ALGORITMA NAÏVE BAYES**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

ADITYA PUTRA PAMUNGKAS

20.11.3609

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

**ANALISIS PENGARUH GAYA HIDUP TERHADAP RISIKO
TERKENA PENYAKIT DIABETES MELITUS TIPE 2
MENGUNAKAN ALGORITMA NAÏVE BAYES**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

ADITYA PUTRA PAMUNGKAS

20.11.3609

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS PENGARUH GAYA HIDUP TERHADAP RISIKO TERKENA
PENYAKIT DIABETES MELITUS TIPE 2 MENGGUNAKAN
ALGORITMA NAÏVE BAYES**

yang disusun dan diajukan oleh

Aditya Putra Pamungkas

20.11.3609

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 19 Desember 2025

Dosen Pembimbing,



Anna Baita, S.Kom., M.Kom.

NIK. 190302290

HALAMAN PENGESAHAN
SKRIPSI
ANALISIS PENGARUH GAYA HIDUP TERHADAP RISIKO TERKENA
PENYAKIT DIABETES MELITUS TIPE 2 MENGGUNAKAN
ALGORITMA NAÏVE BAYES

yang disusun dan diajukan oleh

Aditya Putra Pamungkas

20.11.3609

Telah dipertahankan di depan Dewan Penguji
pada tanggal 19 Desember 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Haryoko, S.Kom., M.Cs.
NIK. 190302286

Andriyan Dwi Putra, S.Kom., M.Kom.
NIK. 190302270

Anna Baita, S.Kom., M.Kom.
NIK. 190302290

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 19 Desember 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Aditya Putra Pamungkas
NIM : 20.11.3609

Menyatakan bahwa Skripsi dengan judul berikut:

ANALISIS PENGARUH GAYA HIDUP TERHADAP RISIKO TERKENA PENYAKIT DIABETES MELITUS TIPE 2 MENGGUNAKAN ALGORITMA NAÏVE BAYES

Dosen Pembimbing : Anna Baita, S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 19 Desember 2025

Yang Menyatakan,

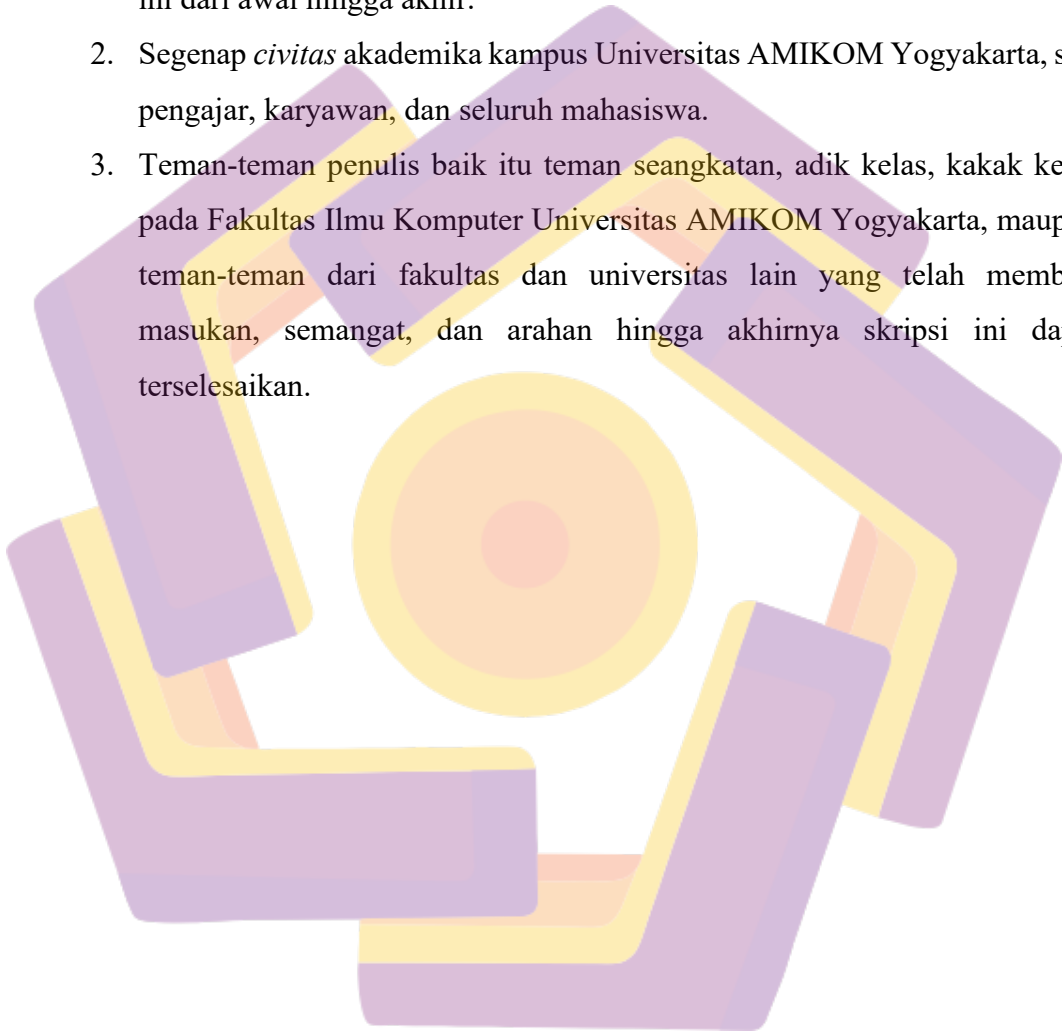


Aditya Putra Pamungkas

HALAMAN PERSEMBAHAN

Dengan rasa syukur yang mendalam kehadiran Allah SWT atas segala rahmat dan karunia-Nya, skripsi ini saya persembahkan kepada :

1. Keluarga besar penulis yang senantiasa membantu menyelesaikan skripsi ini dari awal hingga akhir.
2. Segenap *civitas* akademika kampus Universitas AMIKOM Yogyakarta, staf pengajar, karyawan, dan seluruh mahasiswa.
3. Teman-teman penulis baik itu teman seangkatan, adik kelas, kakak kelas pada Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta, maupun teman-teman dari fakultas dan universitas lain yang telah memberi masukan, semangat, dan arahan hingga akhirnya skripsi ini dapat terselesaikan.



KATA PENGANTAR

Dengan memanjatkan puja dan puji syukur kehadiran Allah SWT yang telah melimpahkan rahmat, taufik, dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan judul **“Analisis Pengaruh Gaya Hidup Terhadap Risiko Terkena Penyakit Diabetes Melitus Tipe 2 Menggunakan Algoritma Naïve Bayes”**, sebagai salah satu syarat untuk menyelesaikan Program Sarjana (S1) Jurusan Informatika Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.

Penulis menyadari bahwa skripsi ini tidak mungkin terselesaikan tanpa adanya dukungan, bantuan, bimbingan, dan nasehat dari berbagai pihak selama penyusunan skripsi ini. Pada kesempatan ini penulis menyampaikan terima kasih kepada :

1. Ibu Prof. Dr. Kusrini, M.Kom., selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
2. Ibu Anna Baita, M.Kom., selaku dosen pembimbing. Atas segala bimbingan, arahan serta saran yang diberikan kepada penulis sehingga skripsi ini dapat diselesaikan dengan baik.
3. Bapak Muhammad Fairul Filza, S.Kom., M.Kom., selaku dosen wali yang telah membantu penulis dalam mengikuti dan menyelesaikan studi di Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
4. Seluruh staff pengajar Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta yang telah memberikan ilmu pengetahuan tak ternilai selama penulis menempuh pendidikan di Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
5. Kedua orang tua penulis, Bapak Budi Heri Siswoko dan Ibu Dwi Kusumawardani, yang selalu memberikan kasih sayang, doa, nasehat, serta atas kesabarannya yang luar biasa dalam setiap langkah hidup penulis, yang merupakan anugerah terbesar dalam hidup. Penulis berharap dapat menjadi anak yang dapat dibanggakan.

6. Dalam perjalanan menyusun skripsi ini, tidak sedikit tantangan, kebingungan, bahkan kelelahan yang penulis hadapi. Namun, di balik proses itu, ada yang memiliki peran besar dan tidak tergantikan dalam keberhasilan penyusunan karya ini yaitu kakak tercinta penulis. Terima kasih yang tak terhingga penulis sampaikan atas segala bentuk dukungan, baik secara moral maupun intelektual yang telah kakak berikan. Kakak bukan hanya menjadi tempat bertanya dan berbagi keluh kesah, tetapi juga menjadi sumber semangat di saat penulis hampir menyerah. Nasihat, motivasi, bahkan waktu dan tenaga yang kakak luangkan demi membantu penulis dalam memahami materi, menyusun ide, dan menyelesaikan setiap bab, adalah hal yang sangat berarti dan takkan terlupakan.
7. Kepada seseorang yang pernah bersama penulis dan tidak bisa penulis sebut namanya. Terimakasih untuk patah hati yang diberikan saat proses penyelesaian skripsi ini. Ternyata perginya anda dari kehidupan penulis memberikan cukup banyak motivasi untuk terus maju dan berproses menjadi pribadi yang mengerti apa itu arti pendewasaan, sabar dan menerima arti kehilangan sebagai bentuk proses penempaan menghadapi dinamika hidup. Terimakasih telah menjadi bagian menyenangkan sekaligus menyakitkan dari pendewasaan ini. Pada akhirnya setiap orang ada masanya dan setiap masa ada orangnya.

Yogyakarta, 17 Desember 2025

Aditya Putra Pamungkas

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH	xiv
INTISARI	xv
<i>ABSTRACT</i>	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	2
1.5 Manfaat Penelitian	2
1.6 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Dasar Teori	13
2.2.1 Gaya Hidup	13
2.2.2 Penyakit Diabetes.....	13
2.2.3 Data Mining	13
2.2.4 Data Selection	14
2.2.5 Data Preprocessing.....	14
2.2.6 Data Transformation	14
2.2.7 SMOTE	14
2.2.8 Naïve Bayes	14
2.2.9 Gaussian Naïve Bayes.....	15
2.2.10 Model Evaluation.....	16
BAB III METODE PENELITIAN	18

3.1 Objek Penelitian.....	18
3.2 Alur Penelitian	18
3.2.1 Data Collecting	18
3.2.2 Data Preprocessing.....	19
3.2.3 Modelling Naïve Bayes.....	19
3.2.4 Evaluation	20
3.3 Alat dan Bahan.....	20
3.3.1 Data Penelitian	20
3.3.2 Alat/instrumen.....	20
BAB IV HASIL DAN PEMBAHASAN	21
4.1 Gambaran Umum Dataset.....	21
4.1.1 Jumlah Sampel dan Dimensi Data	21
4.2 Pra-Pemrosesan Data	24
4.2.1 Standarisasi fitur	24
4.2.2 Data Splitting	25
4.2.3 Penyeimbangan Data (Oversampling dengan SMOTE)	25
4.3 Penerapan Algoritma Naïve Bayes	26
4.3.1 Pemilihan Algoritma	26
4.3.2 Proses Pelatihan	27
4.3.3 Penyimpanan Mode dan Scaler.....	27
4.4 Evaluasi Kinerja Model	27
4.5 Confusion Matrix	28
4.6 Pembahasan	30
BAB V PENUTUP	31
5.1 Kesimpulan	31
5.2 Saran	31
REFERENSI	32
LAMPIRAN.....	35

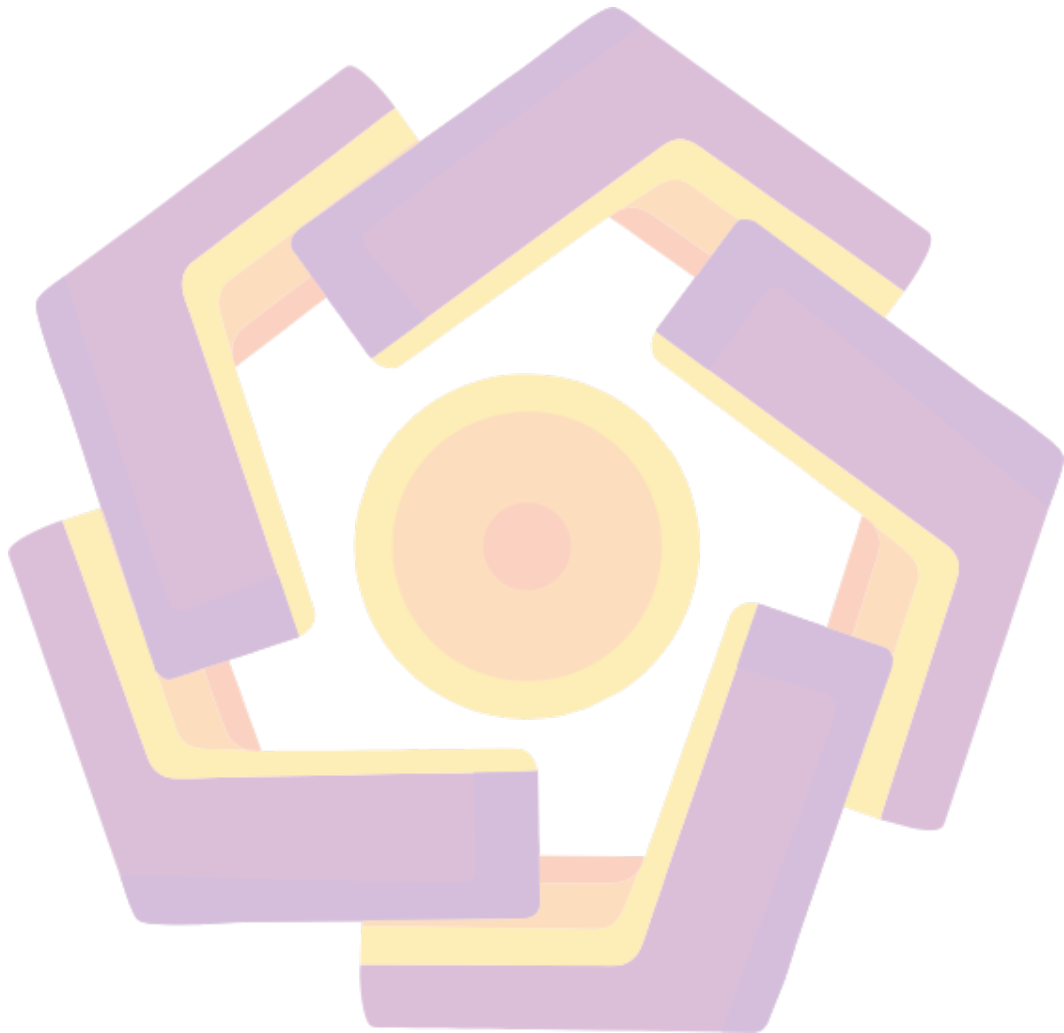
DAFTAR TABEL

Tabel 2.1. Keaslian Penelitian	7
Tabel 4.1. Kelas Target	21
Tabel 4.2. Fitur	22
Tabel 4.3. Data Sebelum Standarisasi	24
Tabel 4.4. Data Sesudah Standarisasi	24
Tabel 4.5. Data Splitting	25
Tabel 4.6. Hasil SMOTE	26
Tabel 4.7. Hasil Evaluasi 20% testing	27
Tabel 4.8. Hasil Evaluasi 30% testing	28
Tabel 4.9. Hasil Evaluasi 40% testing	28



DAFTAR GAMBAR

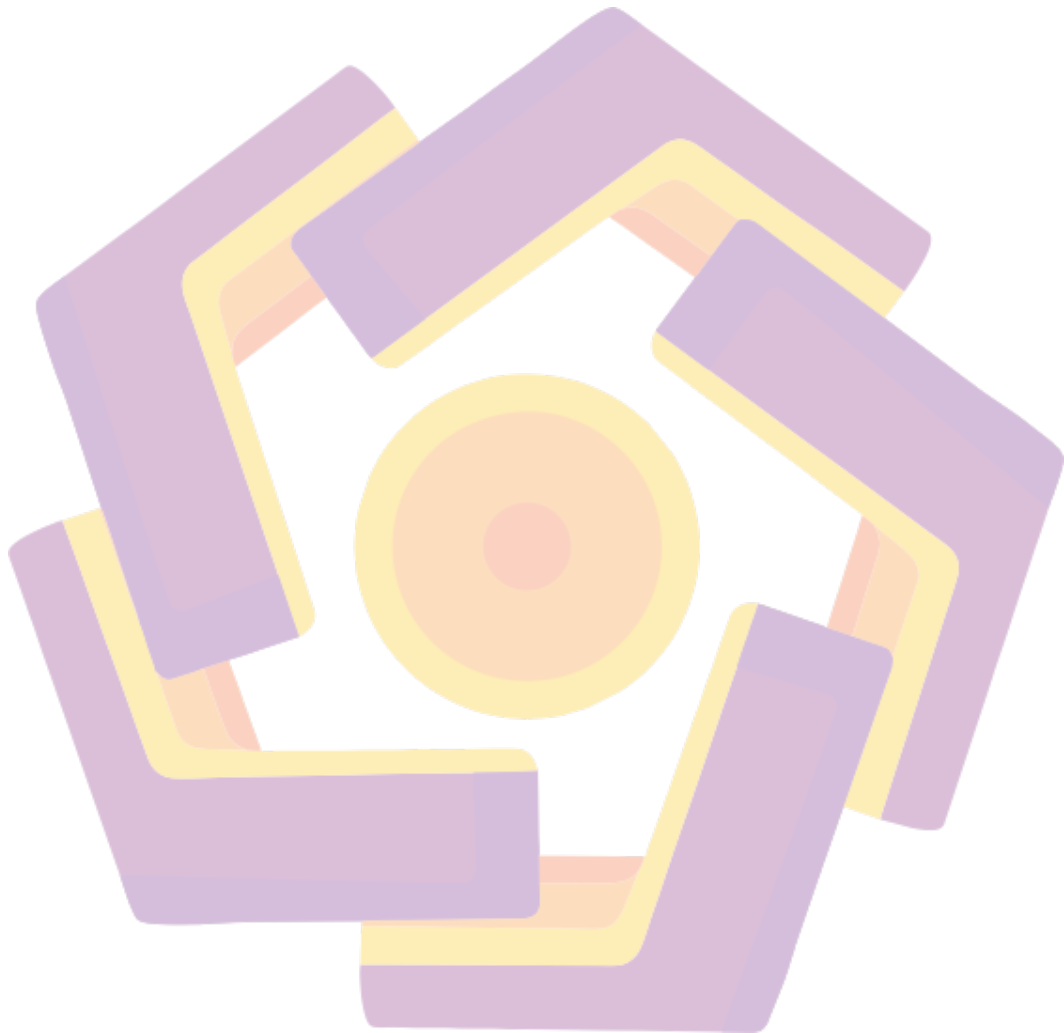
Gambar 3.1. Diagram Alur	18
Gambar 4.1. Confusion Matrix	29



DAFTAR LAMPIRAN

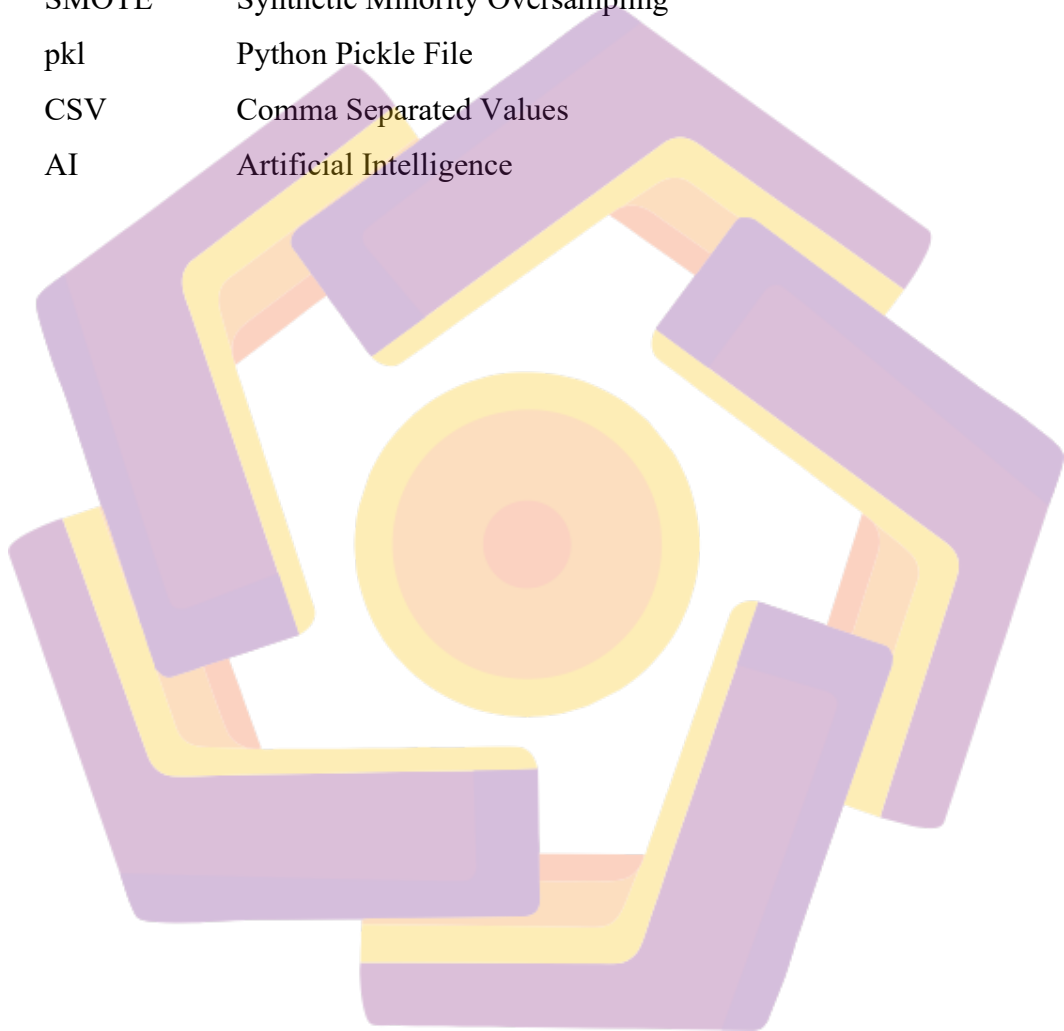
Lampiran 1. Kode pembuatan model

35



DAFTAR LAMBANG DAN SINGKATAN

σ	Standar deviasi
μ	Rata-rata
DM	Diabetes Melitus
SMOTE	Synthetic Minority Oversampling
pkl	Python Pickle File
CSV	Comma Separated Values
AI	Artificial Intelligence



DAFTAR ISTILAH



Diabetes Melitus	penyakit gula atau kencing manis
Mining	proses mencari pola dari sekumpulan data
Artificial Intelligence	teknologi yang dirancang seperti manusia
Selection	data yang dianalisis dari sekumpulan data
Preprocessing	mengubah data mentang menjadi siap olah
Error	adanya ketidaksesuaian dengan hasil
Transformation	mengubah format agar bisa dibaca algoritma
Standard Scaler	menyamakan skala agar seragam
Threshold	nilai minimum agar proses bisa dimulai
Precision	seberapa akurat model mempredik positif
Recall	menemukan kembali data positif
F1-Score	metrik evaluasi dari precision dan recall

INTISARI

DM Tipe 2 merupakan salah satu penyakit tidak menular yang prevalensinya terus meningkat secara global, termasuk di Indonesia. Gaya hidup yang tidak sehat, seperti pola makan tinggi gula dan lemak, kurangnya aktivitas fisik, merokok, serta konsumsi alkohol, menjadi faktor risiko utama yang dapat memicu timbulnya penyakit ini. Penelitian ini bertujuan untuk menganalisis pengaruh gaya hidup terhadap risiko terkena DM Tipe 2 dengan menggunakan algoritma Naïve Bayes sebagai metode klasifikasi. Data yang digunakan dalam penelitian ini diperoleh dari website Kaggle, mencakup variabel gaya hidup seperti frekuensi konsumsi makanan cepat saji, olahraga, kebiasaan merokok, konsumsi alkohol, indeks massa tubuh (IMT), dan riwayat keluarga.

Metode Naïve Bayes dipilih karena kemampuannya dalam mengolah data kategorikal dan menghasilkan prediksi dengan tingkat akurasi yang cukup baik meskipun dengan asumsi independensi antar variabel. Hasil pengujian model menunjukkan bahwa algoritma Naïve Bayes mampu mengklasifikasikan risiko terkena DM Tipe 2 dengan akurasi sebesar 67,93% setelah tiga kali melakukan proses data splitting. Selain itu, variabel gaya hidup seperti aktivitas fisik dan pola makan terbukti memiliki pengaruh paling signifikan terhadap peningkatan risiko penyakit ini.

Penelitian ini menyimpulkan bahwa penerapan metode data mining, khususnya Naïve Bayes, efektif dalam mengidentifikasi individu berisiko tinggi terhadap DM Tipe 2. Hasil ini dapat digunakan sebagai dasar dalam program pencegahan dan edukasi masyarakat mengenai pentingnya gaya hidup sehat.

Kata Kunci: Gaya hidup, Diabetes Melitus Tipe 2, Naïve Bayes, Klasifikasi, Data Mining.

ABSTRACT

DM Type 2 is one of the most prevalent non-communicable diseases globally, with increasing incidence rates, including in Indonesia. An unhealthy lifestyle such as high intake of sugary and fatty foods, lack of physical activity, smoking, and alcohol consumption is a major risk factor contributing to the development of this disease. This study aims to analyze the impact of lifestyle on the risk of developing DM Type 2 using the Naïve Bayes algorithm as a classification method. The data were collected from Kaggle website, covering lifestyle variables such as frequency of fast food consumption, exercise habits, smoking, alcohol intake, body mass index (BMI), and family medical history.

Naïve Bayes was chosen for its effectiveness in processing categorical data and its ability to produce accurate predictions despite the assumption of feature independence. The results showed that the Naïve Bayes algorithm achieved a classification accuracy of 67,93% after using 3 steps data splitting in identifying individuals at risk of DM Type 2. Among all variables, physical activity and dietary patterns had the most significant influence on the risk level.

This study concludes that applying data mining techniques, particularly the Naïve Bayes algorithm, is effective in detecting individuals at high risk of DM Type 2. These findings can serve as a foundation for prevention strategies and public health education to promote healthier lifestyles.

Keywords: *Lifestyle, Type 2 Diabetes Mellitus, Naïve Bayes, Classification, Data Mining.*