

**KLASIFIKASI PENYAKIT DIABETES MELLITUS
MENGUNAKAN ALGORITMA SVM**

SKRIPSI

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



disusun oleh

AHMED IBNU APRIANTO

21.11.4361

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

**KLASIFIKASI PENYAKIT DIABETES MELLITUS
MENGUNAKAN ALGORITMA SVM**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

AHMED IBNU APRIANTO

21.11.4361

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026**

HALAMAN PERSETUJUAN

SKRIPSI

**KLASIFIKASI PENYAKIT DIABETES MELLITUS
MENGUNAKAN ALGORITMA SVM**

yang disusun dan diajukan oleh

Ahmed Ibnu Aprianto

21.11.4361

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 21 Januari 2026

Dosen Pembimbing,

Ike Verawati, M.Kom

NIK. 190302237

HALAMAN PENGESAHAN
SKRIPSI
KLASIFIKASI PENYAKIT DIABETES MELLITUS
MENGUNAKAN ALGORITMA SVM

yang disusun dan diajukan oleh

Ahmed Ibnu Aprianto

21.11.4361

Telah dipertahankan di depan Dewan Penguji
pada tanggal 21 Januari 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Afrig Aminuddin, S.Kom., M.Eng., Ph.D.
NIK. 190302351



Nur Aini, A.Md., S.Kom., M.Kom.
NIK. 190302066



Ike Verawati, S.Kom., M.Kom.
NIK. 190302237



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 29 Januari 2026

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Ahmed Ibnu Aprianto
NIM : 21.11.4361

Menyatakan bahwa Skripsi dengan judul berikut:

Klasifikasi Penyakit Diabetes Menggunakan Algoritma SVM

Dosen Pembimbing : Ike Verawati, 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, 29 Januari 2026

Yang Menyatakan,



Ahmed Ibnu Aprianto

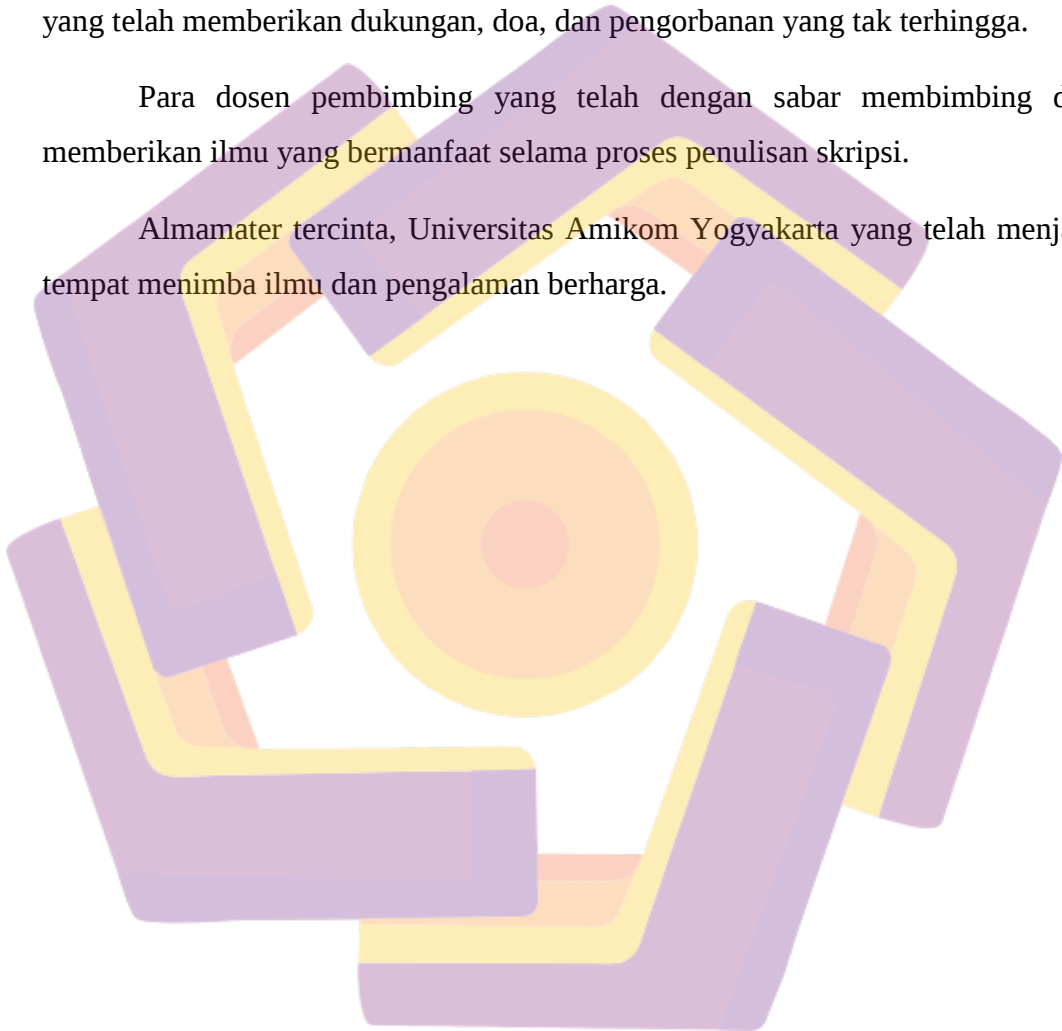
HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur kepada Allah SWT, karya ilmiah ini saya persembahkan kepada:

Kedua orang tua saya tercinta, Bapak Sigit Aryanto dan Ibu Rina Ambarsari yang telah memberikan dukungan, doa, dan pengorbanan yang tak terhingga.

Para dosen pembimbing yang telah dengan sabar membimbing dan memberikan ilmu yang bermanfaat selama proses penulisan skripsi.

Almamater tercinta, Universitas Amikom Yogyakarta yang telah menjadi tempat menimba ilmu dan pengalaman berharga.



KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah Yang Maha Esa atas limpahan rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul "Klasifikasi Penyakit Diabetes Mellitus Tipe I dan Penyakit Diabetes Mellitus Tipe II Menggunakan Algoritma SVM di Puskesmas Polanharjo, Klaten, Jawa Tengah" dengan lancar dan tepat waktu. Skripsi ini disusun guna memenuhi salah satu syarat untuk memperoleh gelar S.Kom pada Program Studi S1 Informatika di Universitas Amikom Yogyakarta.

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Hanafi, S.Kom., M.Eng., Ph.D., selaku dosen pembimbing yang telah memberikan bimbingan, arahan, dan motivasi selama proses penyusunan skripsi ini.
2. Tim Dosen Penguji yang telah memberikan masukan dan saran yang sangat berharga untuk perbaikan skripsi ini.
3. Kedua orang tua yang senantiasa memberikan dukungan moril dan materiil serta doa yang tiada henti.
4. Universitas Amikom Yogyakarta yang telah memberikan kesempatan dan dukungan dalam penelitian ini. Penulis menyadari bahwa skripsi ini masih jauh dari sempurna.

Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun dari berbagai pihak untuk perbaikan di masa mendatang.

Akhir kata, semoga skripsi ini dapat memberikan manfaat bagi pembaca dan pihak-pihak yang berkepentingan.

Yogyakarta, 20 Desember 2026



Ahmed Ibnu Aprianto

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH.....	xiv
INTISARI	xvi
<i>ABSTRACT</i>	xvii
BAB I PENDAHULUAN.....	xvii
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	2
1.4 Tujuan Penelitian.....	2
1.5 Manfaat Penelitian.....	3
1.5.1 Manfaat segi teoritis Dan praktis	3
1.5.2 Manfaat bagi Penelitian Selanjutnya.....	3
1.6 Sistematika Penulisan.....	3

BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Dasar Teori	11
2.2.1 Diabetes Melitus	11
a) Diabetes Melitus Tipe 1	11
b) Diabetes Melitus Tipe 2	12
2.2.2 Google Collab	12
2.2.3 Machine Learning	13
2.2.4 Preprocessing	13
2.2.5 Analisis Korelasi Fitur	13
2.2.6 K- Fold Cross Validation	14
2.2.7 Support Vector Machine (SVM)	15
2.2.8 Confusion Matrix	16
BAB III METODE PENELITIAN	18
3.1 Objek Penelitian	18
3.2 Alur Penelitian	18
a) Studi Literatur	20
b) Pengumpulan Data	20
c) Preprocessing	20
d) Analisis Korelasi Fitur	21
e) Split Data	21
f) Support Vector Machine (SVM)	21
g) Evaluasi Hasil	25
h) Hasil	26
BAB IV HASIL DAN PEMBAHASAN	28

4.1	Pengumpulan Data	28
4.2	Preprocessing	29
a)	Labeling	29
b)	IQR (Interquartile Range).....	29
c)	Splitting Data	30
d)	SMOTE.....	31
4.3	Analisis Korelasi Fitur	32
4.4	Splitting Data.....	33
a)	Splitting 30:70	33
b)	Splitting 20:80	33
c)	Splitting 10:90	34
4.5	SVM	34
4.6	Evaluasi Hasil.....	34
4.7	Hasil	37
BAB V PENUTUP		38
5.1	Kesimpulan.....	38
5.2	Saran.....	38
REFERENSI		39
LAMPIRAN.....		41

DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	8
Tabel 2.2 Persamaan Setiap Kernel SVM	16
Tabel 2.3 Confusion Matrix	17
Tabel 3.1 Perhitungan dataset studi kasus SVM	22
Tabel 3.2 Contoh confusion matrix	25
Tabel 3.3 data sesudah dilakukan penyesuaian	27
Tabel 4.1 visualisasi data akhir	29
Tabel 4.2 Confusion matrix 30:70	35
Tabel 4.3 Confusion matrix 20:80	36
Tabel 4.4 Confusion matrix 10:90	37
Tabel 4.5 perbandingan hasil penelitian	37

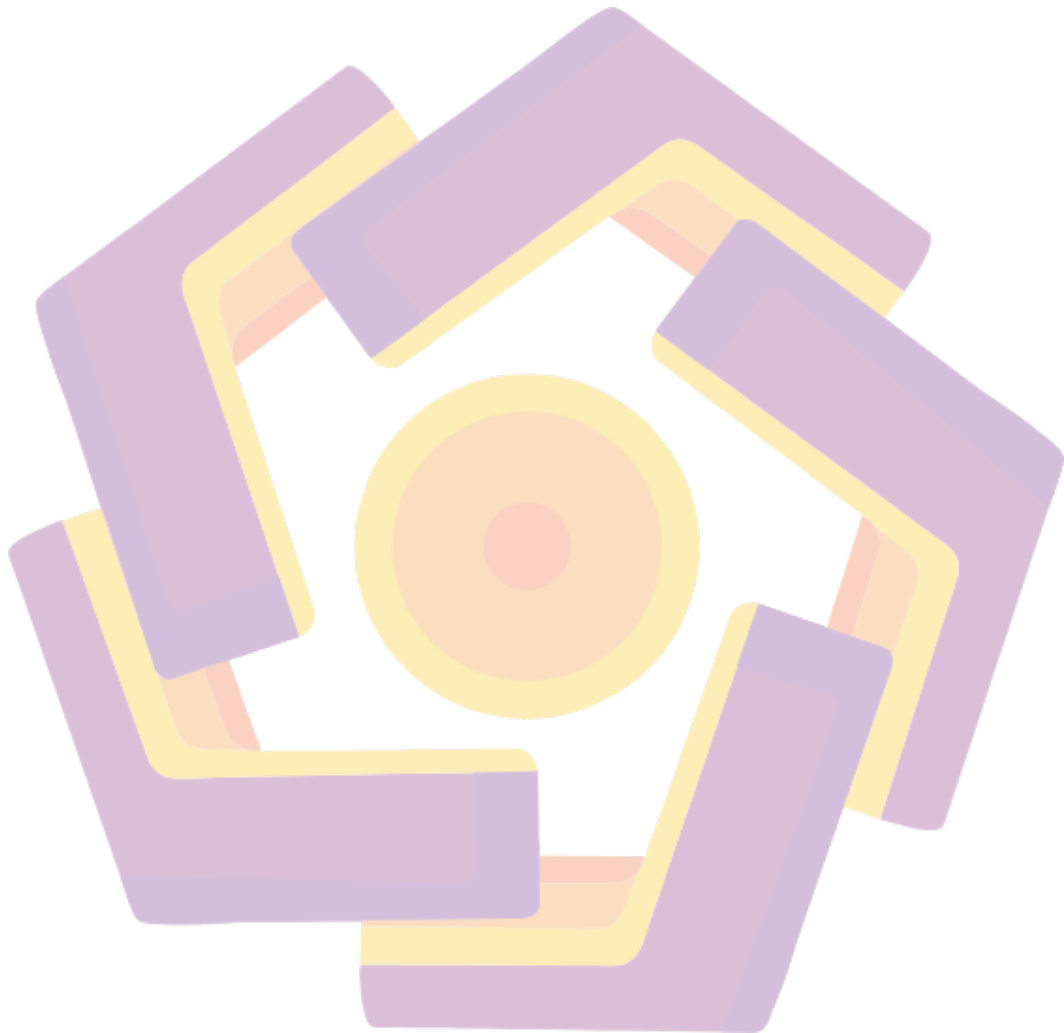
DAFTAR GAMBAR

Gambar 2.1 Contoh ilustrasi k-fold cross validation.	15
Gambar 2.2 Hyperplane Support Vector Machine	15
Gambar 3.1 Alur Penelitian	19
Gambar 3.2 Flowchart SVM	22
Gambar 4.1 Hasil label unik	29
Gambar 4.2 Hasil label encoding	29
Gambar 4.3 Hasil IQR	30
Gambar 4.4 Hasil splitting 30:70	30
Gambar 4.5 Hasil splitting 20:80	30
Gambar 4.6 Hasil splitting 10:90	30
Gambar 4.7 Hasil SMOTE 30:70	31
Gambar 4.8 Hasil SMOTE 20:80	31
Gambar 4.9 Hasil SMOTE 10:90	32
Gambar 4.10 Hasil analisis korelasi fitur	33
Gambar 4.11 Hasil Splitting 30:70	34
Gambar 4.12 Hasil Splitting 20:80	34
Gambar 4.13 Hasil Splitting 10:90	34
Gambar 4.14 Hasil modeling SVM	35
Gambar 4.15 Evaluasi Hasil 30:70	36
Gambar 4.16 Evaluasi Hasil 20:80	36
Gambar 4.17 Evaluasi Hasil 10:90	37

DAFTAR LAMPIRAN

Lampiran 1. Source Code :

42



DAFTAR LAMBANG DAN SINGKATAN

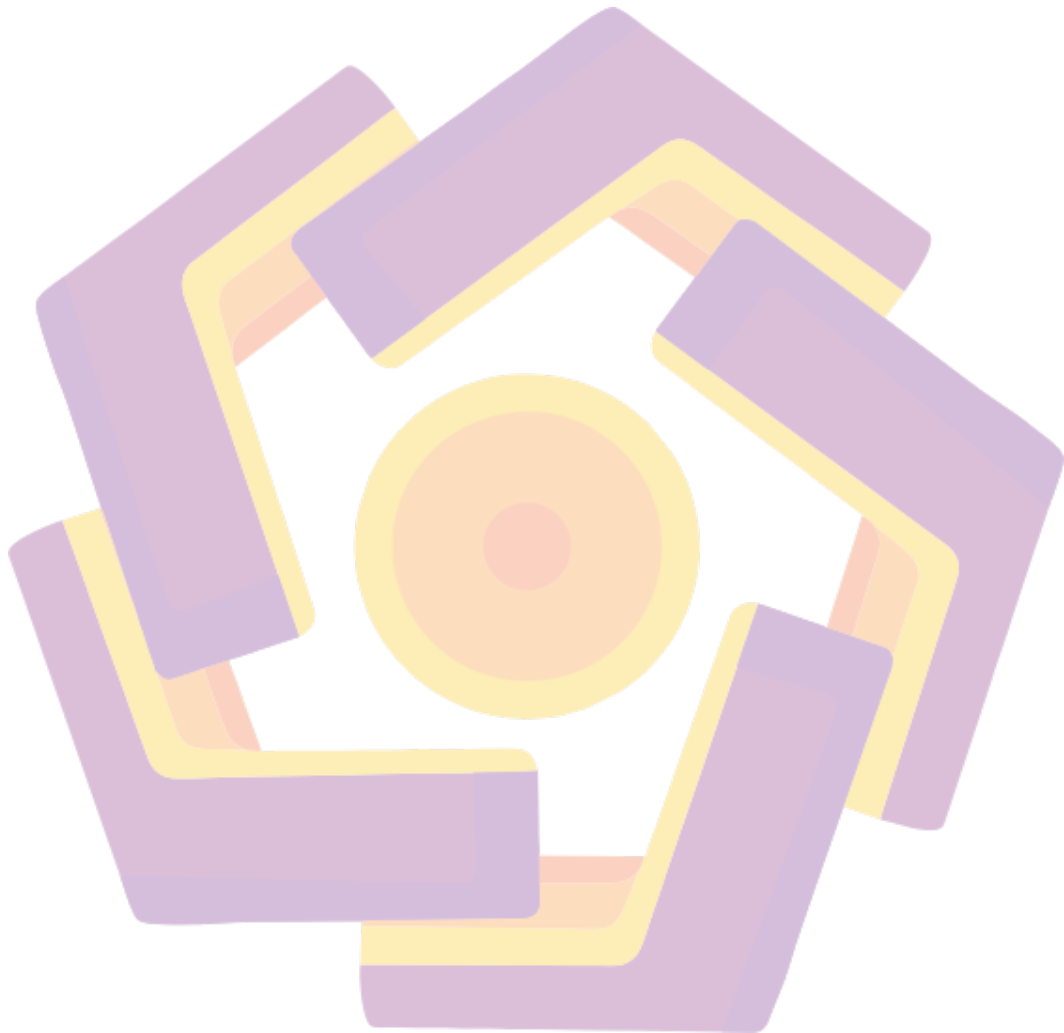


SVM	Support Vector Machines
SMOTE	Synthetic Minority Over-sampling Technique
IQR	Interquartile Range
NLP	Natural Language Processing
CNN	Convolutional Neural Network
EDA	Exploratory Data Analysis
GLCM	Gray Level Co-occurrence Matrix
RBF	Radial Basis Function
TB	Tinggi Badan
BB	Berat Badan
TP	True Positive
FP	False Positive
FN	False Negative
TN	True Negative
MAPE	Mean Absolute Percentage Error
RSME	Root Mean Squared Error
DM	Diabetes Melitus
TPU	Tensor Processing Unit
ML	Machine Learning
SRM	Structural Risk Minimization
RAM	Random Access Memory
α	Alpha
w	Weight
b	Bias

DAFTAR ISTILAH

Confusion Matrix	Tabel evaluasi performa classifier berbasis true/false positif dan negatif
Kernel	Fungsi dalam SVM untuk memetakan data ke dimensi yang lebih tinggi
Linear Kernel	Jenis kernel dalam SVM yang paling sederhana (garis lurus)
Cross Validation	Teknik evaluasi model, membagi data menjadi beberapa subset
Percentage Split	Metode membagi data menjadi data latih dan uji berdasarkan persentase tertentu
Train-Test Split	Membagi data latih dan uji secara acak berdasarkan rasio tertentu
Preprocessing	Teknik penyiapan data sebelum pemodelan
Label Encoding	Proses mengubah label kategorikal menjadi numerik
Standardization	Teknik menskalakan fitur agar memiliki mean 0 dan deviasi standar 1
Overfitting	Ketika model terlalu cocok dengan data latih, tetapi gagal pada data uji
Underfitting	Ketika model terlalu sederhana sehingga gagal menangkap pola pada data
Feature Scaling	Transformasi nilai fitur untuk membantu algoritma bekerja optimal
Outlier	Nilai yang jauh dari pola umum lainnya
Interquartile Range	Teknik untuk mendeteksi outlier dengan melihat rentang kuartil
Learning Model	Model yang dibangun menggunakan algoritma pembelajaran mesin

K-Fold	Metode cross validation dengan pembagian data menjadi k bagian
Bias	Selisih antara prediksi model dan nilai sebenarnya
Variance	Tingkat perubahan prediksi jika data latih berubah sedikit



INTISARI

Diabetes Mellitus (DM) merupakan salah satu penyakit kronis dengan tingkat prevalensi yang terus meningkat di dunia, termasuk di Indonesia. Penyakit ini terbagi menjadi dua tipe utama, yaitu DM tipe 1 dan tipe 2, yang masing-masing membutuhkan penanganan yang berbeda. Oleh karena itu, klasifikasi yang akurat antara kedua tipe DM sangat penting untuk mendukung diagnosis dini dan intervensi medis yang tepat.

Penelitian ini bertujuan untuk membangun model klasifikasi penyakit Diabetes Mellitus menggunakan algoritma Support Vector Machine (SVM). Dataset penelitian terdiri dari sejumlah atribut klinis pasien yang telah melalui tahap preprocessing, termasuk penanganan missing values, penghilangan outlier menggunakan Interquartile Range (IQR), normalisasi dengan StandardScaler, dan konversi label menggunakan Label Encoding. Karena jumlah data antar kelas tidak seimbang, diterapkan metode Synthetic Minority Over-sampling Technique (SMOTE) untuk mengatasi masalah ketidakseimbangan kelas.

Evaluasi model dilakukan menggunakan metrik akurasi, precision, recall, F1-score, dan confusion matrix. Hasil evaluasi menunjukkan bahwa model SVM dengan kernel linier memberikan akurasi cukup, yaitu sebesar 73%, dengan nilai F1-score seimbang untuk kedua kelas. Hal ini menunjukkan bahwa model mampu melakukan klasifikasi dengan performa yang baik dan tanpa bias terhadap salah satu kelas.

Kata Kunci: Diabetes Mellitus, Support Vector Machine, SMOTE, Preprocessing Data, Klasifikasi, Machine Learning.

ABSTRACT

Diabetes Mellitus (DM) is a chronic disease with an increasing global prevalence, including in Indonesia. It is categorized into two main types: type 1 and type 2 DM, each requiring different medical interventions. Therefore, accurate classification between these two types is crucial for early diagnosis and effective treatment.

This study aims to develop a classification model for Diabetes Mellitus using the Support Vector Machine (SVM) algorithm. The dataset used consists of several clinical attributes of patients and has undergone preprocessing steps such as handling missing values, removing outliers using the Interquartile Range (IQR), data normalization using StandardScaler, and label encoding. Due to class imbalance in the dataset, the Synthetic Minority Over-sampling Technique (SMOTE) was applied.

The model was evaluated using metrics such as accuracy, precision, recall, F1-score, and the confusion matrix. The experimental results show that the SVM model with a linear kernel achieved a middle accuracy of 73%, along with balanced F1-scores between both classes. These results indicate that the SVM model can accurately classify DM types without bias towards any particular class.

Keywords: Diabetes Mellitus, Support Vector Machine, SMOTE, Data Preprocessing, Classification, Machine Learning.