



**IMPLEMENTASI SVM DAN NAÏVE BAYES DALAM  
MENDETEKSI PENYAKIT DIABETES**

**SKRIPSI**

Untuk memenuhi persyaratan  
melakukan penelitian dalam rangka penyusunan skripsi

**Oleh**

**GUSTI MAS NUR IRFANSYAH**

**NIM 1811016310005**

**PROGRAM STUDI S-1 ILMU KOMPUTER  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMBUNG MANGKURAT  
BANJARBARU  
JULI 2025**

**SKRIPSI**

**IMPLEMENTASI SVM DAN NAÏVE BAYES DALAM  
MENDETEKSI PENYAKIT DIABETES**

Oleh :

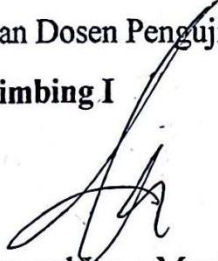
**GUSTI MAS NUR IRFANSYAH**

**NIM 1811016310005**

Telah dipertahankan di depan Dosen Penguji pada tanggal 3 Juli 2025.

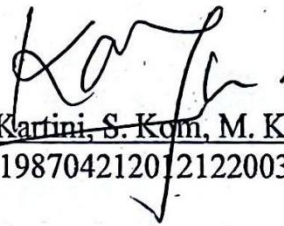
Susunan Dosen Penguji:

**Pembimbing I**



Muhammad Itqan Mazdadi, S.Kom, M.Kom.  
NIP. 199006122019031013

**Dosen Penguji I**



Dwi Kartini, S. Kom, M. Kom.  
NIP. 198704212012122003

**Pembimbing II**



Rudy Herteno, S.Kom, M.Kom.  
NIP. 198809252022031003

**Dosen Penguji II**



Fatma Indriani, S.T, M.I.T, Ph.D  
NIP. 198404202008122004

Banjarbaru, 03 Juli 2025

**Koordinator Program Studi Ilmu Komputer**



Dwi Kartini, S. Kom, M. Kom.  
NIP. 198704212012122003

## **PERNYATAAN**

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan disuatu Perguruan Tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Pustaka.

Banjarbaru, 5 Juli 2025



Gusti Mas Nur Irfansyah  
NIM 1811016310005

## ABSTRAK

### **IMPLEMENTASI SVM DAN NAÏVE BAYES DALAM MENDETEKSI PENYAKIT DIABETES**

(Oleh: Gusti Mas Nur Irfansyah; Pembimbing: Muhammad Itqan Mazdadi, S.Kom, M.Kom. dan Rudy Herteno, S.Kom, M.Kom.; halaman).

Penyakit diabetes merupakan salah satu penyakit kronis yang terus meningkat setiap tahunnya dan membutuhkan penanganan serta deteksi dini untuk meminimalisir komplikasi. Penelitian ini bertujuan untuk membandingkan dua algoritma machine learning, yaitu Support Vector Machine (SVM) dan Naïve Bayes, dalam mendeteksi penyakit diabetes menggunakan dataset Pima Indians Diabetes. Proses preprocessing melibatkan imputasi missing value, balancing data dengan SMOTE, dan normalisasi data menggunakan StandardScaler. Pengujian dilakukan dengan berbagai rasio pembagian data (90:10, 80:20, 75:25, 70:30) serta validasi 10-fold cross validation. Hasil evaluasi menunjukkan bahwa metode SVM dengan kernel RBF memiliki performa terbaik dengan akurasi mencapai 82.25%, diikuti oleh SVM linear dan Naïve Bayes. SVM RBF unggul dalam mengenali pola non-linear dalam data medis, sementara Naïve Bayes tetap menjadi baseline yang stabil meski memiliki keterbatasan dalam asumsi independensi antar fitur. Penelitian ini menunjukkan bahwa pemilihan algoritma yang tepat berpengaruh besar terhadap keberhasilan deteksi dini penyakit diabetes.

Kata Kunci: Diabetes, *Support Vector Machine*, *Naïve Bayes*, *Machine Learning*, Klasifikasi, SMOTE, *Cross Validation*

## **ABSTRACT**

### **IMPLEMENTATION OF SVM AND NAÏVE BAYES IN DETECTING DIABETES**

*(By: Gusti Mas Nur Irfansyah; Supervisors: Muhammad Itqan Mazdadi, S.Kom., M.Kom. and Rudy Herteno, S.Kom., M.Kom.; pages)*

*Diabetes is a chronic disease that continues to increase every year and requires early detection to minimize complications. This study aims to compare two machine learning algorithms, namely Support Vector Machine (SVM) and Naïve Bayes, in detecting diabetes using the Pima Indians Diabetes dataset. The preprocessing process involved imputing missing values, balancing data with SMOTE, and normalizing data using StandardScaler. Testing was conducted using various data split ratios (90:10, 80:20, 75:25, 70:30) and 10-fold cross-validation. Evaluation results show that the SVM method with RBF kernel achieved the best performance with an accuracy of 82.25%, followed by linear SVM and Naïve Bayes. RBF SVM excelled in capturing non-linear patterns in medical data, while Naïve Bayes remained a stable baseline despite limitations in the assumption of feature independence. This research demonstrates that the choice of algorithm significantly impacts the success of early diabetes detection.*

*Keywords: Diabetes, Support Vector Machine, Naïve Bayes, Machine Learning, Classification, SMOTE, Cross Validation*

## PRAKATA

Assalamualaikum Warahmatullahi Wabarakatuh

Puji syukur penulis panjatkan ke hadirat Allah SWT karena atas berkat rahmat dan karunia-Nya penulis dapat menyelesaikan skripsi yang berjudul “Implementasi Support Vector Machine Dan Naive Bayes Dalam Mendeteksi Penyakit Diabetes” untuk memenuhi syarat dalam menyelesaikan Pendidikan S1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat. Tak lupa penulis panjatkan sholawat dan salam kepada Nabi Muhammad SAW, keluarga, para sahabat dan pengikut beliau hingga yaumul qiyamah.

Pada lembar ini penulis ingin menyampaikan ucapan terimakasih kepada pihak- pihak yang sangat mendukung penulis dalam pembuatan dan penyusunan skripsi ini, adapun yang dimaksud adalah sebagai berikut :

1. Keluarga terutama orang tua yang selalu memberikan bantuan, semangat, doa, dan dukungan dalam proses menyelesaikan skripsi ini.
2. Bapak Muhammad Itqan Mazdadi, S.Kom., M.Kom. selaku dosen pembimbing utama yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian skripsi ini.
3. Bapak Rudy Herteno, S.Kom., M.Kom. selaku dosen pembimbing pendamping dan dosen pembimbing akademik yang turut membantu demi kelancaran dalam penyelesaian skripsi ini.
4. Ibu Dwi Kartini, S.Kom., M.Kom. selaku Ketua Program Studi Ilmu Komputer FMIPA ULM dan dosen penguji yang telah memberikan masukan- masukan dalam skripsi ini, atas bantuan dan izin beliau skripsi ini dapat diselesaikan.
5. Ibu Fatma Indriani, S.T., M.I.T., Ph.D selaku dosen penguji yang telah memberikan masukan-masukkan dalam skripsi ini dan seluruh Dosen dan

staf Program Studi Ilmu Komputer FMIPA ULM atas ilmu dan bantuan yang diberikan selama ini yang sangat bermanfaat.

6. Teman-teman dan kerabat khususnya teman-teman Ilmu Komputer 2018, yang sudah ikut serta dalam memberikan dukungan, memberikan doa serta membantu mengerjakan penelitian ini.
7. Semua pihak yang tidak dapat disebutkan satu persatu yang telah turut membantu dalam menyelesaikan skripsi ini.

Penulis berharap skripsi ini dapat bermanfaat memberikan sumbangan pikiran dan ilmu pengetahuan baik bagi penulis sendiri maupun bagi pembaca. Penulis menyadari bahwa skripsi ini jauh dari kata sempurna, oleh karena itu penulis mengharapkan kritik dan saran yang membangun guna kesempurnaan skripsi ini untuk perbaikan dikemudian hari.

Banjarbaru, 15 Juli 2025



Gusti Mas Nur Irfansyah  
NIM. 1811016310005

## DAFTAR ISI

	<b>Halaman</b>
<b>HALAMAN JUDUL</b> .....	<b>i</b>
<b>HALAMAN PENGESAHAN</b> .....	<b>ii</b>
<b>PERNYATAAN</b> .....	<b>iii</b>
<b>ABSTRAK</b> .....	<b>iv</b>
<b>ABSTRACT</b> .....	<b>v</b>
<b>PRAKATA</b> .....	<b>vi</b>
<b>DAFTAR ISI</b> .....	<b>viii</b>
<b>DAFTAR TABEL</b> .....	<b>x</b>
<b>DAFTAR GAMBAR</b> .....	<b>xi</b>
<b>DAFTAR LAMPIRAN</b> .....	<b>xii</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Perumusan Masalah .....	2
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian.....	3
1.5 Batasan Masalah .....	3
<b>BAB II TINJAUAN PUSTAKA</b> .....	<b>4</b>
2.1 Penyakit Diabetes .....	4
2.2 K-Fold Cross Validation .....	7
2.3 StandardScaler.....	7
2.4 Support Vector Machine (SVM) .....	8
2.5 Machine Learning .....	11
2.6 Naive Bayes Classifier .....	13
2.7 Penelitian Terdahulu .....	13
<b>BAB III METODE PENELITIAN</b> .....	<b>17</b>
3.1 Pengumpulan Data.....	17
3.2 Tahapan Penelitian .....	17
3.3 Desain Sistem.....	21
3.3.1 <i>Split Data</i> .....	21

3.3.2 Missing Value.....	21
3.3.3 Balancing Data .....	22
3.3.4 Scaling Data .....	23
3.3.5 StandardScaler.....	23
3.3.6 K-Fold Cross Validation .....	24
3.3.7 Implementasi SVM.....	24
3.3.8 Implementasi Naive Bayes.....	25
3.4 Skenario Uji Coba .....	27
<b>BAB IV PEMBAHASAN.....</b>	<b>30</b>
4.1 Preprocessing .....	30
4.1.1 Missing Value.....	30
4.1.2 Balancing Data .....	32
4.1.3 Scaling Data .....	33
4.2 Split Data .....	33
4.2.1 Pengaruh Rasio Terhadap Model .....	35
4.2.2 Implementasi Metode .....	35
4.2.3 Pemilihan Rasio Akhir.....	37
4.3 Hasil Uji Coba.....	38
4.4 Pembahasan.....	40
4.4.1 Perbandingan Metode Klarifikasi.....	40
4.4.2 Pengaruh Rasio Split Data .....	41
4.4.3 Implikasi Pendidikan .....	42
<b>BAB V KESIMPULAN.....</b>	<b>43</b>
5.1 Kesimpulan .....	43
5.2 Saran .....	44
<b>DAFTAR PUSTAKA.....</b>	<b>45</b>
<b>LAMPIRAN</b>	

## DAFTAR TABEL

Tabel 3.1 Nilai fitur <i>glucose</i> sebelum standarisasi menggunakan standarscaler.	23
Tabel 3.2 Hasil standarisasi fitur <i>glucose</i> menggunakan standarscaler .....	24
Tabel 3.3 Skema Uji Coba .....	28
Tabel 4.1 Jumlah Nilai Nol .....	30
Tabel 4.2 Contoh data sebelum dan sesudah imputasi missing value .....	31
Tabel 4.3 Jumlah data sebelum dan sesudah proses SMOTE .....	32
Tabel 4.4 Contoh data setelah proses SMOTE.....	32
Tabel 4.5 Contoh data sebelum dan sesudah scalling.....	34
Tabel 4.6 Perbandingan jumlah data berdasarkan rasio split .....	35
Tabel 4.7 Hasil Evaluasi Naïve Bayes .....	38
Tabel 4.8 Hasil Pengujian Metode SVM .....	38
Tabel 4.9 Hasil Evaluasi SVM RBF .....	39
Tabel 4.10 Rekapitulasi Perform Akurasi .....	41

## DAFTAR GAMBAR

Gambar 3.1 Tahapan Penelitian.....	17
Gambar 3.2 Missing Value Pada Data .....	22
Gambar 3.3 Implementasi Naive Bayes .....	26
Gambar 4.1 Proses Imputasi Missing Value .....	31
Gambar 4.2 Proses SMOTE .....	33
Gambar 4.3 Proses Scalling.....	34
Gambar 4.4 Persiapan Data .....	36
Gambar 4.5 Preprocessing Data .....	36
Gambar 4.6 Split Data dan Pelatihan Model.....	36
Gambar 4.7 Evaluasi Model.....	37
Gambar 4.8 Kode Python.....	37
Gambar 4.9 Perbandingann Akurasi Model.....	40

## DAFTAR LAMPIRAN

ABSTRAK .....	i
PRAKATA .....	vi
DAFTAR ISI .....	viii
DAFTAR TABEL .....	x
DAFTAR GAMBAR .....	xi
DAFTAR LAMPIRAN .....	xii