

**SISTEM PENUNJANG KEPUTUSAN PREDIKSI PENYAKIT
PARKINSON MENGGUNAKAN SUPPORT VECTOR MA-
CHINE**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Sistem Informasi



disusun oleh

DIEZ DWI YUNITA

21.12.1924

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**SISTEM PENUNJANG KEPUTUSAN PREDIKSI PENYAKIT
PARKINSON MENGGUNAKAN SVM**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Sistem Informasi



disusun oleh

DIEZ DWI YUNITA

21.12.1924

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

HALAMAN PERSETUJUAN

SKRIPSI

**SISTEM PENUNJANG KEPUTUSAN PREDIKSI PENYAKIT
PARKINSON MENGGUNAKAN SUPPORT VECTOR MACHINE**

yang disusun dan diajukan oleh

Diez Dwi Yunita

21.12.1924

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 24 Juli 2025

Dosen Pembimbing,



Krisnawati, S.Si., M.T.
NIK. 1903002038

HALAMAN PENGESAHAN

SKRIPSI

**SISTEM PENUNJANG KEPUTUSAN PREDIKSI PENYAKIT
PARKINSON MENGGUNAKAN SUPPORT VECTOR MACHINE**

yang disusun dan diajukan oleh

Diez Dwi Yunita

21.12.1924

Telah dipertahankan di depan Dewan Penguji
pada tanggal 24 Juli 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Acihmah Sidauruk, S.Kom., M.Kom
NIK. 190302238



Ikmah, S.Kom., M.Kom
NIK. 190302282



Krisnawati, S.Si., M.T
NIK. 190302038



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 24 Juli 2025

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Diez Dwi Yunita
NIM : 21.12.1924

Menyatakan bahwa Skripsi dengan judul berikut:

Sistem Penunjang Keputusan Prediksi Penyakit Parkinson Menggunakan Support Vector Machine

Dosen Pembimbing : Krisnawati, S.S.i, M.T.

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, 24 Juli 2025

Yang Menyatakan,



METERAI TEMPEL
CBAMX425393044

Diez Dwi Yunita

HALAMAN PERSEMBAHAN

Puji Syukur saya panjatkan kepada kehadiran Allah SWT. Atas berkah, Rahmat dan hidayahnya, sehingga saya masih bisa diberikan kesempatan untuk menyelesaikan skripsi ini. Karya tulis ini saya persembahkan untuk orang yang sangat berarti di hidup saya :

1. Orang tua saya, ibu Fahriani dan bapak Iswiyanto yang telah memberikan kasih sayang, doa dan dukungan tiada henti selama proses pendidikan saya. Segala pengorbanan, semangat, dan kepercayaan yang telah bapak dan ibu berikan pada saya, menjadi kekuatan saya hingga saya bisa menyelesaikan pendidikan ini.
2. Adik saya Fahra Tri Septini dan kakak saya Aan Fachrianto Ramadan, yang selalu menjadi sumber semangat, inspirasi dan rasa syukur dalam hidup saya.
3. Ibu Krisnawati, S.Si., M.T. selaku dosen pembimbing, yang telah memberikan bimbingan, dukungan dan arahan saya dalam menyelesaikan karya tulis ini.
4. Teman teman yang selalu menghibur dan menginspirasi saya. Kebersamaan dan dukungan kalian telah menjadi bagian penting yang menemani proses saya hingga titik ini. Kepada Ayu, Sindi, Riya, Yane, Mita, Sayidah, Dhea, Dina, kak Lovina, Isma, kiara, Fanny.

KATA PENGANTAR

Puji Syukur penulis panjatkan ke hadirat Allah SWT, karena atas limpahan Rahmat, Karunia, dan Petunjuk-Nya, penulis dapat menyelesaikan skripsi dengan judul “Sistem Penunjang Keputusan Prediksi Penyakit Parkinson Menggunakan Support Vector Machine” sebagai salah satu syarat untuk memperoleh gelar Sarjana.

Penyusunan skripsi ini bukanlah proses yang mudah, namun berkat bimbingan, semangat dan dukungan dari berbagai pihak, penulis dapat menyelesaikannya dengan baik. Oleh karena itu, dengan segala kerendahan hati, penulis ingin menyampaikan rasa terima kasih kepada :

1. Prof. Dr. M. Suyanto, M.M selaku Rektor Universitas Amikom Yogyakarta
2. Prof. Dr. Kusriani, M.Kom. selaku Dekan Program Fakultas Ilmu Komputer
3. Bapak Anggit Dwi Hartanto, M.Kom. selaku ketua Program Studi Sistem Informasi
4. Ibu Krisnawati, S.Si., M.T. selaku Dosen Pembimbing yang telah memberikan semangat, dukungan serta arahan dalam menyelesaikan karya tulis ini.
5. Ibu Acihmah Sidauruk, S.kom., M.Kom dan ibu Ikma, S.Kom., M.Kom. selaku dosen penguji yang telah memberikan kesempatan, waktu, serta masukan yang sangat berarti demi perbaikan dan penyempurnaan karya tulis ini.
6. Kedua orang tua, keluarga besar, teman teman tercinta yang telah memberikan semangat dan doa.
7. Serta kepada semua pihak yang telah memberikan bantuan, dukungan dan kontribusinya dalam penyusunan tugas akhir ini, baik secara langsung maupun tidak langsung.

Yogyakarta, 28 Juli 2025

Diez Dwi Yunita

DAFTAR ISI

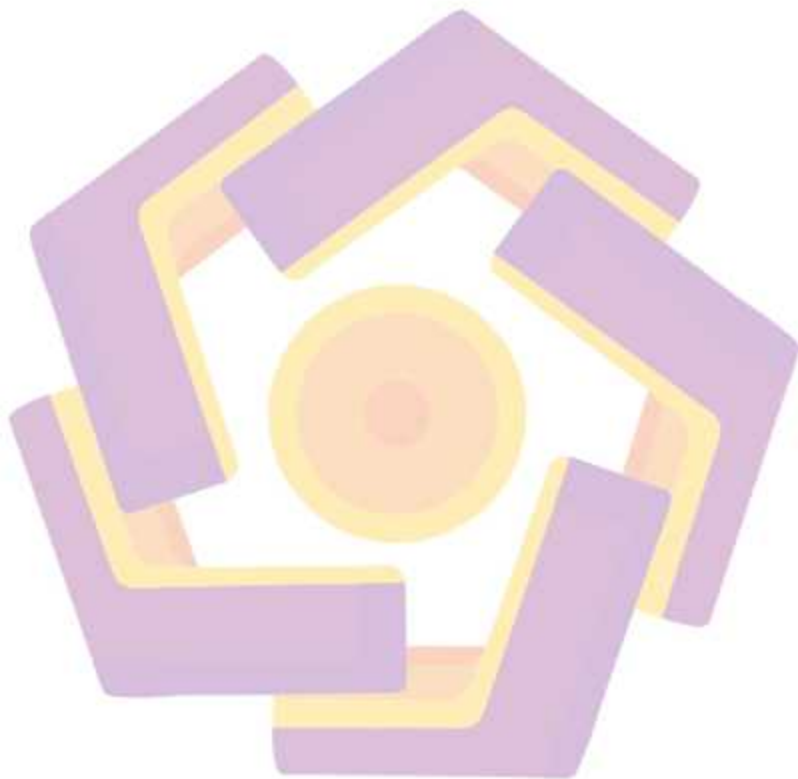
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR.....	vi
DAFTAR ISI.....	vii
DAFTAR TABEL	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH	xvi
INTISARI	xlx
<i>ABSTRACT</i>	xx
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Dasar Teori	32

2.2.1	Penyakit parkinson.....	32
2.2.2	Sistem Penunjang Keputusan.....	33
2.2.3	Data Mining	34
2.2.4	Preprocessing Data.....	35
2.2.5	Data Transformasi.....	36
2.2.6	Synthetic Minority Over Sampling Technique (SMOTE).....	36
2.2.7	Machine Learning	37
2.2.8	Support Vector Machine (SVM).....	38
2.2.9	Evaluasi Model	40
2.2.9.1	Confusion matrix	40
2.2.9.2	Classification Report	41
2.2.9.3	Receiver Operating Characteristic (ROC)	42
BAB III METODE PENELITIAN		43
3.1	Objek Penelitian.....	43
3.2	Alur Penelitian	43
3.3	Alat dan Bahan.....	46
BAB IV HASIL DAN PEMBAHASAN		48
4.1	Pengumpulan Data.....	48
4.2.	Eksplorasi Data Awal	52
4.2.1	Eksplorasi singkat	52
4.2.2	Visualisasi Data Berdasarkan Golongan.....	53
4.2.3	Distribusi kelas.....	55
4.3	Pemrosesan Data (Preprocessing)	57
4.3.1	Pemisahan fitur dan target.....	57
4.3.2	Split data dengan stratify	58
4.3.3	Penanganan Ketidakseimbangan Data	59

4.3.4	Mengecek distribusi kelas setelah SMOTE	60
4.3.5	Standarisasi Data	61
4.3.6	Transformasi data latih dan data uji	62
4.4	Pelatihan Model SVM	63
4.5	Evaluasi Model	64
4.5.1	Prediksi Data Uji	64
4.5.2	Evaluasi dengan confusion Matrix dan Classification Report Thershold 0,5	65
4.5.3	Evaluasi menggunakan ROC Curve dan AUC Score	69
4.5.3.1	Perhitungan AUC Score	69
4.5.3.2	Visualisasi ROC Curve	71
4.5.4	Evaluasi Confusion Matrix dan Classification Report Threshold 0,3 72	72
4.6	Penyimpanan Model	76
4.7	Implementasi Sistem	77
4.7.1	Integrasi Model ke dalam sistem	78
4.7.2	Fitur Utama sistem	78
4.7.3	Proses Prediksi	80
4.7.4	Visualisasi Hasil Prediksi	81
BAB V	PENUTUP	82
5.1	Kesimpulan	82
5.2	Saran	83
REFERENSI		84
LAMPIRAN		88

DAFTAR TABEL

Table 2.1 Keaslian Penelitian	14
Table 2.2 Confusion Matrix	40
Table 4.1 Tampilan 40 sampel dari dataset Parkinson.....	51
Table 4.2 Perbandingan Threshold 0,5 dan 0,3	75

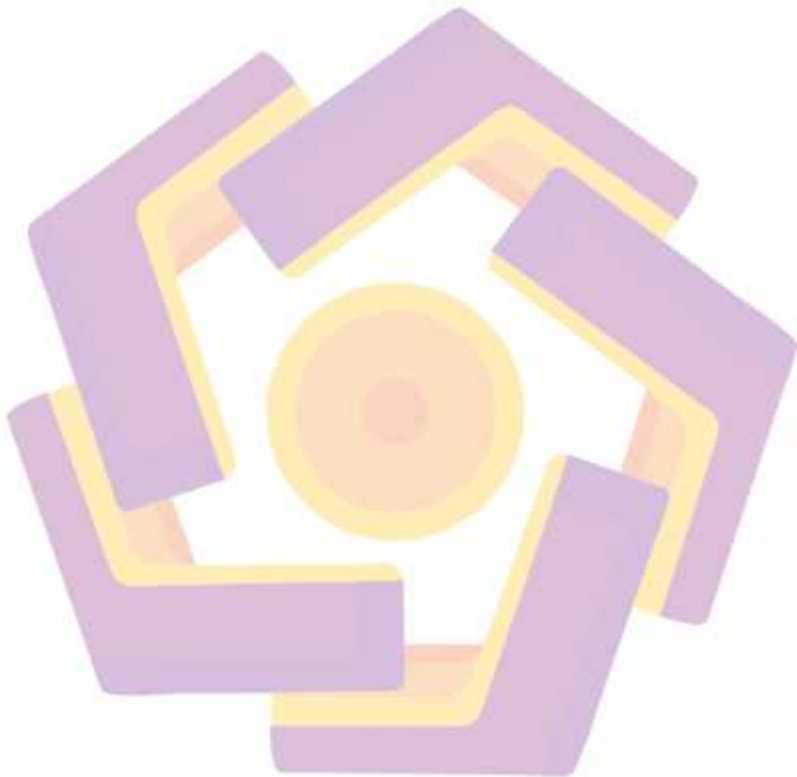


DAFTAR GAMBAR

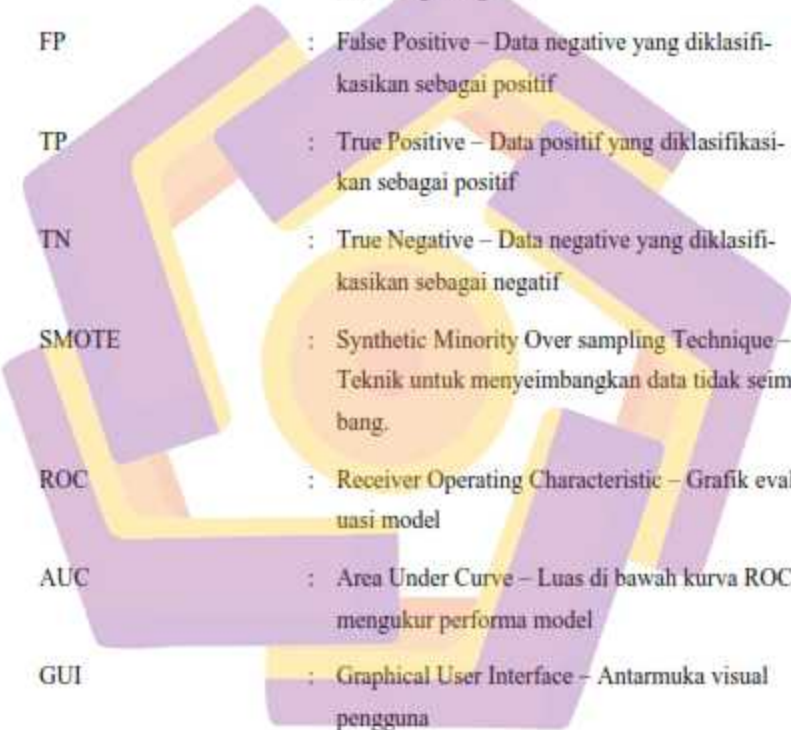
Gambar 3.1 Alur penelitian.....	43
Gambar 4.1 Cuplikan Lima Baris Pertama Dataset	52
Gambar 4.2 Visualisasi Data Pergolongan	54
Gambar 4.3 Distribusi Status Dataset	55
Gambar 4.4 Boxplot Fitur MDVP:F0(Hz).....	56
Gambar 4.5 Pemisahan fitur dan target.....	57
Gambar 4.6 split data menggunakan Stratify.....	58
Gambar 4.7 Penanganan Data Tidak Seimbang.....	60
Gambar 4.8 Cek Distribusi kelas setelah SMOTE.....	60
Gambar 4.9 StandarScaler.....	62
Gambar 4.10 Transformasi data latih dan data uji.....	62
Gambar 4.11 Pelatihan SVM	63
Gambar 4.12 Proses Prediksi	64
Gambar 4.13 Evaluasi thershold 0,5	66
Gambar 4.14 AUC Score	70
Gambar 4.15 Visualisasi Grafik ROC Curve.....	71
Gambar 4.16 Evaluasi thershold 0,3	73
Gambar 4.17 Model Berhasil disimpan	77
Gambar 4.18 Kode model scaler.....	78
Gambar 4.19 Tampilan Fitur Pengisian prediksi	79
Gambar 4.20 Tampilan Upload File .CSV.....	80
Gambar 4.21 Kode Penting Prediksi.....	80
Gambar 4.22 Tampilan Hasil Prediksi.....	81

DAFTAR LAMPIRAN

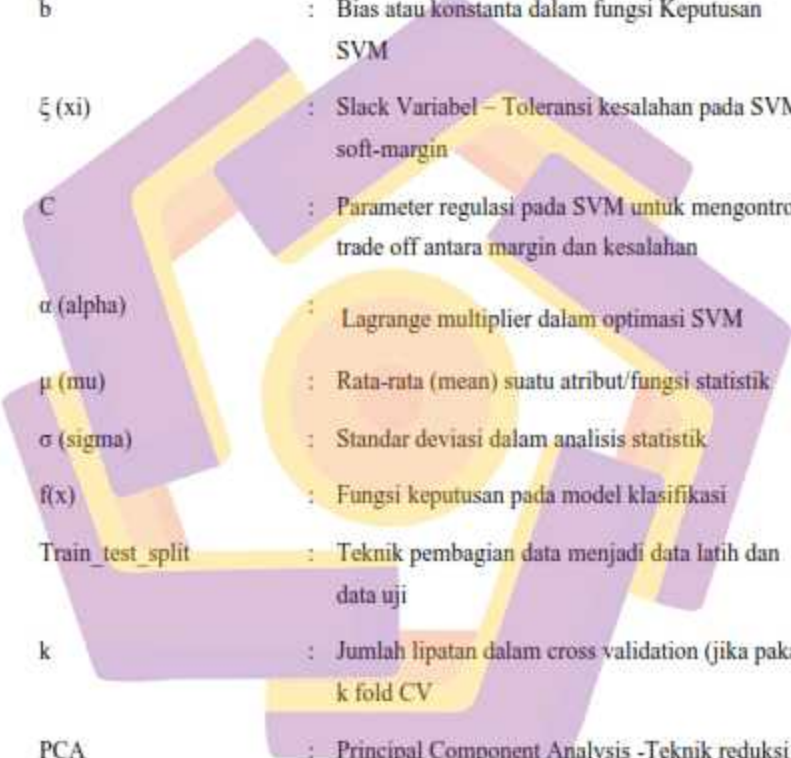
Lampiran 1. Source Code deploy model.....	88
Lampiran 2. web streamlit prediksi penyakit Parkinson.....	88



DAFTAR LAMBANG DAN SINGKATAN

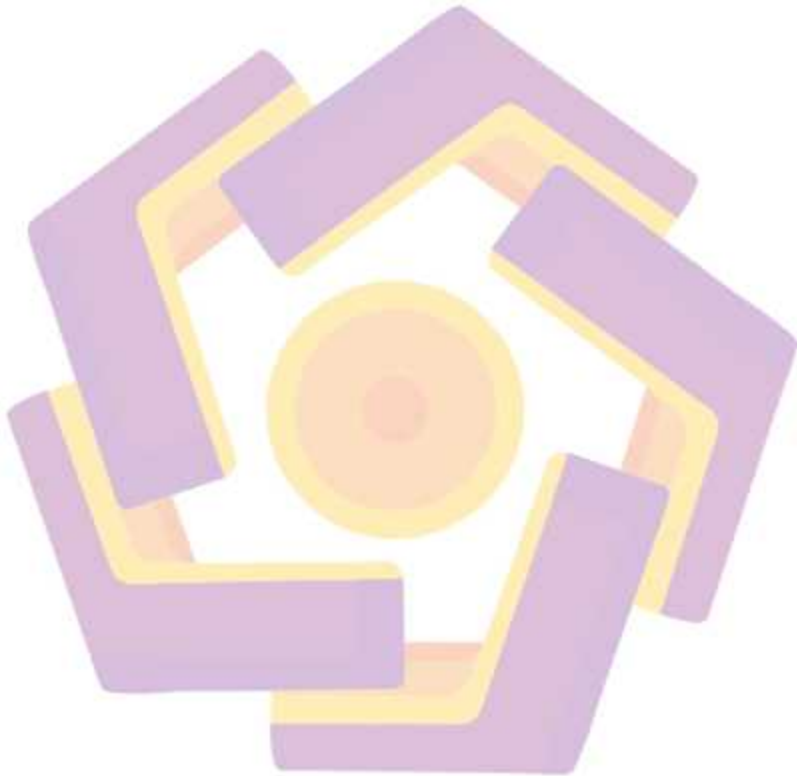


SVM	: Support Vector Machines
SPK	: Sistem Penunjang Keputusan
FN	: False Negative – Data positif yang diklasifikasikan sebagai negatif
FP	: False Positive – Data negatif yang diklasifikasikan sebagai positif
TP	: True Positive – Data positif yang diklasifikasikan sebagai positif
TN	: True Negative – Data negatif yang diklasifikasikan sebagai negatif
SMOTE	: Synthetic Minority Over sampling Technique – Teknik untuk menyeimbangkan data tidak seimbang.
ROC	: Receiver Operating Characteristic – Grafik evaluasi model
AUC	: Area Under Curve – Luas di bawah kurva ROC, mengukur performa model
GUI	: Graphical User Interface – Antarmuka visual pengguna
CSV	: Comma-Separated Values – Format penyimpanan dataset
ML	: Machine learning – Cabang AI untuk pelatihan model dari data
AI	: Artificial Intelligence – Kecerdasan buatan



PD	: Parkinson's Disease - Penyakit Parkinson
X	: Fitur input atau atribut dalam dataset
y	: Label atau target output (Parkinson/sehat)
w	: Vektor bobot (weight vector) pada model SVM
b	: Bias atau konstanta dalam fungsi Keputusan SVM
$\xi (xi)$	: Slack Variabel – Toleransi kesalahan pada SVM soft-margin
C	: Parameter regulasi pada SVM untuk mengontrol trade off antara margin dan kesalahan
α (alpha)	: Lagrange multiplier dalam optimasi SVM
μ (mu)	: Rata-rata (mean) suatu atribut/fungsi statistik
σ (sigma)	: Standar deviasi dalam analisis statistik
$f(x)$	: Fungsi keputusan pada model klasifikasi
Train_test_split	: Teknik pembagian data menjadi data latih dan data uji
k	: Jumlah lipatan dalam cross validation (jika pakai k fold CV)
PCA	: Principal Component Analysis -Teknik reduksi dimensi
FPR	: False Positive Rate
TPR	: True Positive Rate (Recall)
F1	: F1 Score – Harmonik antara precision dan recall

ACC	: Accuracy – Ukuran akurasi model
CM	: Confusion Matrix – Matrix untuk evaluasi hasil klasifikasi
NumPy	: Library Python untuk komputasi numerik



DAFTAR ISTILAH



SMOTE	: Teknik penanganan masalah ketidakseimbangan data dengan menambahkan data sintesis pada kelas minoritas
Preprocessing	: Tahapan membersihkan dan menyiapkan data sebelum pemodelan, termasuk scaling dan balancing.
Normalisasi	: Proses menyamakan skala fitur agar model machine learning bisa belajar dengan lebih baik dan akurat.
Akurasi	: Persentase prediksi yang benar dari total data
Precision	: Persentase prediksi positif yang benar-benar positif
Recall	: Kemampuan model menemukan semua data positif yang benar
F1 Score	: Rata-rata harmonik antara precision dan recall
ROC Curve & AUC (Area Under the Curve)	: Grafik yang menunjukkan performa model pada berbagai threshold; AUC mengukur luas kurva
KDD (Knowledge Discovery in Databases)	: KDD adalah proses menemukan pengetahuan atau pola yang berguna dari kumpulan data yang besar.
Confusion Matrix	: Matriks untuk melihat jumlah prediksi benar dan salah pada tiap kelas.
Classification Report	: laporan yang menunjukkan seberapa baik model machine learning mengklasifikasikan data, dengan menampilkan nilai precision, recall, f1-

score, dan jumlah data (support) untuk setiap kelas.

SVM (Support Vector Machine) : Algoritma machine learning untuk klasifikasi yang mencari garis pemisah terbaik antar kelas.

b-SVM : SVM standar yang digunakan sebagai pembandingan awal (baseline) tanpa teknik tambahan seperti SMOTE atau tuning parameter.

DDS (Decision Support System) : Sistem berbasis komputer yang membantu pengambilan keputusan, khususnya dalam diagnosis medis.

Prediksi : Proses memperkirakan kondisi di masa depan berdasarkan data yang tersedia

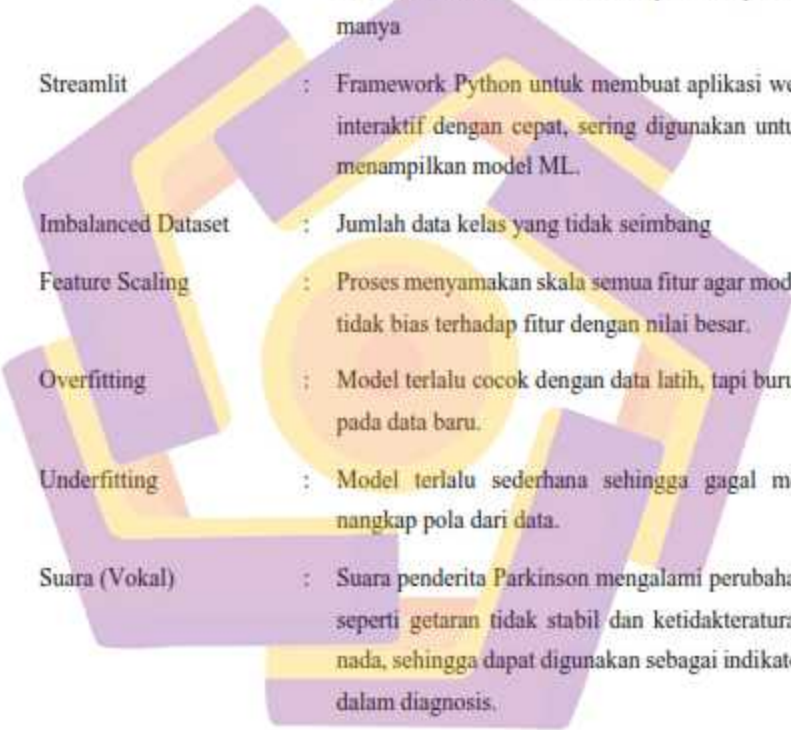
StandarScaler : Teknik normalisasi data dengan mengubah skala fitur menjadi distribusi standar ($\text{mean} = 0$, $\text{std} = 1$). Teknik normalisasi data dengan mengubah skala fitur menjadi distribusi standar ($\text{mean} = 0$, $\text{std} = 1$).

Fitur : Karakteristik atau variabel input dari data, seperti frekuensi suara, jitter, shimmer, dll.

Dataset Parkinson UCI : Dataset berisi fitur suara dari pasien Parkinson dan orang sehat, digunakan untuk klasifikasi.

G Mean (Geometric Mean) : Rata-rata geometrik antara recall dan specificity; cocok untuk data tidak seimbang.

Thershold : Batas nilai probabilitas untuk menentukan kelas prediksi.



Train Test Split	: Proses membagi data menjadi data latih (training) dan data uji (testing).
Model Training	: Proses melatih algoritma machine learning menggunakan data agar bisa mengenali pola
Model Testing	: Proses menguji model yang sudah dilatih dengan data baru untuk melihat seberapa baik performanya
Streamlit	: Framework Python untuk membuat aplikasi web interaktif dengan cepat, sering digunakan untuk menampilkan model ML.
Imbalanced Dataset	: Jumlah data kelas yang tidak seimbang
Feature Scaling	: Proses menyamakan skala semua fitur agar model tidak bias terhadap fitur dengan nilai besar.
Overfitting	: Model terlalu cocok dengan data latih, tapi buruk pada data baru.
Underfitting	: Model terlalu sederhana sehingga gagal menangkap pola dari data.
Suara (Vokal)	: Suara penderita Parkinson mengalami perubahan seperti getaran tidak stabil dan ketidakteraturan nada, sehingga dapat digunakan sebagai indikator dalam diagnosis.
Penyakit Parkinson	: Gangguan saraf progresif yang memengaruhi gerakan tubuh, sering ditandai dengan tremor, kaku otot, dan kesulitan berbicara.
Strafitted	: Teknik pembagian data dengan menjaga proporsi jumlah masing masing label tetap sama antara data latih dan data uji.

INTISARI

Penyakit Parkinson merupakan gangguan neurodegeneratif progresif yang sulit dideteksi secara dini karena gejalanya yang samar pada tahap awal dan hampir menyerupai gejala penyakit neurologis yang lain. Penelitian ini bertujuan untuk membangun sistem penunjang keputusan (SPK) prediksi penyakit Parkinson berdasarkan data suara menggunakan algoritma *Support Vector Machine* (SVM). Dataset yang digunakan berasal dari UCI Machine Learning Repository, terdiri dari 195 sampel suara yang masing-masing memiliki 22 fitur numerik. Data dibagi menjadi 80% data latih dan 20% data uji. Untuk mengatasi ketidakseimbangan kelas pada data latih, diterapkan metode *synthetic minority over-sampling technique* (SMOTE), kemudian dilakukan normalisasi menggunakan *StandardScaler*. Hasil evaluasi menunjukkan bahwa model dengan *threshold* default 0,5 menghasilkan akurasi sebesar 82% namun dengan 6 kasus *false negative*. Setelah dilakukan penyesuaian *threshold* menjadi 0,3, *false negative* berkurang menjadi 3 dan *recall* meningkat dari 0,79 menjadi 0,90, meskipun akurasi turun menjadi 79%. Sistem ini kemudian diimplementasikan ke dalam bentuk aplikasi web menggunakan *framework Streamlit*, yang menyediakan tiga metode input untuk memudahkan pengguna non-teknis. Penelitian ini menunjukkan bahwa SPK berbasis SVM yang dibangun mampu membantu proses deteksi awal penyakit Parkinson secara cepat dan fleksibel.

Kata kunci: Penyakit Parkinson, Support Vector Machine, SPK, SMOTE, prediksi, suara, Streamlit.

ABSTRACT

Parkinson's disease is a progressive neurodegenerative disorder that is difficult to detect at an early stage due to its subtle initial symptoms, which often resemble those of other neurological conditions. This study aims to develop a decision support system (DSS) for predicting Parkinson's disease based on voice data using the Support Vector Machine (SVM) algorithm. The dataset used was obtained from the UCI Machine Learning Repository and consists of 195 voice samples, each containing 22 numerical features. The data was split into 80% for training and 20% for testing. To address class imbalance in the training data, the Synthetic Minority Over-sampling Technique (SMOTE) was applied, followed by normalization using StandardScaler. Evaluation results showed that the model achieved an accuracy of 82% using the default threshold of 0.5, with 6 false negative cases. After adjusting the threshold to 0.3, the number of false negatives was reduced to 3 and the recall for the Parkinson class increased from 0.79 to 0.90, although the accuracy decreased slightly to 79%. The system was then implemented as a web-based application using the Streamlit framework, which provides three input methods to accommodate non-technical users. This study demonstrates that the developed SVM-based DSS can effectively support the early detection of Parkinson's disease in a fast and flexible manner.

Keyword: Parkinson's disease, Support Vector Machine, Decision Support System, SMOTE, prediction, voice, Streamlit