



**PERBANDINGAN KLASIFIKASI PENYAKIT PARKINSON
MENGUNAKAN VARIAN TEKNIK SMOTE DAN *K-NEAREST*
*NEIGHBORS***

SKRIPSI

Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Sarjana Strata-1 Ilmu Komputer

Oleh
FAUZAN NORKAHFI
NIM 1911016110010

**PROGRAM STUDI S1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
JULI 2026**



**PERBANDINGAN KLASIFIKASI PENYAKIT PARKINSON
MENGUNAKAN VARIAN TEKNIK SMOTE DAN *K-NEAREST*
*NEIGHBORS***

SKRIPSI

Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Sarjana Strata-1 Ilmu Komputer

Oleh
FAUZAN NORKAHFI
NIM. 1911016110010

**PROGRAM STUDI S1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
JULI 2026**

SKRIPSI

PERBANDINGAN KLASIFIKASI PENYAKIT PARKINSON MENGGUNAKAN VARIAN TEKNIK SMOTE DAN *K-NEAREST NEIGHBORS*

Oleh :

FAUZAN NORKAHFI

NIM. 1911016110010

telah dipertahankan di depan Dosen Penguji pada tanggal 9 Juli 2026

Susunan Penguji :

Pembimbing Utama

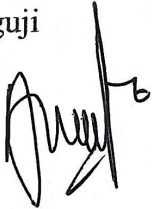

Dwi Kartini, S.Kom., M.Kom.
NIP. 198704212012122003

Pembimbing Pendamping


Friska Abadi, S.Kom., M.Kom.
NIP. 198809132023211010

Penguji

1.


Triando H. Saragih, S.Kom., M.Kom.
NIP. 199308242019031012

2


As'ary Ramadhan, S.Kom., M.Cs.
NIP. 199103292024061001

Banjarbaru, 14 Juli 2026

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 di suatu 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, 9 Juli 2026

Yang Menyatakan,

A handwritten signature in black ink, appearing to be 'Fauzan Norkahfi', written over a horizontal line.

Fauzan Norkahfi
NIM. 1911016110010

ABSTRAK

PERBANDINGAN KLASIFIKASI PENYAKIT PARKINSON MENGUNAKAN VARIAN TEKNIK SMOTE DAN *K-NEAREST NEIGHBORS*

(Oleh : Fauzan Norkahfi; Pembimbing: Dwi Kartini, S.Kom., M.Kom dan Friska Abadi, S.Kom., M.Kom.; 2026; 77 halaman)

Pemanfaatan *machine learning* untuk mendeteksi penyakit Parkinson melalui ekstraksi sinyal suara telah menjadi pendekatan diagnostik medis yang menjanjikan. Walau demikian, pemrosesan data ini sering kali dihadapkan pada dua kendala utama komputasi: tingginya dimensi fitur yang diekstraksi serta ketimpangan proporsi sampel pasien antar kelas. Penelitian ini mengusulkan sebuah kerangka kerja klasifikasi untuk mengatasi kedua masalah tersebut. Algoritma *Principal Component Analysis* (PCA) dengan ambang batas varians 85% diaplikasikan guna memampatkan redundansi fitur akustik tanpa menghilangkan informasi esensial. Data hasil reduksi tersebut kemudian diseimbangkan menggunakan komparasi lima pendekatan sintesis data dari kelompok SMOTE, yaitu SMOTE, ADASYN, *Borderline-SMOTE*, SMOTE-Tomek, dan SVM-SMOTE. Himpunan data yang telah dioptimasi dievaluasi menggunakan pengklasifikasi *K-Nearest Neighbors* (KNN) berbasis metrik *Correlation Distance*. Hasil eksperimen membuktikan bahwa model tanpa penyeimbangan mendapatkan nilai akurasi yang di angka 84,77%. Sebaliknya, penerapan metode SVM-SMOTE secara signifikan berhasil mengungguli seluruh varian lainnya dalam memetakan batas keputusan kelas minoritas. Kombinasi PCA 85%, SVM-SMOTE, dan KNN sukses mendongkrak kinerja model dengan pencapaian akurasi sebesar 90,07%, *recall* 92,04%, presisi 94,55%, dan AUC 93,41%. Penelitian ini menyimpulkan bahwa sinergi antara reduksi dimensi spasial dan teknik *oversampling* berhasil dalam meningkatkan akurasi deteksi sistem.

Kata kunci: Parkinson, Klasifikasi, *K-Nearest Neighbors*, SMOTE, ADASYN, *Borderline-SMOTE*, SMOTE-Tomek, SVM-SMOTE

ABSTRACT

COMPARISON OF PARKINSON'S DISEASE CLASSIFICATION USING SMOTE VARIANTS AND K-NEAREST NEIGHBORS

(Oleh : Fauzan Norkahfi; Pembimbing: Dwi Kartini, S.Kom., M.Kom dan Friska Abadi, S.Kom., M.Kom.; 2026; 77 halaman)

The utilization of machine learning to detect Parkinson's disease through voice signal extraction has become a promising medical diagnostic approach. However, processing this data is often challenged by two main computational constraints: the high dimensionality of the extracted features and the imbalanced proportion of patient samples between classes. This study proposes a classification framework to address both issues. The Principal Component Analysis (PCA) algorithm with an 85% variance threshold is applied to compress acoustic feature redundancy without losing essential information. The reduced data is then balanced by comparing five data synthesis approaches from the SMOTE family, namely SMOTE, ADASYN, Borderline-SMOTE, SMOTE-Tomek, and SVM-SMOTE. The optimized dataset is evaluated using a K-Nearest Neighbors (KNN) classifier based on the Correlation Distance metric. Experimental results demonstrate that the model without balancing (Baseline) yielded an accuracy limited to 84.77%. Conversely, the application of the SVM-SMOTE method significantly outperformed all other variants in mapping the decision boundaries of the minority class. The combination of PCA 85%, SVM-SMOTE, and KNN successfully boosted the model's reliability, achieving 90.07% accuracy, 92.04% recall, 94.55% precision, and 93.41% AUC. This study concludes that the synergy between spatial dimensionality reduction and oversampling techniques successfully improves the system's detection accuracy.

Keywords: Parkinson, Classification, K-Nearest Neighbors, SMOTE, ADASYN, Borderline-SMOTE, SMOTE-Tomek, SVM-SMOTE

PRAKATA

Puji syukur penulis panjatkan ke Tuhan kita Yang Maha Esa karena atas berkat rahmat dan karunia-Nya penulis dapat menyelesaikan skripsi yang berjudul **PERBANDINGAN KLASIFIKASI PENYAKIT PARKINSON MENGGUNAKAN VARIAN TEKNIK SMOTE DAN *K-NEAREST NEIGHBORS*** untuk memenuhi syarat dalam menyelesaikan pendidikan program S1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat.

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

1. Keluarga yang selalu memberikan bantuan, semangat, doa dan dukungan dalam proses penyelesaian skripsi ini.
2. Ibu Dwi Kartini, S.Kom., M.Kom selaku dosen pembimbing utama dan Koordinator Program Studi Ilmu Komputer FMIPA ULM yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian skripsi ini.
3. Bapak Friska Abadi, S.Kom., M.Kom selaku dosen pembimbing pendamping yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian skripsi ini.
4. Seluruh Dosen dan staf Program Studi Ilmu Komputer FMIPA ULM atas ilmu dan bantuan yang diberikan selama ini yang sangat bermanfaat.
5. Ricky Rihanda Syawaliandi dan Falah Isnain yang telah meluangkan waktu dan pikirannya dalam proses pengerjaan skripsi.
6. Teman-teman keluarga Ilmu Komputer angkatan 2019 yang memberikan dukungan dan selalu mengingatkan serta mendoakan dalam proses mengerjakan skripsi.
7. Semua pihak yang tidak dapat disebutkan satu persatu yang telah turut membantu dalam penyelesaian skripsi ini.

Akhir kata penulis menyadari sepenuhnya bahwa penulisan ini jauh dari sempurna, namun penulis mengharapkan bantuan berupa saran dan kritik yang membangun dari semua pihak demi kesempurnaan dan mutu penulisan skripsi ini. Semoga tulisan ini dapat bermanfaat bagi ilmu pengetahuan dan pembaca khususnya serta mendapat keridhaan Allah SWT.

Banjarbaru, 9 Juli 2026

Yang Menyatakan,

A handwritten signature in black ink, consisting of a large, sweeping loop on the left and a series of smaller, connected strokes on the right, ending in a small dot.

Fauzan Norkahfi

NIM. 1911016110010

DAFTAR ISI

HALAMAN JUDUL	ii
HALAMAN PENGESAHAN.....	iii
PERNYATAAN.....	iv
ABSTRAK	v
<i>ABSTRACT</i>	vi
PRAKATA.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN	xiii
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
BAB II TINJAUAN PUSTAKA.....	5
2.1 Kajian Terdahulu	5
2.2 Penyakit Parkinson	11
2.3 Normalisasi <i>Z-Score</i>	11
2.4 SMOTE.....	12
2.5 SMOTE-Tomek	13
2.6 <i>Borderline-SMOTE</i>	14
2.7 ADASYN.....	15
2.8 SVM-SMOTE.....	17
2.9 <i>Principal Component Analysis (PCA)</i>	17
2.10 <i>Grid SearchCV</i>	19
2.11 <i>K-Nearest Neighbor (KNN)</i>	21
2.12 <i>Confusion Matrix</i>	22

2.13	<i>Area Under the Receiver Operating Characteristic Curve</i>	24
BAB III METODE PENELITIAN		26
3.1	Alat Penelitian	26
3.2	Bahan Penelitian	26
3.3	Variabel Penelitian.....	26
3.4	Prosedur Penelitian	27
BAB IV HASIL DAN PEMBAHASAN		30
4.1	Hasil	30
4.1.1	Pengumpulan Data.....	30
4.1.2	Pra-Pemrosesan Data	32
4.1.2.1	Pembersihan Data	32
4.1.2.2	Pembagian Data	32
4.1.2.3	Normalisasi <i>Z-Score</i>	33
4.1.3	Penyeimbangan Kelas.....	36
4.1.3.1	<i>Synthetic Minority Over-sampling Technique</i> (SMOTE)....	37
4.1.3.2	<i>Support Vector Machine</i> SMOTE (SVM-SMOTE)	39
4.1.3.3	<i>Adaptive Synthetic</i> (ADASYN)	42
4.1.3.4	<i>Borderline</i> SMOTE	45
4.1.3.5	<i>SMOTE-Tomek</i>	47
4.1.4	Reduksi Dimensi dengan <i>Principal Component Analysis</i>	49
4.1.5	Klasifikasi Menggunakan <i>K-nearest neighbor</i> (KNN).....	58
4.1.6	Evaluasi.....	63
4.2	Pembahasan.....	65
BAB V PENUTUP.....		72
5.1	Kesimpulan	72
5.2	Saran	72
DAFTAR PUSTAKA		73
LAMPIRAN.....		78
RIWAYAT HIDUP		86

DAFTAR TABEL

Tabel 1. Keaslian Penelitian.....	10
Tabel 2. Perancangan Penelitian	11
Tabel 3. Nama Fitur	30
Tabel 4. Contoh <i>Dataset Parkinson's Disease (PD) Classification</i>	31
Tabel 5. Jumlah data pada <i>dataset Parkinson's Disease (PD) Classification</i>	31
Tabel 6. Dimensi <i>Dataset</i> pada Tahap Pembersihan Data	32
Tabel 7. Data Latih Setelah Proses Normalisasi <i>Z-Score</i>	34
Tabel 8. Data Uji Setelah Proses Normalisasi <i>Z-Score</i>	35
Tabel 9. Nilai rata-rata (<i>X</i>) dan simpangan baku (<i>S</i>) Fitur	36
Tabel 10. Parameter SMOTE pada library <i>imblearn</i>	39
Tabel 11. Parameter SVM-SMOTE pada library <i>imblearn</i>	42
Tabel 12. Parameter ADASYN pada library <i>imblearn</i>	44
Tabel 13. Parameter <i>Borderline</i> SMOTE pada library <i>imblearn</i>	47
Tabel 14. Parameter SMOTE- <i>Tomek</i> pada library <i>imblearn</i>	49
Tabel 15. Kinerja Model Berdasarkan Variance Threshold 85%	50
Tabel 16. Kinerja Model Berdasarkan Variance Threshold 90%	50
Tabel 17. Kinerja Model Berdasarkan Variance Threshold 95%	51
Tabel 18. Hasil Perhitungan Matriks <i>Covariance</i> Data Uji dari SVM-SMOTE	54
Tabel 19. Hasil Perhitungan <i>Eigenvalue</i> Data Uji Hasil SVM-SMOTE	55
Tabel 20. Distribusi Varians Kumulatif pada Komponen Utama	56
Tabel 21 Hasil Reduksi Dimensi PCA.....	57
Tabel 22. Parameter KNN pada library <i>Scikit-learn</i>	58
Tabel 23. Parameter <i>GridSearch</i> pada library <i>Scikit-learn</i>	58
Tabel 25. Komparasi Kinerja Model Klasifikasi KNN pada Berbagai Skenario Uji	64

DAFTAR GAMBAR

Gambar 1. Alur Prosedur Penelitian	27
Gambar 2. Distribusi Kelas Data Latih.....	33
Gambar 3. Distribusi Kelas Data Latih Sesudah Penerapan SMOTE	38
Gambar 4. Distribusi Kelas Data Latih Sesudah Penerapan SVM-SMOTE	41
Gambar 5. Distribusi Kelas Data Latih Sesudah Penerapan ADASYN	43
Gambar 6. Distribusi Kelas Data Latih Sesudah Penerapan <i>Borderline</i> SMOTE	45
Gambar 7. Distribusi Kelas Data Latih Sesudah Penerapan SMOTE- <i>Tomek</i>	48
Gambar 8. Distribusi Varians Kumulatif PCA pada Skenario SVM-SMOTE	56
Gambar 9. Kurva Pergerakan Akurasi Validasi Terhadap Nilai K pada Skenario Non-SMOTE	60
Gambar 10. Kurva Pergerakan Akurasi Validasi Terhadap Nilai K SMOTE	60
Gambar 11. Kurva Pergerakan Akurasi Validasi Terhadap Nilai K ADASYN	61
Gambar 12. Kurva Pergerakan Akurasi Validasi Terhadap Nilai K B-SMOTE	61
Gambar 13. Kurva Pergerakan Akurasi Validasi Terhadap Nilai K SMOTE- <i>Tomek</i>	62
Gambar 14. Kurva Pergerakan Akurasi Validasi Terhadap Nilai K SVM-SMOTE ..	62
Gambar 15. <i>Confusion Matrix</i> KNN dengan SVM-SMOTE.....	64
Gambar 16. Perbandingan Akurasi KNN pada Berbagai Skenario Varian SMOTE ..	66
Gambar 17. Perbandingan <i>Recall</i> KNN pada Berbagai Skenario Varian SMOTE ...	67
Gambar 18. Perbandingan Presisi KNN pada Berbagai Skenario Varian SMOTE ...	68
Gambar 19. Perbandingan <i>F1-Score</i> KNN pada Berbagai Skenario Varian SMOTE	69
Gambar 20. Perbandingan AUC KNN pada Berbagai Skenario Varian SMOTE	70

DAFTAR LAMPIRAN

Lampiran 1 Source Code.....	79
Lampiran 2 Riwayat Hidup.....	86