



**SELEKSI FITUR DENGAN *PARTICLE SWARM OPTIMIZATION* PADA
KLASIFIKASI PENYAKIT PARKINSON MENGGUNAKAN *XGBOOST***

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

Oleh

DENI KURNIA

NIM 1911016210006

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

JUNI 2023



**SELEKSI FITUR DENGAN *PARTICLE SWARM OPTIMIZATION* PADA
KLASIFIKASI PENYAKIT PARKINSON MENGGUNAKAN *XGBOOST***

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

Oleh

**DENI KURNIA
NIM 1911016210006**

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

JUNI 2023

SKRIPSI

SELEKSI FITUR DENGAN *PARTICLE SWARM OPTIMIZATION* PADA KLASIFIKASI PENYAKIT PARKINSON MENGGUNAKAN *XGBOOST*

Oleh:

DENI KURNIA

NIM. 1911016210006

Telah dipertahankan di depan Dosen Penguji pada tanggal 21 Juni 2023.

Susunan Dosen Penguji:

Pembimbing I



Muhammad Itqan Mazdadi, S.Kom, M.Kom

NIP. 199006122019031013

Dosen Penguji I



Radityo Adi Nugroho, S.T., M.Kom

NIP. 198212042008011006

Pembimbing II



Dwi Kartini S.Kom., M.Kom

NIP. 198704212012122003

Dosen Penguji II



Friska Abadi, S.Kom., M.Kom

NIP. 19880913201612110001

Banjarnbaru, 26 Juni 2023

Koordinator Program Studi Ilmu Komputer



Prayudhiman, S.T., M.Kom

NIP. 197703252008121001

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, 21 Juni 2023

Yang Menyatakan,



Deni Kurnia

NIM.1911016210006

ABSTRAK

SELEKSI FITUR DENGAN *PARTICLE SWARM OPTIMIZATION* PADA KLASIFIKASI PENYAKIT PARKINSON MENGGUNAKAN *XGBOOST*

(Oleh : Deni Kurnia; Pembimbing: Muhammad Itqan Mazdadi, S.Kom., M.Kom dan Dwi Kartini S.Kom., M.Kom.; 2023; 59 halaman)

Penyakit Parkinson merupakan gangguan pada sistem saraf pusat yang mempengaruhi sistem motorik. Diagnosis penyakit ini cukup sulit dilakukan karena gejalanya yang serupa dengan penyakit lain. Saat ini diagnosa dapat dilakukan menggunakan *machine learning* dengan memanfaatkan rekaman suara pasien yang diekstraksi. Namun, fitur yang dihasilkan dari ekstraksi rekaman suara tersebut relatif cukup banyak, sehingga penting untuk mempertimbangkan seleksi fitur untuk menghindari penurunan kinerja sebuah model. Pada penelitian ini, *Particle Swarm Optimization* digunakan sebagai seleksi fitur, sedangkan *XGBoost* akan digunakan sebagai model klasifikasi. Selain itu model juga akan diterapkan SMOTE untuk mengatasi masalah ketidakseimbangan kelas data dan *hyperparameter tuning* menggunakan *Random Search* pada *XGBoost* untuk mendapatkan *hyperparameter* yang optimal. Hasil pengujian menunjukkan bahwa nilai AUC pada model *XGBoost* dengan seleksi fitur adalah 0.9325, sedangkan pada model *XGBoost* tanpa seleksi fitur hanya mendapat nilai AUC sebesar 0.9250. Kemudian, ketika metode SMOTE dan *Random Search* diterapkan, maka kinerja model *XGBoost* akan semakin baik dengan mendapatkan nilai AUC sebesar 0.9366. Namun, ketika seleksi fitur diterapkan terlebih dahulu dari kedua metode tadi, maka model akan mendapatkan nilai AUC yang lebih baik yaitu sebesar 0.9483. Dengan demikian, penggunaan seleksi fitur pada model telah terbukti mampu dalam meningkatkan kinerja dan hasil prediksi.

Kata kunci: Klasifikasi, *XGBoost*, PSO, SMOTE, *Hyperparameter Tuning*, *Random Search*

ABSTRACT

FEATURE SELECTION WITH PARTICLE SWARM OPTIMIZATION IN PARKINSON'S DISEASE CLASSIFICATION USING XGBOOST

(By : Deni Kurnia; Suvervisor: Muhammad Itqan Mazdadi, S.Kom., M.Kom and Dwi Kartini S.Kom., M.Kom.; 2023; 59 pages)

Parkinson's disease is a disorder of the central nervous system that affects the motor system. Diagnosis of this disease is quite difficult because the symptoms are similar to other diseases. Currently, diagnosis can be done using machine learning by utilizing extracted patient voice recordings. However, the features generated from the extraction of voice recordings are relatively large, so it is important to consider feature selection to avoid degrading the performance of a model. In this study, Particle Swarm Optimization is used as feature selection, while XGBoost will be used as a classification model. In addition, the model will also be applied SMOTE to overcome the problem of data class imbalance and hyperparameter tuning using Random Search on XGBoost to get the optimal hyperparameter. The test results show that the AUC value on the XGBoost model with feature selection is 0.9325, while the XGBoost model without feature selection only gets an AUC value of 0.9250. Then, when the SMOTE and Random Search methods are applied, the performance of the XGBoost model will be even better by getting an AUC value of 0.9366. However, when feature selection is applied first of the two methods, the model will get a better AUC value of 0.9483. Thus, the use of feature selection in the model has proven to be able to improve the performance and prediction results.

Keywords: *Classification, XGBoost, PSO, SMOTE, Hyperparameter Tuning, Random Search*

PRAKATA

Puji syukur penulis panjatkan ke Tuhan kita Yang Maha Esa karena atas berkat rahmat dan karunia-Nya penulis dapat menyelesaikan skripsi yang berjudul “SELEKSI FITUR DENGAN *PARTICLE SWARM OPTIMIZATION* PADA KLASIFIKASI PENYAKIT PARKINSON MENGGUNAKAN *XGBOOST*” 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 terimakasih 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. 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. Ibu Dwi Kartini, S.Kom., M.Kom. selaku dosen pembimbing pendamping yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian skripsi ini.
4. Bapak Irwan Budiman S.T., M.Kom selaku Ketua Program Studi Ilmu Komputer FMIPA ULM, atas bantuan dan izin beliau skripsi ini dapat diselesaikan.
5. Seluruh Dosen dan staff Program Studi Ilmu Komputer FMIPA ULM atas ilmu dan bantuan yang diberikan selama ini yang sangat bermanfaat.
6. Teman-teman dan sahabat-sahabat 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 serupa 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, 21 Juni 2023

A handwritten signature in black ink, appearing to be 'Deni Kurnia', written in a cursive style.

Deni Kurnia

DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERNYATAAN	iii
ABSTRAK	iv
<i>ABSTRACT</i>	v
PRAKATA.....	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN.....	xiv
BAB I PENDAHULUAN	
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan	3
1.4 Manfaat Penelitian.....	3
1.5 Batasan Masalah	3
BAB II TINJAUAN PUSTAKA	
2.1 Kajian Terdahulu	4
2.2 Landasan Teori	8
2.2.1 Penyakit Parkinson.....	8
2.2.2 <i>Min-Max Scaler</i>	8
2.2.3 <i>Particle Swarm Optimization</i>	9
2.2.4 Ketidakseimbangan Kelas.....	10
2.2.5 SMOTE	11
2.2.6 <i>Decision Tree</i>	15
2.2.7 <i>Ensemble Learning</i>	16
2.2.8 <i>Boosting</i>	17
2.2.9 <i>Extreme Gradient Boosting</i>	17
2.2.10 <i>Hyperparameter Tuning</i>	19

2.2.11	<i>Random Search</i>	19
2.2.12	<i>Cross Validation</i>	20
2.2.13	<i>Area Under Curve (AUC)</i>	20
BAB III METODE PENELITIAN		
3.1	Bahan Penelitian	23
3.2	Alat Penelitian	23
3.3	Variabel Penelitian.....	23
3.4	Prosedur Penelitian	23
BAB IV HASIL DAN PEMBAHASAN		
4.1	Hasil.....	27
4.1.1	Pengumpulan Dataset.....	27
4.1.2	<i>Preprocessing Data</i>	28
4.1.3	Pembagian <i>Data Training</i> dan <i>Data Testing</i>	29
4.1.4	<i>Particle Swarm Optimization</i>	30
4.1.5	Klasifikasi	37
4.1.6	Evaluasi dengan AUC (<i>Area Under Curve</i>)	44
4.2	Pembahasan	48
BAB V PENUTUP		
5.1	Kesimpulan.....	57
5.2	Saran	57
DAFTAR PUSTAKA		
LAMPIRAN		

DAFTAR TABEL

Tabel	Halaman
Tabel 1. Keaslian Penelitian.....	6
Tabel 2. Perencanaan Penelitian	8
Tabel 3. Contoh Dataset Sebelum Dilakukan Proses SMOTE	12
Tabel 4. Contoh Komposisi Dataset Sebelum dan Sesudah SMOTE.....	14
Tabel 5. Contoh Dataset Setelah Dilakukan SMOTE.....	15
Tabel 6. Klasifikasi AUC.....	22
Tabel 7. <i>Hyperparameter XGBoost</i>	26
Tabel 8. Nama Fitur	27
Tabel 9. Contoh <i>Parkinson's Disease Classification Dataset</i>	28
Tabel 10. Jumlah Distribusi Kelas	28
Tabel 11. Dataset Setelah Dilakukan Normalisasi <i>Min-Max Scaler</i>	29
Tabel 12. Hasil Percobaan Jumlah Iterasi Terbaik.....	32
Tabel 13. Hasil Percobaan Jumlah Partikel Terbaik	33
Tabel 14. Hasil Percobaan Bobot Inersia Terbaik	34
Tabel 15. Hasil Percobaan Koefisien Akselerasi Terbaik.....	37
Tabel 16. <i>Confusion matrix</i> Model <i>XGBoost</i>	38
Tabel 17. <i>Confusion Matrix</i> Model <i>XGBoost</i> dengan SMOTE	39
Tabel 18. Kandidat <i>Hyperparameter Random Search</i>	40
Tabel 19. <i>Confusion Matrix</i> Model <i>XGBoost</i> dengan SMOTE dan <i>Random Search</i>	41
Tabel 20. <i>Confusion matrix</i> Model <i>XGBoost</i> dengan PSO	41
Tabel 21. <i>Confusion matrix</i> Model <i>XGBoost</i> dengan PSO dan SMOTE.....	42
Tabel 22. Kandidat <i>Hyperparameter Random Search</i>	43
Tabel 23. <i>Confusion Matrix</i> Model <i>XGBoost</i> dengan PSO, SMOTE dan <i>Random Search</i>	43
Tabel 24. Hasil Evaluasi AUC dengan Model <i>XGBoost</i>	44
Tabel 25. Hasil Evaluasi AUC Model <i>XGBoost</i> dengan SMOTE.....	45
Tabel 26. Hasil Evaluasi AUC Model <i>XGBoost</i> dengan SMOTE dan <i>Random Search</i>	45
Tabel 27. Hasil Evaluasi AUC Model <i>XGBoost</i> dengan PSO	46

Tabel 28. Hasil Evaluasi AUC Model <i>XGBoost</i> dengan PSO dan SMOTE.....	47
Tabel 29. Hasil Evaluasi AUC Model <i>XGBoost</i> dengan PSO, SMOTE, dan <i>Random Search</i>	48

DAFTAR GAMBAR

Gambar	Halaman
Gambar 1. Ilustrasi SMOTE	12
Gambar 2. <i>Flowchart</i> SMOTE.....	12
Gambar 3. Struktur <i>Decission Tree</i>	16
Gambar 4. Ilustrasi <i>Ensemble Learning</i>	16
Gambar 5. Kerja Internal dari Algoritma <i>Boosting</i>	17
Gambar 6. Diagram Skema dari Algoritma <i>XGBoost</i>	18
Gambar 7. Ilustrasi Metode <i>Random Search</i>	19
Gambar 8. Ilustrasi <i>K-Folds Cross Validation</i>	20
Gambar 9. Contoh <i>AUC Curve</i>	21
Gambar 10. Alur Penelitian.....	24
Gambar 11. Diagram Seleksi Fitur PSO	30
Gambar 12. Hasil Pengujian Terhadap Jumlah Iterasi.....	31
Gambar 13. Hasil Pengujian Terhadap Jumlah Partikel	32
Gambar 14. Hasil Pengujian Terhadap Bobot Inersia.....	34
Gambar 15. Hasil Pengujian Terhadap $C1 > C2$	35
Gambar 16. Hasil Pengujian Terhadap $C1 < C2$	36
Gambar 17. Hasil Pengujian Terhadap $C1 = C2$	36
Gambar 18. Perbandingan Kelas Sebelum SMOTE	38
Gambar 19. Perbandingan Kelas Setelah SMOTE	39
Gambar 20. Grafik ROC AUC Model <i>XGBoost</i>	44
Gambar 21. Grafik ROC AUC Model <i>XGBoost</i> dengan SMOTE.....	45
Gambar 22. Grafik ROC AUC Model <i>XGBoost</i> dengan SMOTE dan <i>Random Search</i>	46
Gambar 23. Grafik ROC AUC Model <i>XGBoost</i> dengan PSO	47
Gambar 24. Grafik ROC AUC Model <i>XGBoost</i> dengan PSO dan SMOTE.....	47
Gambar 25. Grafik ROC AUC Model <i>XGBoost</i> dengan PSO, SMOTE, dan <i>Random Search</i>	48
Gambar 26. Perbandingan Nilai AUC Model <i>XGBoost</i> dengan <i>XGBoost + PSO</i>	54

Gambar 27. Perbandingan Nilai AUC Model <i>XGBoost</i> + SMOTE dengan <i>XGBoost</i> + PSO + SMOTE.....	55
Gambar 28. Perbandingan Nilai AUC Model <i>XGBoost</i> + SMOTE + <i>Random Search</i> dengan <i>XGBoost</i> + PSO + SMOTE + <i>Random Search</i>	55

DAFTAR LAMPIRAN

Lampiran

Lampiran 1 Bobot Setiap Fitur Hasil Seleksi Fitur

Lampiran 2 Daftar Fitur Terpilih

Lampiran 3 *Source Code*

Lampiran 4 Riwayat Hidup Penulis