

**KLASIFIKASI POTENSI ATLET PORDA SHORINJI KEMPO
BERDASARKAN DATA ENDURANCE, SPEED DAN POWER
DENGAN PENDEKATAN MACHINE LEARNING**

SKRIPSI

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



disusun oleh
MUH NAUFAL RAFIF
21.11.4155

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

**KLASIFIKASI POTENSI ATLET PORDA SHORINJI KEMPO
BERDASARKAN DATA ENDURANCE, SPEED DAN POWER
DENGAN PENDEKATAN MACHINE LEARNING**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

Muh Naufal Ralfi

21.11.4155

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

KLASIFIKASI POTENSI ATLET PORDA SHORINJI KEMPO BERDASARKAN
DATA ENDURANCE, SPEED DAN POWER DENGAN PENDEKATAN MACHINE
LEARNING

yang disusun dan diajukan oleh

Muh Naufal Rafif

21.11.4155

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 21 Februari 2025

Dosen Pembimbing,


Kusnawi, S.kom., M.Eng
NTK. 190302112

HALAMAN PENGESAHAN

SKRIPSI

KLASIFIKASI POTENSI ATLET PORDA SHORINJI KEMPO
BERDASARKAN DATA ENDURANCE, SPEED DAN POWER DENGAN
PENDEKATAN MACHINE LEARNING

yang disusun dan diajukan oleh

Muh Naufal Rafif

21.11.4155

Telah dipertahankan di depan Dewan Penguji
pada tanggal 21 Februari 2025

Susunan Dewan Penguji

Nama Penguji

Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302096

Fertan Fauzi Abdulloh, S.Kom., M.Kom.
NIK. 190302276

Kusnawi, S.Kom., M.Eng.
NIK. 190302112

Tanda Tangan

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 21 Februari 2025

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Muh Naufal Rafif

NIM : 21.11.4155

Menyatakan bahwa Skripsi dengan judul berikut:

KLASIFIKASI POTENSI ATLET PORDA SHORINJI KEMPO BERDASARKAN DATA ENDURANCE, SPEED DAN POWER DENGAN PENDEKATAN MACHINE LEARNING

Dosen Pembimbing : Kusnawi, S.Kom., M.Eng.

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, 21 Februari 2025

Yang Menyatakan,



Muh Naufal Rafif

HALAMAN PERSEMBAHAN

Puji syukur kehadiran Allah SWT yang telah melimpahkan rahmat, taufik, dan hidayah-Nya sehingga karya ini dapat diselesaikan dengan baik. Karya ini dipersembahkan sebagai bentuk rasa syukur dan penghargaan kepada semua pihak yang telah memberikan dukungan, motivasi, dan bantuan dalam proses penyusunannya.

Karya ini kupersembahkan untuk:

1. **Kedua Orang Tuaku**

Terima kasih atas segala doa, kasih sayang, dan pengorbanan yang tak terilai. Tanpa dukungan kalian, karya ini tidak akan terwujud.

2. **Guru dan Dosen Pembimbing**

Terima kasih atas bimbingan, ilmu, dan motivasi yang telah diberikan. Semoga amal baik Bapak/Ibu diterima oleh Allah SWT.

3. **Sahabat dan Teman-Teman**

Terima kasih atas kebersamaan, dukungan, dan semangat yang selalu kalian berikan. Kalian adalah bagian penting dalam perjalananku.

4. **Pelatih dan Atlit Porda Sleman**

Terima kasih atas kontribusi dan bantuan yang telah diberikan, baik secara langsung maupun tidak langsung.

Semoga karya ini dapat bermanfaat bagi semua pihak dan menjadi amal jariyah bagi penulis serta semua yang terlibat. Aamiin.

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT atas segala rahmat, hidayah, dan karunia-Nya, sehingga penulis dapat menyelesaikan karya ini dengan baik. Shalawat serta salam semoga tercurah kepada Nabi Muhammad SAW, keluarga, sahabat, dan seluruh umatnya hingga akhir zaman.

Karya ini disusun sebagai upaya untuk [tujuan karya, misalnya: memenuhi tugas akhir, menyumbangkan pemikiran, atau memberikan kontribusi dalam bidang tertentu]. Dalam proses penyusunannya, penulis menyadari bahwa karya ini tidak akan terwujud tanpa bantuan, dukungan, dan bimbingan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. **Kedua Orang Tua**

Atas doa, kasih sayang, dan dukungan yang tak pernah putus dalam setiap langkah penulis.

2. **Dosen/Guru Pembimbing**

Terima kasih atas bimbingan, saran, dan motivasi yang telah diberikan selama proses penyusunan karya ini.

3. **Rekan-Rekan dan Sahabat**

Terima kasih atas semangat, diskusi, dan kebersamaan yang telah membantu penulis dalam menyelesaikan karya ini.

4. **Semua Pihak yang Terlibat**

Terima kasih atas kontribusi dan bantuan yang telah diberikan, baik secara langsung maupun tidak langsung.

Penulis menyadari bahwa karya ini masih jauh dari sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun dari para pembaca demi perbaikan dan penyempurnaan karya ini di masa yang akan datang.

Semoga karya ini dapat memberikan manfaat bagi semua pihak yang membacanya. Aamiin.

Yogyakarta, 21 Februari 2025

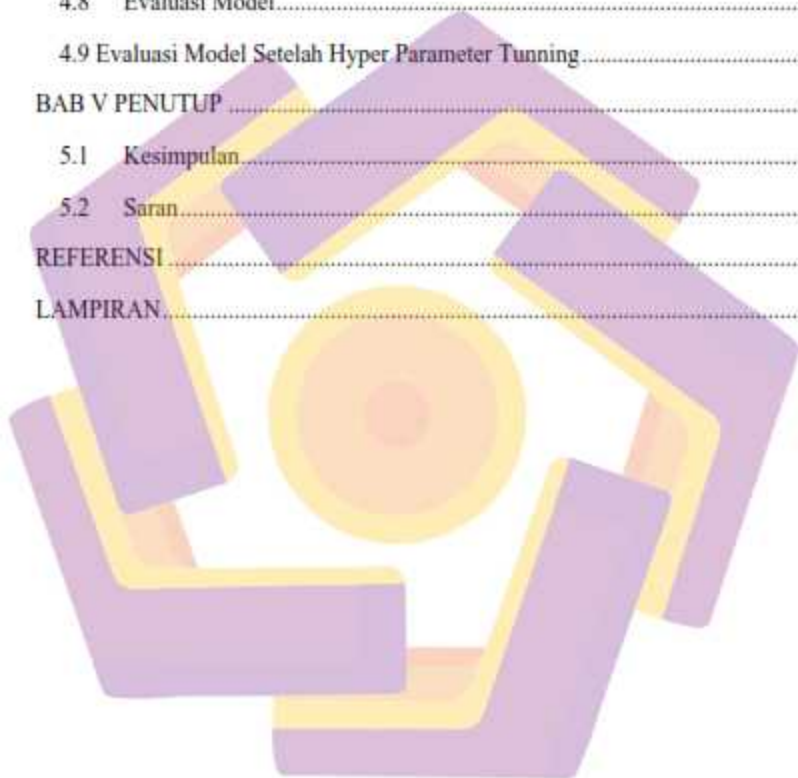
Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	Error! Bookmark not defined.
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMBANG DAN SINGKATAN	xii
DAFTAR ISTILAH.....	xiii
INTISARI	xiv
<i>ABSTRACT</i>	xv
BAB I PENDAHULUAN.....	1
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
1.6 Sistematika Penulisan.....	3
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Dasar Teori.....	20
2.2.1 Teori Performa Olahraga	20
2.2.2 Data Mining	23
2.2.3 Machine Learning	25

2.2.4	Sistem Rekomendasi	26
2.2.5	Klasifikasi	26
2.2.6	Clustering	26
2.2.7	Performance Evaluasi	27
2.2.8	Preprocessing	29
2.2.9	Random Forest	32
2.2.10	Logistik Regresi	33
2.2.11	Python	33
2.2.12	Hyper Parameter Tunning	33
2.2.13	K-Means	34
BAB III METODE PENELITIAN		36
3.1	Objek Penelitian	36
3.2	Alur Penelitian	36
3.3	Alat dan Bahan	41
3.3.1	Data Penelitian	41
3.3.2	Alat/Instrumen	41
BAB IV HASIL DAN PEMBAHASAN		43
4.1	Preprocessing	43
4.1.1	Konversi Format Data	43
4.1.2	Label Encoding	44
4.1.3	Outlier Detecscion	45
4.1.4	Fitur Selection	46
4.1.5	Pra Perhitungan Fitur	47
4.2	Visualisasi Data	47
4.3	Penentuan Label Untuk Klasifikasi	53

4.4	Feature Selection	54
4.5	Proses K-Fold Cross Validation	54
4.6	Pembuatan Model	56
4.7	Pembuatan Model Dengan Hyper Parameter	58
4.8	Evaluasi Model	60
4.9	Evaluasi Model Setelah Hyper Parameter Tuning	70
BAB V PENUTUP		81
5.1	Kesimpulan	81
5.2	Saran	82
REFERENSI		83
LAMPIRAN		88



DAFTAR TABEL

Tabel 4.1 Konversi Format Data	51
Tabel 4.2 Data setelah dikonversi	51
Tabel 4.3 Label Encoding	52
Tabel 4.4 Outlier Detection	53
Tabel 4.5 Pra Perhitungan Fitur	53
Tabel 4.6 Visualisasi VO ₂ Max Berdasarkan Gender	54
Tabel 4.7 Scatter plot Power Otot Tungkai dan Berat Badan	55
Tabel 4.8 Kecepatan Berdasarkan Kelompok	56
Tabel 4.9 Korelasi Fitur	57
Tabel 4.10 Penentuan Kategori untuk Klasifikasi	59
Tabel 4.11 Feature Selection	60
Tabel 4.12 Pembagian Data	61
Tabel 4.13 Model Random Forest	61
Tabel 4.14 Model Logistik Regresi	62
Tabel 4.15 Evaluasi Model Random Forest	63
Tabel 4.16 Evaluasi Model Logistik Regresi	64
Tabel 4.17 Perbandingan Evaluasi Model	65
Tabel 4.21 Perbandingan Evaluasi Model Sebelum Hyper Parameter Tuning	73
Tabel 4.22 Perbandingan Evaluasi Model Sesudah Hyper Parameter Tuning	74

DAFTAR GAMBAR

Gambar 2.1. Table Vo2max	33
Gambar 2.1 Table Vo ₂ Max	33
Gambar 3.1 Alur Penelitian	44
Gambar 3.2 Informasi Dataset	45
Gambar 4.1 Hasil Encoding	52
Gambar 4.2 Outlier	53
Gambar 4.3 Visualisasi VO ₂ Max Berdasarkan Gender	54
Gambar 4.4 Scatter plot Power Otot Tungkai dan Berat Badan	55
Gambar 4.5 Boxplot Kecepatan Berdasarkan Kelompok	56
Gambar 4.6 Heatmap Korelasi Fitur	57
Gambar 4.7 Evaluasi Model Random Forest	63
Gambar 4.8 Confusion Matrix Random Forest	63
Gambar 4.9 Evaluasi Model Logistik Regresi	64
Gambar 4.10 Confusion Matrix Logistik Regresi	64
Gambar 4.11 Hasil Klasifikasi	65
Gambar 4.12 Confusion Matrix Random Forest Stl Hyper Parameter Tuning	74
Gambar 4.13 Evaluasi Model Logistik Regresi Stl Hyper Parameter Tuning	75
Gambar 4.14 Confusion Matrix Logistik Regresi stl Hyper Parameter Tuning	75

DAFTAR LAMBANG DAN SINGKATAN

RF	Random Forest
LR	Logistic Regression
VO ₂ Max	Volume Oksigen Maksimum
KNN	K-Nearest Neighbors
F1-score	Harmonic Mean dari Precision dan Recall



DAFTAR ISTILAH

Overfitting	Kondisi di mana model terlalu cocok dengan data latih dan kurang mampu menggeneralisasi data baru
Confusion Matrix	Matriks yang digunakan untuk mengevaluasi performa klasifikasi model
Precision	Proporsi prediksi positif yang benar terhadap total prediksi positif
Recall	Proporsi data positif yang berhasil diprediksi benar oleh model
Bootstrap	Teknik resampling dalam statistika
Feature Selection	Proses memilih fitur yang paling berpengaruh dalam model
Heatmap	Visualisasi matriks yang menunjukkan korelasi antar fitur dalam data
Logistic Regression	Model statistik yang digunakan untuk klasifikasi

INTISARI

Atlet Porda Shorinji Kempo menghadapi tantangan dalam mengoptimalkan performa mereka, terutama dalam menentukan kategori kemampuan yang sesuai berdasarkan aspek fisik seperti endurance, speed, dan power. Evaluasi yang masih banyak dilakukan secara manual dapat menghasilkan subjektivitas dalam penilaian dan kurangnya ketepatan dalam pembinaan atlet. Oleh karena itu, tujuan penelitian ini adalah untuk mengembangkan sistem klasifikasi berbasis machine learning guna mengelompokkan atlet ke dalam tiga kategori: Beginner, Intermediate, dan Advanced, sehingga pelatih dapat memperoleh rekomendasi yang lebih objektif dan berbasis data.

Penelitian ini menggunakan algoritma Random Forest dan Regresi Logistik Multinomial untuk membangun model klasifikasi. Data atlet yang dikumpulkan, dilakukan preprocessing seperti Konversi format data, Label Encoding, Outlier Detecstion, Fitur selection dan Praperhitungan fitur, setelah itu dilakukan clustering untuk menentukan label yang di tandai menjadi tiga kategori yaitu Beginner, Intermediate dan Advance. Model dikembangkan dan dan dinilai dengan metrik akurasi, precision, recall, dan F1-score. Selain itu akan di lakukan hyper patameter tuning juga untuk meningkatkan perfoma model menggunakan great search hasil yang di tunjukan bahwa Random Forest memberikan akurasi tinggi sebesar 99%, dibandingkan dengan Regresi Logistik yang mencapai 82%. Analisis menunjukkan bahwa faktor Speed dan Endurance memiliki pengaruh dominan dalam klasifikasi atlet.

Penelitian ini dapat dimanfaatkan oleh pelatih dalam menyusun program latihan yang lebih efektif dan berbasis data. Dengan implementasi lebih lanjut, sistem ini dapat di implementasikan menjadi aplikasi berbasis web atau mobile yang memudahkan evaluasi atlet secara real-time.

Kata kunci: Machine Learning, Random Forest, Regresi Logistik, Klasifikasi Atlet, Shorinji Kempo.

ABSTRACT

Shorinji Kempo Porda athletes face challenges in optimizing their performance, particularly in determining the appropriate skill category based on physical aspects such as endurance, speed, and power. Evaluations that are still largely conducted manually can lead to subjectivity in assessment and a lack of precision in athlete development. Therefore, the aim of this research is to develop a machine learning-based classification system to group athletes into three categories: Beginner, Intermediate, and Advanced, enabling coaches to obtain more objective and data-driven recommendations. This research employs the Random Forest and Multinomial Logistic Regression algorithms to build the classification model. Athlete data is collected and undergoes preprocessing steps such as data format conversion, label encoding, outlier detection, feature selection, and feature pre-calculation. Following this, clustering is performed to determine the labels, which are categorized into three groups: Beginner, Intermediate, and Advanced. The model is developed and evaluated using accuracy, precision, recall, and F1-score metrics. The results show that Random Forest achieves a high accuracy of 99%, compared to Logistic Regression, which reaches 82%. The analysis indicates that Speed and Endurance are the dominant factors influencing athlete classification. This research can be utilized by coaches to design more effective and data-driven training programs. With further implementation, this system can be developed into a web or mobile-based application, facilitating real-time athlete evaluation.

Keywords: Machine Learning, Random Forest, Logistic Regression, Athlete Classification, Shorinji Kempo