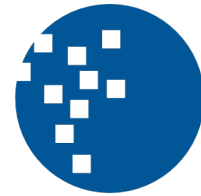


**PREDIKSI JUARA DUNIA PEMBALAP FORMULA 1 TAHUN
2025 MENGGUNAKAN XGBOOST**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

**AULIYAA VISHWAKARMA HESTIA
00000059515**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

**PREDIKSI JUARA DUNIA PEMBALAP FORMULA 1 TAHUN
2025 MENGGUNAKAN XGBOOST**



**AULIYAA VISHWAKARMA HESTIA
00000059515**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Auliyaa Vishwakarma Hestia

Nomor Induk Mahasiswa : 00000059515

Program Studi : Informatika

Skripsi dengan judul:

Prediksi Juara Dunia Pembalap Formula 1 Tahun 2025 Menggunakan XGBoost

merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.

Tangerang, 03 Juli 2025



(Auliyaa Vishwakarma Hestia)

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

HALAMAN PENGESAHAN

Skripsi dengan judul

PREDIKSI JUARA DUNIA PEMBALAP FORMULA 1 TAHUN 2025 MENGUNAKAN XGBOOST

oleh

Nama : Auliyaa Vishwakarma Hestia
NIM : 00000059515
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Jumat, 11 Juli 2025

Pukul 08.00 s/s 10.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

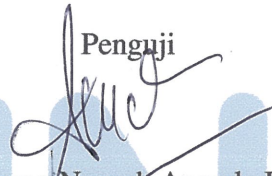
Ketua Sidang



(Eunike Endariahna Surbakti, S.Kom.,
M.T.I.)

NIDN: 0322099401

Penguji



(Anak Agung Ngurah Ananda Kusuma,
B.Eng., M.Eng., Ph.D.)

NIDK: 8984101024

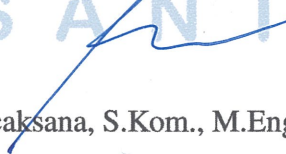
Pembimbing



(Aditiyawan, S.Kom., M.Si.)

NIDK: 8994550022

Ketua Program Studi Informatika,



(Arya Wicaksana, S.Kom., M.Eng.Sc., OCA)

NIDN: 0315109103

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini:

Nama : Auliyaa Vishwakarma Hestia
NIM : 00000059515
Program Studi : Informatika
Jenjang : S1
Judul Karya Ilmiah : Prediksi Juara Dunia Pembalap
Formula 1 Tahun 2025 Menggunakan
XGBoost

Menyatakan dengan sesungguhnya bahwa saya bersedia (**pilih salah satu**):

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- ☐ Lainnya, pilih salah satu:
 - Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - Embargo publikasi karya ilmiah dalam kurun waktu tiga tahun.

Tangerang, 03 Juli 2025

Yang menyatakan



Auliyaa Vishwakarma Hestia

**Jika tidak bisa membuktikan LoA jurnal/HKI, saya bersedia mengizinkan penuh karya ilmiah saya untuk dipublikasikan ke KC UMN dan menjadi hak institusi UMN.

HALAMAN PERSEMBAHAN / MOTTO

"You cannot deliver to your maximum if you're thinking 'if I do an error, it might close a door to somewhere' or whatever."

Charles Leclerc



KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan Skripsi ini dengan judul: Prediksi Juara Dunia Pembalap Formula 1 Tahun 2025 Menggunakan XGBoost dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan laporan tugas akhir ini, sangatlah sulit bagi saya untuk menyelesaikan laporan tugas akhir ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Arya Wicaksana, S.Kom., M.Eng.Sc., OCA, selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Bapak Aditiyawan, S.Kom., M.Si., sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.
6. Formula 1 sebagai sumber inspirasi dan hiburan yang telah menumbuhkan minat penulis dalam bidang analisis data olahraga serta menjadi motivasi dalam menyelesaikan penelitian ini.
7. Teman-teman saya Acil, Adhy, Chika, Michele, Indira, Nabilah, Ragil, dan Renita yang telah memberikan semangat, dukungan, dan keceriaan selama masa perkuliahan hingga penyusunan tugas akhir ini. Terima kasih atas setiap motivasi, kebersamaan, dan tawa yang menjadi bagian tak terlupakan dalam perjalanan akademik ini.

8. Pihak-pihak lain yang tidak dapat disebutkan satu per satu, namun telah memberikan bantuan, dukungan, maupun inspirasi secara langsung maupun tidak langsung dalam proses penyusunan laporan ini.

Disadari bahwa tugas akhir ini masih jauh dari sempurna, baik dari segi isi maupun penyajiannya. Semoga karya ilmiah ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan, khususnya dalam bidang analisis data dan penerapan algoritma *machine learning* di dunia olahraga, serta dapat menjadi referensi bagi penelitian selanjutnya.

Tangerang, 03 Juli 2025



Auliya Vishwakarma Hestia



PREDIKSI JUARA DUNIA PEMBALAP FORMULA 1 TAHUN 2025 MENGUNAKAN XGBOOST

Auliyaa Vishwakarma Hestia

ABSTRAK

Formula 1 (F1) merupakan salah satu olahraga balap mobil paling bergengsi di dunia yang memadukan kecepatan, strategi, dan teknologi tinggi. Prediksi juara dunia atau *World Driver Champion* (WDC) menjadi tantangan menarik dalam era analitik olahraga modern. Penelitian ini bertujuan membangun model prediksi WDC musim 2025 dengan menggunakan algoritma *Extreme Gradient Boosting* (XGBoost), yang dikenal unggul dalam menangani data kompleks dan tabular. Dataset diperoleh dari Jolpi API, mencakup data historis F1 dari musim 1950 hingga 2025, termasuk hasil balapan, sesi kualifikasi, klasemen pembalap, dan data tim. Data kemudian diproses melalui tahapan *merging*, *cleaning*, *feature engineering*, dan transformasi, lalu dibagi menjadi data latih dan data uji. Model dikembangkan dan disesuaikan secara optimal melalui proses *tuning hyperparameter* menggunakan pustaka Optuna, untuk menemukan kombinasi parameter terbaik yang meningkatkan akurasi prediksi. Evaluasi model dilakukan menggunakan metrik *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan *R² Score*. Hasil evaluasi pada data musim 2025 menunjukkan bahwa model memiliki nilai MAE sebesar 1.0610, RMSE sebesar 1.3634, dan skor *R²* sebesar 0.9441. Hasil tersebut menunjukkan bahwa model XGBoost mampu memprediksi posisi akhir pembalap dengan akurasi tinggi. Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan sistem prediksi berbasis data untuk olahraga otomotif, khususnya Formula 1.

Kata kunci: Formula 1, Juara Dunia Pembalap, Prediksi, XGBoost

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

FORMULA 1 WORLD DRIVER CHAMPION 2025 PREDICTION USING XGBOOST

Auliyaa Vishwakarma Hestia

ABSTRACT

Formula 1 (F1) is one of the most prestigious motorsport events in the world, combining speed, strategy, and high technology. Predicting the World Driver Champion (WDC) is an intriguing challenge in the era of modern sports analytics. This research aims to build a model to predict the WDC for the 2025 season using the Extreme Gradient Boosting (XGBoost) algorithm, known for its superiority in handling complex and tabular data. The dataset is obtained from the Jolpi API, covering historical F1 data from the 1950 to 2025 seasons, including race results, qualifying sessions, driver standings, and team data. The data is then processed through merging, cleaning, feature engineering, and transformation stages, and then split into training and testing data. The model is developed and optimally adjusted through hyperparameter tuning using the Optuna library to find the best parameter combinations that increase prediction accuracy. The model evaluation was conducted using the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2 Score metrics. The evaluation results for the 2025 season data indicate that the model has an MAE value of 1.0610, an RMSE of 1.3634, and an R^2 score of 0.9441. These results demonstrate that the XGBoost model is capable of predicting drivers' final positions with high accuracy. This research is expected to contribute to the development of data-driven prediction systems for motorsports, particularly Formula 1.

Keywords: *Formula 1, Prediction, World Driver Champion, XGBoost*

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	iv
HALAMAN PERSEMBAHAN/MOTO	v
KATA PENGANTAR	vi
ABSTRAK	viii
ABSTRACT	ix
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR KODE	xiv
DAFTAR RUMUS	xv
DAFTAR LAMPIRAN	xvi
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Permasalahan	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB 2 LANDASAN TEORI	6
2.1 Formula 1	6
2.2 Machine Learning	10
2.3 Extreme Gradient Boosting (XGBoost)	11
2.4 Evaluation Metrics	15
2.4.1 Mean Absolute Error (MAE)	15
2.4.2 Root Mean Square Error (RMSE)	16
2.4.3 R ² Score (Koefisien Determinasi)	16
2.4.4 Cross-Validation	16
BAB 3 METODOLOGI PENELITIAN	18
3.1 Studi Literatur	19
3.2 Pengumpulan Data	20
3.3 Analisis Data	21
3.4 Pre-processing Data	21
3.5 Perancangan Model	23
3.5.1 Train-test Split	24
3.5.2 Cross-Validation	25
3.5.3 Hyperparameter Tuning	25
3.6 Pengujian dan Evaluasi Model	26
BAB 4 HASIL DAN DISKUSI	28
4.1 Spesifikasi Sistem	28
4.2 Deskripsi Dataset	28
4.3 Import dan Merge Dataset	29
4.4 Pre-processing Data	31
4.4.1 Data Merging	31
4.4.2 Data Cleaning dan Feature Engineering	35
4.4.3 Feature Importance dan Feature Selection	39

4.4.4	Feature Transformation	41
4.5	Perancangan Model	47
4.5.1	Train-test Split Data	47
4.5.2	Initialize Base Model dan Training Model	48
4.5.3	Cross-Validation	51
4.5.4	Hyperparameter Tuning	52
4.6	Pengujian dan Evaluasi Model	54
BAB 5	SIMPULAN DAN SARAN	59
5.1	Simpulan	59
5.2	Saran	60
DAFTAR PUSTAKA		61



DAFTAR TABEL

Tabel 2.1	<i>Feature race</i> dan <i>sprint race</i> points	6
Tabel 2.2	Evolusi sistem poin Formula 1 dari masa ke masa	7
Tabel 2.3	Sistem poin ganda pada <i>Grand Prix</i> Abu Dhabi 2014	8
Tabel 3.1	Ringkasan studi terdahulu prediksi juara dunia pembalap Formula 1 dan <i>research gap</i>	19
Tabel 4.1	Pembagian <i>train data</i> dan <i>test data</i> berdasarkan skenario	48
Tabel 4.2	Dataset Sederhana untuk Simulasi	49
Tabel 4.3	Gradient dan Hessian Awal	49
Tabel 4.4	<i>Gradient</i> dan <i>Hessian</i> Iterasi Kedua	50
Tabel 4.5	Hasil evaluasi pengujian dengan <i>best model</i>	54
Tabel 4.6	Perbandingan hasil evaluasi pengujian data <i>train</i>	57



DAFTAR GAMBAR

Gambar 2.1	Hubungan antara <i>Artificial Intelligence</i> , <i>Machine Learning</i> , dan <i>Deep Learning</i>	11
Gambar 2.2	Proses pelatihan model pada XGBoost (<i>Extreme Gradient Boosting</i>)	12
Gambar 3.1	<i>Flowchart</i> metodologi penelitian	18
Gambar 3.2	<i>Flowchart pre-processing data</i>	22
Gambar 3.3	<i>Flowchart</i> perancangan model	23
Gambar 4.1	<i>Output</i> mengambil data klasemen akhir pembalap dari setiap musim	32
Gambar 4.2	<i>Output</i> penggabungan data klasemen akhir pembalap dengan informasi pembalap	33
Gambar 4.3	<i>Output</i> penggabungan data poin konstruktor dengan informasi pembalap	34
Gambar 4.4	<i>Output</i> hasil akhir proses <i>data merging</i>	35
Gambar 4.5	<i>Output</i> pembersihan dan penambahan fitur usia dan performa pembalap	36
Gambar 4.6	<i>Output</i> fitur pengalaman pembalap	37
Gambar 4.7	<i>Output</i> akhir pembersihan data (<i>cleaned dataset</i>)	38
Gambar 4.8	Dataset akhir setelah proses <i>data cleaning dan feature engineering</i>	38
Gambar 4.9	Matriks Korelasi Numerik	39
Gambar 4.10	<i>Output</i> standarisasi fitur numerik menggunakan <i>z-score</i> per musim	42
Gambar 4.11	<i>Output</i> transformasi fitur posisi klasemen	42
Gambar 4.12	<i>Output</i> transformasi fitur <i>driver_id</i>	44
Gambar 4.13	Hasil prediksi model sebelum didiskretisasi	55
Gambar 4.14	Hasil prediksi model setelah diskretisasi	56

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR KODE

Kode 4.1	<i>Import dataset</i>	30
Kode 4.2	<i>Merge dataset</i>	30
Kode 4.3	Mengambil data klasemen akhir pembalap dari setiap musim . .	32
Kode 4.4	Menggabungkan data klasemen akhir dengan informasi pembalap	32
Kode 4.5	Mengambil data poin konstruktor pada ronde terakhir	33
Kode 4.6	Menggabungkan perhitungan statistik performa pembalap	34
Kode 4.7	Menghitung usia dan performa pembalap dari data kualifikasi dan balapan	36
Kode 4.8	Menghitung pengalaman pembalap berdasarkan tahun debut . . .	37
Kode 4.9	Pembersihan data duplikat dan nilai tidak valid	37
Kode 4.10	Standardisasi fitur numerik menggunakan <i>z-score</i> per musim . .	41
Kode 4.11	Transformasi fitur posisi klasemen dengan ordinal encoding . . .	42
Kode 4.12	Konversi fitur <i>driver_id</i> menggunakan Label Encoding	43
Kode 4.13	Normalisasi semua fitur numerik menggunakan <i>Standard Scaler</i> .	44
Kode 4.14	<i>Split dataset</i>	47
Kode 4.15	Inisialisasi model XGBoost dengan parameter <i>default</i>	48
Kode 4.16	Proses <i>cross-validation</i>	51
Kode 4.17	Proses tuning hyperparameter	52
Kode 4.18	Pelatihan model dengan parameter terbaik	54



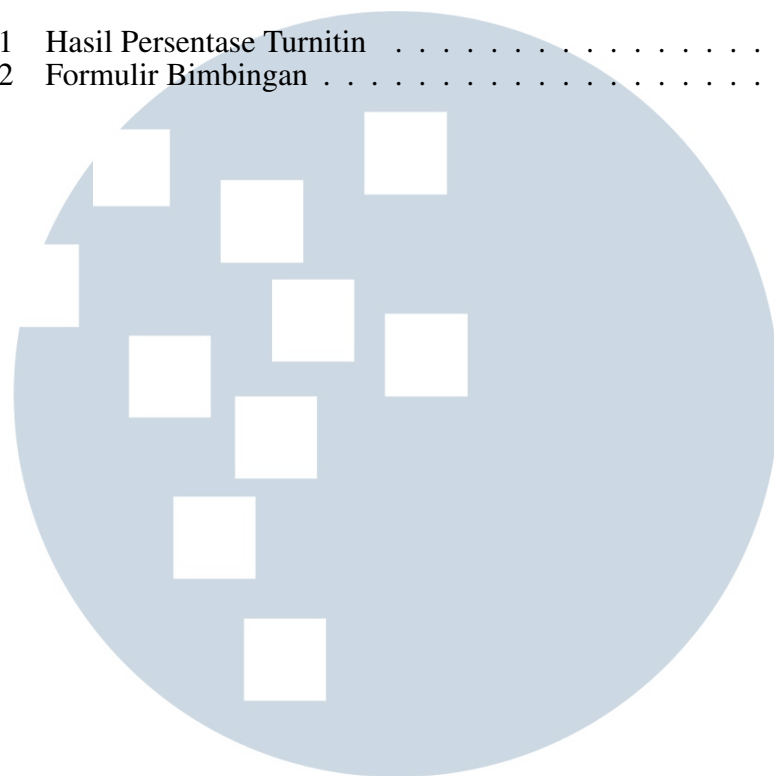
DAFTAR RUMUS

Rumus 2.1	Fungsi prediksi pada XGBoost	13
Rumus 2.2	Fungsi objektif pada XGBoost	13
Rumus 2.3	Fungsi regulasi pada XGBoost	14
Rumus 2.4	Perhitungan Gain pada XGBoost	14



DAFTAR LAMPIRAN

Lampiran 1	Hasil Persentase Turnitin	65
Lampiran 2	Formulir Bimbingan	75



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA