

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Kejuaraan balap mobil tingkat tertinggi yang menggabungkan kecepatan, strategi, dan teknologi mutakhir atau disebut dengan Formula 1 (F1) [1]. Sejak dimulainya pada tahun 1950, F1 telah berkembang menjadi salah satu olahraga paling kompleks dan berbasis data di dunia [2]. Balapan F1 tidak hanya ditentukan di lintasan saja, tetapi juga di balik layar-mulai dari desain mobil hingga strategi tim dan analisis data *real-time*. Seiring meningkatnya digitalisasi, tim F1 mengandalkan ribuan titik data dari sensor kendaraan untuk memantau performa secara *real-time* dan membuat keputusan strategis selama balapan [3, 4]. Proses ini mencakup pemantauan ban, konsumsi bahan bakar, keausan komponen, serta pengaturan aerodinamika yang semuanya dianalisis secara komputasional untuk mengoptimalkan hasil balapan [5].

Kompetisi di Formula 1 berlangsung dalam satu musim penuh, dengan pembalap mengumpulkan poin berdasarkan posisi *finish* di setiap balapan [6]. Pembalap dengan akumulasi poin tertinggi pada akhir musim akan menjadi juara dunia. Banyaknya variabel yang memengaruhi hasil akhir-mulai dari performa kendaraan, strategi tim, hingga keberlangsungan teknis-upaya untuk memprediksi siapa yang menjadi juara dunia menjadi tantangan menarik di era *big data* [7].

Regulasi teknis yang ditetapkan oleh FIA (*Fédération Internationale de l'Automobile*) juga memainkan peran besar dalam membentuk dinamika kompetisi. Misalnya, implementasi *Drag Reduction System* (DRS) terbukti meningkatkan peluang menyalip, sehingga membuat balapan lebih kompetitif dan hasilnya semakin sulit diprediksi [8, 6].

Sejak pertama kali diselenggarakan pada tahun 1950, Formula 1 telah tumbuh menjadi olahraga balap mobil paling prestisius di dunia [9]. Perkembangannya tidak hanya mencakup teknologi kendaraan dan strategi balapan, tetapi juga aspek manajerial dan analisis data yang semakin maju [10]. Seiring waktu, F1 telah menjadi ladang kompetisi yang tak hanya mengandalkan kemampuan pembalap tetapi juga kekuatan data dan prediksi performa [11].

Dominasi Max Verstappen dalam beberapa musim terakhir menjadi bukti nyata dari kompleksitas dan dinamika kompetisi modern. Ia berhasil meraih empat

gelar juara dunia secara berturut-turut hingga musim 2024, menjadikannya salah satu pembalap paling sukses dalam sejarah F1 modern [12]. Rekor ini mendekati pencapaian Michael Schumacher, yang meraih lima gelar beruntun pada era awal 2000-an [13]. Dengan performa luar biasa yang konsisten, Verstappen kini menjadi kandidat kuat peraih gelar kelima di musim 2025, menciptakan peluang besar untuk eksplorasi prediksi berbasis *machine learning*.

Tren penggunaan data juga meluas di industri olahraga lainnya, seperti yang terlihat dalam penjualan klub basket legendaris Los Angeles Lakers oleh keluarga buss kepada Mark Walter, dengan valuasi mencapai 10 miliar dollar AS [14]. Transaksi ini mencerminkan pentingnya analisis data dalam pengambilan keputusan bisnis dan strategi jangka panjang dalam dunia olahraga profesional.

Fenomena-fenomena tersebut memperlihatkan bagaimana teknologi dan data kini menjadi elemen kunci dalam dunia olahraga, baik dalam meraih kemenangan di lintasan maupun dalam mengelola organisasi olahraga global. Oleh karena itu, penelitian ini bertujuan untuk membangun model prediksi juara dunia pembalap Formula 1 tahun 2025 menggunakan algoritma XGBoost, yang mampu menangani data kompleks dan menghasilkan prediksi yang akurat.

Dalam lingkungan kompetitif yang kompleks ini, pendekatan *data-driven* menjadi sangat penting [2, 11]. Berbagai studi telah menunjukkan bahwa algoritma *machine learning*, khususnya *Extreme Gradient Boosting* (XGBoost), memiliki kinerja tinggi dalam memprediksi hasil kompetisi olahraga. Misalnya, XGBoost telah digunakan secara efektif dalam memprediksi hasil pertandingan kriket [15], tenis [16], dan strategi taruhan balapan kuda [17]. Penelitian Patil et al. (2020) menunjukkan bahwa model statistik dan analitik dapat digunakan untuk mengidentifikasi variabel teknis mobil F1 yang berkorelasi kuat dengan total poin pembalap di akhir musim. Dengan pendekatan *principal component analysis* dan regresi, studi ini berhasil mengurangi kompleksitas data dan mengungkap faktor paling relevan dalam performa pembalap [18].

Seiring meningkatnya peran data dalam mendukung strategi tim dan pengambilan keputusan di dunia olahraga, kebutuhan akan model prediktif yang akurat dan andal menjadi semakin penting. Informasi hasil prediksi dapat dimanfaatkan oleh berbagai pihak, mulai dari penggemar dan analis olahraga hingga sponsor dan media, untuk menyusun strategi komunikasi, konten, dan promosi yang lebih relevan. Di sisi lain, pengembangan model prediksi juga mendorong penerapan algoritma *machine learning* seperti XGBoost ke dalam domain kompetitif, sehingga membuka peluang inovasi pada analitik olahraga.

Maka dari itu, penelitian ini bertujuan untuk memanfaatkan algoritma XGBoost untuk memprediksi pembalap yang berpotensi menjadi juara dunia Formula 1 musim 2025, berdasarkan data hasil kualifikasi, posisi balapan, klasemen pembalap, serta atribut tim dan kendaraan. Fokus penelitian ini adalah pada prediksi sementara hingga ronde ke-10 musim 2025, menggunakan data dari Jolpi API yang mengakses *database* resmi Formula 1.

1.2 Rumusan Masalah

1. Bagaimana membangun model prediksi menggunakan algoritma XGBoost untuk menentukan pembalap yang berpotensi menjadi juara dunia Formula 1 pada musim 2025?
2. Bagaimana performa model XGBoost dalam memprediksi hasil klasemen pembalap berdasarkan metrik evaluasi *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan R^2 Score?

1.3 Batasan Permasalahan

- Penelitian ini hanya memprediksi sementara klasemen juara dunia pembalap Formula 1 pada *round* ke-10 musim 2025.
- Data yang digunakan bersumber dari Jolpi API (<https://api.jolpi.ca/ergast/>), yang menyediakan data historis dan *real-time* terkait balapan Formula 1.
- Model hanya memanfaatkan data numerik dan kategorikal dari hasil balapan, kualifikasi, klasemen pembalap, serta informasi tim dan pembalap, tanpa mempertimbangkan faktor eksternal seperti kondisi cuaca, penalti, atau insiden teknis selama balapan.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Membangun model prediksi menggunakan algoritma XGBoost untuk memperkirakan pembalap yang berpotensi menjadi juara dunia Formula 1 musim 2025.
2. Mengevaluasi performa model prediksi menggunakan metrik *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan R^2 Score.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

- Memberikan informasi prediksi pembalap yang berpotensi menjadi juara dunia Formula 1 musim 2025, yang dapat dimanfaatkan oleh penggemar, media olahraga, sponsor, hingga penyelenggara konten untuk kebutuhan hiburan, strategi pemasaran, atau pembuatan narasi prediktif berbasis data yang lebih menarik dan relevan.
- Menambah kontribusi keilmuan dalam bidang penerapan algoritma XGBoost dan *data science* untuk analisis dan prediksi dalam konteks olahraga balap mobil.
- Menjadi dasar pengembangan model prediktif lanjutan yang lebih kompleks dan akurat dengan mempertimbangkan lebih banyak variabel dan data historis.

1.6 Sistematika Penulisan

Sistematika penulisan laporan adalah sebagai berikut:

- Bab 1 PENDAHULUAN
Bab ini berisi Latar Belakang Penelitian, Rumusan Masalah, Tujuan Penelitian, Manfaat Penelitian, Batasan Permasalahan, dan Sistematika Penulisan.
- Bab 2 LANDASAN TEORI
Bab ini berisi teori-teori yang relevan dengan topik penelitian, seperti penjelasan mengenai ajang balap Formula 1, konsep dasar *Machine Learning*, serta algoritma *Extreme Gradient Boosting* (XGBoost) yang digunakan dalam penelitian ini.
- Bab 3 METODOLOGI PENELITIAN
Bab ini menjelaskan metodologi penelitian yang digunakan, meliputi tahapan studi literatur, pengumpulan data, analisis dan *pre-processing* data, perancangan model prediksi, serta pengujian dan evaluasi performa model yang digunakan.

- Bab 4 HASIL DAN DISKUSI

Bab ini hasil dari pemodelan prediksi yang telah dilakukan. Termasuk di dalamnya adalah hasil visualisasi data, pemilihan fitur, *hyperparameter tuning*, serta hasil prediksi dari model XGBoost yang dibangun.

- Bab 5 KESIMPULAN DAN SARAN

Bab ini berisi simpulan dari hasil penelitian yang telah dilakukan serta saran untuk pengembangan atau penelitian selanjutnya.

