



SOMAIYA
VIDYAVIHAR UNIVERSITY

K J Somaiya School of Engineering
(formerly K J Somaiya College of Engineering)

Somaiya
TRUST

Advanced ML Algorithms in Formula 1

A Comprehensive Literature Survey

vedansh savla 16010122323

Introduction & Motivation

- Machine learning is essential in F1, where marginal performance gains determine victory or defeat.
- Vast amounts of data from telemetry, simulations, and competitor analysis require advanced ML to extract meaningful insights.
- ML is leveraged to optimize aerodynamics, predict tire strategies, analyze driver behavior, and refine overall race performance.
- The ability to interpret patterns and predict outcomes provides a crucial strategic edge in this technology-driven sport.



Evolution in F1

The sport has evolved from rudimentary statistical analysis to employing sophisticated, real-time AI-driven approaches.

- Deep learning techniques refine car designs and optimize mechanical components.
- Cloud computing enables seamless data sharing and collaborative analysis across global teams.

Key Challenges

Implementing ML in F1 is not without obstacles:

- **Data Volume:** Managing and interpreting the sheer volume and complexity of data generated.
- **Real-time Processing:** Split-second decisions require instantaneous data processing.
- **Secrecy:** Data security is paramount, as insights are a core strategic edge.
- **Regulation:** FIA rules impose constraints on how ML can be applied.

Methodology: Four Key Research Areas

Methodology 1: Aerodynamics



ML for Motorsport Aerodynamics

"Machine learning for Motorsport Aerodynamics" (Watts et al., 2019)

This study investigates integrating Computational Fluid Dynamics (CFD) simulations with ML algorithms to enhance aerodynamic efficiency.



Key Techniques

- **Principal Component Analysis (PCA):** Used to reduce the dimensionality of aerodynamic parameters, focusing on the most critical factors.
- **Artificial Neural Networks (ANNs):** Employed to predict aerodynamic efficiency based on various input parameters from simulations.

Methodology 2: Tire Strategy



Predicting F1 Compound Decisions

"From Data to Podium" (Leischner, 2023)

This paper explores developing models to predict the ideal tire compound choices for different race scenarios, a critical aspect of race strategy.



Key Techniques

Models were trained on historical data, weather, and track conditions using:

- **Decision Trees & Random Forests:** Used for classification to determine the optimal tire.
- **Gradient Boosting Machines (GBM):** Employed for improved accuracy in prediction.

Methodology 3 & 4: Drivers & Learning

Paper 3: Cluster Analysis of F1 Drivers

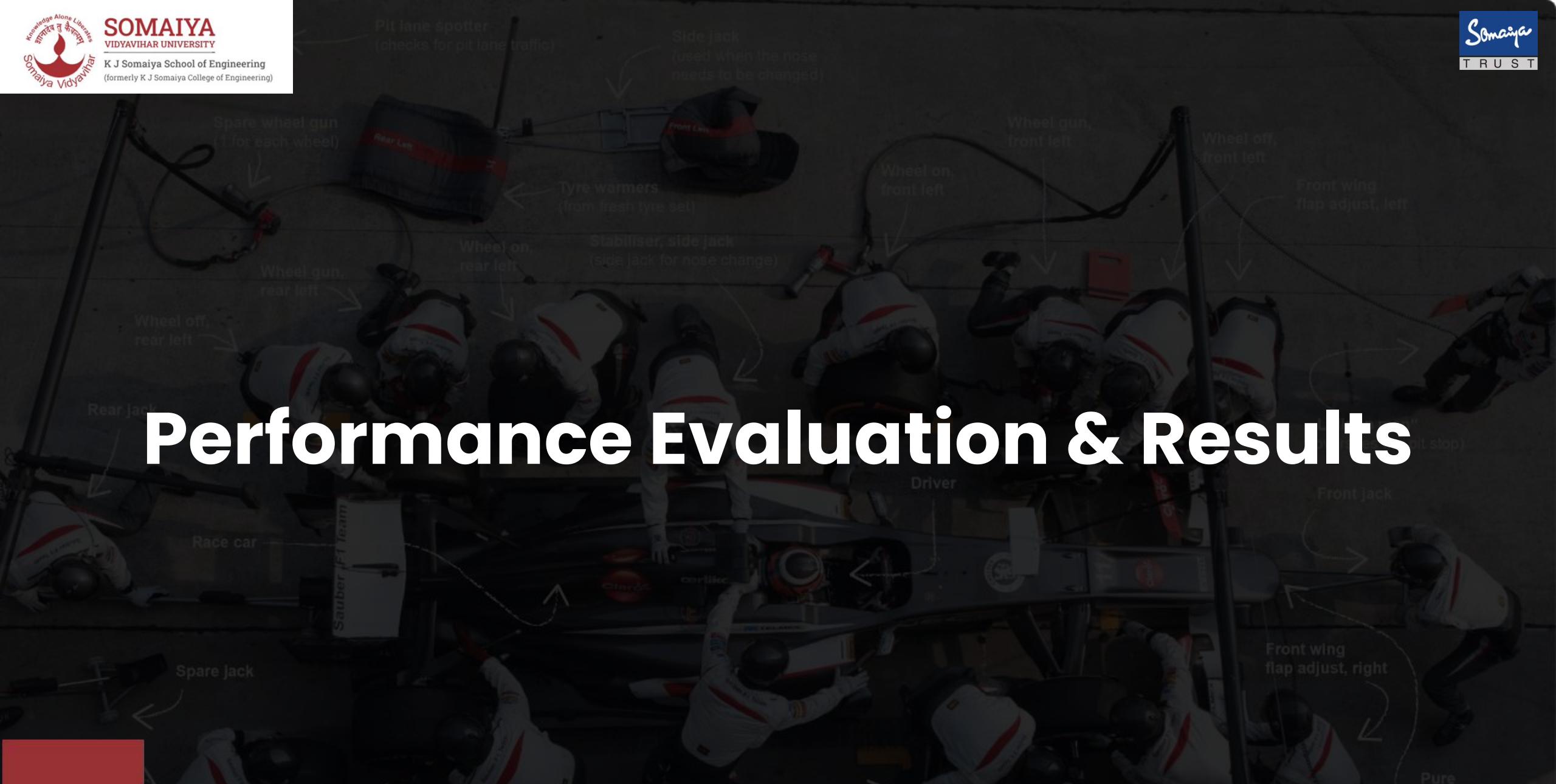
"Racing Your Rival" (Syracuse University, 2021)

- **K-Means & Hierarchical Clustering:** Used to group drivers based on performance metrics, racing styles, and consistency.
- This enables teams to assess threats and develop targeted strategies.

Paper 4: Competitive Learning

"Learning from... Success and Failure" (INFORMS, 2022)

- **Bayesian Learning:** Updates belief probabilities based on race outcomes to refine strategies.
- **Q-learning (Reinforcement):** Applied to optimize race strategy and decision-making under uncertainty.



Performance Evaluation & Results

Results: Aerodynamics & Tire Strategy

Aerodynamics

Machine learning significantly enhances aerodynamic efficiency.

- PCA successfully reduced data dimensionality, allowing engineers to focus on critical variables.
- Clustering identified optimal aerodynamic configurations, leading to better car performance and fuel efficiency.

Tire Decisions

The models proved highly effective in predicting tire choice.

- Random Forest and XGBoost models achieved an **accuracy of 85%** in predicting optimal compounds.
- Feature analysis revealed **track temperature and pit stop strategy** as the most influential factors.

Results: Driver Analysis & Learning

Driver Clustering

Analysis revealed distinct groups of drivers (e.g., elite, midfield, strugglers).

- Top-tier drivers exhibit unique and superior patterns in race execution, tire management, and overtaking.
- This aids in talent scouting and refining strategies based on driver tendencies.

Competitive Learning

Teams that learn from data consistently perform better.

- NLP of race debriefs and radio shows that teams systematically analyze competitor strategies.
- Data-driven learning and dynamic adaptation contribute to long-term success.

Comparative Analysis: Discussion

Strengths & Applications

- **Strengths:** High predictive accuracy (Paper 2), effective dimensionality reduction (Paper 1), provides clear, actionable segments (Paper 3).
- **Applications:**
 - Real-time strategy calls (Tires).
 - Optimizing pre-race car setup (Aero).
 - Talent scouting & rival analysis (Drivers).

Limitations & Improvements

- **Limitations:** Difficulty accessing proprietary team data, models may not be fast enough for real-time decisions, reliance on historical data.
- **Potential Improvements:**
 - Integrate live telemetry data streams.
 - Use Reinforcement Learning for dynamic strategy.
 - Develop XAI models for better interpretability.

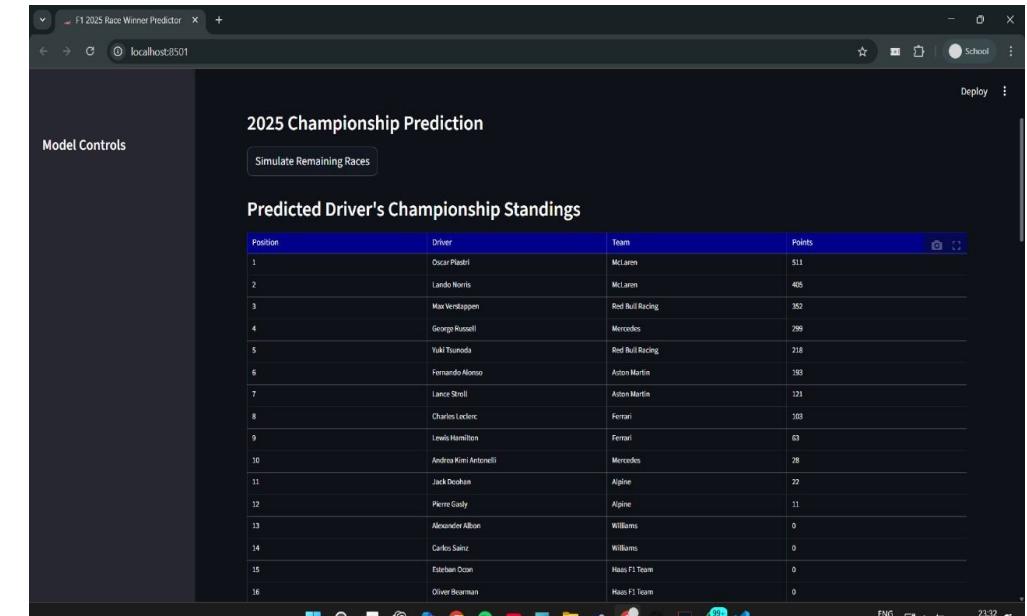
Prototype: F1 Race Winner Predictor

Model Performance

A prototype application was built to predict F1 race winners based on historical data.

- Test Accuracy: 96.87%
- Test ROC AUC: 98.46%
- Validation Accuracy: 99.09%
- Validation ROC AUC: 99.85%

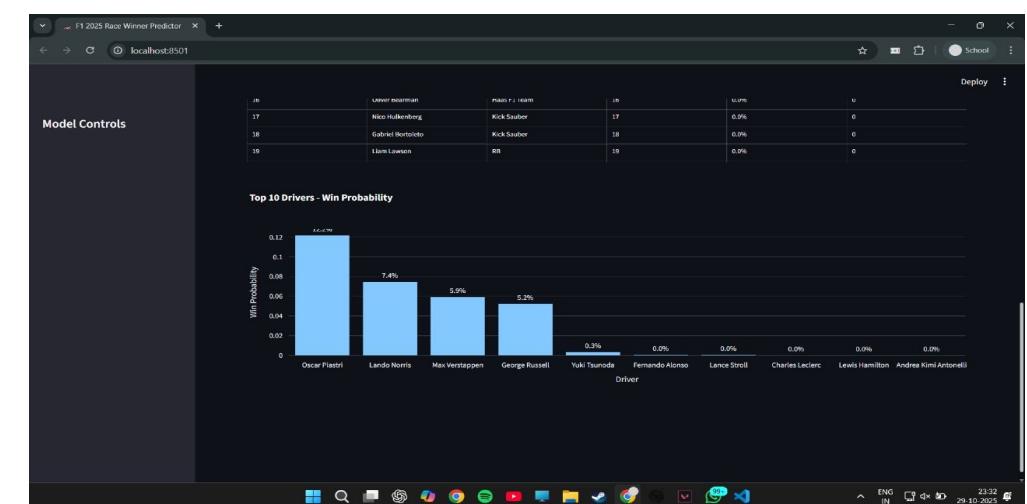
The model demonstrates high accuracy and robust performance, providing a strong foundation for predictive analysis.



2025 Championship Prediction

Predicted Driver's Championship Standings

Position	Driver	Team	Points
1	Oscar Piastri	McLaren	511
2	Lando Norris	McLaren	405
3	Max Verstappen	Red Bull Racing	352
4	George Russell	Mercedes	299
5	Yuki Tsunoda	Red Bull Racing	218
6	Fernando Alonso	Aston Martin	181
7	Lance Stroll	Aston Martin	121
8	Charles Leclerc	Ferrari	103
9	Lewis Hamilton	Ferrari	63
10	Andrea Kimi Antonelli	Mercedes	28
11	Jack Doohan	Alpine	22
12	Pierre Gasly	Alpine	11
13	Alexander Albon	Williams	0
14	Carlos Sainz	Williams	0
15	Esteban Ocon	Haas F1 Team	0
16	Oliver Bearman	Haas F1 Team	0



Model Controls

Top 10 Drivers - Win Probability

Rank	Driver	Team	Win Probability
1	Oscar Piastri	Kick Sauber	0.96
2	Nico Hulkenberg	Kick Sauber	0.96
3	Gabriel Monteiro	Kick Sauber	0.96
4	Liam Lawson	NS	0.96
5			0.96
6			0.96
7			0.96
8			0.96
9			0.96
10			0.96

Win Probability

Driver

0.12
0.1
0.08
0.06
0.04
0.02
0

Oscar Piastri Lando Norris Max Verstappen George Russell Yuki Tsunoda Fernando Alonso Lance Stroll Charles Leclerc Lewis Hamilton Andrea Kimi Antonelli

Model Performance

Accuracy Metrics

Train Accuracy

98.68%

Validation Accuracy

99.09%

Test Accuracy

96.87%

ROC AUC Scores

Train ROC AUC

99.69%

Validation ROC AUC

99.85%

Test ROC AUC

98.46%

Test Set Metrics

Test Precision

69.57%

Test Recall

66.67%

Test F1 Score

68.09%

Model Performance

Accuracy Metrics

Train Accuracy

98.68%

Validation Accuracy

99.09%

Test Accuracy

96.87%

ROC AUC Scores

Train ROC AUC

99.69%

Validation ROC AUC

99.85%

Test ROC AUC

98.46%

Test Set Metrics

Test Precision

69.57%

Test Recall

66.67%

Test F1 Score

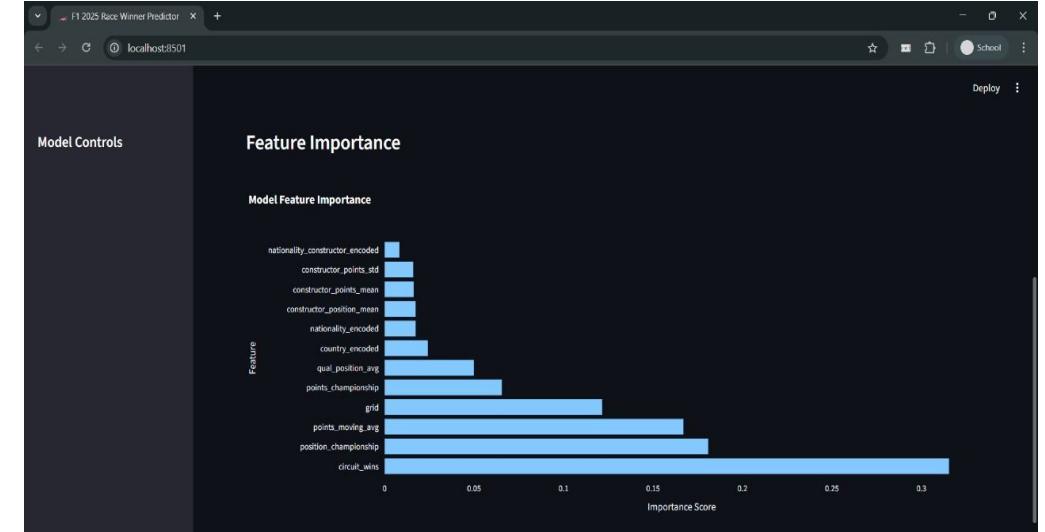
68.09%

Prototype: Predictions & Feature Analysis



The screenshot shows the 'Model Controls' section with the title 'F1 2025 Race Winner Predictor'. It displays 'Predicting Formula 1 race winners using machine learning'. Under 'Model Training', '2025 Predictions' is selected. In the '2025 Race Predictions' section, the 'Select Circuit' dropdown is set to 'São Paulo Grand Prix - São Paulo (9 Nov)'. A checkbox 'Modify Grid Positions' is checked. Below it is a button 'Predict Race Results'. The table 'Predicted Race Results for São Paulo Grand Prix - São Paulo (9 Nov)' lists the top 5 drivers with their teams and grid positions, along with their win probability and championship points.

Position	Driver	Team	Grid	Win Probability	Championship Points
1	Oscar Piastri	McLaren	8	13.2%	25
2	Lando Norris	McLaren	7	7.4%	24.5
3	Max Verstappen	Red Bull Racing	1	5.9%	24
4	George Russell	Mercedes	5	5.2%	23
5	Yuki Tsunoda	Red Bull Racing	2	0.9%	0



Race Outcome Prediction

The tool predicts race results, win probability, and expected points for specific circuits, like São Paulo.

Feature Importance

Analysis shows 'Circuit Wins' and 'Championship Position' are the most significant factors in predicting a race winner.

Future Scope & Innovations

The Future of ML in F1



Real-Time Aero

Using live telemetry to dynamically adjust wing angles, DRS configurations, and airflow management during a race.



AI-Driven CFD

Using deep neural networks as "surrogate models" to approximate CFD outputs, rapidly testing new aero designs in minutes, not days.



Adaptive Strategy

Multi-agent reinforcement learning systems that simulate rival team responses to optimize strategies in real-time.

Conclusion

- ML has fundamentally revolutionized F1, enhancing aerodynamics, tire strategy, and driver analysis.
- Predictive modeling optimizes car design, while dynamic models refine in-race strategy.
- Clustering and learning models help teams assess talent and strategically plan against rivals.
- While challenges in data privacy and computation remain, data-driven decision-making is undeniably at the core of competitive success in Formula 1.



Thank you