

Can I Predict an F1 Winner?

Lekha Karanam



Asking the Big Question

Formula 1 isn't just speed; it's strategy, pit stops, weather, and surprise DNFs.

Can machine learning predict a winner based on patterns in the data?

Goal: Use historical data to simulate race outcomes and make it visual.



Defining My Role

I acted as the full-stack project owner, from data scraping to model deployment.

Tasks included:



Data collection & preprocessing

Feature engineering

Model development
(Random Forest)

Dashboard design in Streamlit

Gathering the Data

- Source: Ergast API (2018–2024 seasons)
- Extracted:



Race results



Drivers



Constructors



Circuits

- Created 15,000+ driver-race rows after merging & cleaning

Designing the Right Features

Engineered 5 key features:



Grid Position:
Starting
advantage



Average Finish:
Season
consistency



Win Ratio:
Recent form



Constructor Avg.
Points:
Car/team strength



Circuit-Specific
History:
Track familiarity

Training the Model

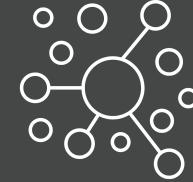
Model: Random Forest Classifier



Handles nonlinear
interactions



Built-in feature
importance



Robust to outliers

- Achieved approx. 85% accuracy
- Top predictors: Grid position, constructor stats, circuit history

Making It Interactive

The screenshot shows a web application for predicting F1 race winners. On the left, there's a sidebar titled "Race Info" with dropdowns for "Select Circuit" (Shanghai) and "Weather" (Dry). Below it is a section titled "Driver Stats" with a dropdown for "Select Driver" (Lewis Hamilton). Three horizontal sliders show "Qualifying Position" (5), "Average Finish This Season" (2.01), and "Season Win Ratio" (0.09). The main content area has a title "F1 Race Winner Predictor". Underneath, there's a "Prediction" section showing "Driver: Lewis Hamilton" and "Circuit: Shanghai". A green bar indicates a "Predicted Probability of Winning: 20.38%". A yellow warning box says "Low chance of winning". A dashed box labeled "Minimal, intuitive design" covers the top right of the sidebar. Another dashed box labeled "Predicts win probability" covers the bottom right of the main content. A third dashed box labeled "Real-time output with visual feedback" covers the bottom center.

Select driver + race

Race Info

Select Circuit: Shanghai

Weather: Dry

Driver Stats

Select Driver: Lewis Hamilton

Qualifying Position: 5

Average Finish This Season: 2.01

Season Win Ratio: 0.09

Minimal, intuitive design

F1 Race Winner Predictor

Prediction

Driver: Lewis Hamilton

Circuit: Shanghai

Predicted Probability of Winning: 20.38%

Low chance of winning

Predicts win probability

Real-time output with visual feedback

Reflecting on the Process

Key Takeaways

Feature engineering is 50% of model success

Prediction ≠ certainty (rookies, weather, etc.)

Streamlit makes deployment accessible

Building explainable tools > chasing perfect accuracy