



F1 Predictor



Giovanni Manfredi
0512112926

Teodoro Grauso
0512111084

Cos'è F1 Predictor?

F1 Predictor è un'applicazione web che permette di simulare un'intera gara di F1 prevedendo, sulla base di alcuni parametri modificabili, l'ordine di arrivo dei piloti e il relativo tempo impiegato.



Quali sono i parametri?

- **Circuito:** in cui i piloti gareggeranno;
- **Numero di giri:** da 1 a 100 giri;
- **Condizioni meteo:** cielo sereno, nuvoloso e pioggia
- **Piloti:** è possibile selezionare fino a 20 piloti diversi.





Panoramica Generale

Contesto

Specifica PEAS

Proprietà dell'ambiente

Contesto

La formula 1 è soggetta a continui cambiamenti regolamentari, uno dei più importanti è stato quello del 2022, il quale ha rivoluzionato completamente le monoposto.

Abbiamo deciso quindi di lavorare sul penultimo cambio regolamentare che va dal 2014 al 2022 per avere una quantità molto più significativa di dati

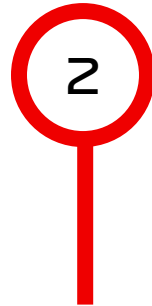
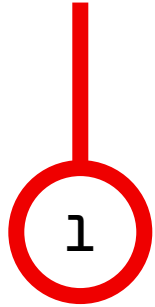
Specifica PEAS

Performance	Environment	Actuators	Sensors
Predizione accurata del risultato di una gara in un certo circuito/condizioni meteo.	Rappresentato con la web-app F1 Predictor.	Monitor per la visualizzazione della predizione effettuata.	Puntatore per selezionare il circuito, le condizioni atmosferiche e i piloti in gara.



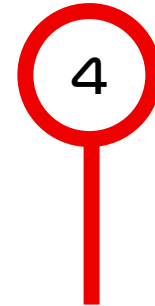
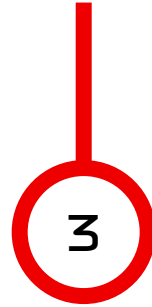
Proprietà dell'ambiente

**Completamente
osservabile**



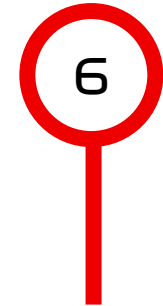
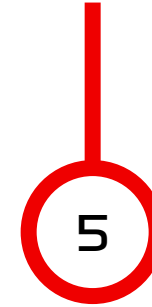
Deterministico

Sequenziale



Statico

Discreto



**Singolo
agente**

Data Understanding & Data Preparation

Dataset

Data cleaning

Data balancing

Meteo



Dataset

Il dataset utilizzato per il training del modello è disponibile sulla piattaforma Kaggle.

In questa pagina sono disponibili diversi datasets, ognuno rappresentante informazioni specifiche, come ad esempio: risultati, dati sui piloti, gare, qualifiche, costruttori e circuiti.

Dataset

Dopo l'acquisizione dei dataset, è stata eseguita un'analisi preliminare per comprenderne la struttura e identificare quali sono i dataset rilevanti che poi verranno uniti.

Datasets Mantenuti

- circuits
- drivers
- lap times
- races
- results
- status

Datasets Eliminati

- constructor results
- constructor standings
- constructors
- driver standings
- pit stops
- qualifying
- seasons
- sprint results

Dataset

Dopo aver eliminato i dataset irrilevanti, abbiamo unito i rimanenti per formare un unico dataset, rimosse le colonne irrilevanti/duplicate ai fini della predizione:

raceId	driverId	driver_name	circuitId	<i>circuit_name</i>
race_date	time_race	lap	position_lap	time_lap
statusId	status	<i>weather_code</i>	weather_descriptio n	

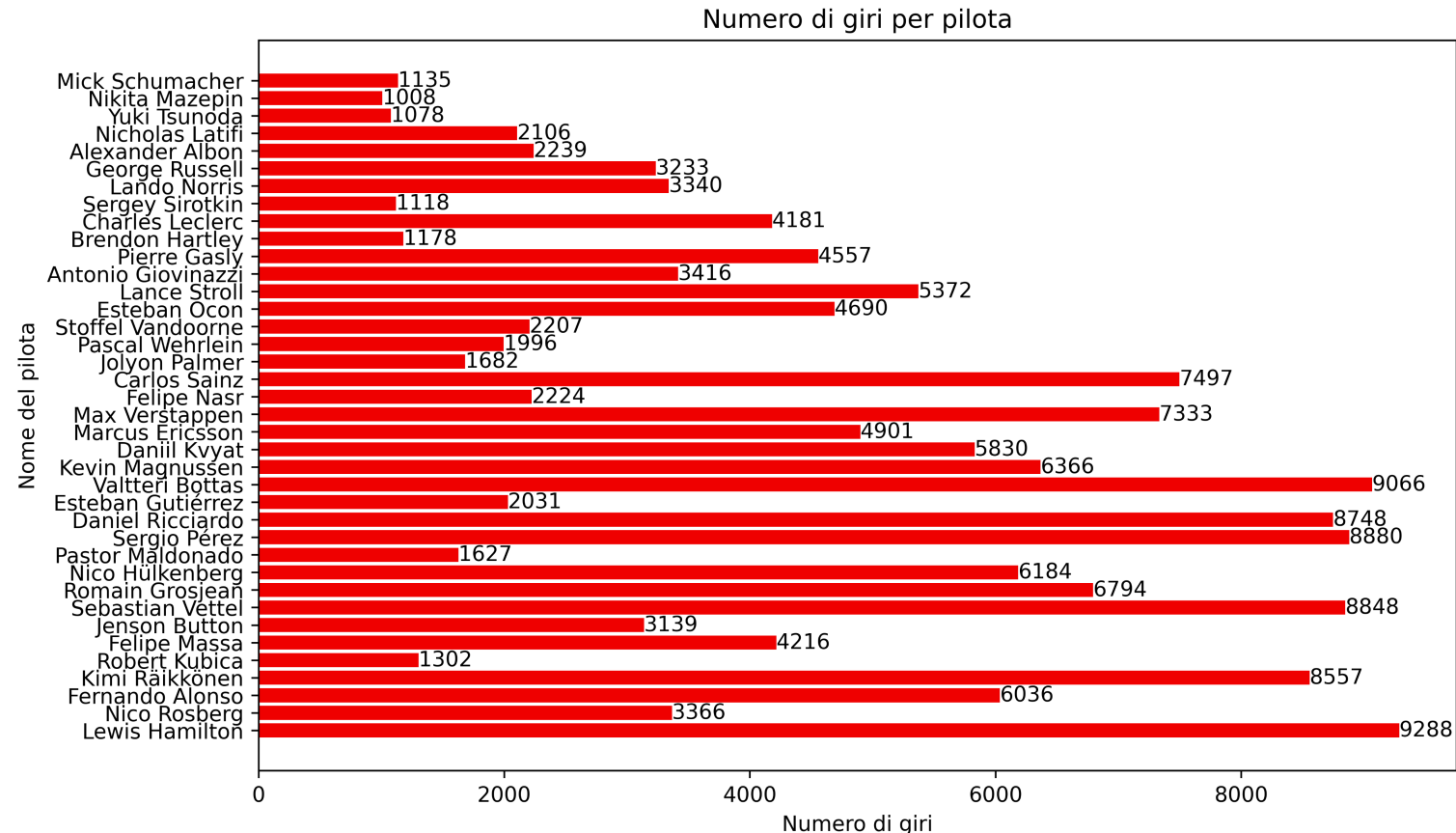
Data cleaning

Dopo la corretta fusione dei dataset, abbiamo eseguito una serie di operazioni per pulire e manipolare i dati al fine di renderli più coerenti e adatti all'analisi successiva. Le operazioni sono:

- Rimozione dei duplicati.
- Rimozione dei valori nulli.
- Rimozione dei piloti con meno di 1000 giri.
- Valutazione dal penultimo cambio regolamentare.

Data balancing

Analizzando il dataset questa è situazione dei giri effettuati dai piloti.



Mmmh...

Abbiamo del lavoro da fare qui!

Come possiamo fare per sistemare la situazione?

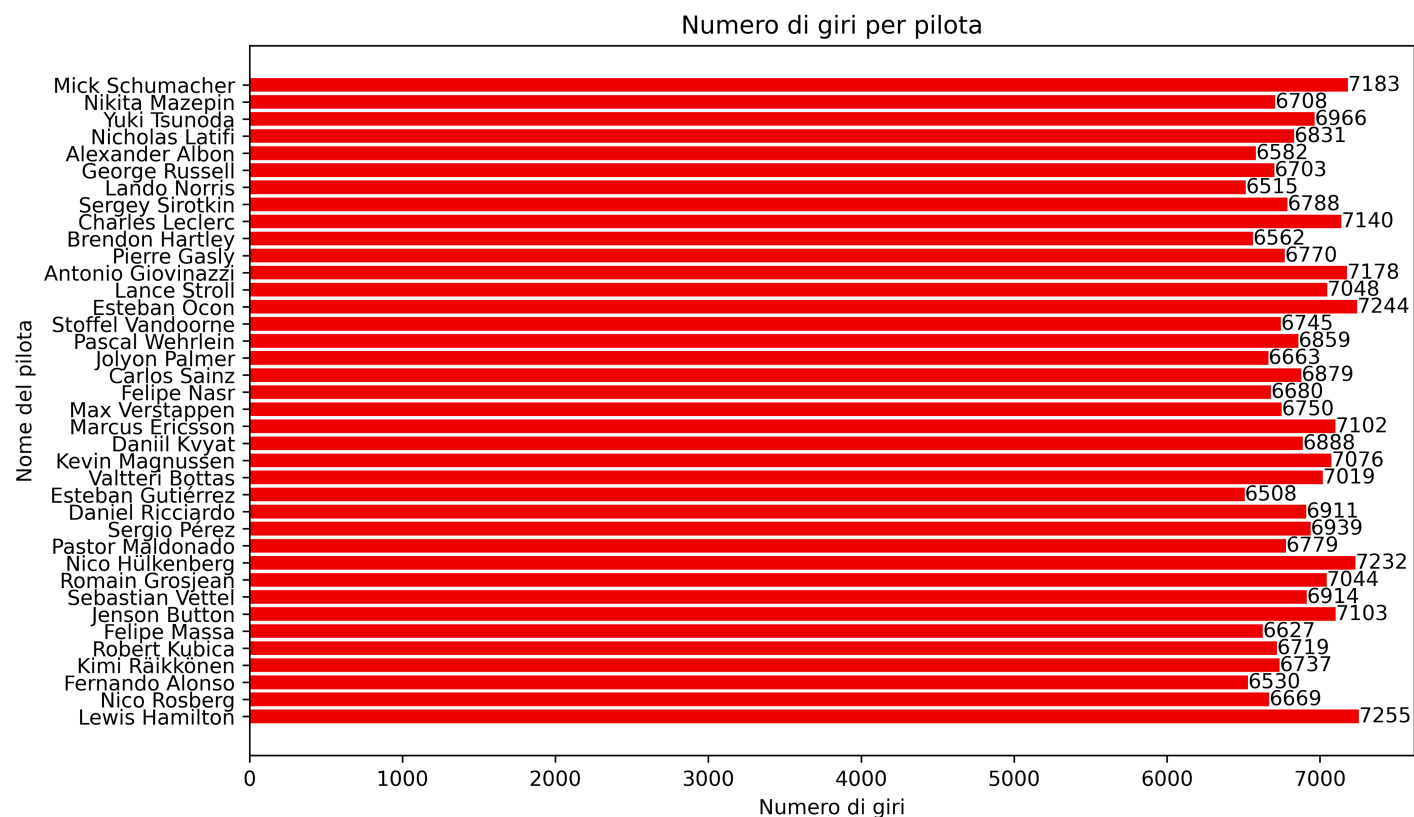


OVERSAMPLING



Data balancing

Ecco la situazione aggiornata!



Meteo

Il dataset non fornisce informazioni riguardo le condizioni meteo tuttavia, è possibile notare che, oltre che la data e l'ora dello svolgimento della gara, abbiamo anche la posizione geografica.

Utilizzando «Open-Meteo Weather API» siamo riusciti a modificare il dataset aggiungendo le seguenti colonne:

- weather code
- weather description





Data Modelling

Scelta dell'algoritmo
Addestramento

La regressione

L'obiettivo della applicazione è predire un giro, cioè una variabile numerica, quindi è un problema di regressione.

Data la presenza di molteplici modelli di regressione, abbiamo posto particolare attenzione sui seguenti:

- Random Forest
- Decision Tree
- Gradient Boosting
- Extra Trees

Addestramento

Abbiamo suddiviso il nostro dataset in due gruppi:

- Training set: 78%
- Test set: 22%

Inoltre è stato sviluppato uno script per l'oversampling che: calcola il massimo numero di giri svolti da ogni pilota e, sulla base di questo valore, aggiunge giri ai piloti fino al raggiungimento di una soglia vicina al massimo.



Evaluation

Mettriche

Quale modello scegliere?





Metriche

Decision Tree Regressor	
Mean Absolute Error	7.487555514758704
Accuracy	0.23752717047597838

Gradient Boosting	
Mean Absolute Error	11.791618631878723
Accuracy	0.10723481572237503

Random Forest Regressor	
Mean Absolute Error	6.86429203701672
Accuracy	0.4079788133234922

Extra Trees Regressor	
Mean Absolute Error	7.487555514758704
Accuracy	0.23752717047597838

A dark, blurred background image of a Formula 1 race track with several cars in motion. The cars are out of focus, creating a sense of speed and motion. The track surface and white boundary lines are visible in the foreground.

Deployment

F1 Predictor

F1 Predictor

Ok! È l'ora di scendere in pista e provarla!





GRAZIE PER LA VOSTRA

ATTENZIONE!

