



# CLASIFICACIÓN DE POSICIONES FINALES DE PILOTOS EN LA FORMULA 1



Anderson Acuña – 2191965  
Deciré Jaimes – 2211835  
Diego Sepúlveda – 2210063





# INTRODUCCIÓN

Nuestro proyecto utiliza los distintos métodos vistos en clase para analizar datos históricos e intentar predecir mediante clasificación la posición final del piloto.





# OBJETIVOS

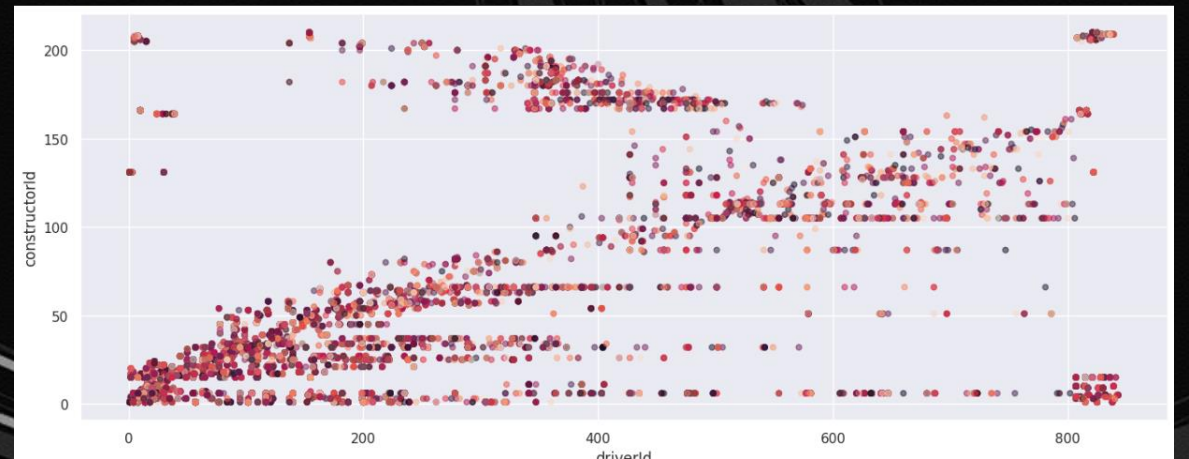
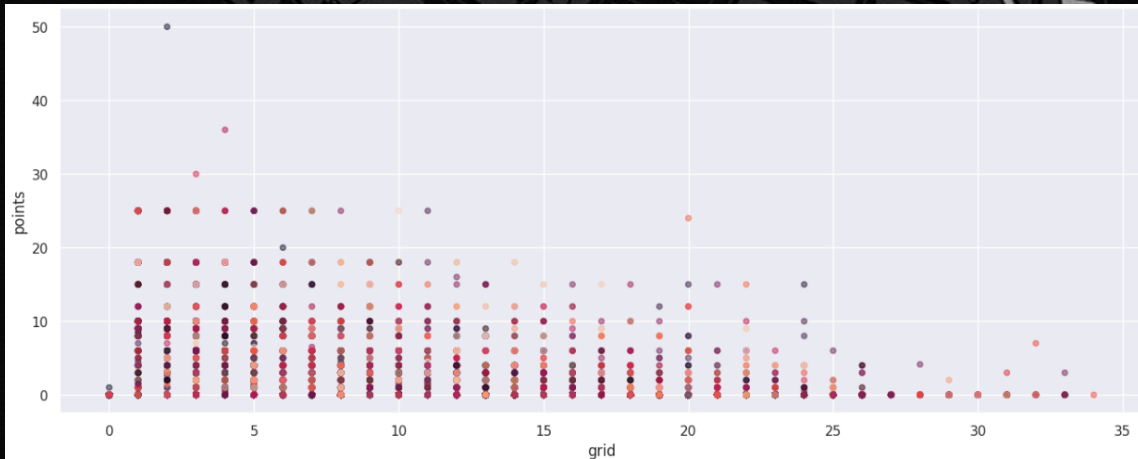
1. Aplicar los distintos métodos de clasificación vistos en clase.
2. Comparar los distintos métodos de clasificación para determinar el que mejor se ajuste a nuestro dataset.



3. Obtener un “Accuracy” aceptable de al menos 50%
4. Observar el “accuracy” de un dataset destinado al “aprendizaje supervisado” sometiéndolo a métodos de “Aprendizaje no supervisado”

# DATASET

resultId	raceId	driverId	constructorId	grid	positionOrder	points	laps	milliseconds	fastestLap	fastestLapSpeed	statusId	lapPit
0	1	18	1	1	1	10.0	58	5690616.0	39.0	218.3	1	1.0
1	2	18	2	2	5	8.0	58	5696094.0	41.0	217.586	1	1.0
2	3	18	3	3	7	6.0	58	5698779.0	41.0	216.719	1	11.0
3	4	18	4	4	11	5.0	58	5707797.0	58.0	215.464	1	12.0
4	5	18	5	1	3	4.0	58	5708630.0	43.0	218.385	1	13.0





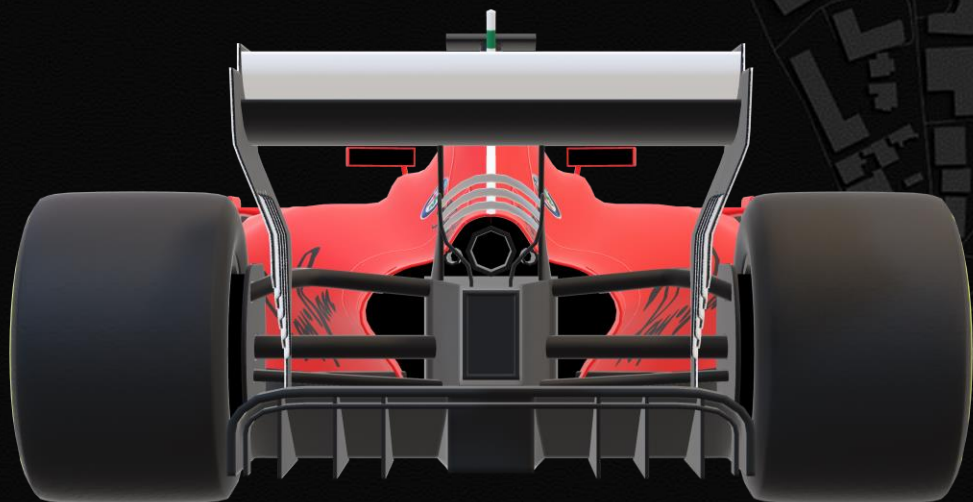
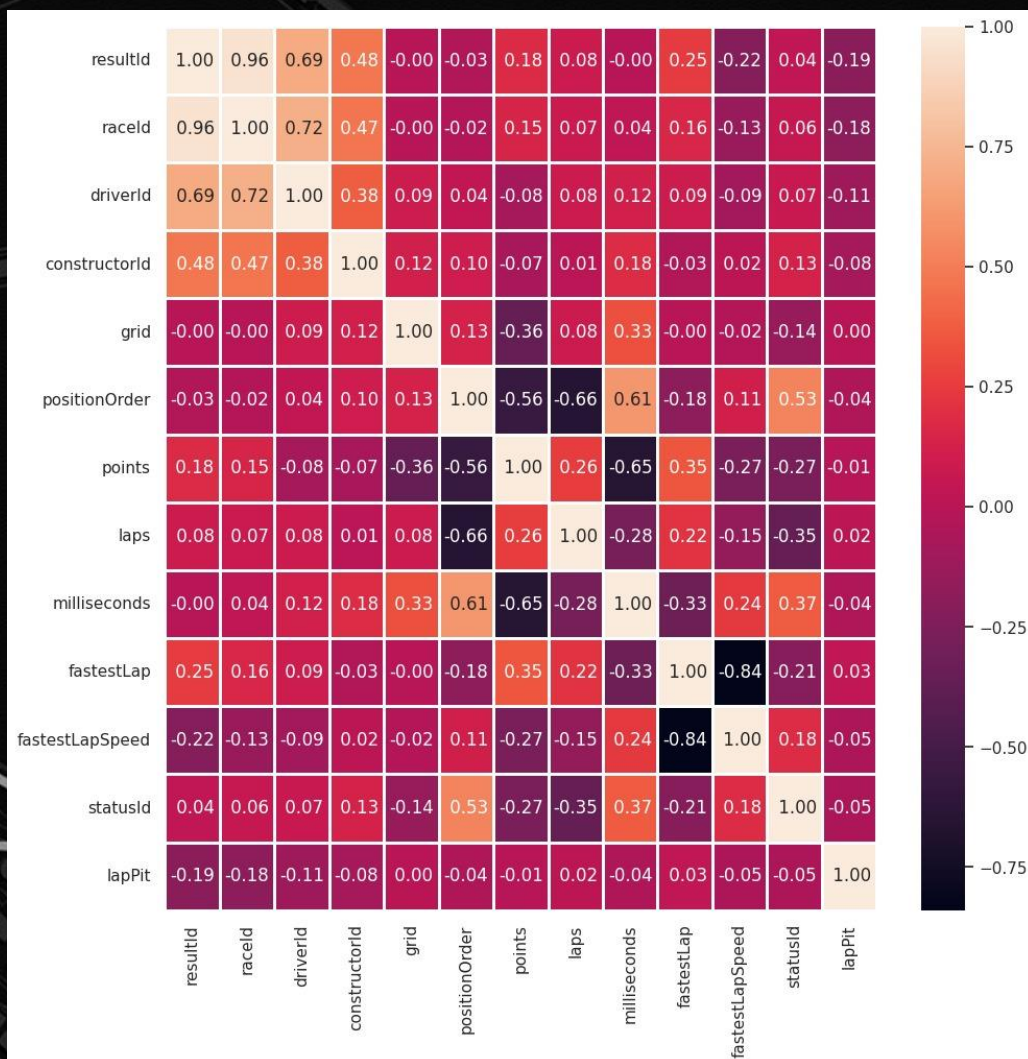
Data columns (total 13 columns):

#	Column	Non-Null Count	Dtype
---	-----	-----	-----
0	resultId	23777 non-null	int64
1	raceId	23777 non-null	int64
2	driverId	23777 non-null	int64
3	constructorId	23777 non-null	int64
4	grid	23777 non-null	int64
5	positionOrder	23777 non-null	int64
6	points	23777 non-null	float64
7	laps	23777 non-null	int64
8	milliseconds	6003 non-null	float64
9	fastestLap	5383 non-null	float64
10	fastestLapSpeed	5383 non-null	object
11	statusId	23777 non-null	int64
12	lapPit	6251 non-null	float64

dtypes: float64(4), int64(8), object(1)

memory usage: 2.4+ MB

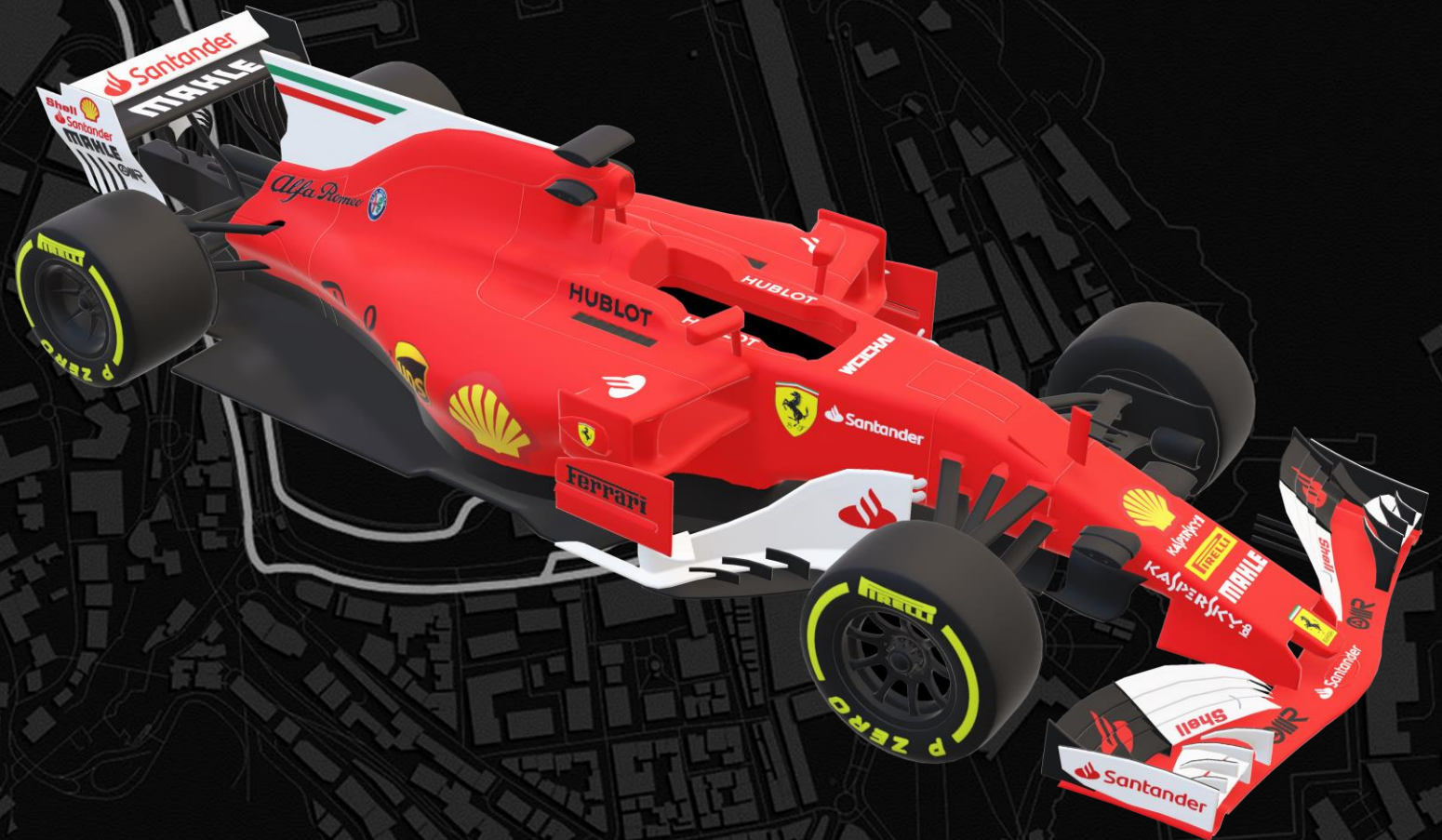
# DESARROLLO





# MÉTODOS UTILIZADOS

- ⌘ Decision Tree
- ⌘ Random Forest
- ⌘ SVC
- ⌘ DNN
- ⌘ Kmeans
- ⌘ DBSCAN





# RESULTADOS

- ⌘ Se aplicaron con éxito los métodos de clasificación vistos en clase.
- ⌘ Se obtuvo que el método de clasificación que mejor se ajusta a los datos es el DecisionTree.
- ⌘ No se logro el “accuracy” esperado (50%), se obtuvo un 38.5% al usar DecisionTree.
- ⌘ El “accuracy” obtenido con métodos de aprendizaje no supervisado es muy inferior a los obtenidos con aprendizaje supervisado.

