



TEAM ROCKET

Peleas Pokémon

Integrantes:

- Lautaro Terreno
- Matias Mattioli
- Nazareno Magallanes



Nociones básicas sobre Pokémon

- Peleas 1 vs 1
- Estadísticas
- Tipos de Pokémon
- Niveles



1

Etapas de trabajo



Presentación
del problema



Exploración
del dataset



Preparación
de los datos



Entrenamos
modelo y
ajustamos

2

¿Se podrá predecir al ganador de una batalla Pokémon?



“Depende”



—Ash Ketchum
(probablemente)

DATASETS

Combats.csv

Batallas de
2 Pokemones
y el ganador
*

50.000 Filas
3 Columnas

* El ganador (supuestamente) está
definido a partir de las
estadísticas básicas y el tipo del
Pokémon

Pokemon.csv

Información
sobre los
Pokemones

800 Filas
12 Columnas



Fuente:

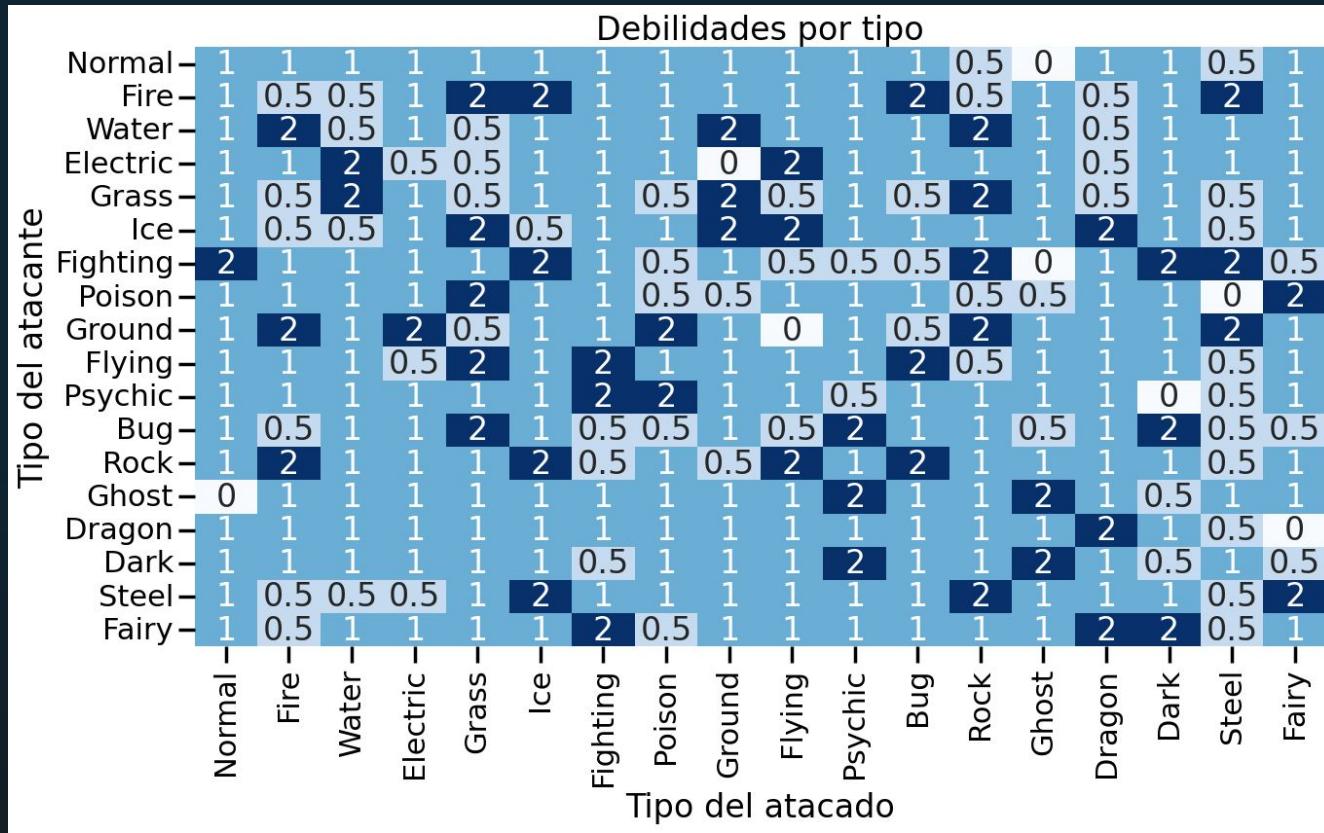
<https://www.kaggle.com/datasets/terminus7/pokemon-challenge?select=pokemon.csv>

Columnas del dataset *Pokemon.csv*

- ★ Número de pokémon (índice)
- ★ Nombre de pokémon
- ★ Primer tipo del pokémon
- ★ Segundo tipo del pokémon
- ★ HP (Vida)
- ★ Ataque
- ★ Defensa
- ★ Ataque especial
- ★ Defensa especial
- ★ Velocidad *
- ★ Generación
- ★ Legendario



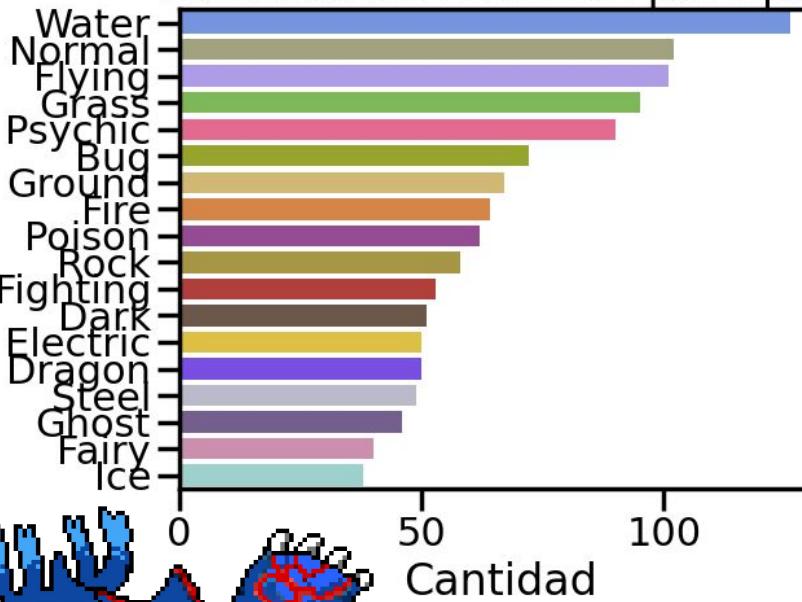
EXPLORACIÓN DE DATOS



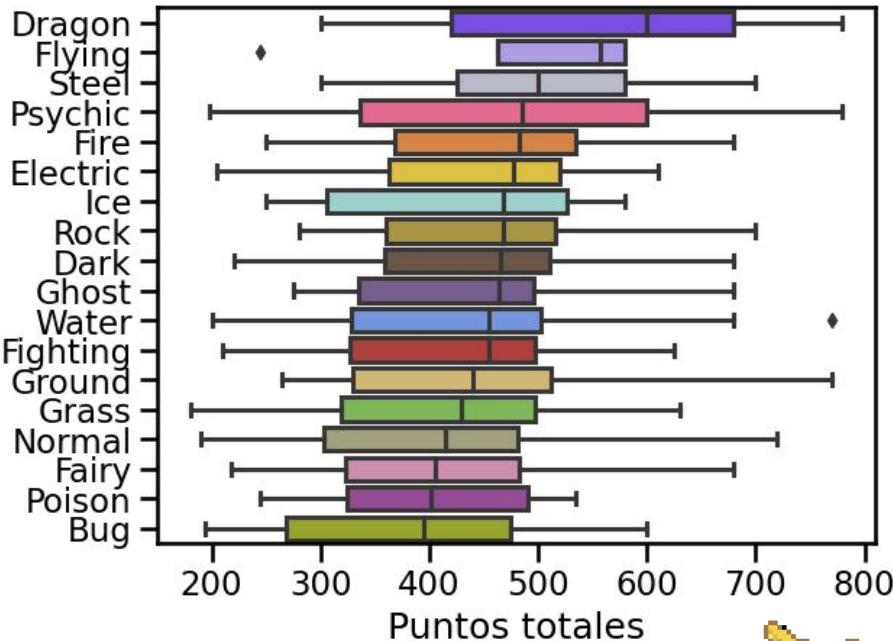
EXPLORACIÓN DE DATOS

Tipos de Pokemon

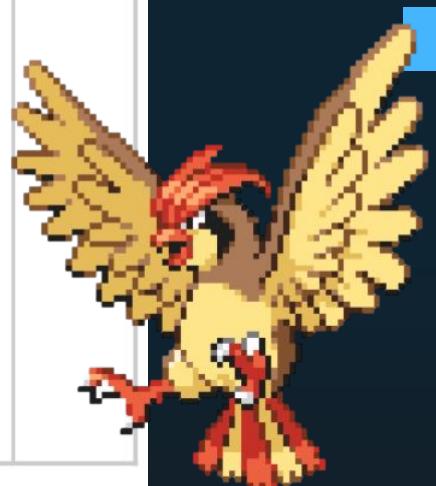
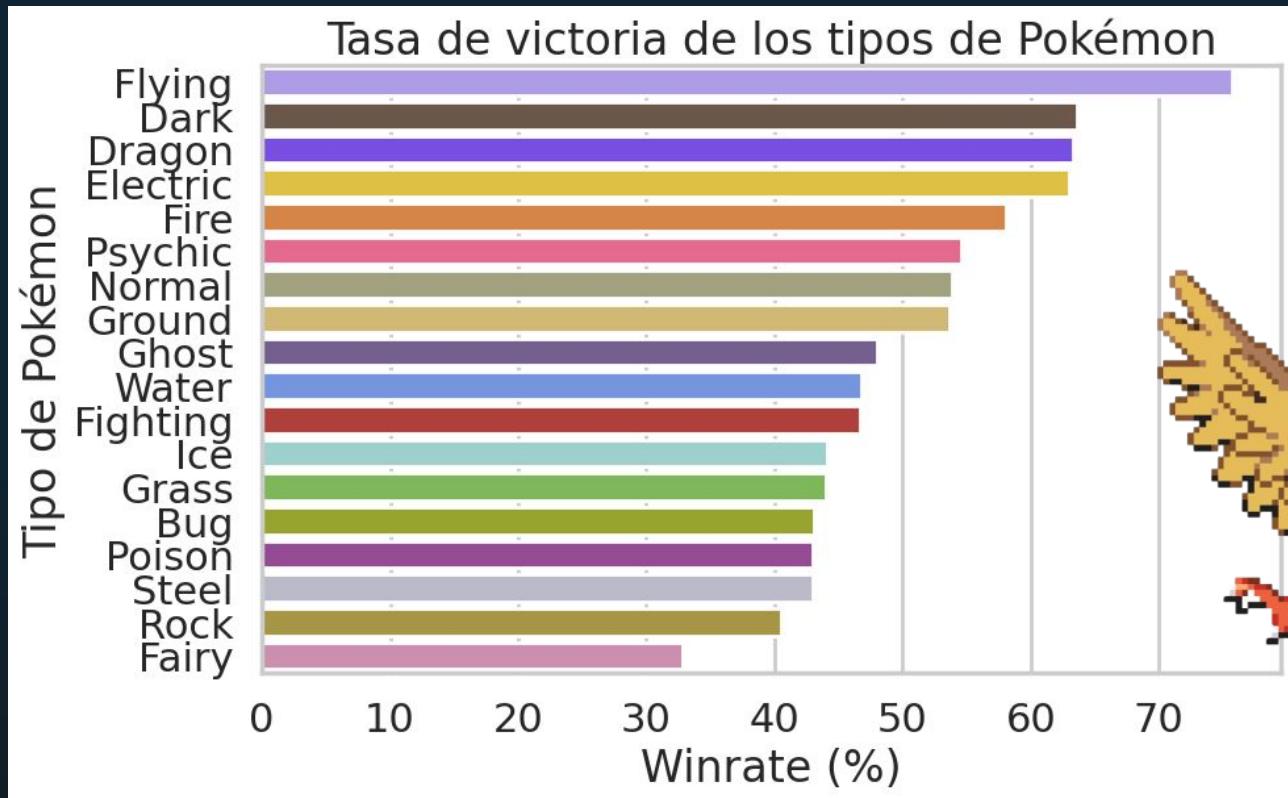
Cantidad de Pokemon por tipos



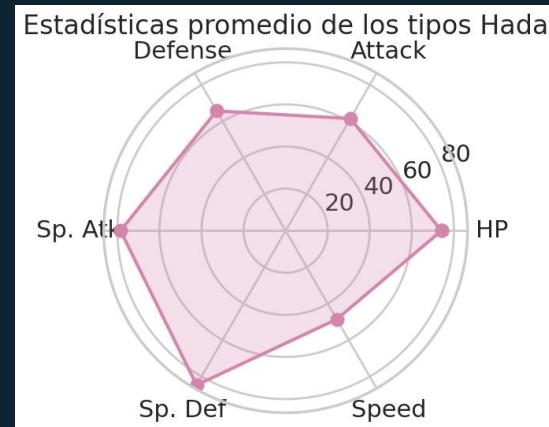
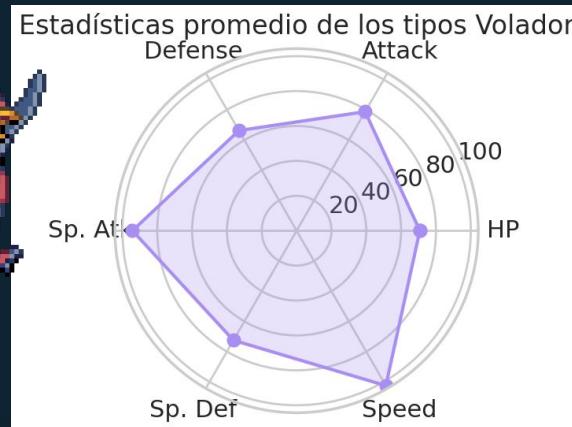
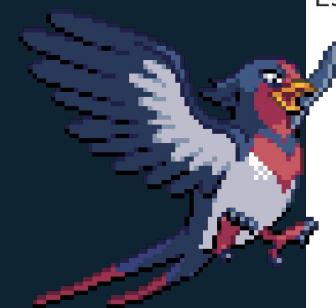
Tipo de pokemon



EXPLORACIÓN DE DATOS



EXPLORACIÓN DE DATOS



Hay diferencia en la Velocidad



Pre-procesado de los datos

- Cambio de índice a nombres en combats.
- Creamos un dataframe con toda la información combinada.
- Cambiamos la variable Winner.
- One-hot encoding de tipos.
- Droeamos features.



Modelado y entrenamiento de datos



VAMO A MODELAR

PREPARACIÓN DE DATOS

- Train Test Split (%70, %30)
- X → *dropeamos nombre de Pokémons y winner dejando las demás columnas.*
- Y → *ganador de pelea*



MODELADO

- Regresión Logistica
 - Grid Search CV, k-folds=5
 - Penalty: L1, L2 y ElasticNet
- Arbol de decisión
 - Grid Search, CV, k-folds=5
 - Max depth entre 5 y 8,
 - Min samples split entre 20 y 40
 - Min samples leaf entre 25 y 30



REGRESOR LOGISTICO



	param_C	param_l1_ratio	mean_test_score
7	0.00001	0.5	0.914057
6	0.00001	0.25	0.902714
8	0.00001	0.75	0.896943
14	0.0001	1.0	0.896400
13	0.0001	0.75	0.894600
12	0.0001	0.5	0.893171
11	0.0001	0.25	0.891314
23	0.01	0.75	0.890829
22	0.01	0.5	0.890657
24	0.01	1.0	0.890657

REGRESOR LOGISTICO

$$C = 0.00001$$

log.coef_



ARBOL

- `Max_depth = 8`
- `Min_samples_leaf = 25`
- `Min_samples_split = 30`
- `Cross_val_score = 0.94`



Accuracy:
0.94

LOG

- `Penalty = elasticnet`
- `C = 0.01`
- `L1_ratio = 0.5`
- `Solver = 'saga'`
- `Cross_val_score = 0.89`

Accuracy:
0.89



17

JÁ PELEAR!

GOLEM



PIKACHU



GANADOR: ¡&PIKACHU?!

GOLEM



PIKACHU



GANADOR: ¡&PIKACHU?!

GOLEM

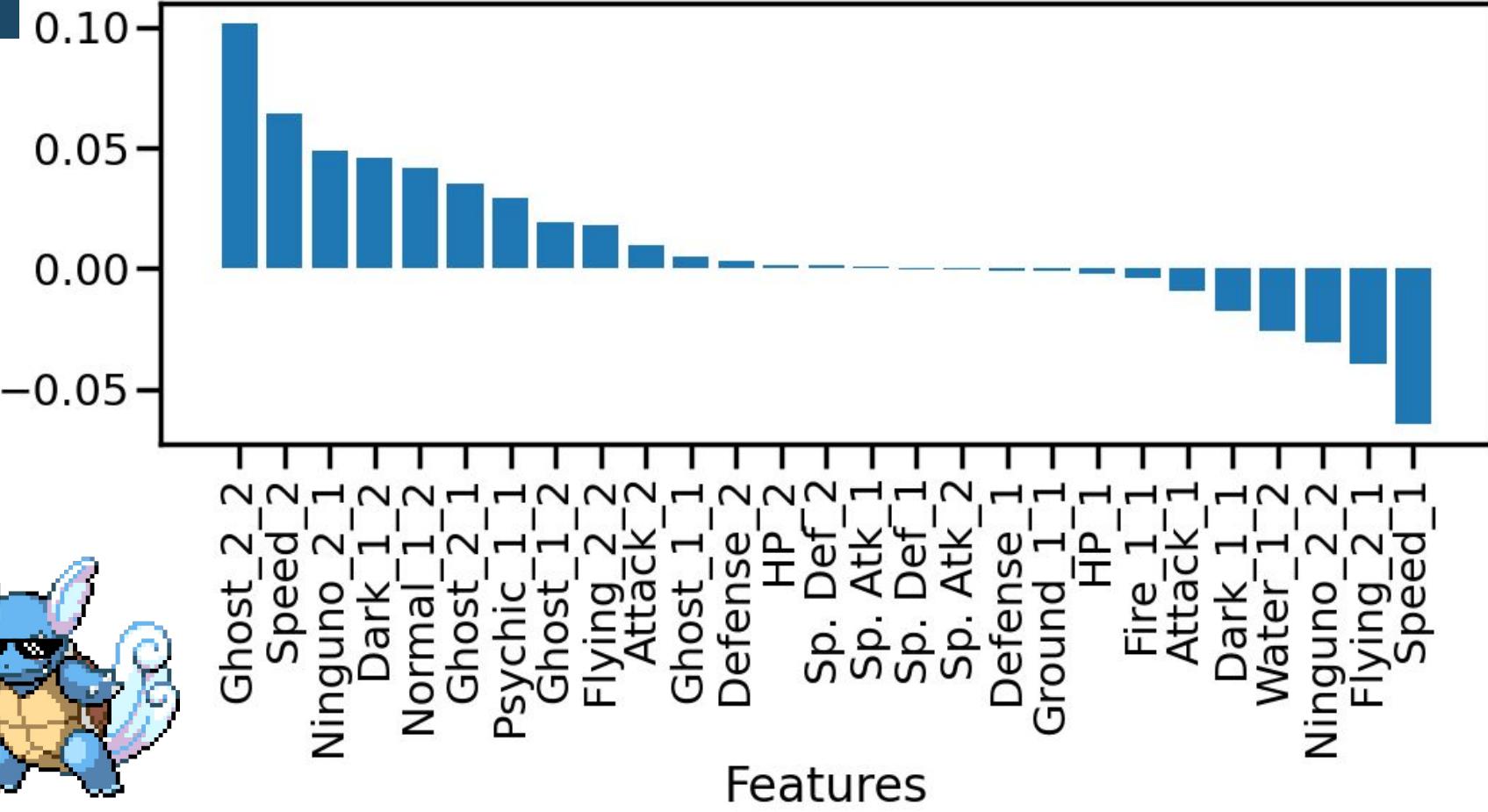


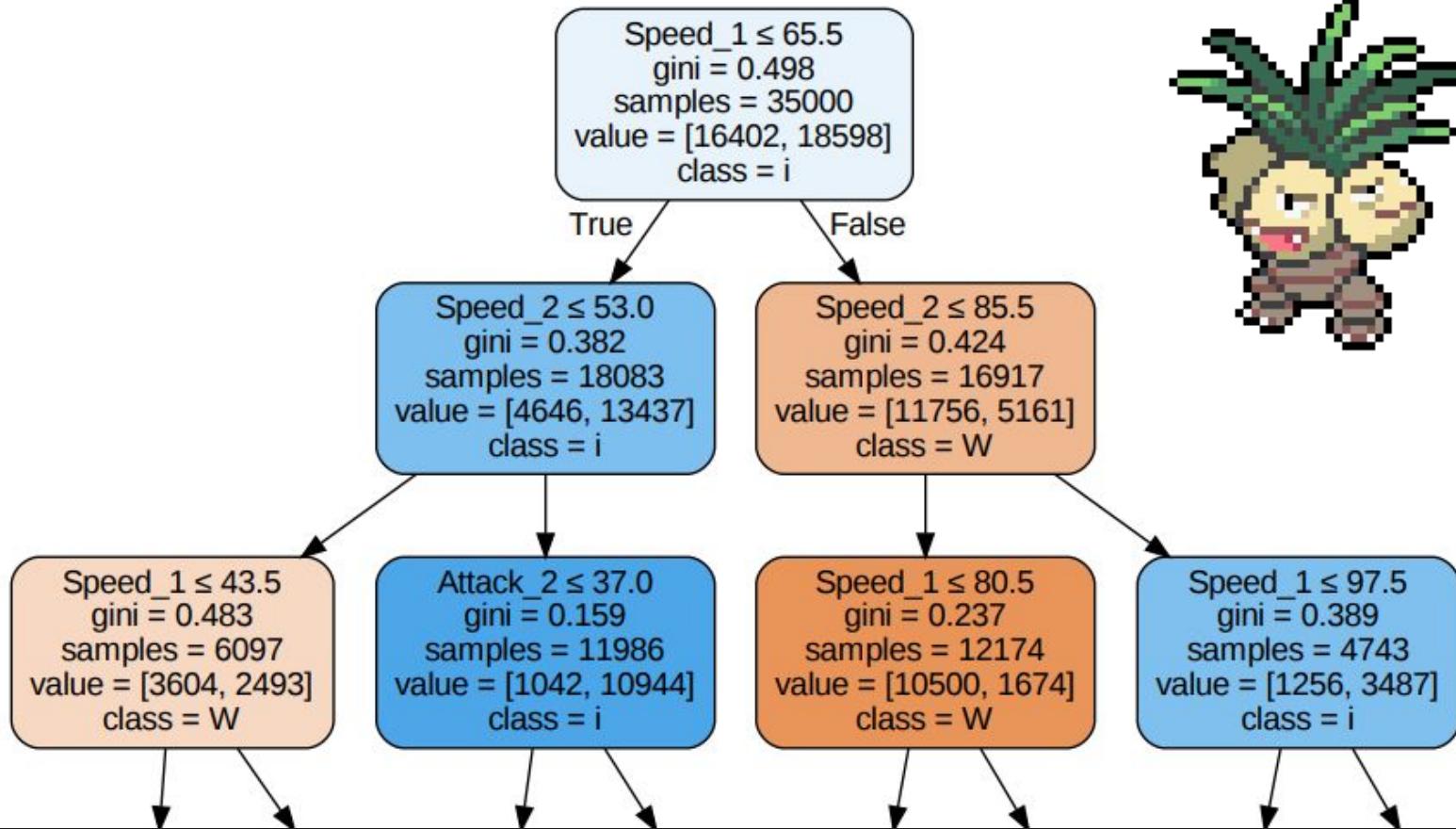
PIKACHU



Coeficientes de cada feature

Coeficientes





COMPARACION ARBOL Y REGRESOR

pok1	pok2	Ganador_Arbol	Ganador_Log
Pikachu	Golem	Pikachu	Pikachu
Charmander	Articuno	Articuno	Articuno
Golem	Pikachu	Pikachu	Pikachu
Charizard	Bulbasaur	Charizard	Charizard
Bulbasaur	Squirtle	Bulbasaur	Squirtle
Bulbasaur	Onix	Onix	Onix
Crobat	Pikachu	Crobat	Crobat
Pikachu	Crobat	Crobat	Crobat
Venusaur	Onix	Venusaur	Venusaur



CONCLUSIONES

MODELO Y DATASET

- Combats.csv tiene errores que se reflejan en nuestro modelo.

STATS IMPORTANTES

- Pokemones tipo fantasma.
- Velocidad de pokémon.

ARBOL Y LOGISTICO

- Ambos tienen un poder de predicción similar





Muchas Gracias!

¿Preguntas?



luterreno@estudiantes.unsam.edu.ar

mmattioli@estudiantes.unsam.edu.ar

nazarenomm98@gmail.com

