



# Prácticas Inteligencia Artificial:

## Web Scraping

## Machine Learning

## Interpretabilidad

David San José  
Mario Uceda  
Ignacio Gil  
Teodora Nikolaeva

# Práctica 1 - Web Scraping

Basada en la búsqueda de información e interacción con páginas web, a través de un programa diseñado en Python.

La práctica 1 fue de carácter individual, por lo que los archivos son dependientes de cada proyecto

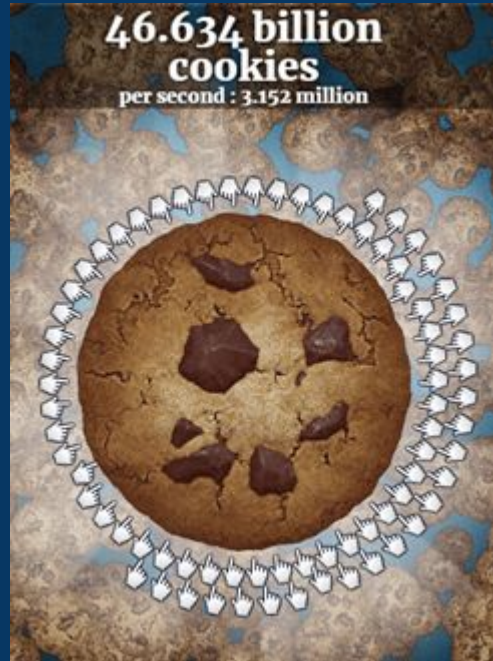
## Librerías Utilizadas:

- Selenium
- BeautifulSoup



# Cookie Clicker

Mediante web scraping se consigue hacer un bot que juega solo



2ª parte de la práctica de web scraping:  
utilización de librerías de Twitter para  
extraer tweets de un tema concreto  
para analizarlos.

## Librería utilizada:

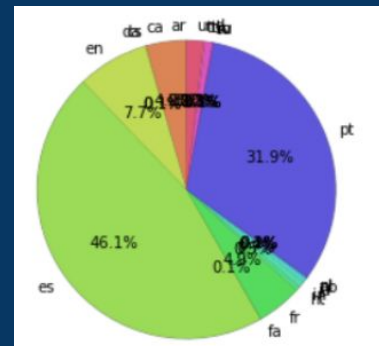
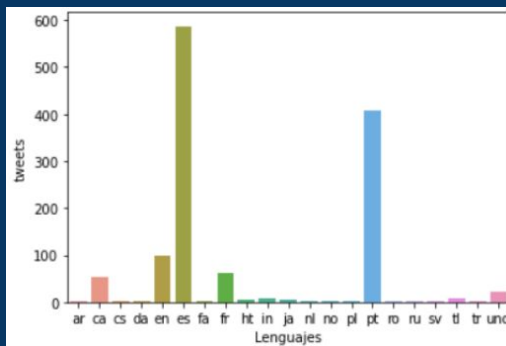
- Tweepy

## Representación de Datos:

- Pandas
- Numpy
- Matplot

```
twitter_stream = Stream(auth, MyListener())  
twitter_stream.filter(track=['madrid'])
```

```
1 {"created_at": "Fri Oct 29 14:37:41 +0000 2021", "id": "1454095098636120070", "id_str": "1454095098636120070", "text": "Apoio a convoca\u00e7\u00e3o do Coutinho, mas o Vi  
2  
3 {"created_at": "Fri Oct 29 14:37:43 +0000 2021", "id": "1454095103161733120", "id_str": "1454095103161733120", "text": "RT @Confillegal: El magistrado Luis Vacas obliga al  
4  
5 {"created_at": "Fri Oct 29 14:37:43 +0000 2021", "id": "1454095104797511687", "id_str": "1454095104797511687", "text": "RT @spaincrisis: Mucha rabia y dolor ha causado qu  
6  
7 {"created_at": "Fri Oct 29 14:37:45 +0000 2021", "id": "1454095111814582280", "id_str": "1454095111814582280", "text": "RT @EmilioBelgadoOr: Quien presid\u00eda en ese mo  
8  
9 {"created_at": "Fri Oct 29 14:37:45 +0000 2021", "id": "1454095112749920258", "id_str": "1454095112749920258", "text": "El Gobierno \u00c9 socialcomunista\u00c9 pag\u00fa 00F3  
10  
11 {"created_at": "Fri Oct 29 14:37:45 +0000 2021", "id": "1454095113366421504", "id_str": "1454095113366421504", "text": "RT @incondelpeta: en CONEJO en medio de MADRID ht  
12  
13 {"created_at": "Fri Oct 29 14:37:45 +0000 2021", "id": "1454095114385625091", "id_str": "1454095114385625091", "text": "RT @ChezNieto: Si Casado acepta y apoya la estrate  
14  
15 {"created_at": "Fri Oct 29 14:37:45 +0000 2021", "id": "1454095115165786112", "id_str": "1454095115165786112", "text": "RT @juncalssolano Por mencionarte algunos que egresar  
16  
17 {"created_at": "Fri Oct 29 14:37:46 +0000 2021", "id": "1454095115639791618", "id_str": "1454095115639791618", "text": "RT @GersonKorlinga: Saiu a convoca\u00e7\u00e3o do Coutinho, mas o Vi  
18  
19 {"created_at": "Fri Oct 29 14:37:46 +0000 2021", "id": "1454095116096974849", "id_str": "1454095116096974849", "text": "RT @SquawkaNews: Eden Hazard isn't a Real Madrid s  
20  
21 {"created_at": "Fri Oct 29 14:37:46 +0000 2021", "id": "1454095119632699405", "id_str": "1454095119632699405", "text": "RT @KratosCule: Messi saludando los t\u00fa\u00e9dulos q  
22  
23 {"created_at": "Fri Oct 29 14:37:47 +0000 2021", "id": "1454095120962400262", "id_str": "1454095120962400262", "text": "RT @SanchezRM_: Si el penalti es a favor del Real  
24  
25 {"created_at": "Fri Oct 29 14:37:47 +0000 2021", "id": "1454095121121693702", "id_str": "1454095121121693702", "text": "Benzema out? no problem Real Madrid got the serbia  
26
```



# Práctica 2 - Machine Learning

## Herramientas para aprendizaje automático y métricas

- Sklearn

## Gestión de datos

- Pandas
- Numpy

## Representación de datos

- Matplot
- Seaborn

## Archivos utilizados

- Uno por algoritmo con el mismo nombre
- Combates.csv y Pokemon.csv

Elección de una base de datos de Kaggle.com (Pokemon) para entrenar algoritmos de Machine Learning:

- Regresión lineal
- Regresión logística
- Árbol de clasificación
- Random forest
- K - Nearest Neighbour
- K - Fold
- Bayes (Bernoulli y modelo Gaussiano)
- Support Vector Machine

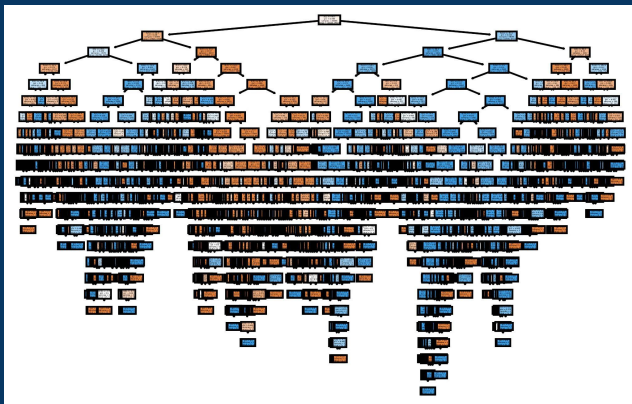


# Mapa de calor de correlaciones

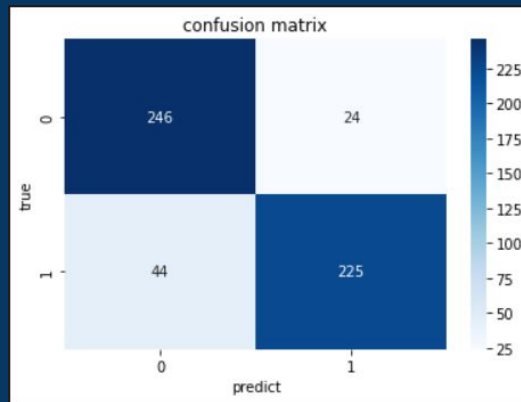
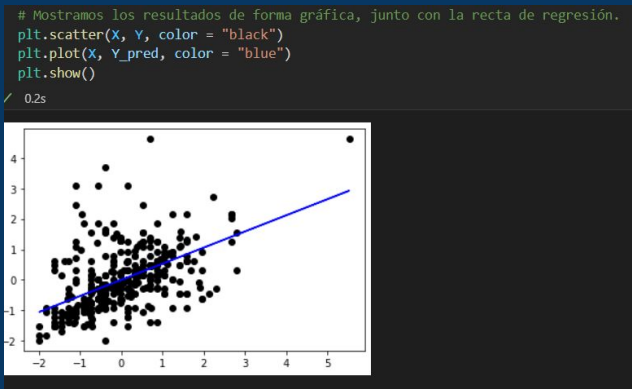
	#	HP	Attack	Defense	Sp. Atk	Sp. Def	Speed	Generation	Legendary	Bug	Dark	Dragon	Electric	Fairy	Fighting	Fire	Flying	Ghost	Grass	Ground	Ice	Normal	Poison	Psychic	Rock	Steel	Water
#	1.000000	0.098000	0.103000	0.095000	0.089000	0.086000	0.012000	0.983000	0.154000	-0.040000	0.099000	0.141000	0.016000	0.050000	0.048000	-0.009000	-0.033000	0.169000	0.044000	-0.047000	0.071000	-0.053000	-0.215000	0.003000	0.025000	0.093000	-0.102000
HP	0.098000	1.000000	0.422000	0.240000	0.362000	0.379000	0.176000	0.059000	0.274000	-0.154000	0.010000	0.138000	-0.061000	-0.007000	0.055000	0.010000	0.031000	-0.062000	-0.046000	0.075000	0.082000	0.112000	-0.076000	0.027000	-0.036000	-0.043000	0.030000
Attack	0.103000	0.422000	1.000000	0.439000	0.396000	0.264000	0.381000	0.051000	0.345000	-0.070000	0.143000	0.213000	-0.075000	-0.123000	0.211000	0.046000	0.014000	-0.016000	-0.063000	0.127000	0.021000	-0.075000	-0.073000	-0.070000	0.101000	0.108000	-0.071000
Defense	0.095000	0.240000	0.439000	1.000000	0.224000	0.511000	0.015000	0.042000	0.246000	-0.028000	-0.024000	0.106000	-0.061000	-0.025000	0.002000	-0.035000	-0.068000	0.061000	-0.007000	0.134000	0.020000	-0.175000	-0.094000	0.011000	0.298000	0.351000	0.006000
Sp. Atk	0.089000	0.362000	0.396000	0.224000	1.000000	0.506000	0.473000	0.036000	0.449000	-0.185000	0.037000	0.194000	0.127000	0.026000	-0.048000	0.182000	0.041000	0.035000	0.001000	-0.107000	0.070000	-0.184000	-0.044000	0.237000	-0.106000	-0.006000	0.026000
Sp. Def	0.086000	0.379000	0.264000	0.511000	0.506000	1.000000	0.259000	0.028000	0.364000	-0.082000	-0.021000	0.139000	0.017000	0.093000	0.014000	0.026000	-0.009000	0.045000	-0.005000	-0.077000	0.060000	-0.112000	-0.047000	0.190000	0.019000	0.103000	-0.025000
Speed	0.012000	0.176000	0.381000	0.015000	0.473000	0.259000	1.000000	-0.023000	0.327000	-0.071000	0.068000	0.123000	0.130000	-0.098000	0.078000	0.073000	0.237000	-0.061000	-0.096000	-0.096000	-0.005000	0.052000	-0.030000	0.108000	-0.165000	-0.098000	-0.049000
Generation	0.983000	0.059000	0.051000	0.042000	0.036000	0.028000	-0.023000	1.000000	0.080000	-0.019000	0.094000	0.102000	0.006000	0.066000	0.051000	0.006000	-0.036000	0.156000	0.073000	-0.051000	0.038000	-0.038000	-0.192000	-0.022000	0.009000	0.079000	-0.105000
Legendary	0.154000	0.274000	0.345000	0.246000	0.449000	0.364000	0.327000	0.080000	1.000000	-0.094000	-0.021000	0.226000	0.018000	-0.005000	-0.006000	0.047000	0.094000	-0.014000	-0.067000	-0.007000	0.041000	-0.086000	-0.086000	0.169000	-0.013000	0.019000	-0.066000
Bug	-0.040000	-0.154000	-0.070000	-0.028000	-0.185000	-0.082000	-0.071000	-0.019000	-0.094000	1.000000	-0.082000	-0.081000	-0.045000	-0.072000	-0.049000	-0.061000	0.065000	-0.059000	-0.034000	-0.064000	-0.070000	-0.120000	0.121000	-0.112000	-0.004000	0.047000	-0.124000
Dark	0.099000	0.010000	0.143000	-0.024000	0.037000	-0.021000	0.068000	0.094000	-0.021000	-0.082000	1.000000	-0.004000	-0.067000	-0.060000	-0.008000	-0.020000	-0.022000	0.001000	-0.048000	-0.023000	-0.010000	-0.100000	-0.018000	-0.044000	-0.033000	-0.024000	-0.029000
Dragon	0.141000	0.138000	0.213000	0.106000	0.194000	0.139000	0.123000	0.102000	0.226000	-0.081000	-0.004000	1.000000	-0.024000	-0.036000	-0.069000	-0.038000	0.026000	-0.019000	-0.079000	0.052000	0.015000	-0.099000	-0.056000	-0.027000	-0.032000	-0.044000	-0.083000
Electric	0.016000	-0.061000	-0.075000	-0.061000	0.127000	0.017000	0.130000	0.006000	0.018000	-0.045000	-0.067000	-0.024000	1.000000	-0.036000	-0.069000	-0.057000	-0.020000	-0.042000	-0.079000	-0.059000	-0.033000	-0.068000	-0.075000	-0.092000	-0.072000	-0.001000	-0.069000
Fairy	0.050000	-0.007000	-0.123000	-0.025000	0.026000	0.093000	-0.098000	0.066000	-0.005000	-0.072000	-0.060000	-0.036000	-0.036000	1.000000	-0.061000	-0.068000	-0.053000	-0.057000	-0.049000	-0.069000	-0.051000	-0.002000	-0.066000	0.027000	0.002000	0.013000	-0.068000
Fighting	0.048000	0.055000	0.211000	0.002000	-0.048000	0.014000	0.076000	0.051000	-0.006000	-0.049000	-0.008000	-0.069000	-0.069000	-0.061000	1.000000	0.051000	-0.086000	-0.066000	-0.051000	-0.081000	-0.059000	-0.072000	-0.040000	0.001000	-0.055000	-0.005000	-0.074000
Fire	-0.009000	0.010000	0.046000	-0.035000	0.182000	0.026000	0.073000	0.006000	0.047000	-0.061000	-0.020000	-0.038000	-0.057000	-0.068000	0.051000	1.000000	-0.029000	-0.013000	-0.108000	-0.023000	-0.066000	-0.085000	-0.085000	-0.061000	-0.065000	-0.056000	-0.115000
Flying	-0.033000	0.031000	0.014000	-0.068000	0.041000	-0.009000	0.237000	-0.036000	0.094000	0.065000	-0.022000	0.026000	-0.020000	-0.053000	-0.029000	1.000000	-0.062000	-0.081000	-0.061000	-0.049000	0.126000	-0.068000	-0.064000	-0.048000	-0.081000	-0.092000	-0.092000
Ghost	0.169000	-0.062000	-0.016000	0.061000	0.035000	0.045000	-0.061000	0.156000	-0.014000	-0.059000	0.001000	-0.019000	-0.042000	-0.057000	-0.066000	-0.013000	0.062000	1.000000	0.075000	-0.036000	-0.030000	-0.094000	0.009000	-0.071000	-0.069000	0.026000	-0.077000
Grass	0.044000	-0.046000	-0.063000	-0.007000	0.001000	-0.005000	-0.096000	0.073000	-0.067000	-0.034000	-0.048000	-0.079000	-0.079000	-0.049000	-0.051000	-0.108000	-0.081000	0.075000	1.000000	-0.097000	-0.027000	-0.117000	0.110000	-0.094000	-0.073000	-0.062000	-0.127000
Ground	-0.047000	0.075000	0.127000	0.134000	-0.107000	-0.077000	-0.096000	-0.051000	-0.007000	-0.064000	-0.023000	0.052000	-0.059000	-0.069000	-0.081000	-0.023000	-0.061000	-0.036000	-0.097000	1.000000	-0.065000	-0.111000	1.000000	-0.079000	0.072000	-0.021000	-0.074000
Ice	0.071000	0.082000	0.021000	0.020000	0.070000	0.060000	-0.005000	0.038000	0.041000	-0.070000	-0.010000	0.015000	-0.033000	-0.051000	-0.059000	-0.068000	-0.049000	0.030000	-0.027000	-0.004000	1.000000	-0.085000	-0.065000	-0.042000	-0.017000	-0.057000	0.000000
Normal	-0.053000	0.112000	-0.075000	-0.175000	-0.184000	-0.112000	0.052000	-0.036000	-0.086000	-0.120000	-0.100000	-0.099000	-0.068000	-0.002000	-0.072000	-0.085000	0.126000	-0.094000	-0.117000	-0.102000	-0.085000	1.000000	-0.111000	-0.112000	-0.107000	-0.098000	-0.155000
Poison	-0.215000	-0.076000	-0.073000	-0.094000	-0.044000	-0.047000	-0.030000	-0.192000	-0.086000	0.121000	-0.018000	-0.056000	-0.075000	-0.066000	-0.040000	-0.085000	-0.068000	0.009000	0.110000	-0.054000	-0.065000	-0.100000	1.000000	-0.103000	-0.081000	-0.074000	-0.074000
Psychic	0.003000	0.027000	-0.070000	0.011000	0.237000	0.190000	0.108000	-0.022000	0.169000	-0.112000	-0.044000	-0.027000	-0.092000	0.027000	0.001000	-0.061000	-0.064000	-0.071000	-0.084000	-0.079000	-0.042000	-0.112000	-0.103000	1.000000	-0.069000	0.025000	-0.100000
Rock	0.025000	-0.036000	0.101000	0.298000	-0.106000	0.019000	-0.165000	0.009000	-0.013000	-0.004000	-0.033000	-0.032000	-0.072000	0.002000	-0.055000	-0.065000	-0.048000	-0.069000	-0.073000	0.072000	-0.017000	-0.107000	-0.081000	1.000000	0.049000	0.011000	-0.096000
Steel	0.093000	-0.043000	0.108000	0.351000	-0.006000	0.103000	-0.098000	0.079000	0.019000	0.047000	-0.024000	-0.044000	-0.001000	0.013000	-0.005000	-0.056000	-0.081000	0.026000	-0.062000	-0.021000	-0.057000	-0.098000	-0.074000	0.025000	0.049000	1.000000	-0.096000
Water	-0.102000	0.030000	-0.071000	0.006000	0.026000	-0.025000	-0.049000	-0.105000	-0.066000	-0.124000	-0.029000	-0.083000	-0.069000	-0.068000	-0.074000	-0.115000	-0.092000	-0.077000	-0.127000	-0.007000	0.000000	-0.155000	-0.074000	-0.100000	0.011000	-0.096000	1.000000

# Varios Resultados

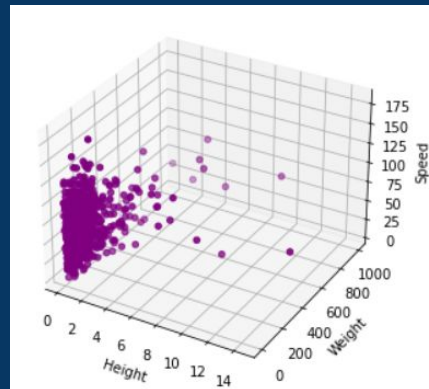
Árbol de  
clasificación:  
92'56%



Regresión  
lineal:  
61%



SVM:  
89'98%



K-Fold:  
3%

# Resultados de los algoritmos para la predicción de ganador

## COMPROBACIÓN DE RESULTADOS

Objetivo/Algoritmo	Árboles	Random Forest	Support Vector Machine
Predecir el ganador (Atributos de ataque)	91.23%	94.08%	89.9%
Predecir el ganador (Tipos de Pokemon)	63.63%	64.59%	67.16%
Predecir el ganador (Todos los datos)	92.5633%	93.23%	89.23%



# Práctica 3 - Interpretabilidad



## Herramientas para aprendizaje automático y métricas

- Sklearn



## Gestión de datos

- Pandas
- Numpy
- XgBoost



## Representación de datos

- Matplot
- Seaborn

## Interpretabilidad

- |        |          |
|--------|----------|
| - Eli5 | - PDP    |
| - SHAP | - Skater |
| - LIME | - FairML |

## Objetivo:

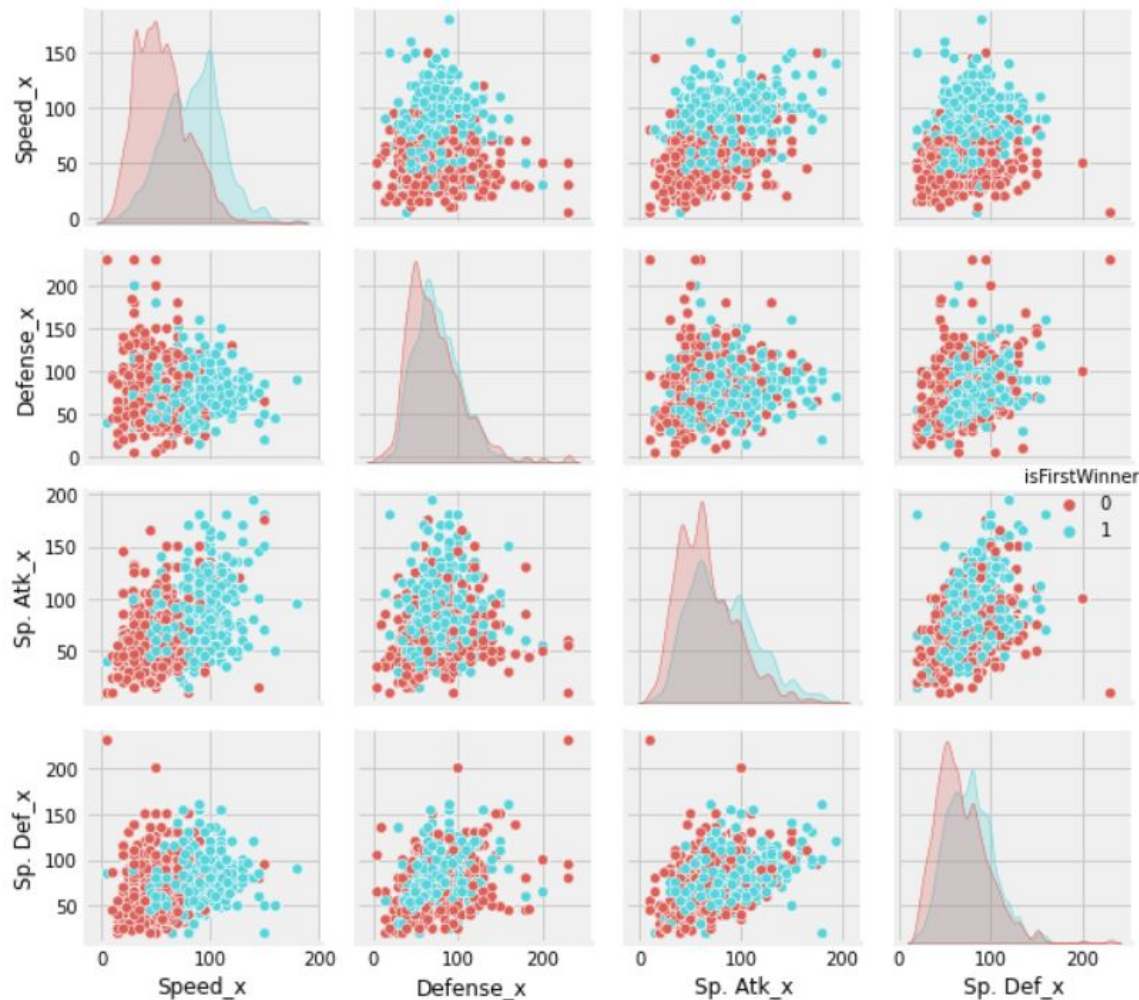
Utilización de técnicas de interpretabilidad de modelos de machine learning para analizar los resultados de la práctica 2

## Archivos:

- Combates.csv, pokemon.csv
- Interpretabilidad.ipynb

# Análisis Univariante

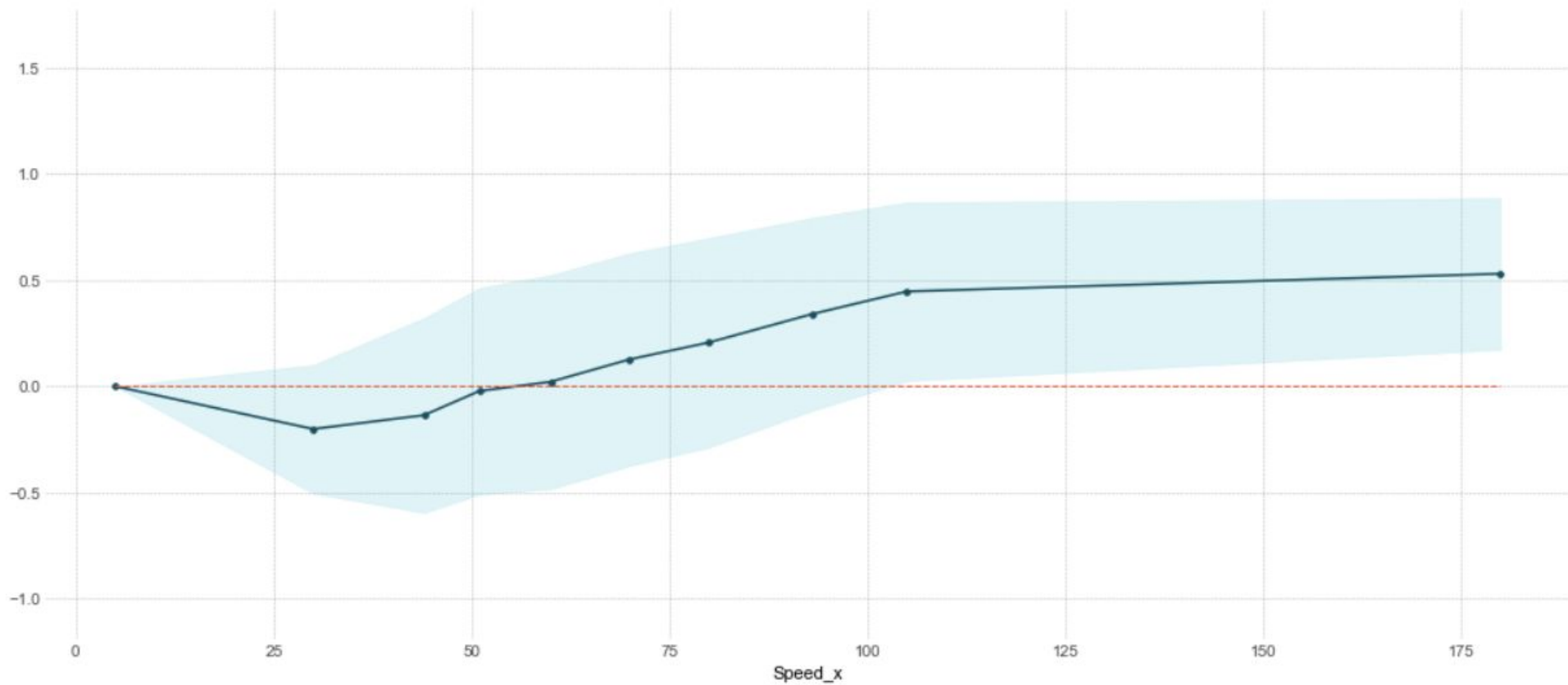
Weight	Feature
0.1383	Speed_x
0.1364	Speed_y
0.0463	Normal_y
0.0367	Normal_x
0.0334	Ghost_y
0.0299	Psychic_x
0.0257	Dragon_x
0.0252	Electric_x
0.0241	Poison_x
0.0231	Ghost_x
0.0228	Fighting_y
0.0227	Attack_y
0.0226	Attack_x
0.0221	Psychic_y
0.0214	Ground_x
0.0203	Ground_y
0.0197	Steel_y
0.0186	Fighting_x
0.0170	Rock_y
0.0164	Fire_x
... 32 more ...	



## PDP for feature "Speed\_x"

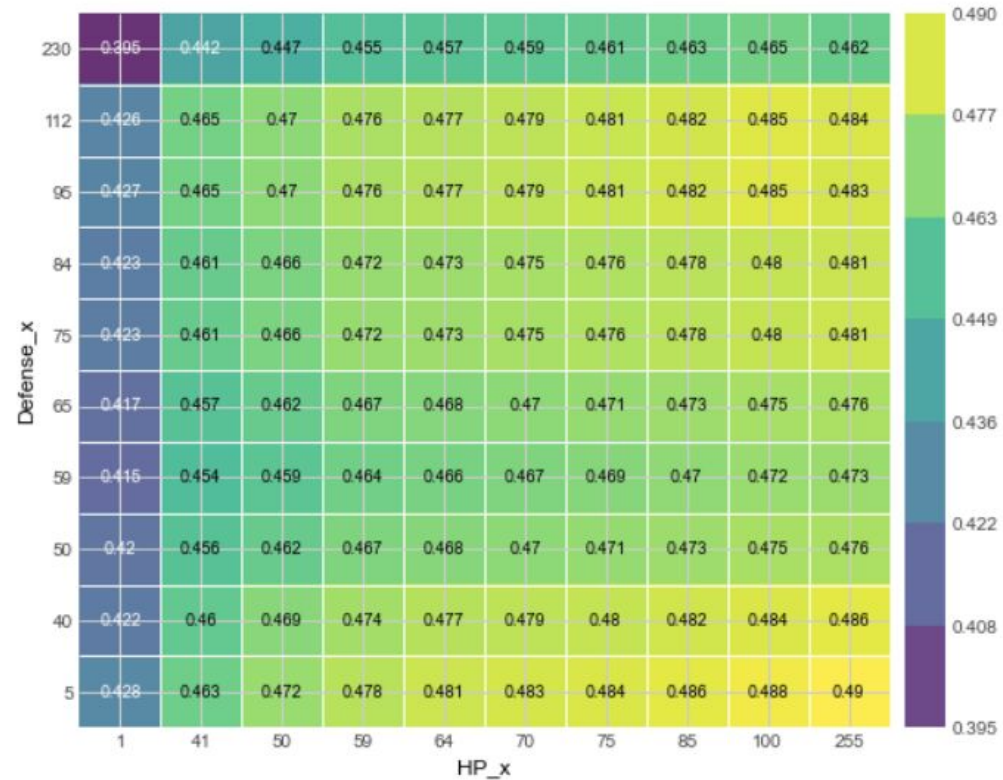
Number of unique grid points: 10

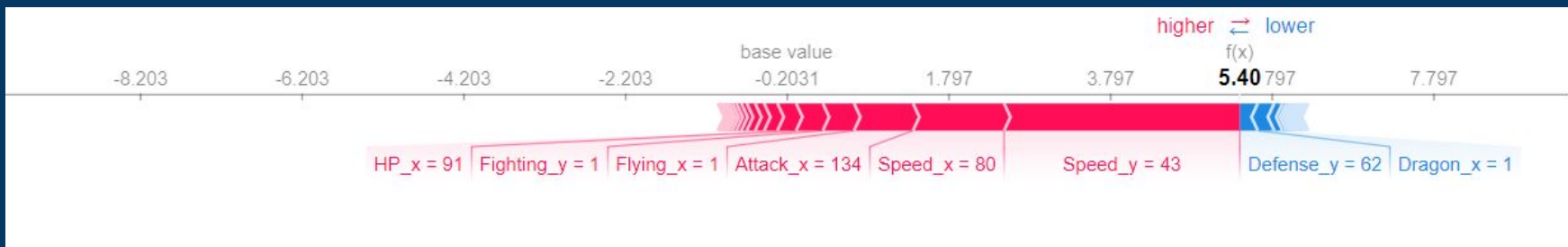
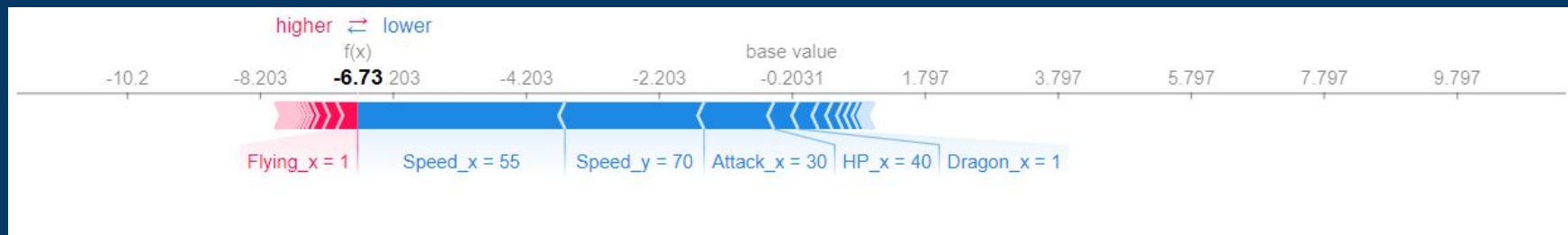
**PDP**



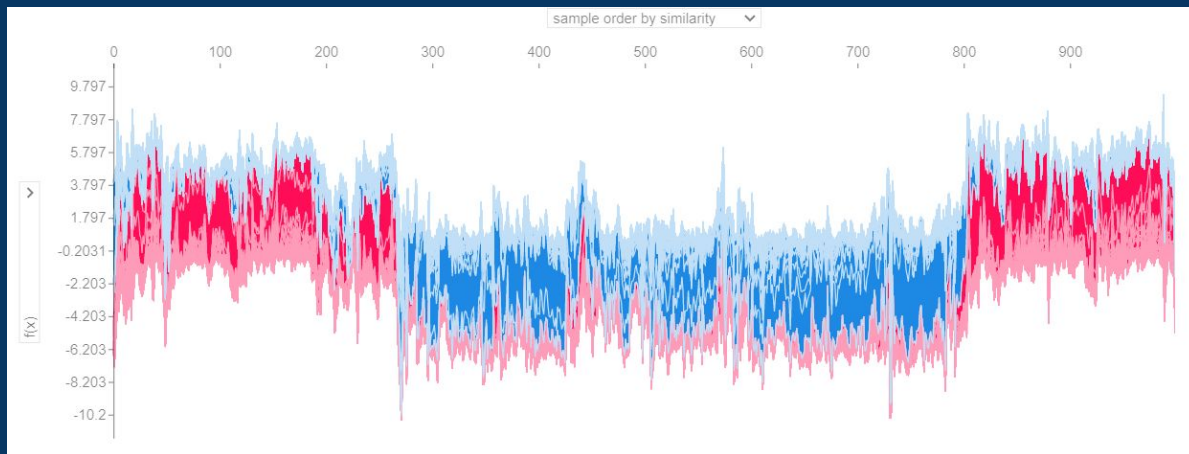
## PDP interact for "HP\_x" and "Defense\_x"

Number of unique grid points: (HP\_x: 10, Defense\_x: 10)

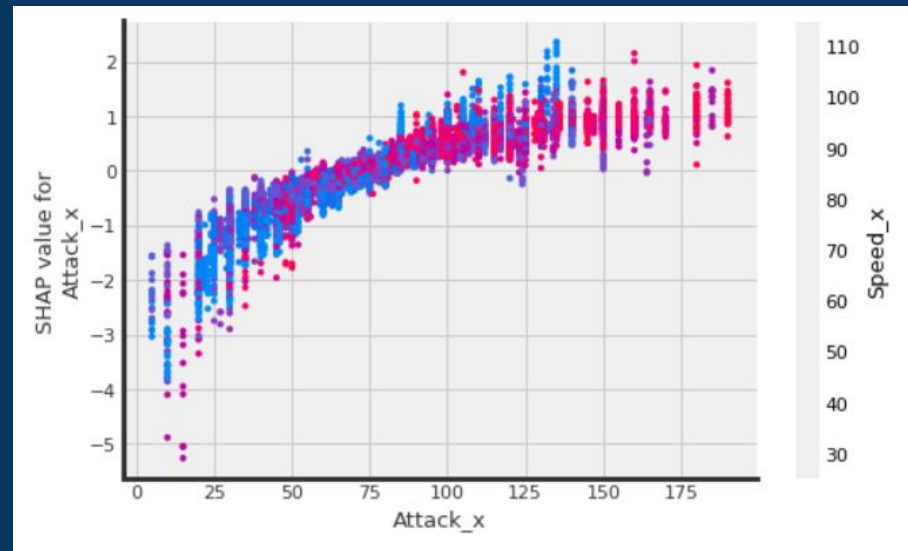
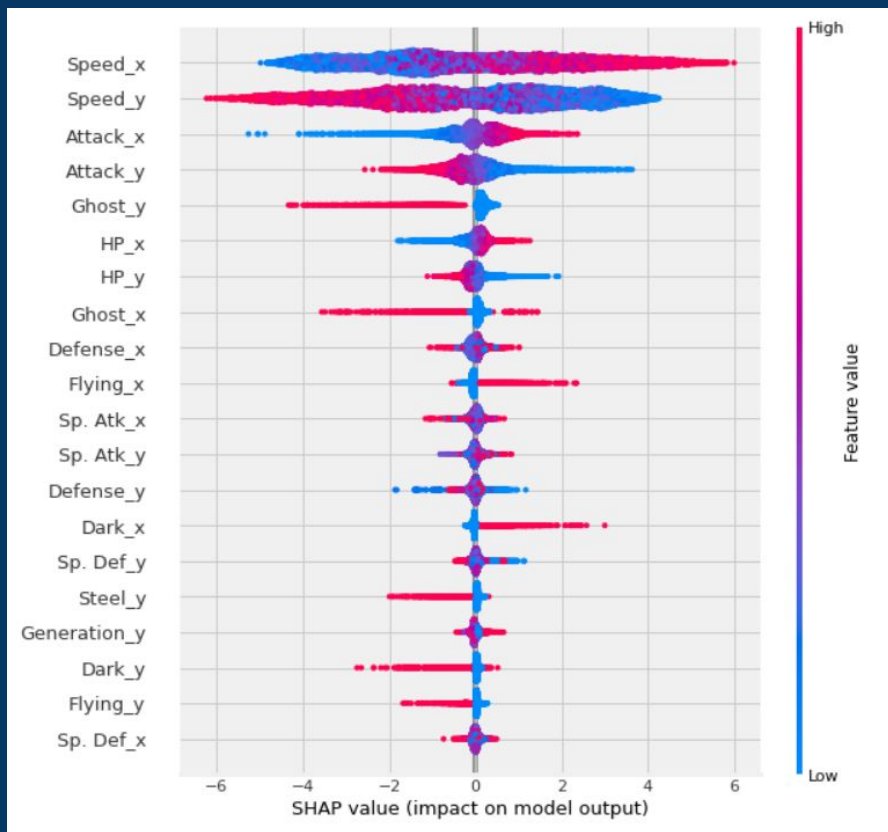




SHAP







## Conclusiones y siguientes pasos

**Práctica 1:** Web Scraping es una técnica muy sencilla para la obtención de datos y muy personalizable.

**Práctica 2:** Hemos aprendido a entrenar a los diferentes algoritmos con conjuntos de datos tratándolos previamente.

Hemos descubierto otra fuente de datos abierta como es Kaggle.

**Práctica 3:** Hemos conseguido reforzar el análisis de los datos de la Práctica 2, e interpretar el por qué de los resultados. Es una mejora respecto a la anterior porque observar un dataframe con muchos datos para sacar información es un trabajo engorroso.

Ya que las prácticas han estado más enfocadas hacia juegos y ocio, a futuro se podrían ampliar los objetivos, como por ejemplo, juegos más complejos, análisis del meta de juegos a nivel competitivo, y simuladores que puedan llevarse al entorno profesional.

# GRACIAS!

¿Alguna pregunta?