

Introducción a Machine Learning:

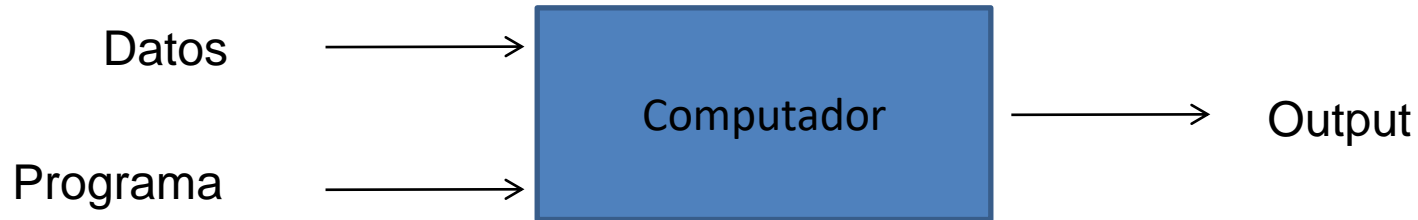
Aplicaciones con scikit-learn



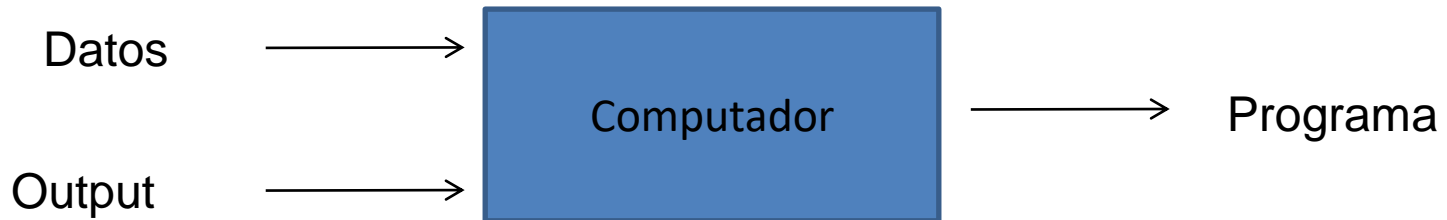
Víctor Fica León

¿Que es machine learning?

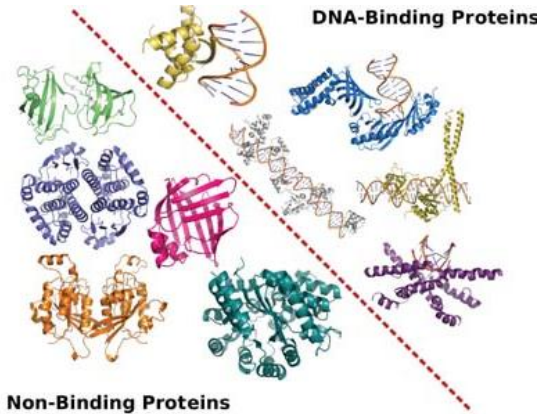
Programación Tradicional



Machine Learning



Aéreas de aplicación



Clasificación de proteínas

{2 → 2, 5 → 5, 4 → 8, 0 → 0, 2 → 2, 7 → 7, 5 → 5, 1 → 1,
3 → 3, 0 → 0, 3 → 3, 4 → 9, 6 → 6, 2 → 2, 8 → 8, 2 → 2,
0 → 0, 6 → 6, 6 → 6, 1 → 1, 1 → 1, 7 → 7, 8 → 8, 5 → 5,
0 → 0, 4 → 4, 7 → 7, 6 → 6, 0 → 0, 2 → 2, 5 → 5,
3 → 3, 1 → 1, 5 → 5, 6 → 6, 7 → 7, 5 → 5, 4 → 4, 1 → 1,
9 → 9, 3 → 3, 6 → 6, 8 → 8, 0 → 0, 9 → 9, 3 → 3,
0 → 0, 3 → 3, 7 → 7, 4 → 4, 4 → 4, 3 → 3, 8 → 8, 0 → 0,
4 → 4, 1 → 1, 3 → 3, 7 → 7, 6 → 6, 4 → 4, 7 → 7, 2 → 2,
7 → 7, 2 → 2, 5 → 5, 2 → 2, 0 → 0, 9 → 9, 8 → 8, 4 → 9,
8 → 8, 1 → 1, 6 → 6, 4 → 4, 8 → 8, 5 → 5, 8 → 8,
0 → 0, 6 → 6, 7 → 7, 4 → 4, 5 → 5, 8 → 8, 4 → 4,
3 → 3, 1 → 1, 5 → 5, 1 → 1, 9 → 9, 9 → 9, 9 → 9, 2 → 2,
4 → 4, 7 → 7, 3 → 3, 1 → 1, 9 → 9, 2 → 2, 9 → 9, 6 → 6}]

Reconocimiento de caracteres/ texto



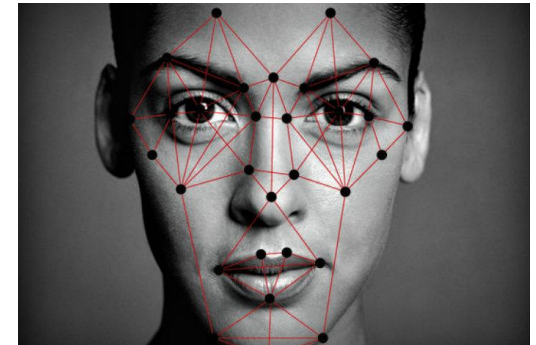
Vehículos autónomos



Reconocimiento del habla



Motores de recomendación



Reconocimiento Facial

Google Ads

Publicidad y Negocios

Tipos de aprendizaje

Supervisado

- Aprender desde datos etiquetados
- Ej. Clasificación de Spam

No supervisado

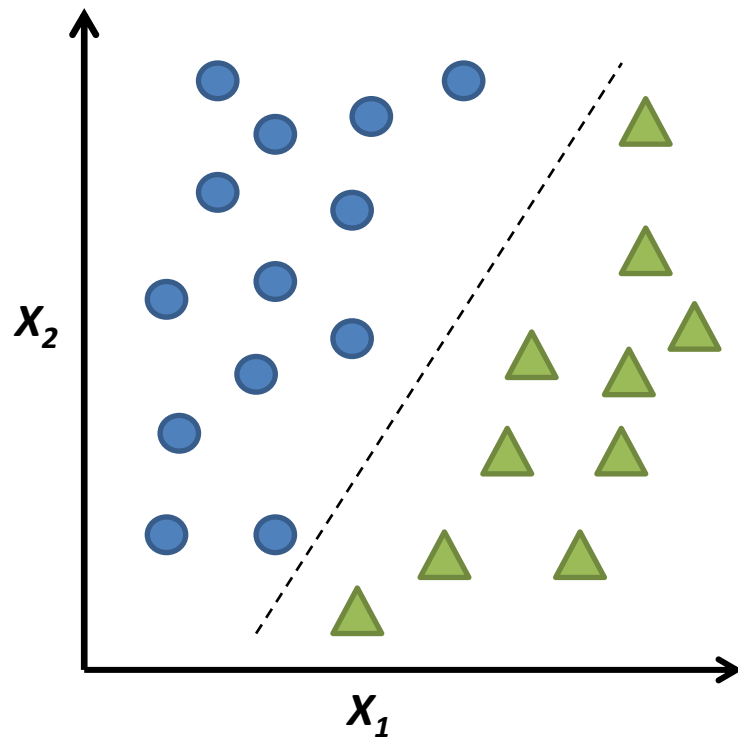
- Descubrir estructuras en datos no etiquetados
- Ej. Agrupamiento de documentos

Aprendizaje reforzado

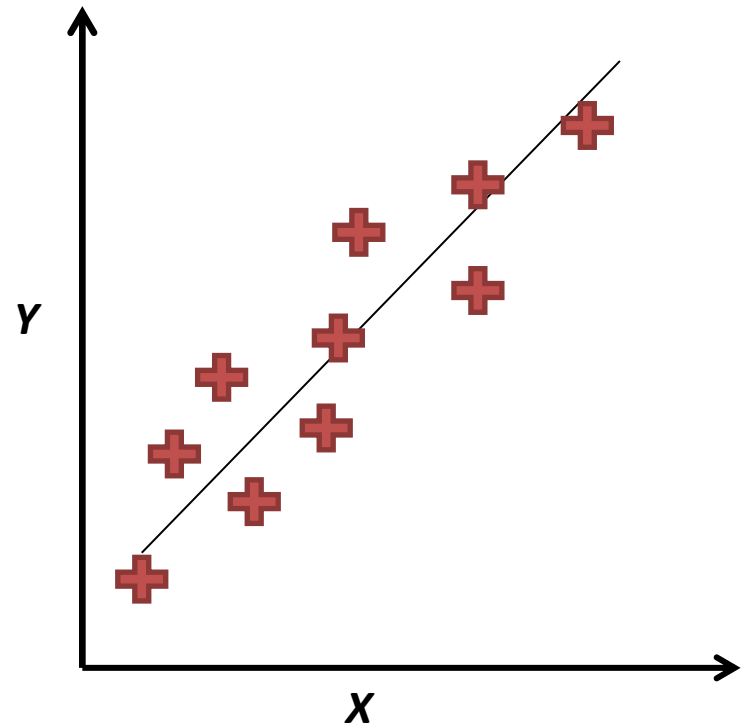
- Aprender interactuando con el ambiente
- Ej. Ajedrez

Aprendizaje Supervisado

Clasificación

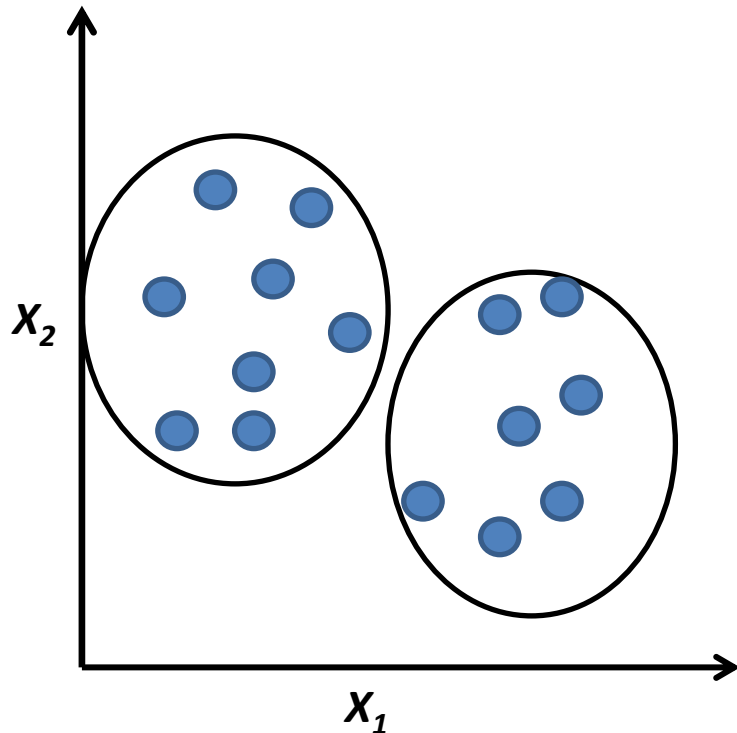


Regresión

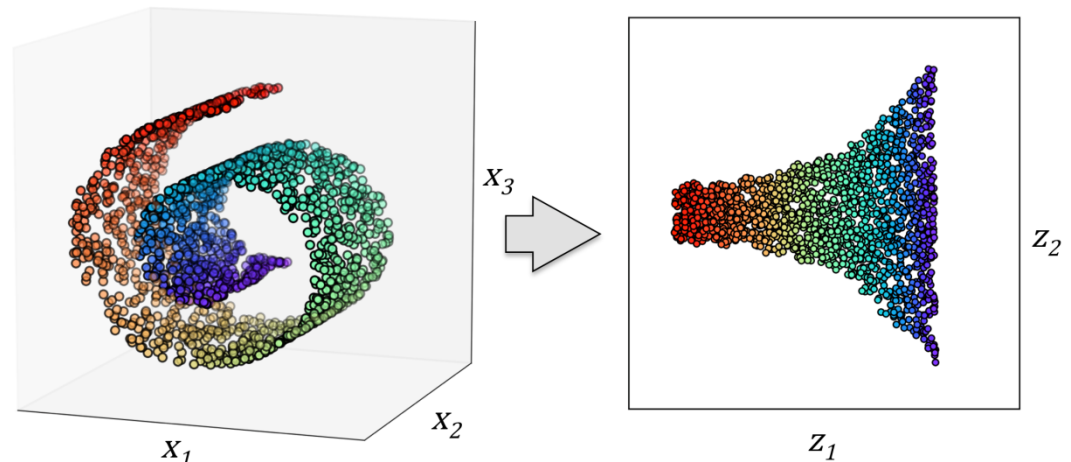


Aprendizaje No Supervisado

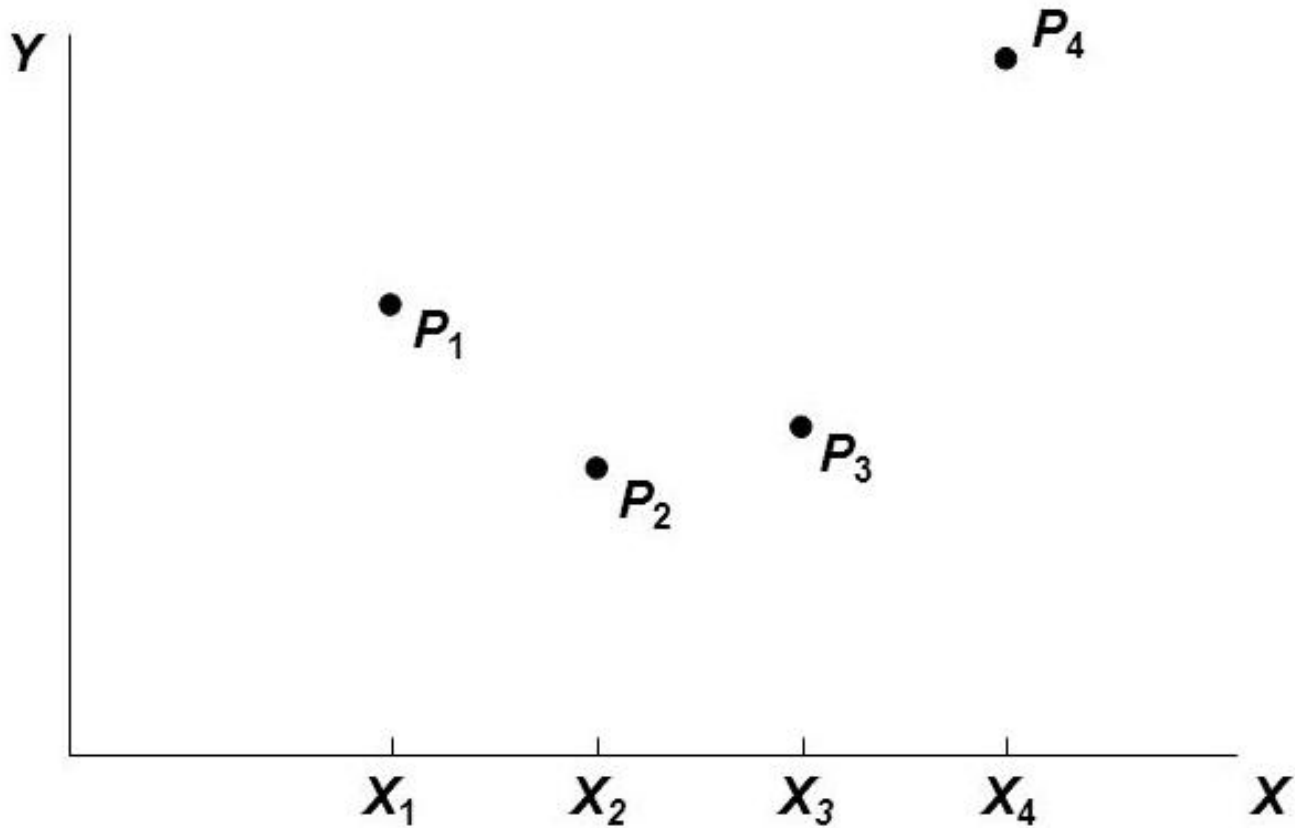
**Agrupamiento
/Clustering**



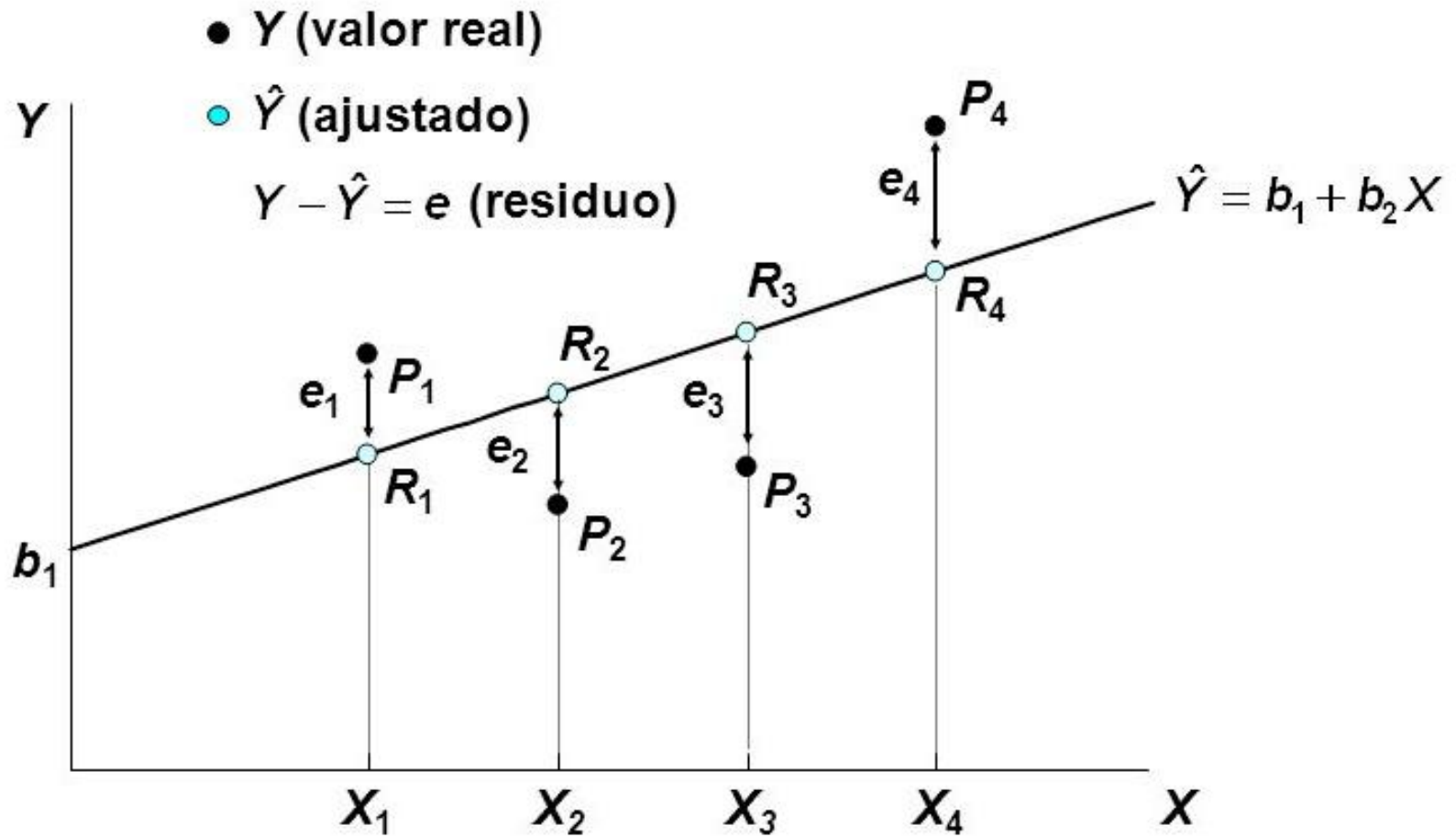
**Reducción
Dimensional**



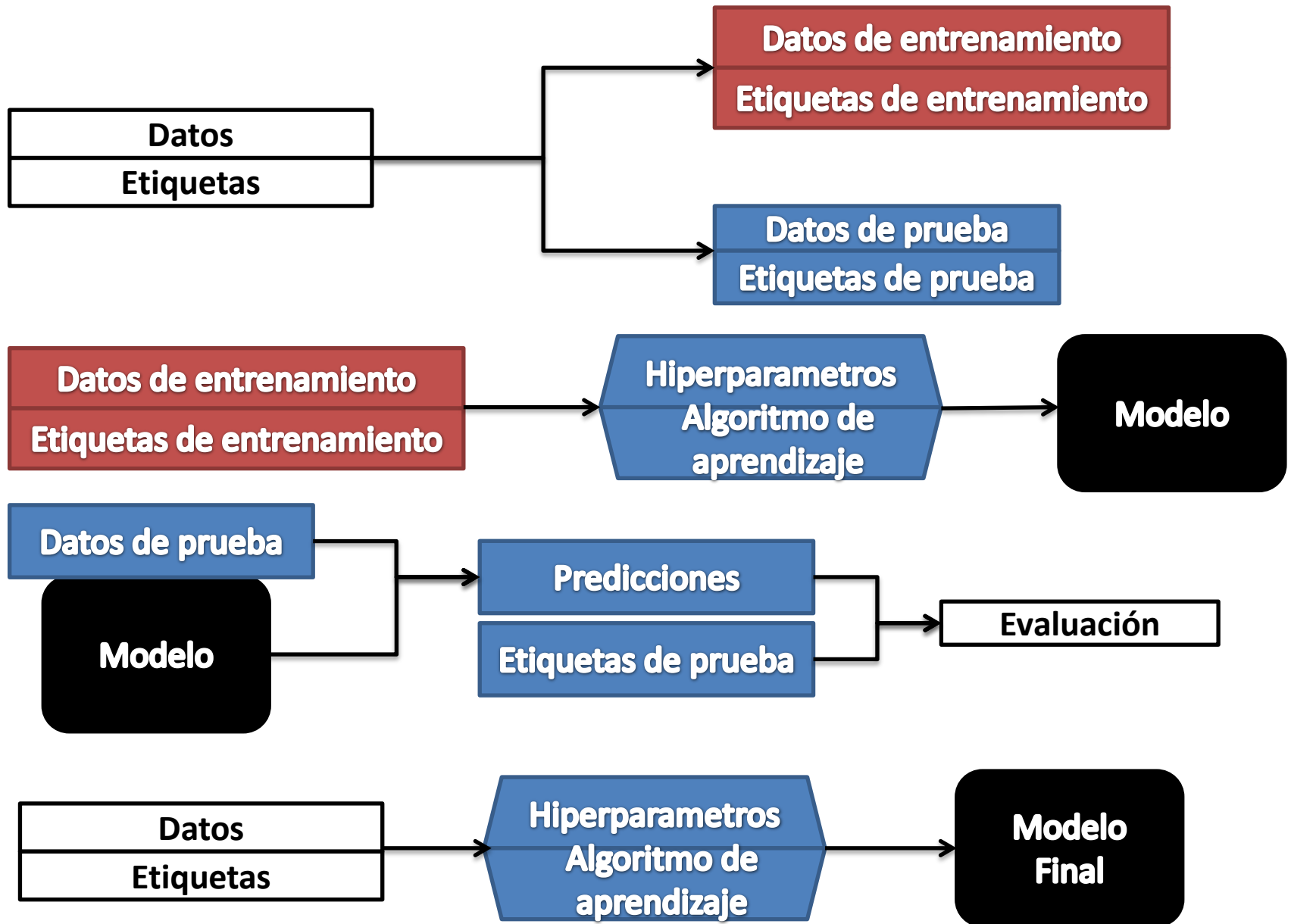
Modelo de Regresión Lineal Simple



Modelo de Regresión Lineal Simple

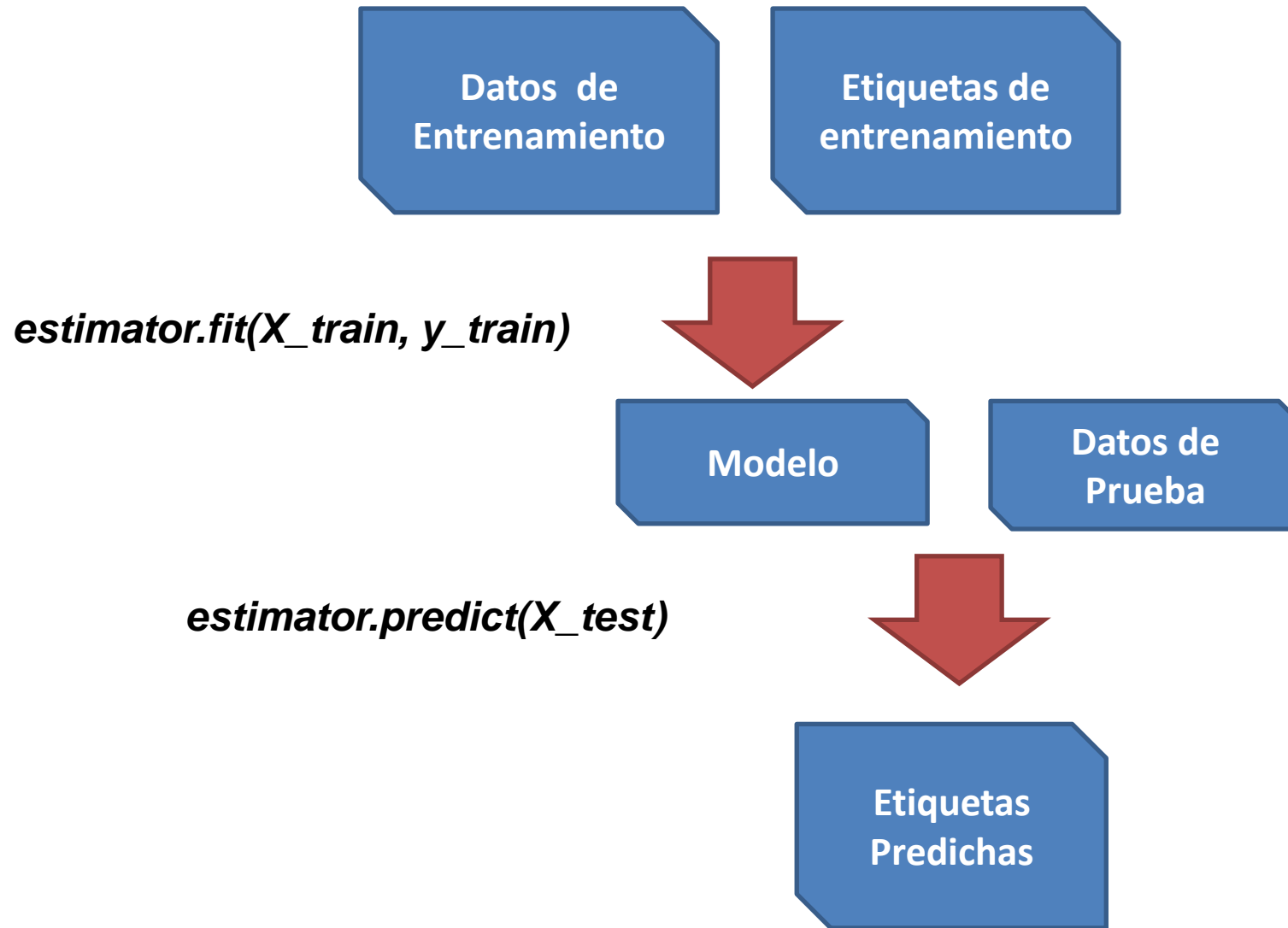


Esquema “básico” de aprendizaje supervisado



Generando Modelos

Como funciona scikit-learn



Clasificación de una Planta

Conjunto de datos Iris



Iris-setosa



Iris-versicolor



Iris-virginica

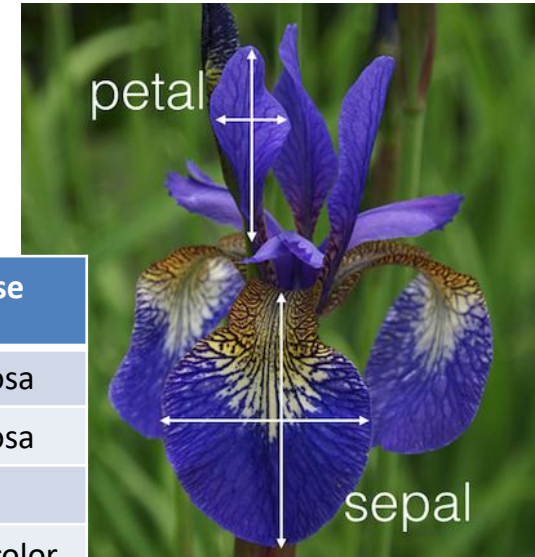
Clasificación de una Planta

Representación de los datos

Muestras/
Objetos

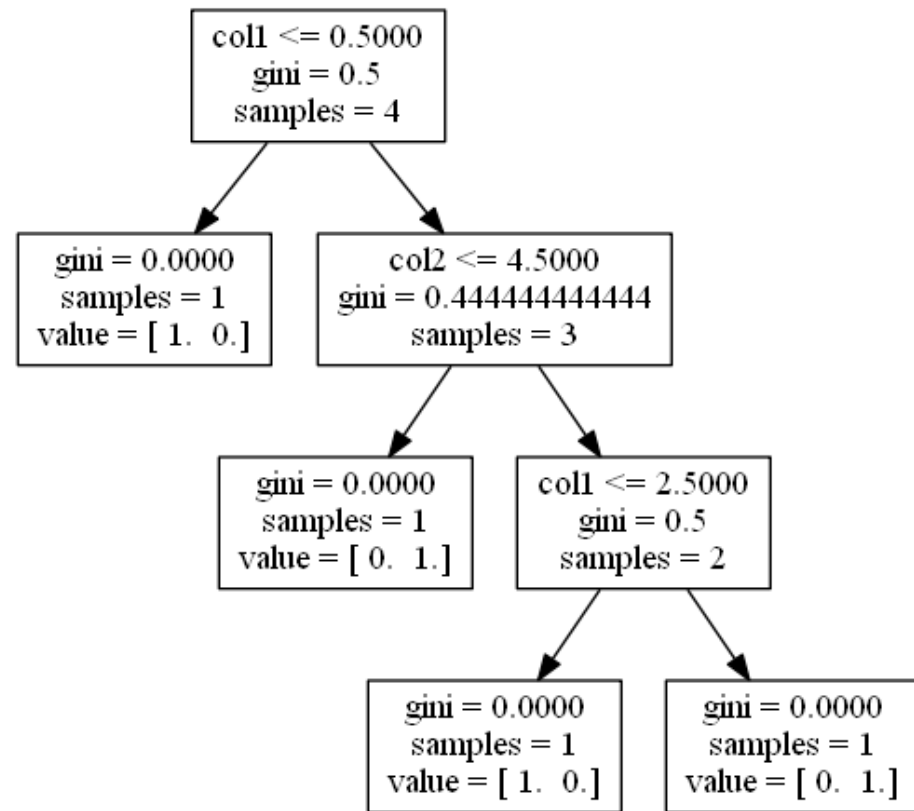
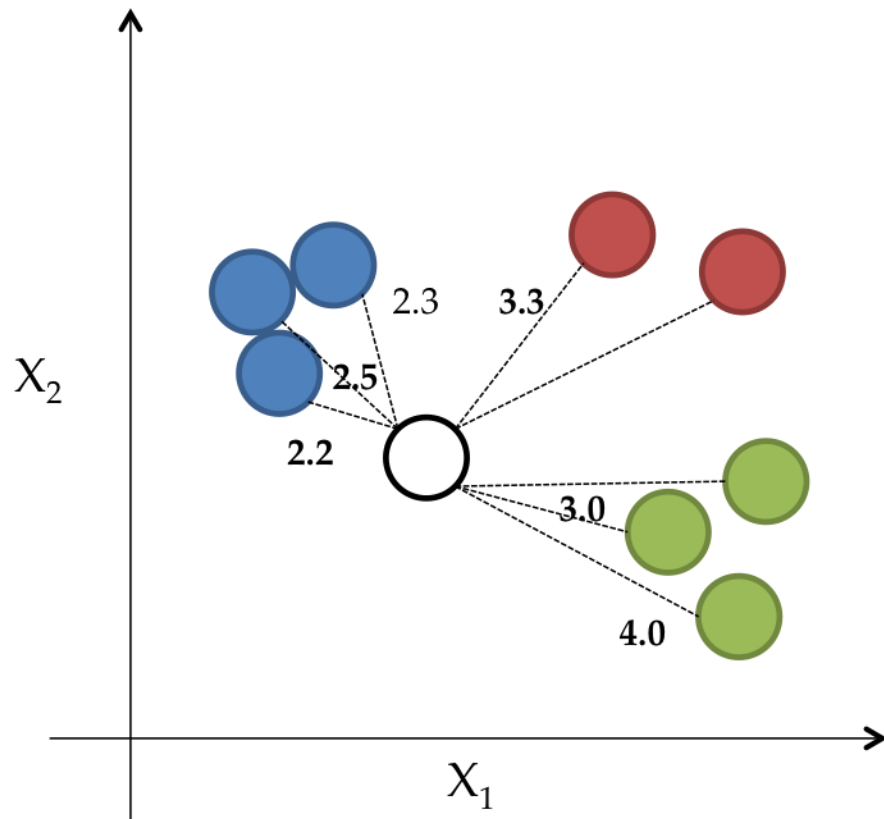
	Sepal_lenght	Sepal_width	Petal_lenght	Petal_width	clase
1	5.1	3.5	1.4	0.2	Setosa
2	4.9	3.0	1.4	0.2	Setosa
...
50	6.4	3.2	4.5	1.5	Versicolor
...
150	5.9	3.0	5.1	1.8	Virginica

Características
/ Atributos



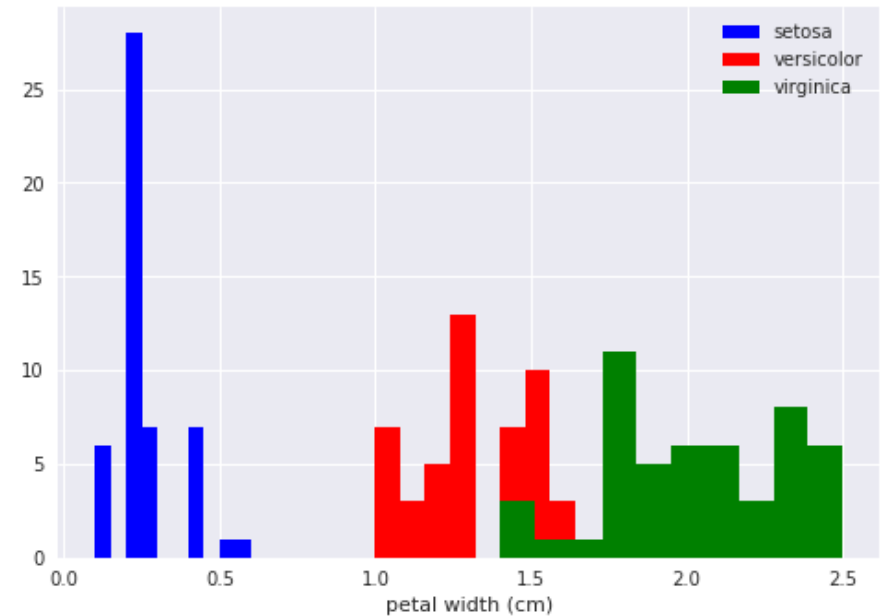
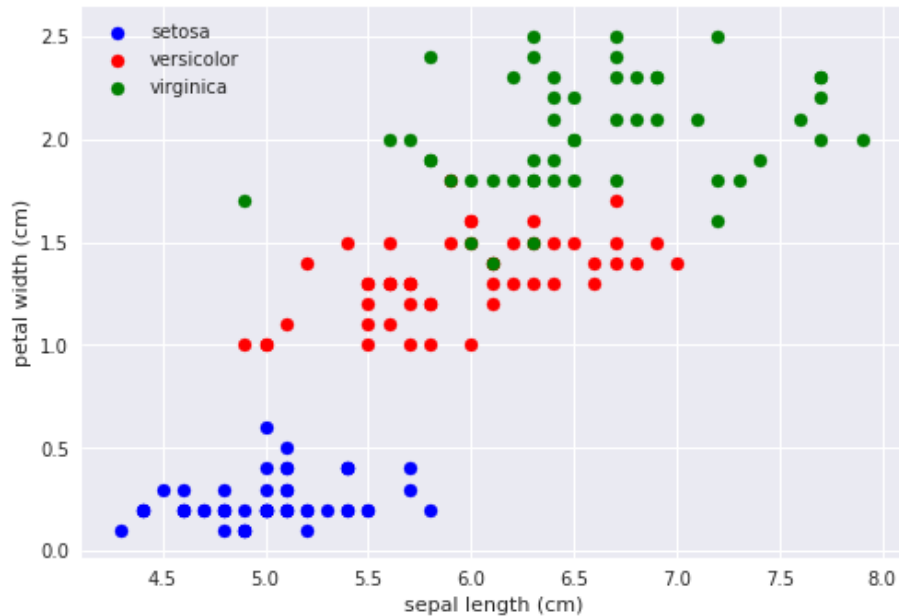
Etiqueta/Clase

Ejemplos de modelos de clasificación



Visualización de los datos

Representación de los datos

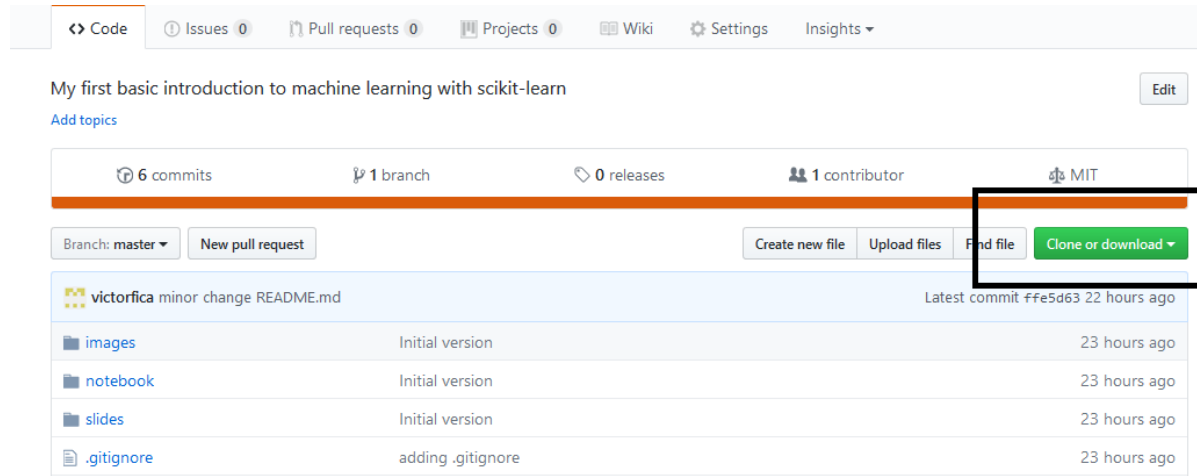


Para el tutorial de hoy...

Recomiendo que revisen el repositorio github donde tengo todo el material de hoy.

<https://github.com/victorfica/ML-tutorial2017-vfica>

Están las instrucciones ahí, básicamente los que tienen github instalado pueden clonar el repositorio, y los que no, pueden descargar todo el material aquí...



The screenshot shows the GitHub interface for the repository 'victorfica/ML-tutorial2017-vfica'. The top navigation bar includes links for Code, Issues (0), Pull requests (0), Projects (0), Wiki, Settings, and Insights. Below this, the repository description reads 'My first basic introduction to machine learning with scikit-learn'. The repository statistics show 6 commits, 1 branch, 0 releases, and 1 contributor. The 'Clone or download' button is highlighted with a black box. The file list shows the following files and their commit history:

File	Commit	Time
images	Initial version	23 hours ago
notebook	Initial version	23 hours ago
slides	Initial version	23 hours ago
.gitignore	adding .gitignore	23 hours ago

Requerimientos

Recomiendo utilizar miniconda. Fácil de instalar.

Una vez instalado ejecutar el comando:

conda install numpy scipy matplotlib scikit-learn notebook seaborn