


Inteligencia Artificial con Python y scikit-learn

[Install](#) [User Guide](#) [API](#) [Examples](#) [Community](#) [More](#)

scikit-learn

Machine Learning in Python

[Getting Started](#) [Release Highlights for 1.6](#)

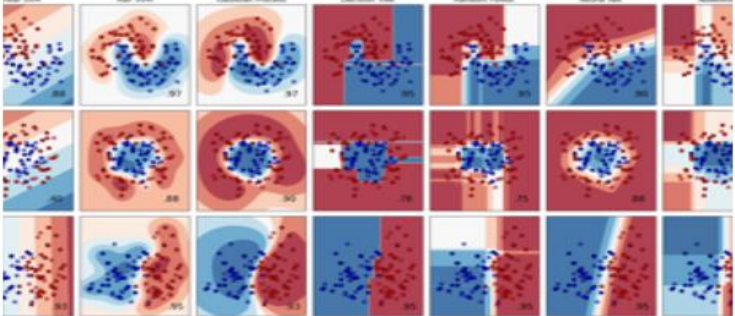
- Simple and efficient tools for predictive data analysis
- Accessible to everybody, and reusable in various contexts
- Built on NumPy, SciPy, and matplotlib
- Open source, commercially usable - BSD license

Classification

Identifying which category an object belongs to.

Applications: Spam detection, image recognition.

Algorithms: [Gradient boosting](#), [nearest neighbors](#), [random forest](#), [logistic regression](#), and [more...](#)

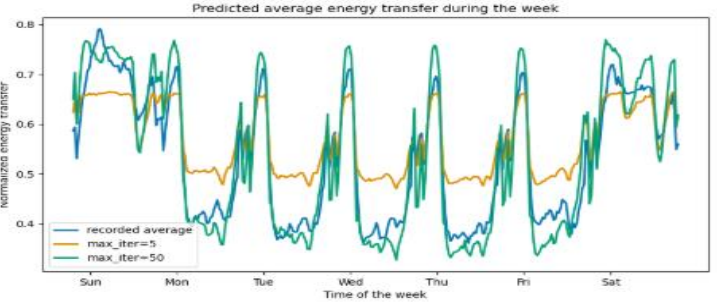


Regression

Predicting a continuous-valued attribute associated with an object.

Applications: Drug response, stock prices.

Algorithms: [Gradient boosting](#), [nearest neighbors](#), [random forest](#), [ridge](#), and [more...](#)




Clustering

Automatic grouping of similar objects into sets.

Applications: Customer segmentation, grouping experiment outcomes.

Algorithms: [k-Means](#), [HDBSCAN](#), [hierarchical clustering](#), and [more...](#)

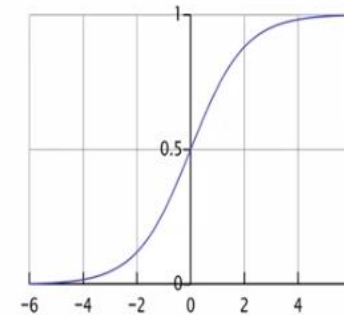
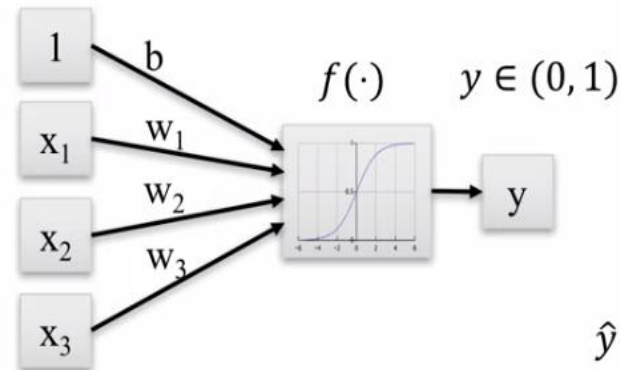


Aprendizaje automático aplicado

Logistic Regression

Regresión Logística

Input features



$$\hat{y} = \text{logistic}(\hat{b} + \hat{w}_1 \cdot x_1 + \cdots \hat{w}_n \cdot x_n)$$

$$= \frac{1}{1 + \exp[-(\hat{b} + \hat{w}_1 \cdot x_1 + \cdots \hat{w}_n \cdot x_n)]}$$

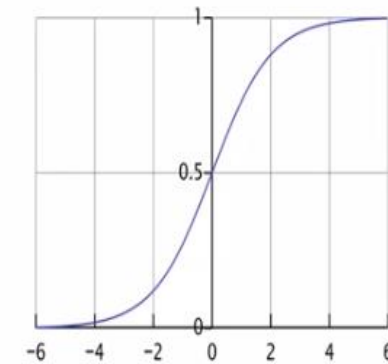
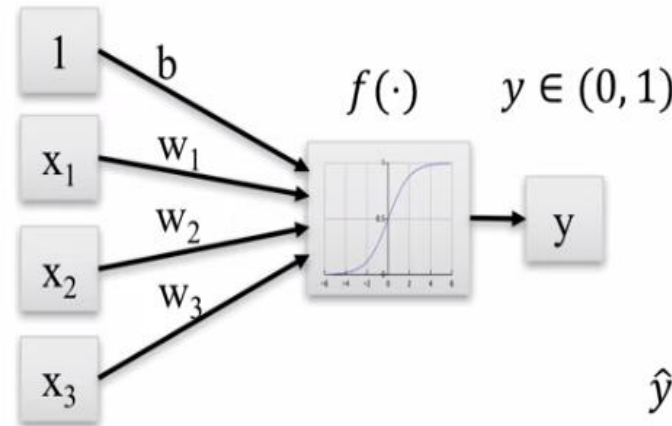
Regresión logística

- La **regresión logística** es una técnica estadística y de aprendizaje automático utilizada para modelar relaciones entre una variable dependiente categórica y una o más variables independientes.
- Es particularmente **útil cuando el resultado que se desea predecir es binario** (dos categorías, como "sí/no" o "1/0").
- A pesar de ser llamado medida de regresión, en realidad **se utiliza para la clasificación**

Regresión logística

- La regresión logística toma un conjunto de variables de entrada, las características (variables), y estima un valor objetivo.
- La regresión logística es similar a la regresión lineal, la diferencia consiste en que el modelo **de regresión logística además calcula una suma ponderada** de la entrada de las características x_i y el término de intercepción b , pero **ejecuta este resultado a través de una función no lineal f** , para producir la salida y (predicción).

Input features



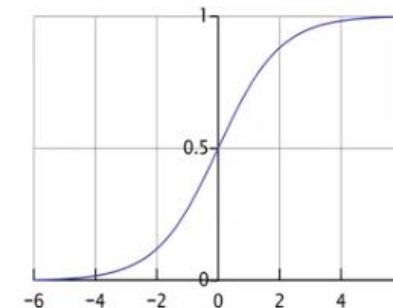
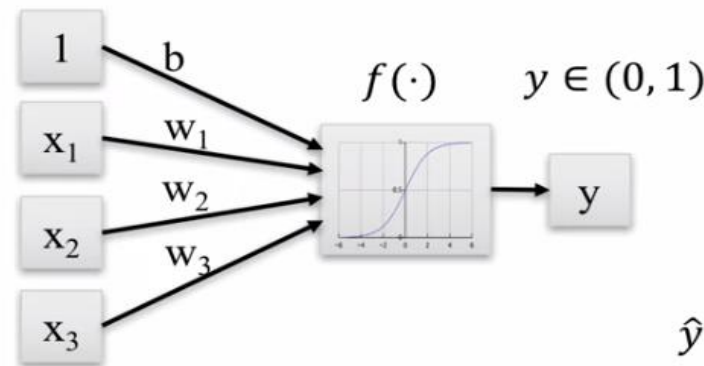
$$\hat{y} = \text{logistic}(\hat{b} + \hat{w}_1 \cdot x_1 + \dots \hat{w}_n \cdot x_n)$$

$$= \frac{1}{1 + \exp[-(\hat{b} + \hat{w}_1 \cdot x_1 + \dots \hat{w}_n \cdot x_n)]}$$

Regresión logística

- El efecto de **aplicar la función logística** es **comprimir la salida de la función lineal** para que se **limite a un rango entre 0 y 1**.
- Es una función en **forma de S** que se acerca a **1** a medida que **el valor de entrada aumenta por encima de 0** y más cerca de **0** a medida que **el valor de entrada disminuye** muy por debajo de 0.
- Primero calcula la misma combinación lineal de las entradas **x_i** , los pesos de las características o variables predictoras **w_i** , y el valor de intercept **b** , pero lo ejecuta a través de **aplicar la función logística** para producir **y** .

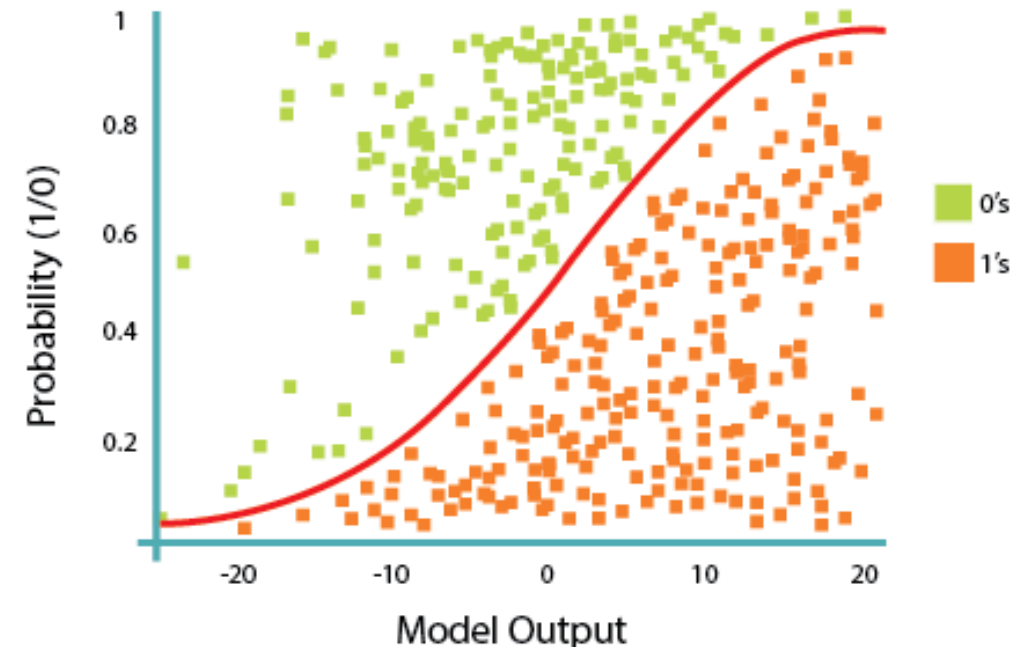
Input features



$$\hat{y} = \text{logistic}(\hat{b} + \hat{w}_1 \cdot x_1 + \cdots \hat{w}_n \cdot x_n)$$
$$= \frac{1}{1 + \exp[-(\hat{b} + \hat{w}_1 \cdot x_1 + \cdots \hat{w}_n \cdot x_n)]}$$

Regresión logística

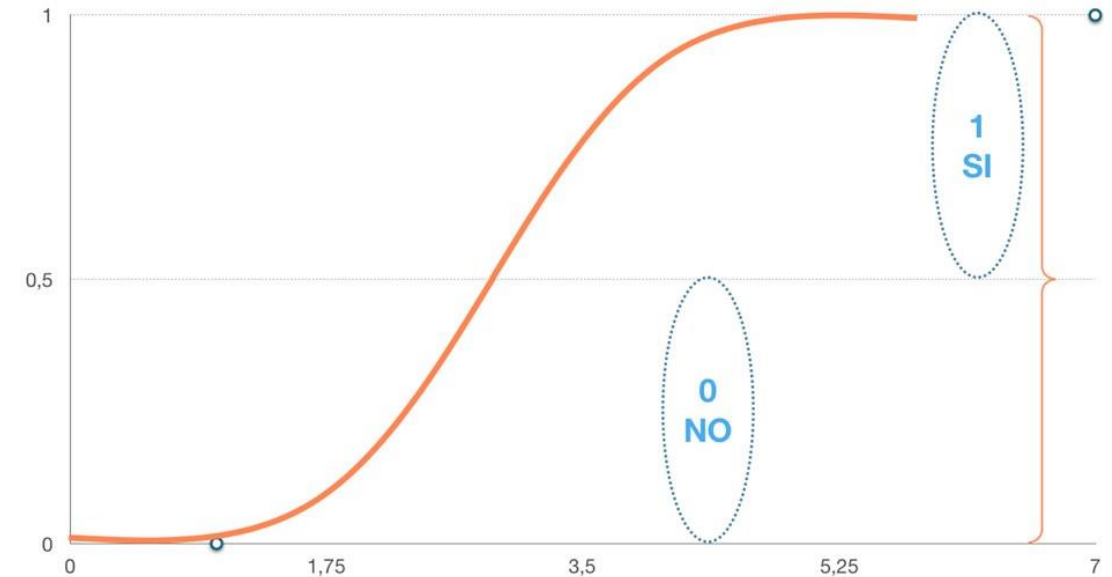
- A diferencia de la regresión lineal ordinaria, en su forma más básica **el valor es una variable binaria** en lugar de un valor continuo.
- La regresión logística también se puede **usar en casos donde el valor objetivo que se predice es una variable categórica de clase múltiple**, no solo binaria.
- Se establece un **umbral para clasificar los valores predichos**. Por defecto, este umbral suele ser 0.5.
- Si la **probabilidad** predicha **es mayor o igual a 0.5**, se clasifica como una **categoría** (por ejemplo, "1"), y si es menor, se clasifica como la otra ("0").



Regresión logística

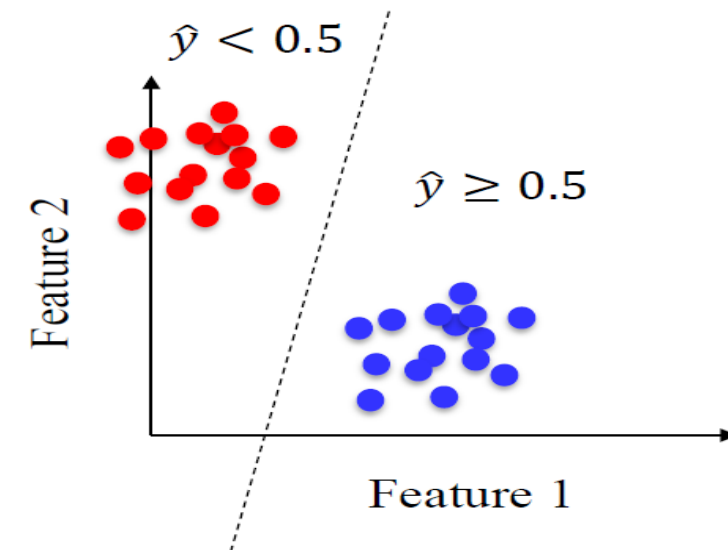
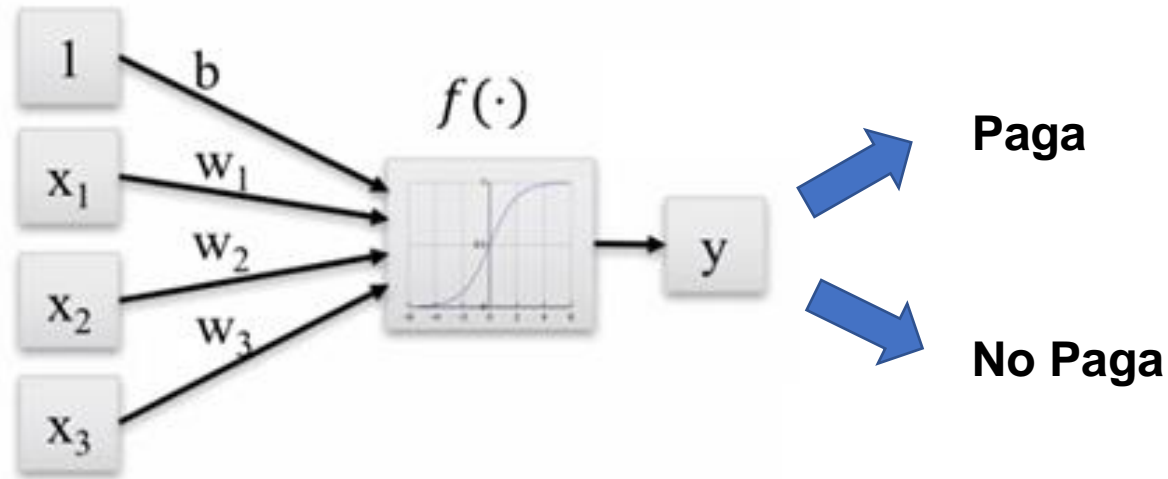
Aplicaciones comunes

- Predicción de enfermedades (diagnóstico médico).
- Clasificación de correos electrónicos (spam/no spam).
- Predicción de comportamiento del cliente (compra/no compra).
- Análisis de riesgos (default/no default en préstamos).



Practica

- Construir un modelo de clasificación para **predecir si un préstamo se pagará en su totalidad.**
- **El propósito del modelo es identificar los préstamos que probablemente se pagarán en su totalidad.**



Referencias

Python Intermedio

<https://python-intermedio.readthedocs.io/es/latest/>

Pandas_Cheat_Sheet.

https://pandas.pydata.org/Pandas_Cheat_Sheet.pdf

NearestNeighborsClassification

https://scikit-learn.org/stable/auto_examples/neighbors/plot_classification.html

Confusionmatrix

https://scikit-learn.org/stable/modules/generated/sklearn.metrics.confusion_matrix.html