


# 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](#)

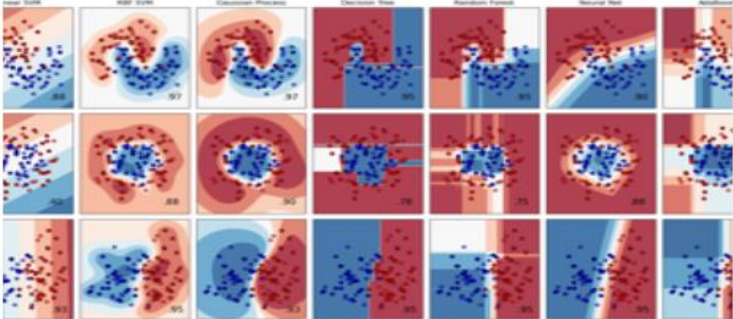
- Simple and efficient tools for predictive data analysis
- Accessible to everybody, and reusable in various contexts
- Built on NumPy, SciPy, and matplotlib
- Open source, commercially usable - BSD license

### Classification

Identifying which category an object belongs to.

**Applications:** Spam detection, image recognition.

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

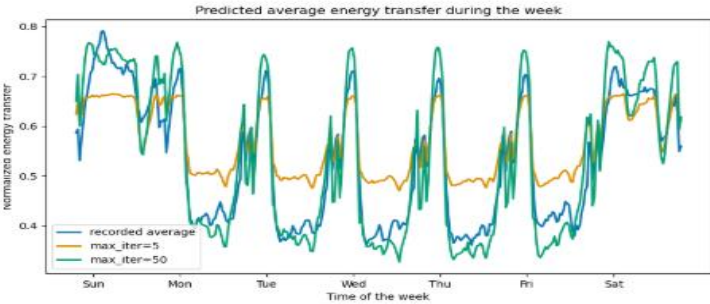


### Regression

Predicting a continuous-valued attribute associated with an object.

**Applications:** Drug response, stock prices.

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

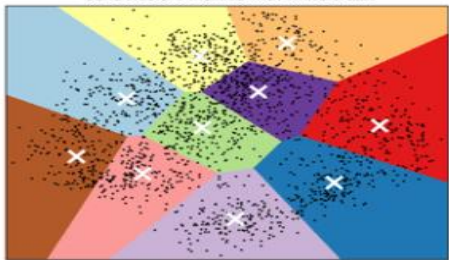


### Clustering

Automatic grouping of similar objects into sets.

**Applications:** Customer segmentation, grouping experiment outcomes.

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

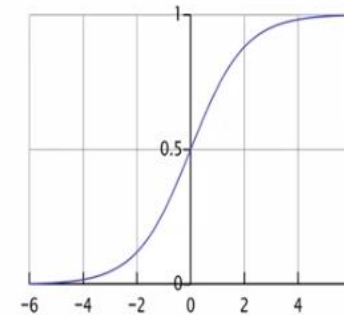
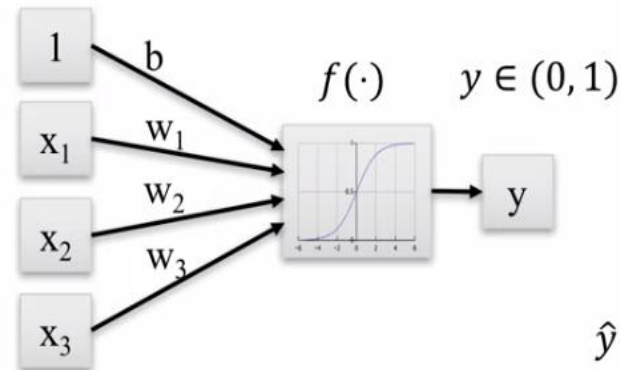


# Aprendizaje automático aplicado

## Logistic Regression

### Regresión Logística

Input features



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

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

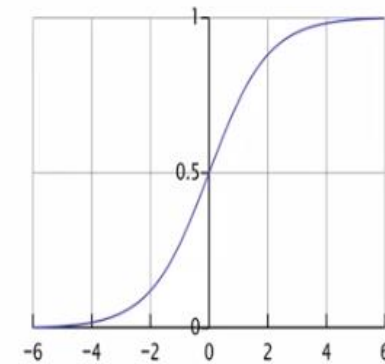
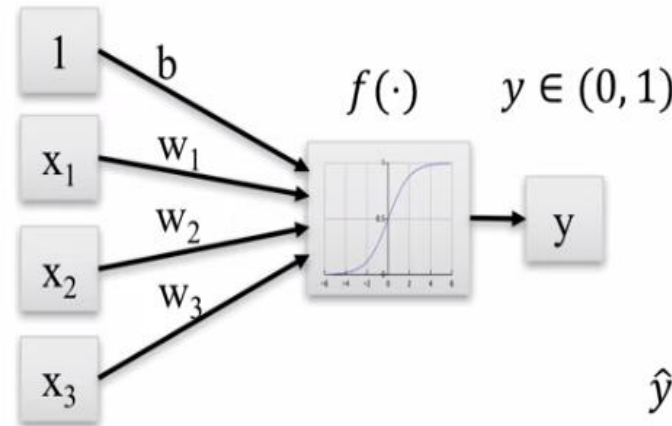
# Regresión logística

- 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



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

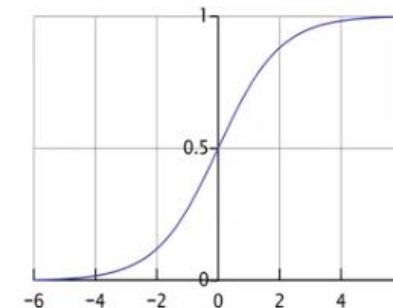
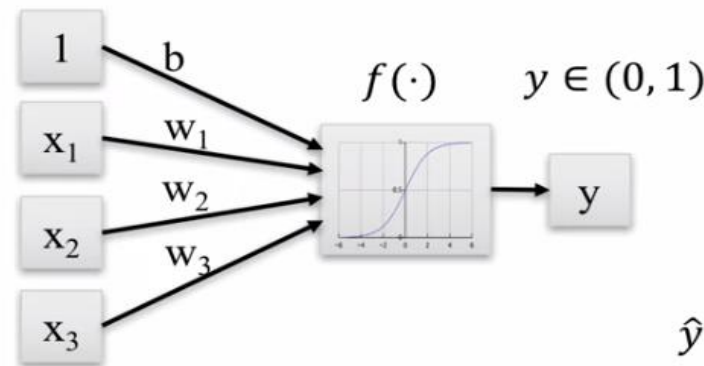
$$= \frac{1}{1 + \exp[-(\hat{b} + \hat{w}_1 \cdot x_1 + \cdots \hat{w}_n \cdot x_n)]}$$



# Regresión logística

- 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  **$y$** .

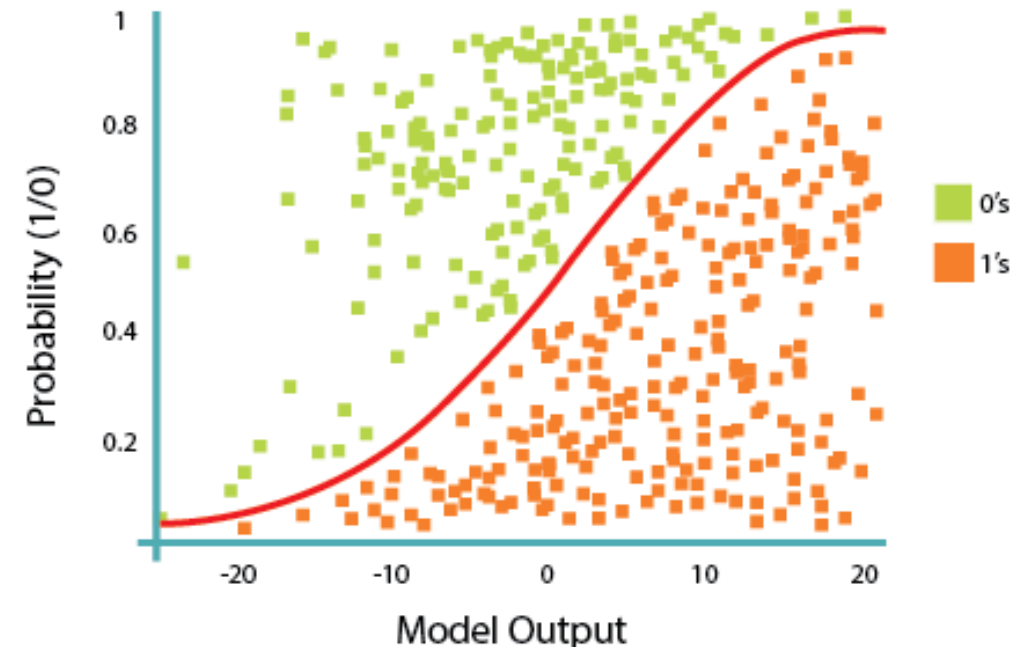
Input features



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

# Regresión logística

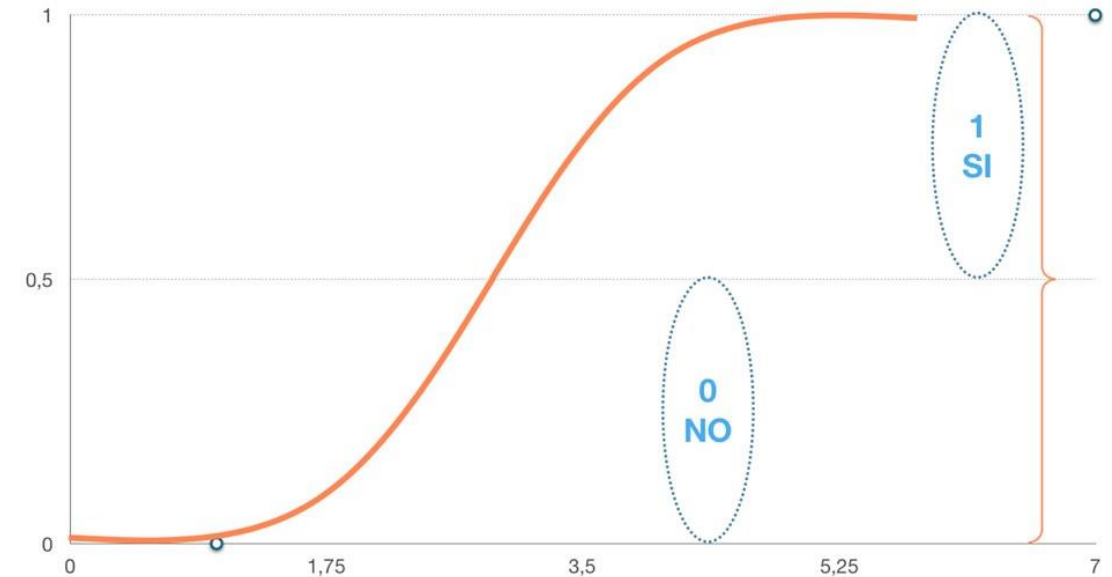
- A diferencia de la regresión lineal ordinaria, en su forma más básica **el valor es una variable binaria** en lugar de un valor continuo.
- La regresión logística también se puede **usar en casos donde el valor objetivo que se predice es una variable categórica de clase múltiple**, no solo binaria.
- 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).





**EDUCACIÓN**  
SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO  
NACIONAL DE MÉXICO

**TECNOLÓGICO NACIONAL DE MÉXICO**  
**INSTITUTO TECNOLÓGICO DE DURANGO**



# Práctica Python



## Club de préstamos LendingClub

- Es una compañía estadounidense de **préstamos entre pares**, con sede en San Francisco, California.
- Es la plataforma de préstamos peer-to-peer más grande del mundo.
- Permite a los **prestatarios solicitar préstamos personales no garantizados** entre 1,000 y 40,000.



**LendingClub** What We Offer ▾ Resources ▾ About Us ▾ Help [Sign In](#)

### Personal Loans Up to \$40,000

Check your rate. It won't impact your credit score. Privacy & security PROTECTION

\$ How much do you need? What's the money for? ▾ [Check Your Rate](#)

[Respond to Mail Offer](#)




#### Do More With a Personal Loan

Pay down high-interest debt, consolidate credit card balances – and save more than \$1,000.\* With single monthly payments, lower fixed interest rates, and no prepayment penalties, a personal loan from LendingClub Bank can help you simplify your finances and take control of your debt.

[Check Your Rate](#)

## Club de préstamos LendingClub

### P2P Lending: Key Players

			
<b>Overview and market share</b>	<ul style="list-style-type: none"><li>• Largest P2P lending platform</li><li>• Market leader— own about 70% share of market</li></ul>	<ul style="list-style-type: none"><li>• Second largest P2P lending player</li><li>• About 30% share of market</li></ul>	<ul style="list-style-type: none"><li>• Emerging competitor</li><li>• Likely &lt;1% market share</li></ul>
<b>Founding and total loans issued</b>	<ul style="list-style-type: none"><li>• Opened to investors in 2007</li><li>• \$5.0B in loans to date</li></ul>	<ul style="list-style-type: none"><li>• First entrant into P2P lending; founded in 2006</li><li>• \$2.0B in loans to date</li></ul>	<ul style="list-style-type: none"><li>• Founded in 2012</li><li>• Likely about 5,000 loans funded to date</li></ul>
<b>Borrower profile</b>	<ul style="list-style-type: none"><li>• Has historically attracted higher-quality borrowers, with approval rates at ~10%</li></ul>	<ul style="list-style-type: none"><li>• Initially catered to more sub-prime borrowers; suffered a 22% total charge-off rate in 2010</li></ul>	<ul style="list-style-type: none"><li>• Targeted towards younger borrowers with less established credit histories</li></ul>
<b>Minimum FICO</b>	<ul style="list-style-type: none"><li>• 660</li></ul>	<ul style="list-style-type: none"><li>• 640 (imposed after high post-recession charge offs)</li></ul>	<ul style="list-style-type: none"><li>• 640, but do accept insufficient credit information</li></ul>
<b>Loan terms</b>	<ul style="list-style-type: none"><li>• 3 and 5 year terms</li></ul>	<ul style="list-style-type: none"><li>• 3 and 5 year terms</li></ul>	<ul style="list-style-type: none"><li>• 3 year term</li></ul>
<b>APR range</b>	<ul style="list-style-type: none"><li>• 6.78 – 29.99%</li></ul>	<ul style="list-style-type: none"><li>• 6.73 – 35.36%</li></ul>	<ul style="list-style-type: none"><li>• 6 – 24.60%</li></ul>
<b>Most common use of loan</b>	<ul style="list-style-type: none"><li>• Credit card debt consolidation / refinancing</li></ul>	<ul style="list-style-type: none"><li>• Credit card debt consolidation / refinancing</li></ul>	<ul style="list-style-type: none"><li>• Credit card debt consolidation / refinancing</li></ul>





# Club de préstamos LendingClub

## Lending Club: Investing

### Browse Notes

Summary | Invest | **Browse Notes** | Alert | Transfer | Trading Account | Automated Investing

Available: \$410.82

Build a Portfolio

Per Note: \$25

Filter Notes Save | Open

CREDIT Score

660 850

660 850

Delinquencies (Last 2 yrs)

☒ Any

☐ 0

☐ 1-3

☐ 4+

Interest Rate

Keyword

Months Since Last Delinquency

Loan Term

Exclude Loans already invested in

More Filters

Update Results

Add to Order

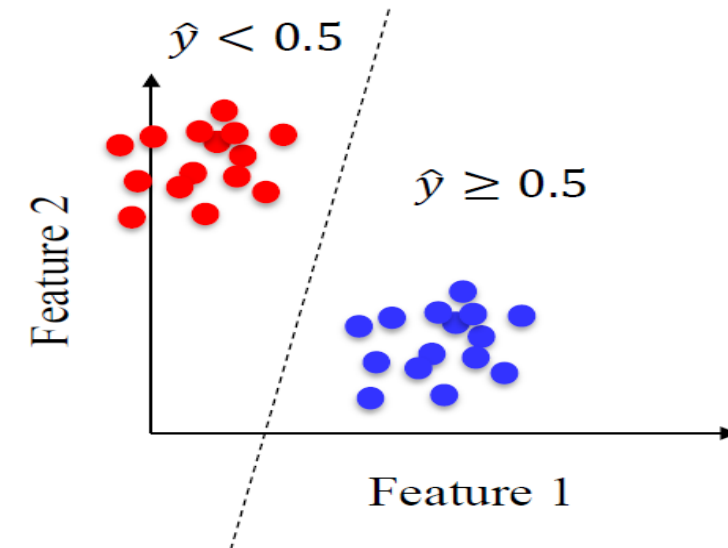
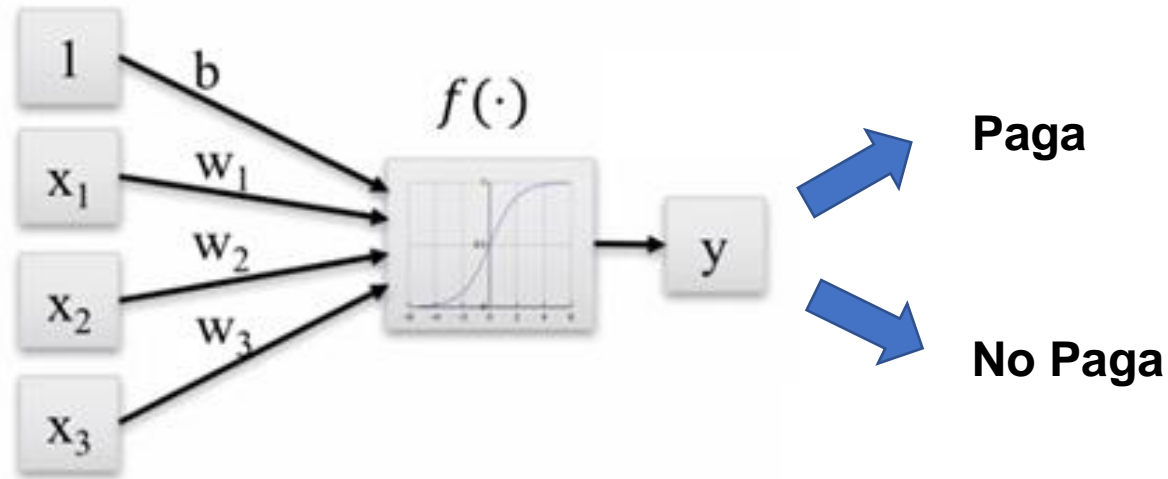
Showing Notes 1 - 15 of 1509

Investment	Rate	Term	FICO®	Amount	Purpose	% Funded	Amount / Time Left
<input type="checkbox"/> \$0	A 4 7.69%	36	815-819	\$3,000	Loan Refinancing & Consolidation	99%	\$25 / 12 days
<input type="checkbox"/> \$0	B 4 11.67%	36	660-664	\$4,000	Credit Card Payoff	98%	\$50 / 12 days
<input type="checkbox"/> \$0	A 5 8.39%	36	695-699	\$10,200	Loan Refinancing & Consolidation	95%	\$475 / 11 days
<input type="checkbox"/> \$0	A 6 8.39%	36	725-729	\$4,900	Loan Refinancing & Consolidation	94%	\$275 / 11 days
<input type="checkbox"/> \$0	B 0 10.99%	36	690-694	\$4,000	Credit Card Payoff	89%	\$425 / 11 days
<input type="checkbox"/> \$0	B 1 9.17%	36	665-669	\$6,000	Home Improvement	83%	\$375 / 11 days
<input type="checkbox"/> \$0	B 4 11.67%	36	685-689	\$6,000	Loan Refinancing & Consolidation	90%	\$600 / 11 days
<input type="checkbox"/> \$0	A 4 7.69%	36	665-669	\$5,000	Loan Refinancing & Consolidation	85%	\$725 / 11 days
<input type="checkbox"/> \$0	B 2 10.15%	36	715-719	\$10,150	Loan Refinancing & Consolidation	82%	\$775 / 11 days
<input type="checkbox"/> \$0	C 1 12.99%	60	690-694	\$10,625	Credit Card Payoff	67%	\$3,475 / 10 days



# 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](https://pandas.pydata.org/Pandas_Cheat_Sheet.pdf)

## **LogisticRegression**

[https://scikit-learn.org/stable/modules/generated/sklearn.linear\\_model.LogisticRegression.html](https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html)

## **Confusionmatrix**

[https://scikit-learn.org/stable/modules/generated/sklearn.metrics.confusion\\_matrix.html](https://scikit-learn.org/stable/modules/generated/sklearn.metrics.confusion_matrix.html)