



## SOCIEDAD ECUATORIANA DE ESTADÍSTICA

### FUNDAMENTOS DE MACHINE LEARNING

RESUMEN NO. 1: CONCEPTOS BÁSICOS DEL APRENDIZAJE AUTOMÁTICO

Andrés Merino • Febrero 2026

#### DEFINICIÓN 1: Aprendizaje Automático.

El Aprendizaje Automático (Machine Learning) es una subdisciplina de la Inteligencia Artificial que se centra en el desarrollo de algoritmos y modelos que permiten a las máquinas aprender a partir de datos sin ser explícitamente programadas para cada tarea específica.

### 1. TIPOS DE APRENDIZAJE

#### DEFINICIÓN 2: Aprendizaje Supervisado.

Sea  $X \subseteq \mathbb{R}^n$  y  $Y \subseteq \mathbb{R}$ , al conjunto:

$$\{(x, y) : x \in X, y \in Y\},$$

se lo denomina conjunto de datos de entrenamiento, a los elementos  $(x_i, y_i)$  se los denomina **instancias** o ejemplos de entrenamiento, donde  $x_i \in \mathbb{R}^n$  son las **características** de las instancias y  $y_i \in Y$  son las **etiquetas** asociadas (variable objetivo). El **aprendizaje supervisado** busca determinar una función

$$f: \mathbb{R}^n \rightarrow \mathbb{R}$$

que aproxime la etiqueta  $y$  de una instancia  $x \in \mathbb{R}^n$ , es decir,  $f(x) \approx y$ .

#### DEFINICIÓN 3: Aprendizaje No Supervisado.

Sea  $X \subseteq \mathbb{R}^n$ , al conjunto:

$$\{x : x \in X\},$$

se lo denomina conjunto de datos de entrenamiento, a los elementos  $x_i$  se los denomina **instancias** o ejemplos de entrenamiento, donde  $x_i \in \mathbb{R}^n$  son las **características** de las instancias. El **aprendizaje no supervisado** busca determinar una función

$$f: \mathbb{R}^n \rightarrow \{1, 2, \dots, k\}$$

que asigne a cada instancia  $x \in \mathbb{R}^n$  un grupo o clúster.

## 2. TIPOS DE TAREAS

### DEFINICIÓN 4: Agrupamiento.

El **agrupamiento** es una tarea de **aprendizaje no supervisado** donde se busca agrupar las instancias en función de sus características.

### DEFINICIÓN 5: Regresión.

La **regresión** es una tarea de **aprendizaje supervisado** donde el conjunto de etiquetas es un conjunto continuo.

### DEFINICIÓN 6: Clasificación.

La **clasificación** es una tarea de **aprendizaje supervisado** donde el conjunto de etiquetas es un conjunto finito y discreto.



En la clasificación, si el conjunto de etiquetas es  $Y$ , entonces se dice que es una clasificación **binaria** si  $|Y| = 2$  y **multiclas** si  $|Y| > 2$ .

En la clasificación, si el conjunto de etiquetas es  $Y$ , se puede tener tres tipos de modelos:

- $f: \mathbb{R}^n \rightarrow Y$ : modelo que asigna una etiqueta a cada instancia.
- $f: \mathbb{R}^n \rightarrow [0, 1]^{|Y|}$ : modelo que asigna una probabilidad a cada etiqueta para cada instancia.
- $f: \mathbb{R}^n \rightarrow \mathbb{R}^{|Y|}$ : modelo que asigna un puntaje a cada etiqueta para cada instancia.

### 2.1 Modelos principales

Métodos	Supervis.		No super.
	Clasificación	Regresión	
k-NN	×		
SVM	×	×	
Árboles de decisión	×	×	
Redes neuronales	×	×	
Agrupamiento jerárquico			×
k-means			×
DBSCAN			×

### 3. MÉTRICAS

#### DEFINICIÓN 7.

Dado un conjunto  $X$ , una métrica o distancia es una función

$$d: X \times X \rightarrow \mathbb{R}$$

que cumple las siguientes propiedades para todo  $x, y, z \in X$ :

- I.  $d(x, y) \geq 0$  (no negatividad)
- II.  $d(x, y) = 0$  si y solo si  $x = y$  (identidad)
- III.  $d(x, y) = d(y, x)$  (simetría)
- IV.  $d(x, z) \leq d(x, y) + d(y, z)$  (desigualdad triangular)

#### DEFINICIÓN 8: Distancia Euclíadiana.

En  $\mathbb{R}^n$ , la **distancia euclíadiana** entre dos puntos  $x = (x_1, x_2, \dots, x_n)$  y  $y = (y_1, y_2, \dots, y_n)$  se define como:

$$d_2(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}.$$

#### DEFINICIÓN 9: Distancia de Manhattan.

En  $\mathbb{R}^n$ , la **distancia de Manhattan** entre dos puntos  $x = (x_1, x_2, \dots, x_n)$  y  $y = (y_1, y_2, \dots, y_n)$  se define como:

$$d_1(x, y) = \sum_{i=1}^n |x_i - y_i|.$$

#### DEFINICIÓN 10: Distancia de Minkowski.

En  $\mathbb{R}^n$ , la **distancia de Minkowski** entre dos puntos  $x = (x_1, x_2, \dots, x_n)$  y  $y = (y_1, y_2, \dots, y_n)$  se define como:

$$d_p(x, y) = \left( \sum_{i=1}^n |x_i - y_i|^p \right)^{1/p},$$

donde  $p \geq 1$ .



La distancia de Minkowski generaliza tanto la distancia euclíadiana ( $p = 2$ ) como la distancia de Manhattan ( $p = 1$ ).



La distancia euclíadiana puede representarse por medio del producto de matrices como:

$$d_2(x, y) = \sqrt{(x - y)^T (x - y)}.$$

**DEFINICIÓN 11: Distancia de Mahalanobis.**

En  $\mathbb{R}^n$ , dado un conjunto de datos con media  $\mu$  y matriz de covarianza  $S$ , la **distancia de Mahalanobis** se define como:

$$d_M(x, y) = \sqrt{(x - y)^T S^{-1} (x - y)}.$$

**DEFINICIÓN 12: Hamming.**

En  $\mathbb{R}^n$ , la **distancia de Hamming** entre dos puntos  $x = (x_1, x_2, \dots, x_n)$  y  $y = (y_1, y_2, \dots, y_n)$  se define como:

$$d_H(x, y) = \sum_{i=1}^n \delta(x_i, y_i),$$

donde

$$\delta(x_i, y_i) = \begin{cases} 0 & \text{si } x_i = y_i, \\ 1 & \text{si } x_i \neq y_i. \end{cases}$$

Es decir, cuenta el número de posiciones en las que los elementos correspondientes son diferentes.

**DEFINICIÓN 13: Distancia Coseno.**

En  $\mathbb{R}_{\geq 0}^n$ , la **distancia coseno** entre dos puntos  $x = (x_1, x_2, \dots, x_n)$  y  $y = (y_1, y_2, \dots, y_n)$  se define como:

$$d_{\cos}(x, y) = 1 - \frac{x \cdot y}{\|x\| \|y\|}.$$



La distancia coseno no es una métrica en el sentido estricto, ya que no cumple con la identidad; por esto se la llama una **pseudo-métrica**.