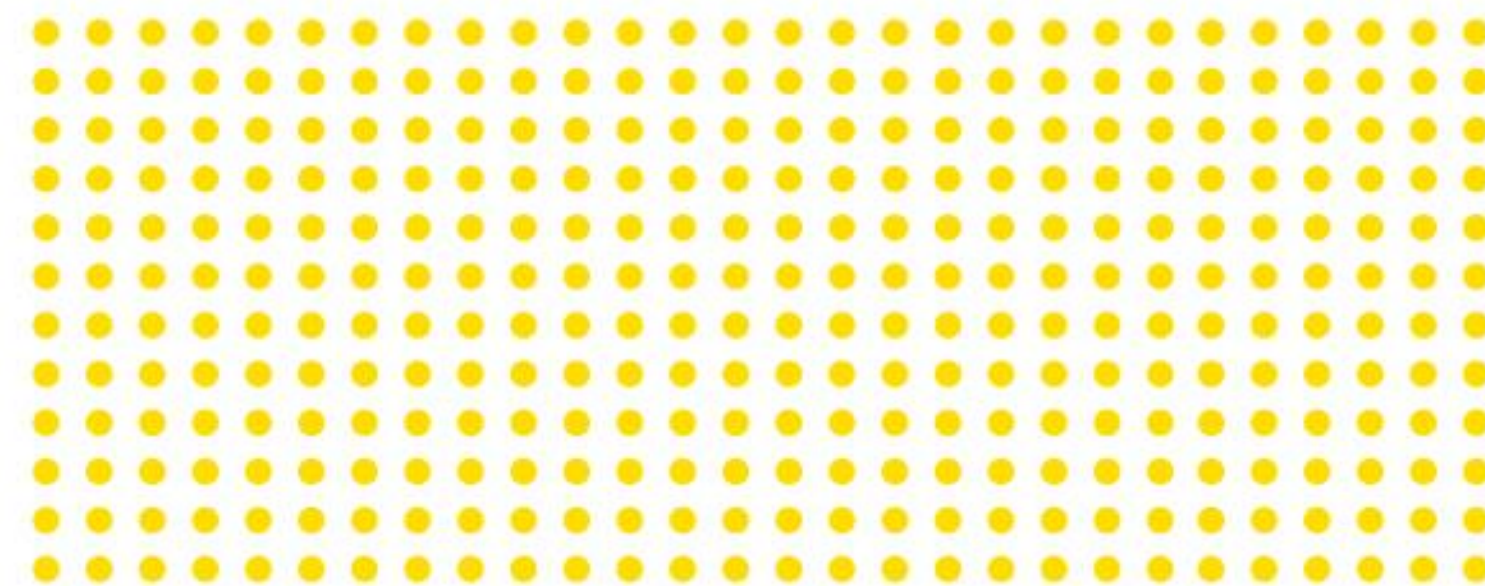


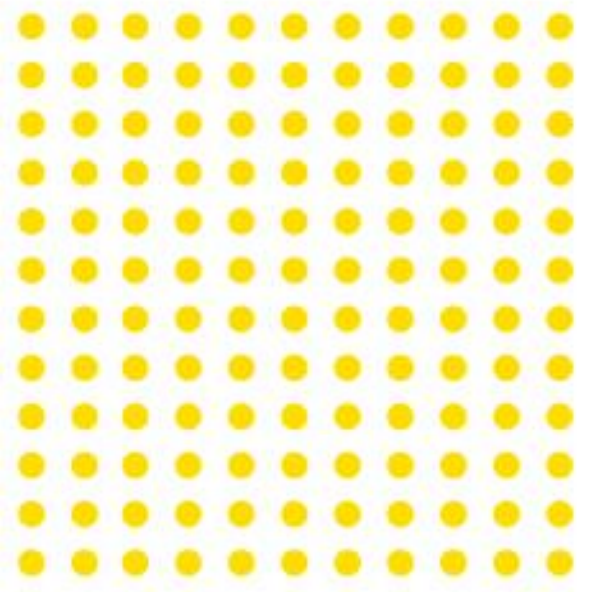


Universidad de
los Andes

Educación
Continua
Vicerrectoría Académica



Contenido

- 
- **Introducción**
 - **Aprendizaje Supervisado**
 - **Aprendizaje no supervisado**
 - **Interpretabilidad/Predicción**



Al final de la clase de hoy

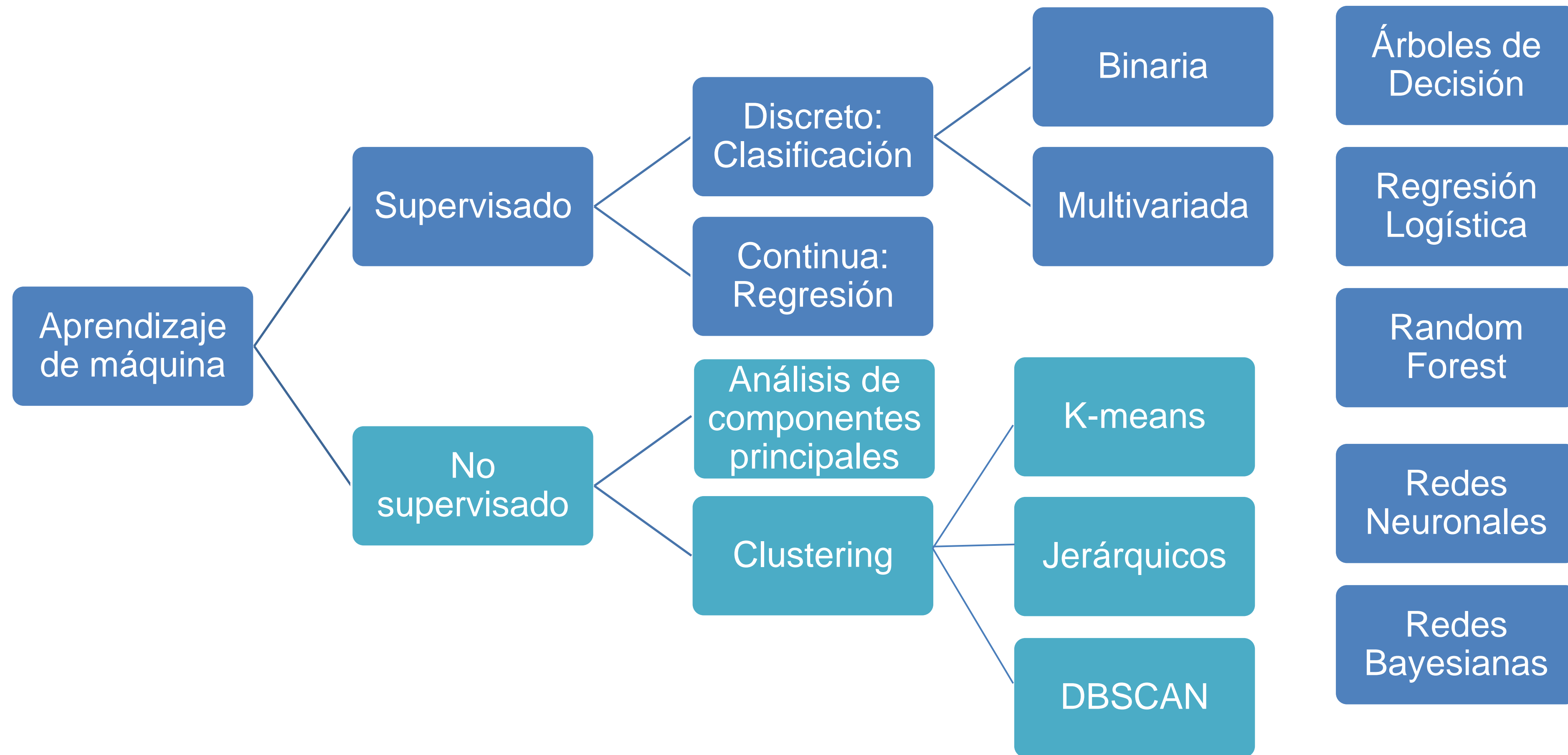


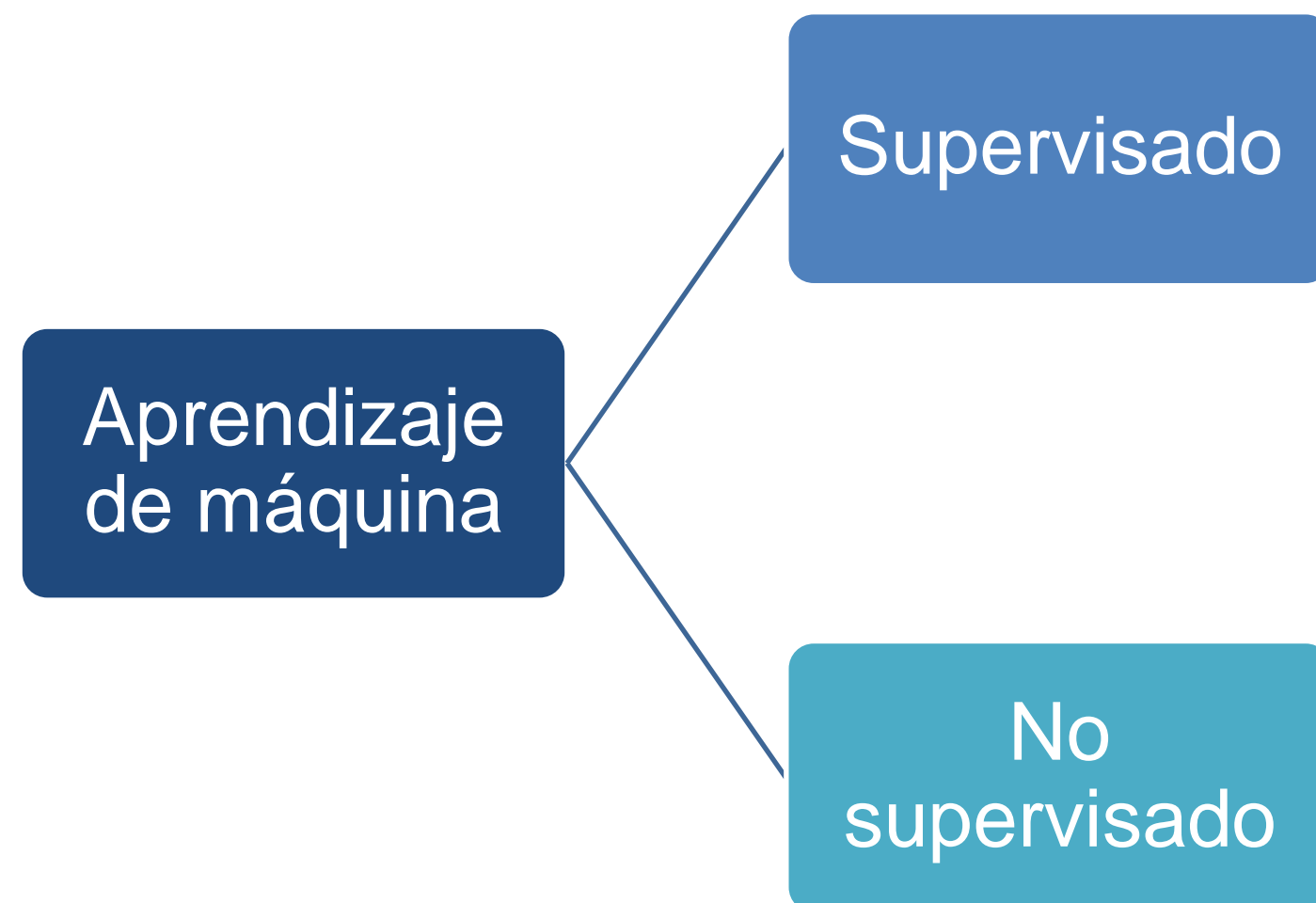
1. Conoceremos los conceptos base de aprendizaje de máquinas y sus procesos.
2. Entenderemos cuándo usar aprendizaje supervisado y no supervisado.
3. Predecir categorías binarias a partir de la probabilidad predicha de cada una.
4. Distinguir entre los datos de prueba y los datos de entrenamiento.

Recuerden que estamos entrando en el mundo de la predicción. Aquí vamos a poner menos atención a los parámetros y más al desempeño de los modelos.

Aprendizaje supervisado y no supervisado

INTRODUCCIÓN





- **Supervisado:** Conocemos una variable de respuesta. Ejemplo: Explicar el salario con base en características de la persona.
- **No supervisado:** No conocemos la categoría o propiedad de cada elemento. Ejemplo: Agrupar grupos de personas con base en sus características demográficas.

APRENDIZAJE SUPERVISADO

Aprendizaje
de máquina

Supervisado

Hablamos de aprendizaje supervisado cuando tenemos un conjunto de datos, con una etiqueta o valor de respuesta

Discreto:
Clasificación

Continua:
Regresión

Binaria

Multivariada

Cuando la respuesta es una **etiqueta**, hablamos de **clasificación**. Cuando la variable respuesta es un **continua**, hablamos de **regresión**.

Ej: Clasificar el sentimiento de un texto (positivo, negativo)

Ej: Explicar el ingreso de una persona en función de género, edad carrera.



Cuando hablamos de clasificación, usualmente buscaremos minimizar una estadística de error de clasificación. Usualmente se habla de **exactitud (*accuracy*)** de la clasificación y seleccionaremos el modelo con la mayor exactitud (el que minimiza el error).

Dependiendo del caso de uso que estemos analizando, le daremos más importancia a otras estadísticas como la **precisión** o la **exhaustividad (*recall*)**.

¿Cuáles son los modelos usuales en clasificación?

Regresión
Logística

Árboles de
Decisión

Random
Forest

Redes
Bayesianas

Redes
Neuronales

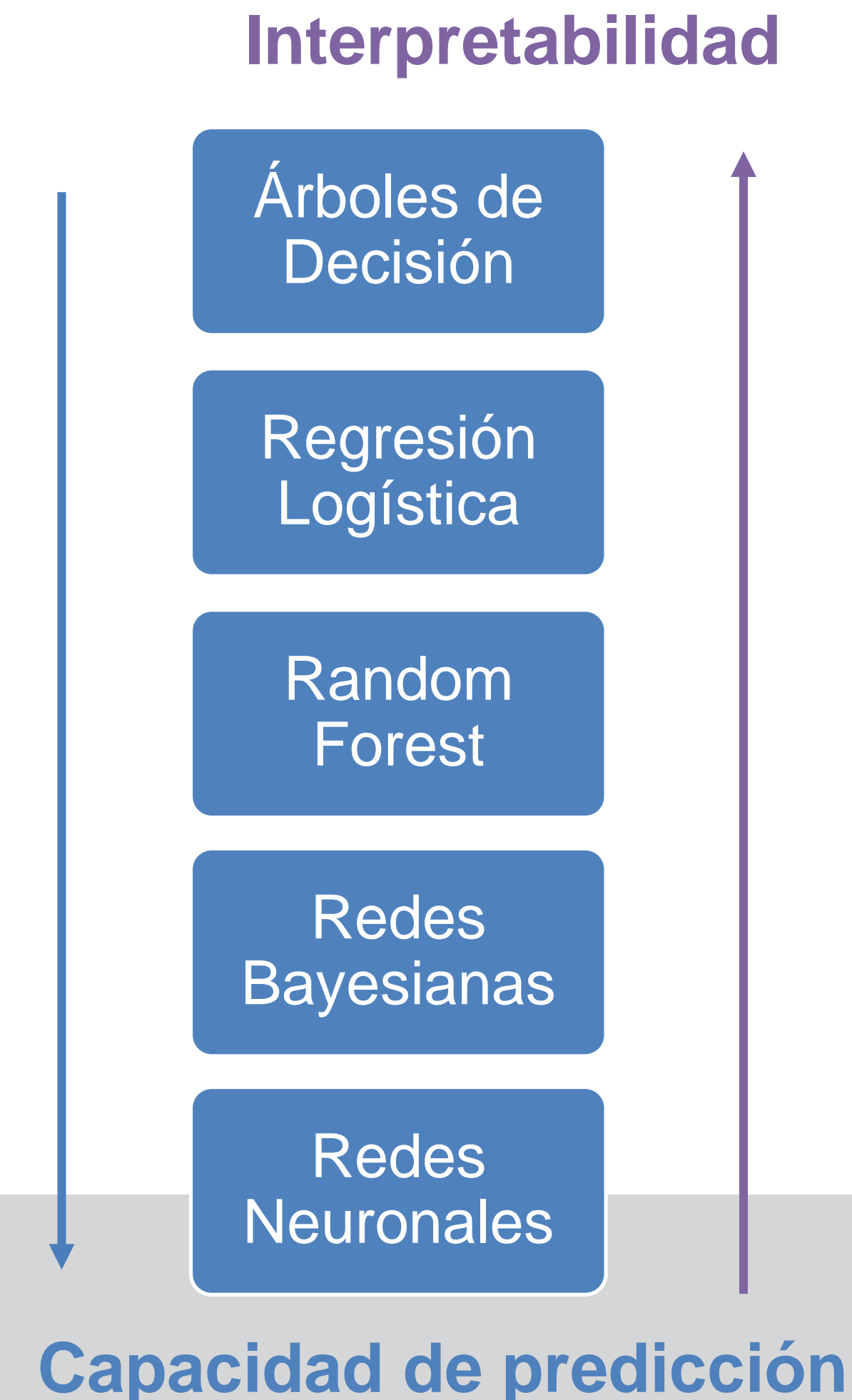
¿Cómo escoger un modelo?

Interpretabilidad: El modelo me permite entender cómo se relacionan las variables explicativas, con la variable de respuesta.

Ejemplo: Entender los factores que afectan el riesgo de cartera.

Capacidad de predicción: El modelo me permite predecir con alta exactitud, la variable de respuesta, al conocer las variables explicativas.

Ejemplo: Poner etiquetas en una imagen.



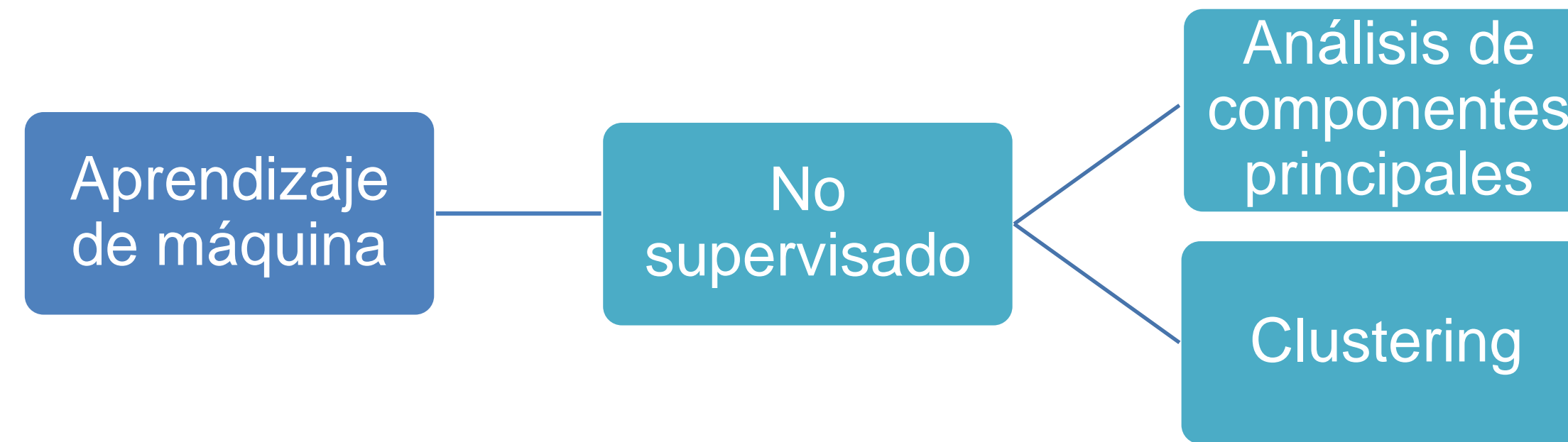
Para discutir:

- Si usted debe implementar un modelo judicial para determinar la pena recomendada para una persona, con base a las características de la sentencia y la persona (delito, edad, género), y cuenta con una base histórica.

¿Qué tipo de modelo recomendaría? ¿Por qué? ¿Qué tendría en cuenta sobre predictibilidad e interpretabilidad?

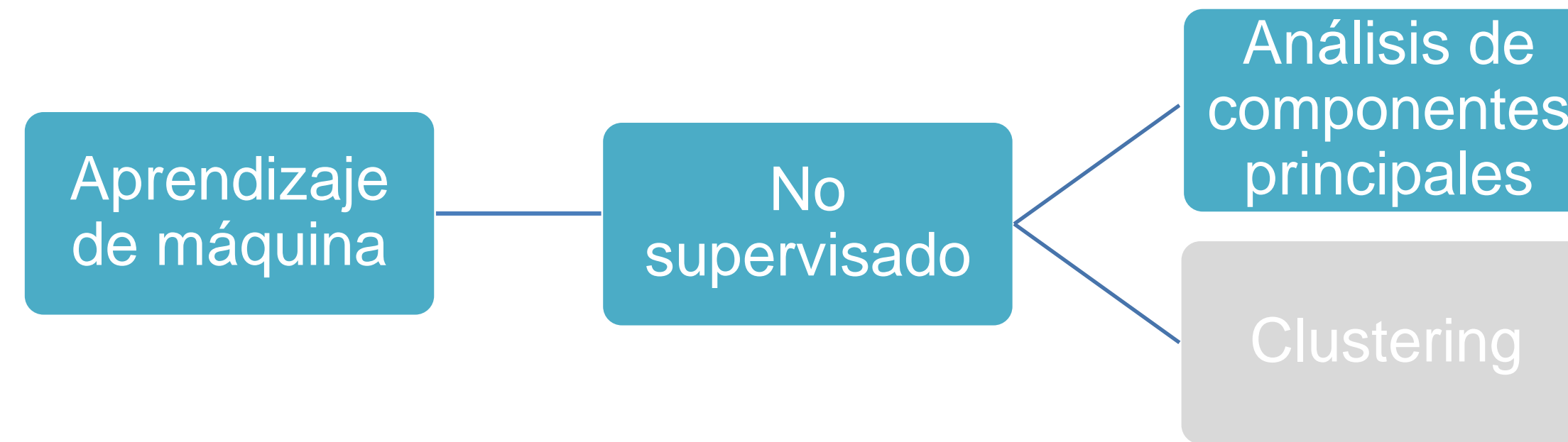
Tres minutos...

APRENDIZAJE NO SUPERVISADO



Hablamos de **aprendizaje no supervisado** cuando **no tenemos un valor o etiqueta**, que queramos explicar a través de otras variables.

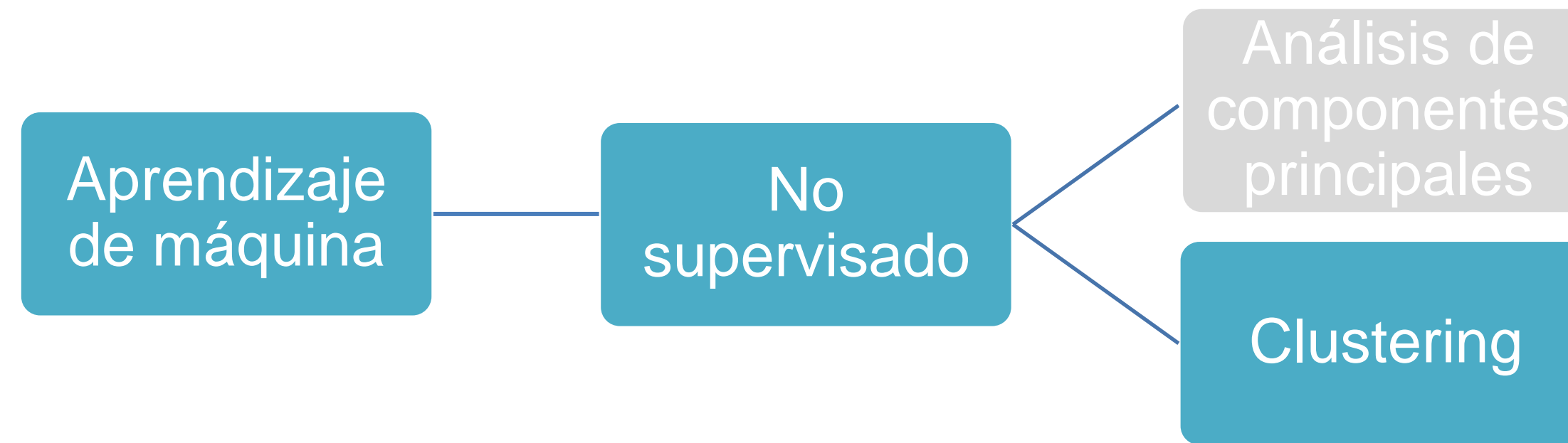
Es usado en escenarios como segmentación de clientes.



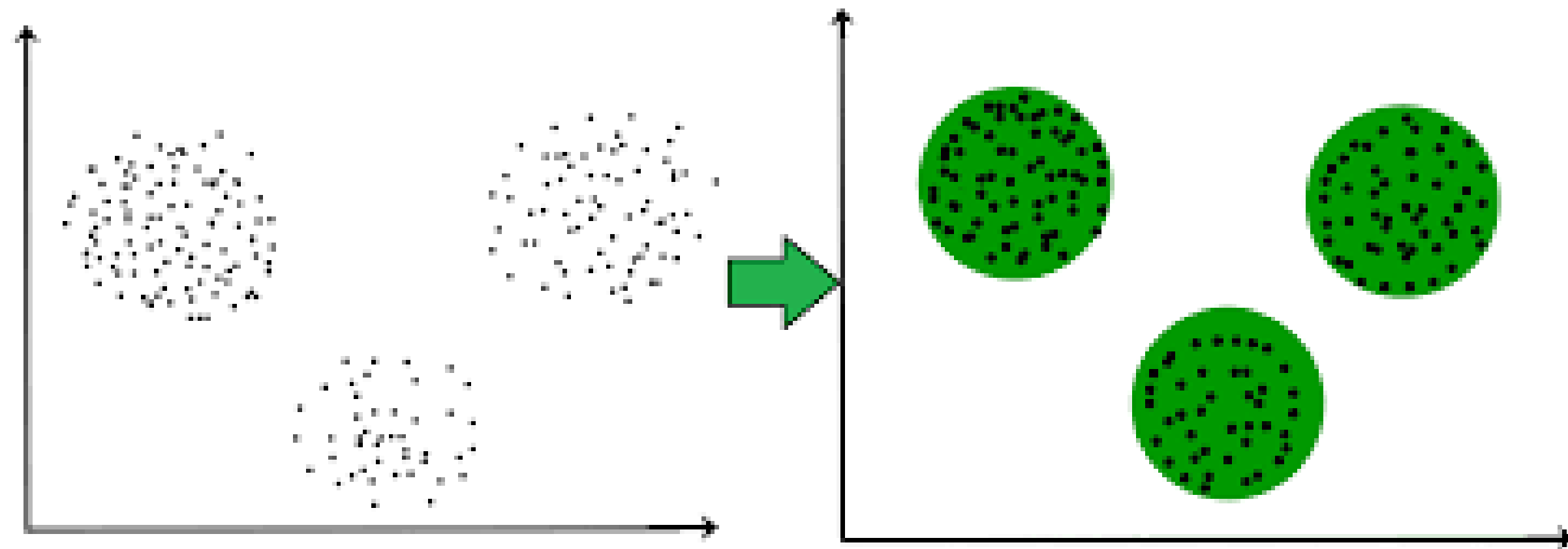
El **análisis de componentes principales (PCA)** nos permite entender o representar gráficamente la relación entre diferentes variables cuantitativas.

Es usado frecuentemente para hacer **reducción dimensional**, es decir, reducir la cantidad de variables que necesitamos para trabajar con otro modelo posterior como clasificación o Clustering.

NO LO INCLUIREMOS EN
ESTE CURSO



Los modelos de **clustering o agrupación** buscan encontrar grupos homogéneos (parecidos) de individuos a partir de las variables que los definen.



Pero aplicado a variables dicótomas

MODELO LINEAL

El modelo lineal puede usarse para predecir

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + e$$

Enchufamos x

...

Salen una y

Ejercicio

Digamos que nuestro modelo lineal nos dio los siguientes valores entrenados:

$$\text{TieneHijos} = 0.3 + 0.2 x + e$$

Y queremos predecir y para el siguiente dato:

Nombre	x
Don Marcos	2.1

¿Se predice que Don Marcos tiene hijos o no? ¿Cuál es la predicción?

Ejercicio

Tenemos:

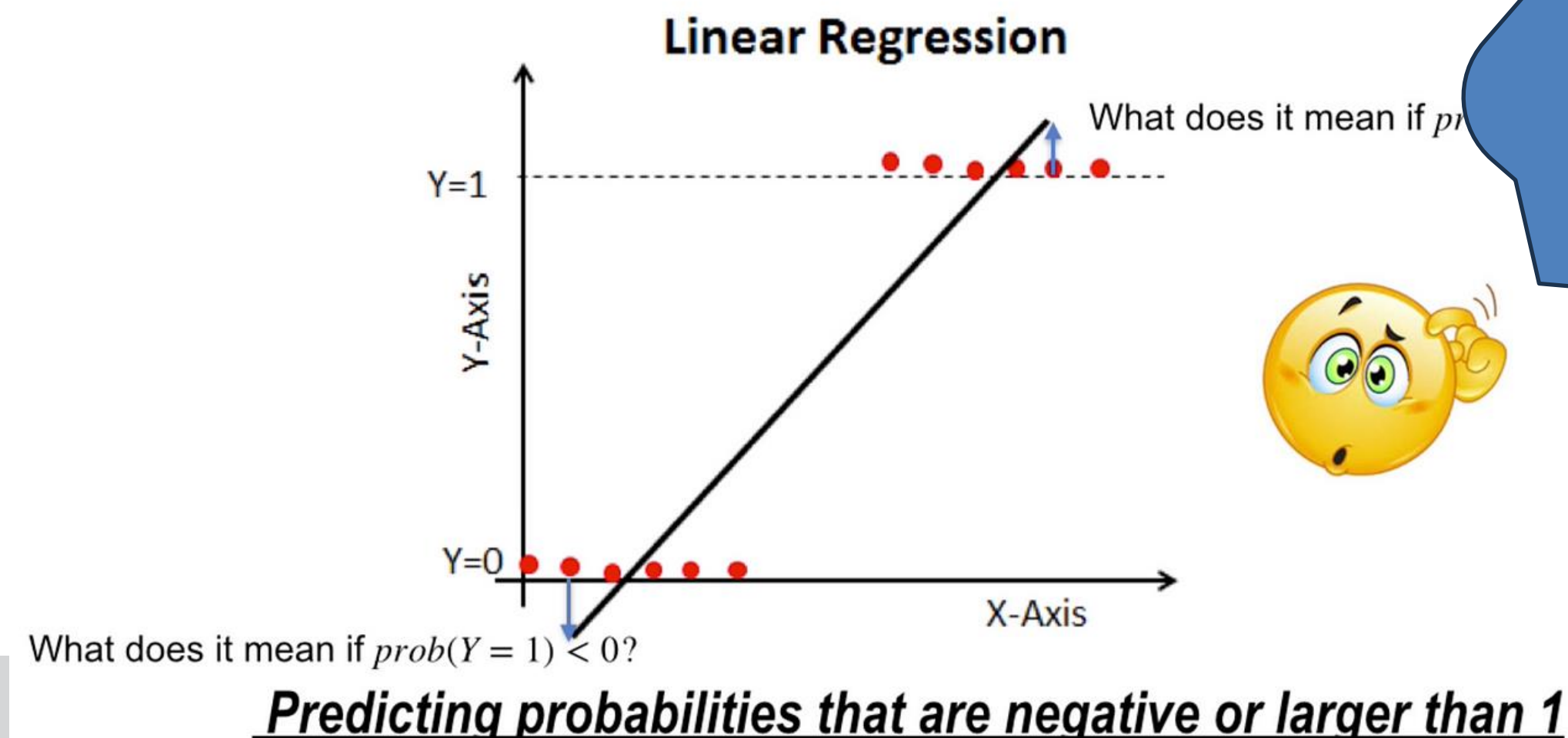
$$\text{Predicción} = 0.72 = 0.3 + 0.2(2.1)$$

Esto se puede interpretar como la probabilidad de don Marcos de tener hijos, dadas sus característica x .

Nombre	x
Don Marcos	2.1

Si tenemos una variable y con valores de 1 y 0 (dicotómica), se puede demostrar que la predicción de y cuando le enchufamos las x se puede interpretar como una probabilidad.

Why not Linear Probability Model?



Con un
pequeño...
problema...

Nos vamos a encargar de
esto en las siguientes
clases

En aprendizaje de máquinas

CÓMO CLASIFICAMOS

Clasificación binaria

- Nuestros modelos (lineal, con su problema; y logit), nos dan una predicción de la **probabilidad de que la variable explicada sea 1**.
- Para predecir nuestras categorías simplemente determinamos un umbral, por ejemplo:

$$y = \begin{cases} 1 & \text{si } P(Y|X) > 0.5 \\ 0 & \text{de lo contrario} \end{cases}$$

¡Gracias!

Aprendiendo juntos a lo largo de la vida

educacioncontinua.uniandes.edu.co

Síguenos: **EdcoUniandes**     



**Educación
Continua**
Vicerrectoría Académica

Universidad de los Andes | Vigilada Mineducación. Reconocimiento como Universidad: Decreto 1297 del 30 de mayo de 1964. Reconocimiento personería jurídica: Resolución 28 del 23 de febrero de 1949 Minjusticia.

