



FACULTAD
DE CIENCIAS
ECONÓMICAS

FAMAF

Facultad de Matemática, Astronomía,
Física y Computación



UNC

Universidad
Nacional
de Córdoba

DIPLOMATURA

**CIENCIA DE DATOS, INTELIGENCIA
ARTIFICIAL Y SUS APLICACIONES
EN ECONOMÍA Y NEGOCIOS**

A solid blue triangle is located in the bottom-left corner of the slide, pointing towards the center.

**A modo de
síntesis**



Clasificación de las técnicas multivariadas

Objetivo	Enfoque	
	Descriptivo /Exploratorio	Inferencial / Explicativo y predictivo
Reducción de dimensionalidad	<ul style="list-style-type: none">• Análisis de componentes principales• Análisis de correspondencias	
Clasificación supervisada		<ul style="list-style-type: none">• Regresión logística• Árboles de clasificación
Construcción de grupos (clasificación no supervisada)	<ul style="list-style-type: none">• Análisis de conglomerados	

Fuente: elaboración propia en base a Peña (2002)



Reducción de dimensionalidad

Técnica	Características	Objetivo	Variables
Análisis de Componentes Principales	Dadas n observaciones de p variables, analiza si es posible representar adecuadamente es información con un número menor de variables (componentes) que son construidas como combinaciones lineales de las originales.	<ul style="list-style-type: none">Reducción de la dimensionalidadEvitar la correlación entre las variables originales	Métricas



Reducción de dimensionalidad

Técnica	Características	Objetivo	Variables
Análisis de Correspondencias	Técnica descriptiva para representar la información contenida en una tabla de contingencia en un espacio de dimensión menor.	<ul style="list-style-type: none">■ Reducción de la dimensionalidad■ Crear un mapa de la posición relativa de las variables cualitativas estudiadas con cada uno de sus valores posibles.■ Identificar las asociaciones de mayor peso entre las modalidades de variables cualitativas	No métricas



Clasificación supervisada

Técnica	Características	Objetivo	Variables
Regresión logística	Construcción de un modelo que explique los valores de la variable de clasificación.	<ul style="list-style-type: none">▪ Determinar que variables caracterizan a cada categoría de la variable dependiente▪ Estimar la probabilidad de que se produzca el suceso definido por la variable dependiente, con valores de las variables conocidas (regresoras)	Variable de respuesta: no métrica, dicotómica Variables regresoras: métricas y no métricas



Clasificación supervisada

Técnica	Características	Objetivo	Variables
Árboles de clasificación	Dar con un esquema de múltiples bifurcaciones, anidadas en forma de árbol, de manera que siguiendo cada una de esas ramas obtengamos finalmente una predicción para la clase de pertenencia.	Clasificar un nuevo elemento, con valores de las variables conocidas, en uno de los grupos.	Variable criterio: no métrica, con k categorías Variables predictoras: métricas y no métricas



Clasificación no supervisada

Técnica	Características	Objetivo	Variables
Clustering	Dado un conjunto de individuos (de n elementos) caracterizados por la información de p variables se busca clasificarlos en grupos, en función de medidas de distancia o similitud entre observaciones. Los grupos se crean en función de algoritmos jerárquicos (aglomerativos: centroide, vecino mas cercano o lejano, vinculación promedio y Ward / desagregativos) o no jerárquicos (K-means)	Formar grupos de manera que los individuos pertenecientes a un grupo (cluster) sean tan similares entre sí como sea posible, siendo los distintos grupos entre ellos tan disimilares como sea posible.	Variables: métricas y no métricas