00

[Tema 2 – Machine learning y técnicas cuantitativas. 2](#_Toc129071032)

[1. Introducción y nomenclatura 2](#_Toc129071033)

[2. Generalización del problema de la regresión. 4](#_Toc129071034)

[3. Predicción: error reducible e irreducible. 4](#_Toc129071035)

[4. Inferencia 4](#_Toc129071036)

[5. Enfoques paramétricos y no paramétricos. 4](#_Toc129071037)

[6. Paramétrico: modelo para ajustar la función a los datos de entrenamiento. 4](#_Toc129071038)

[7. Flexibilidad e interpretabilidad. 4](#_Toc129071039)

[8. Evaluación de la precisión del modelo: MSE para diferentes ajustes. 4](#_Toc129071040)

[9. The bias – Variance Trade – off. 4](#_Toc129071041)

# Tema 2 – Machine learning y técnicas cuantitativas.

## Introducción y nomenclatura

Tenemos un conjunto de datos que consisten en;

* Precio de algunas casas en una zona.



* Pies cuadrados (metros cuadrados) de cada uno de ellos.

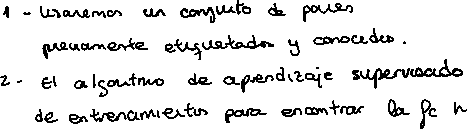
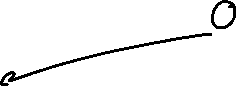
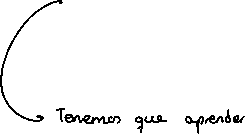


Gráfico, Gráfico de dispersión

Descripción generada automáticamente



El aprendizaje supervisado es un enfoque de aprendizaje automático en el que el objetivo es aprender a predecir una salida o etiqueta (Y) a partir de una entrada (X), dadas una serie de muestras de entrenamiento. El objetivo final es encontrar una función matemática (hipótesis) que pueda predecir con precisión la salida para nuevas entradas desconocidas.



Cuando la variable objetivo que estamos tratando de predecir es continua 🡪 problema de regresión.

Cuando puede asumir solo un pequeño número de valores discretos 🡪 problema de clasificación.

**El aprendizaje supervisado es un conjunto de algoritmos de aprendizaje automático que se   
pueden entrenar conociendo pares de entrada y salida.**

## Generalización del problema de la regresión.

## Predicción: error reducible e irreducible.

## Inferencia

## Enfoques paramétricos y no paramétricos.

## Paramétrico: modelo para ajustar la función a los datos de entrenamiento.

Es el método que usamos para ajustar un modelo lineal con una sola variable a unos datos de entrenamiento.

Tabla

Descripción generada automáticamente

## Flexibilidad e interpretabilidad.

## Evaluación de la precisión del modelo: MSE para diferentes ajustes.

## The bias – Variance Trade – off.

Big Data Stack

¿Qué son los sistemas “Stack”? 🡪 Grupo de herramientas e infraestructuras integradas. Se puede configurar dependiendo de la necesidad final del proyecto.

El entorno debe estar formado por hardware, software de infraestructura, operativo, de administración, interfaces de programación de apps (API) e incluso herramientas de desarrollo de software.

Esta arquitectura tiene que ser capaz de abordar los **6 requisitos fundamentales:**

* Ingerir
* Integrar
* Organizar y almacenar
* Análisis
* Actuar o entregar o visualizar
* Asegurar el sistema