



# Elementos de Machine Learning

Diplomado de Análisis Computacional Estadístico de Datos con Python

Impartido por: Ronaldo Canizales

# Acerca de mí

- Doctorando en Ciencias de la Computación, CSU.
- Docente del Depto. de Electrónica e Informática, UCA.
- Asistente de Investigación, Computer Science Dept, CSU.
- Becario o patrocinado por:



National  
Science  
Foundation



Association for  
Computing Machinery



IEEE COMPUTER SOCIETY  
**TCPP**

Technical Committee on Parallel Processing



UNIVERSIDAD  
DE GRANADA



**BID**  
Banco Interamericano  
de Desarrollo

- Ponente en congresos:



San Francisco, CA USA  
May 27-31, 2024

**ParSocial**  
2024



- Áreas de interés:
  - Explainable IA.
  - Representation Learning.
  - Formal Methods.
  - GraphML & HPC.
  - CS Education.

# Agenda del Módulo

01

Intro. teórica  
y práctica

02

Aprendizaje  
supervisado

03

Deep Learning

04

Otras  
aplicaciones

05

¡Al infinito y  
más allá!

Link

Enlace: [github.com/  
armandocodigos/  
Intro-ML-2025-UCA/](https://github.com/armandocodigos/Intro-ML-2025-UCA/)



# Bloque A

Desmitificando el  
Machine Learning:  
conceptos relacionados.

# ¿Qué es el Machine Learning?

- “Machine learning is the science (and art) of programming computers so they can learn from data.” [Aurelien Geron, 2023](#).
- “[Machine learning is the] field of study that gives computers the ability to learn without being explicitly programmed.” [Arthur Samuel, 1959](#).
- “A computer program is said to learn from experience  $E$  with respect to some task  $T$  and some performance measure  $P$ , if its performance on  $T$ , as measured by  $P$ , improves with experience  $E$ .”  
[Tom Mitchell, 1997](#).

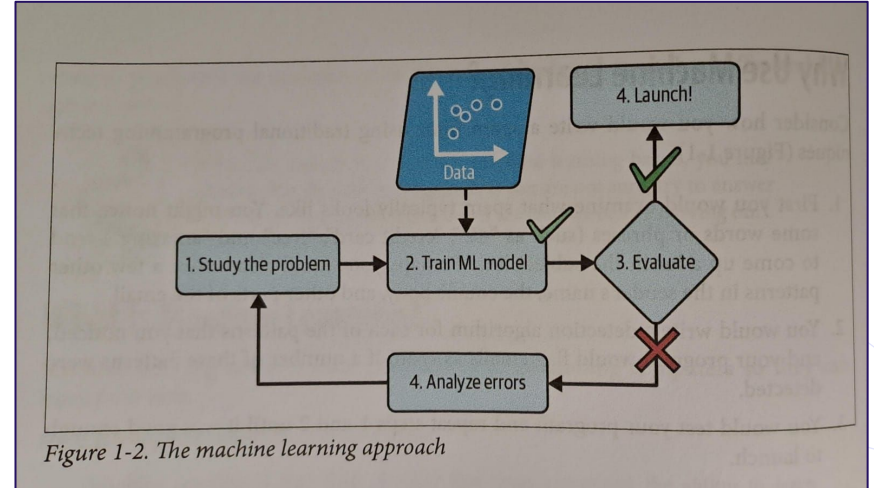


Figure 1-2. The machine learning approach

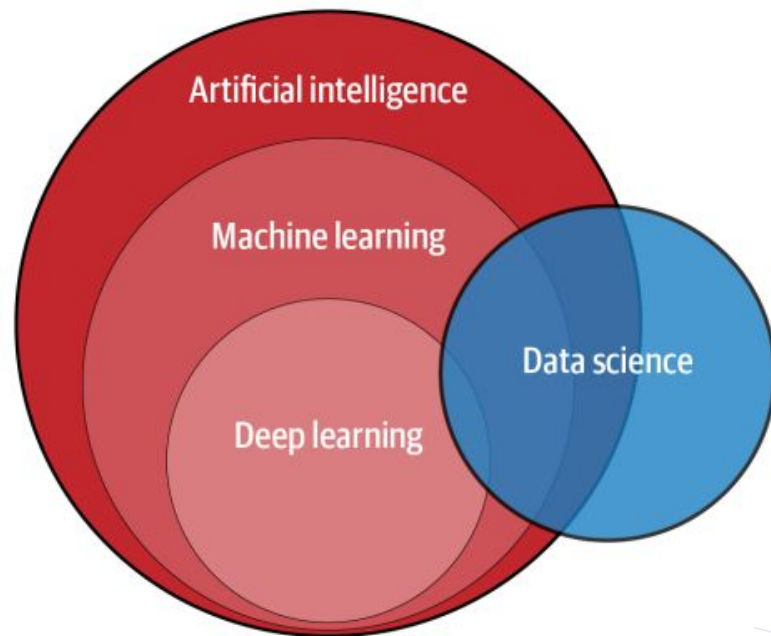
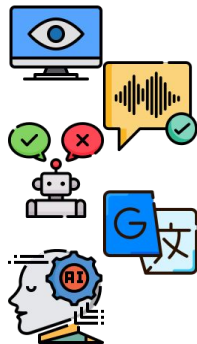


# DS vs AI vs ML vs DL

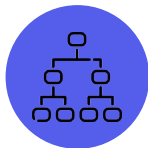


La **inteligencia artificial** es la ciencia que estudia la forma de hacer que **una computadora desarrolle la capacidad de realizar con éxito tareas complejas** que generalmente requieren inteligencia humana.

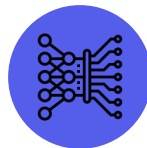
- Percepción visual
- Reconocimiento de voz
- Toma de decisiones
- Traducción entre idiomas
- ¡Muchas más!



# DS vs AI vs ML vs DL



El **aprendizaje de máquina (ML)** es una aplicación de la IA. Es un conjunto de algoritmos que poseen la capacidad de **aprender automáticamente del entorno**.



El **aprendizaje profundo (DL)** es un subconjunto del ML que implica el estudio de algoritmos relacionados con **redes neuronales artificiales**.  
Permite resolver problemas más complejos.

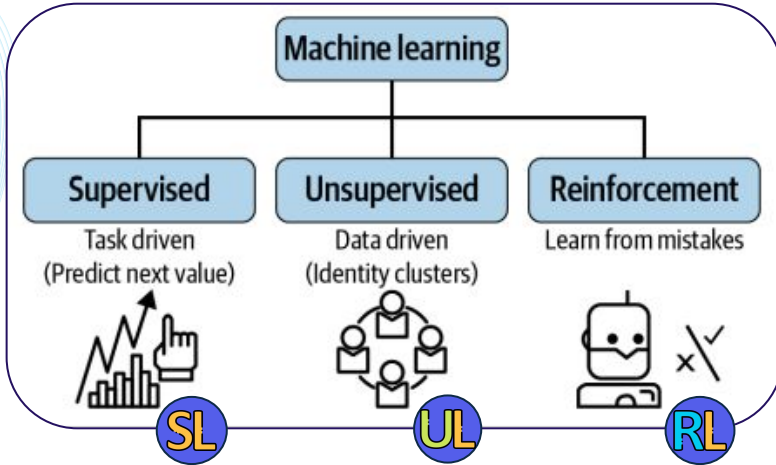


La **ciencia de datos (DS)** es un campo **interdisciplinario** similar a la minería de datos.  
Se encarga de **extraer información de los datos** en diversas formas, ya sea estructuradas o no.



# Tipos de Aprendizaje de Máquina

SL



El objetivo principal del **aprendizaje supervisado** es entrenar un modelo a partir de **datos etiquetados** que nos permita hacer predicciones sobre datos futuros o no vistos.

Aquí, el término **supervisado** se refiere a un conjunto de muestras donde **ya se conocen** las señales de salida deseadas (etiquetas).

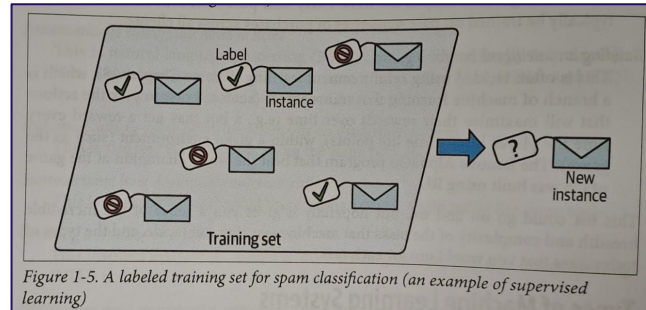
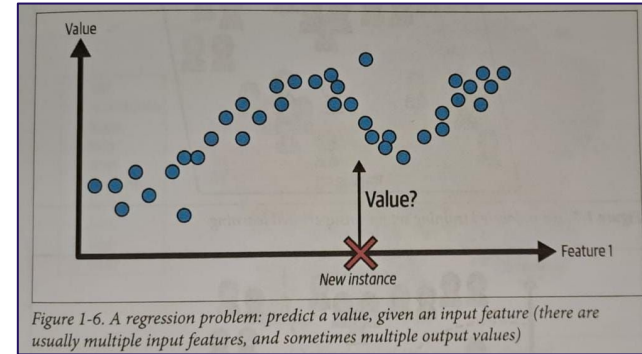


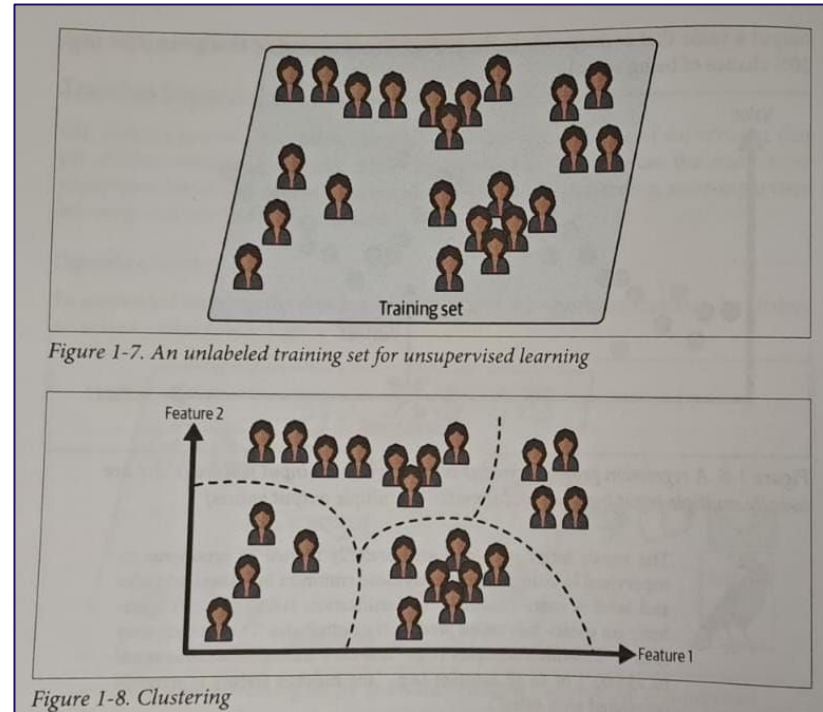
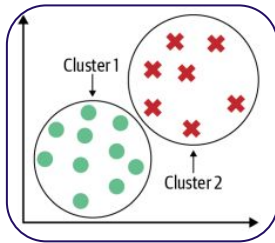
Figure 1-5. A labeled training set for spam classification (an example of supervised learning)





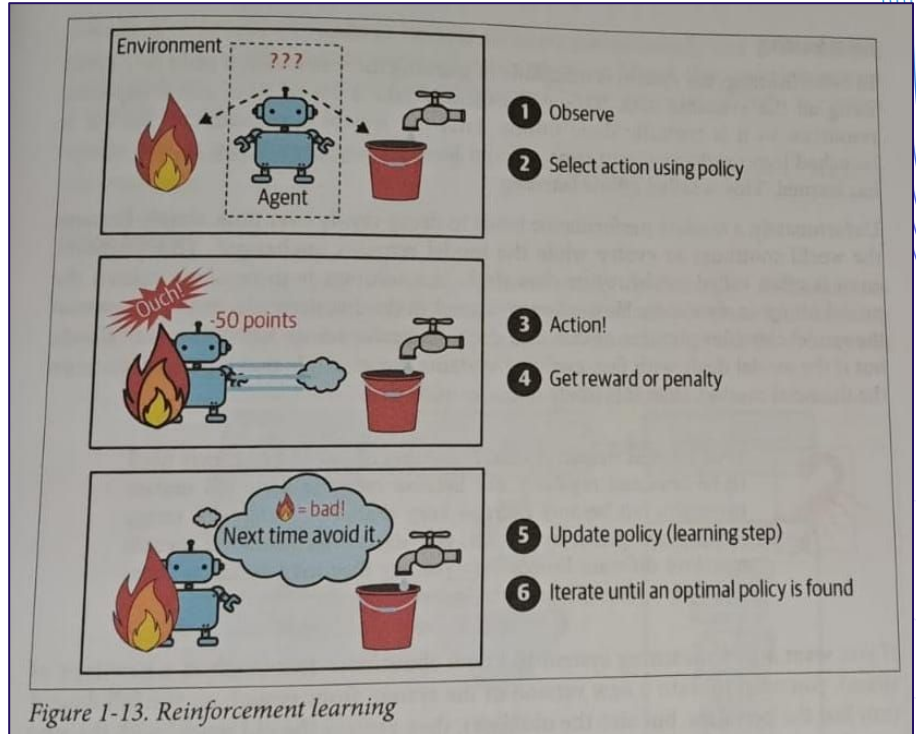
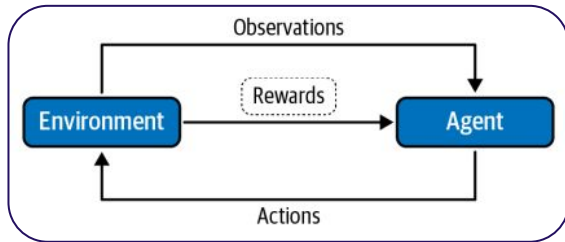
# Tipos de Aprendizaje de Máquina

El **aprendizaje no supervisado** es un tipo de aprendizaje automático que se utiliza para **extraer inferencias** de conjuntos de datos que consisten en datos de entrada sin respuestas etiquetadas.

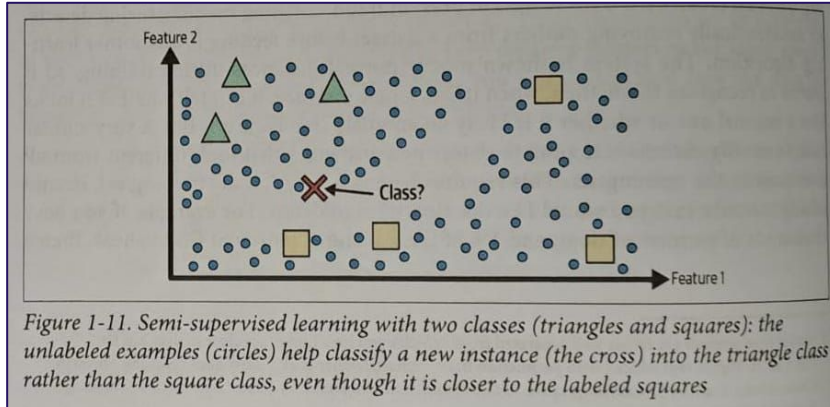


# Tipos de Aprendizaje de Máquina

Aprender de la experiencia, las recompensas o castigos es el concepto detrás del **aprendizaje por refuerzo**. Se trata de tomar las acciones adecuadas para maximizar la recompensa en situaciones particulares donde **no hay una respuesta explícita**.

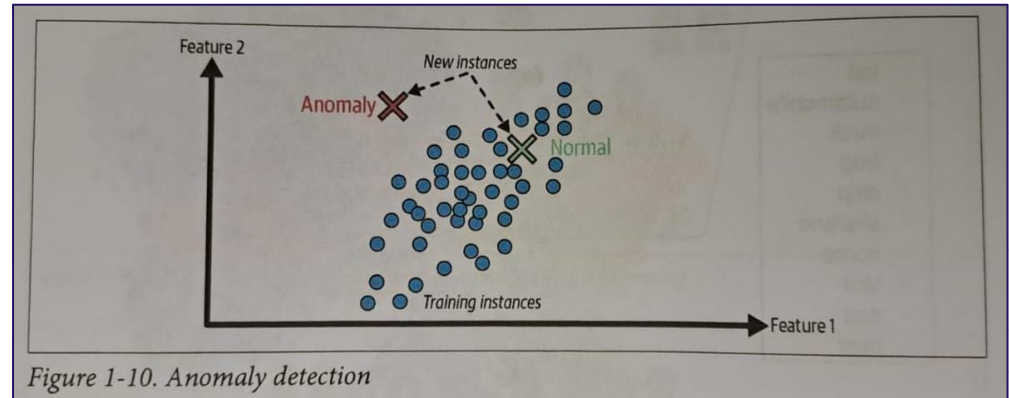



# Tipos de Aprendizaje de Máquina



Aprendizaje semi-supervisado

Detección de anomalías





# Resolución de problemas con IA

Mi experiencia



## Automatización

Ambientes tectónicos, fuentes sísmicas (redes neuronales).



## Sismología

Mapa de respuesta sísmica homogénea de El Salvador (métodos de ensamble).



## Turismo

Uso de reconocimiento de imágenes en aplicación móvil de RA para volcanes.



## Redes sociales

Percepción cultura digital, Análisis del discurso en pandemia (sentimientos, modelado de topics).



## Noticias y DDHH

Observatorio Universitario de DDHH (web scraping y LLMs).



## Recomendación

Exploración de mercado laboral para recién graduados UCA, uso de grafos bipartitos cliente-restaurantes.



## Educación

Evaluaciones cualitativas (árboles de decisión).



## Síntesis Progs.

Estimación de recursos en ambientes heterogéneos (ML con grafos).



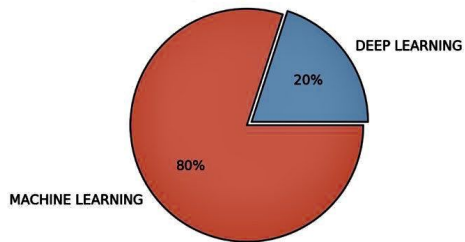
## Realidad Virtual

Estimación de estrés cognitivo y emocional en tratamiento ante fobias (redes neuronales).

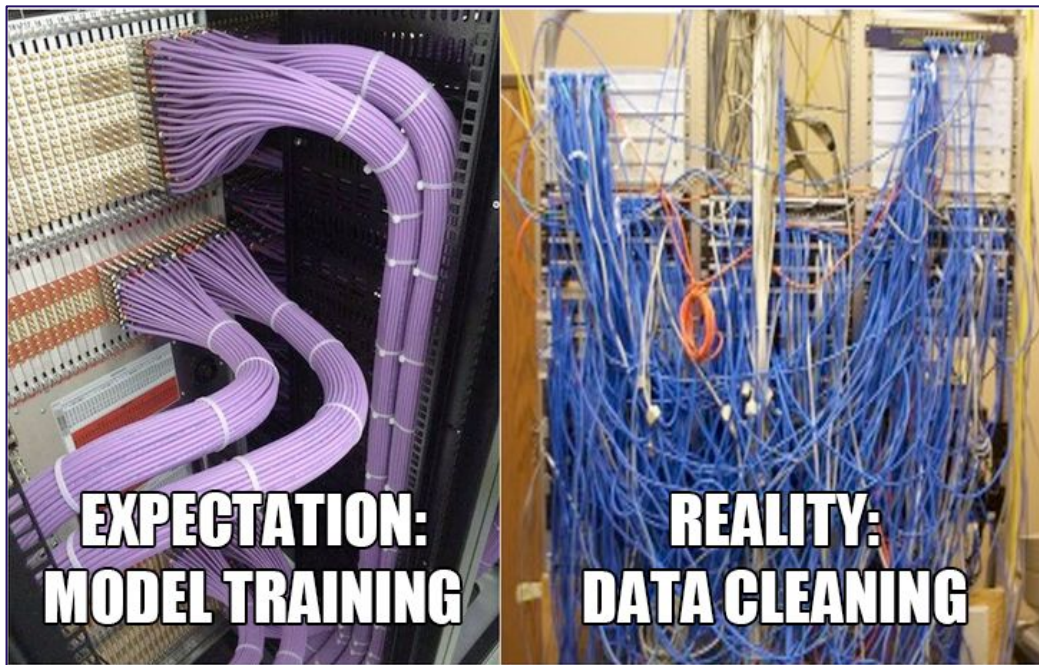
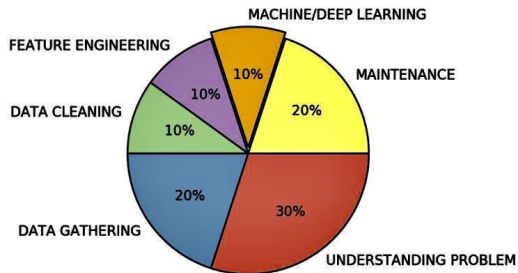


# Etapas de un proyecto de DS y ML

## DATA SCIENTIST JOB - EXPECTATION

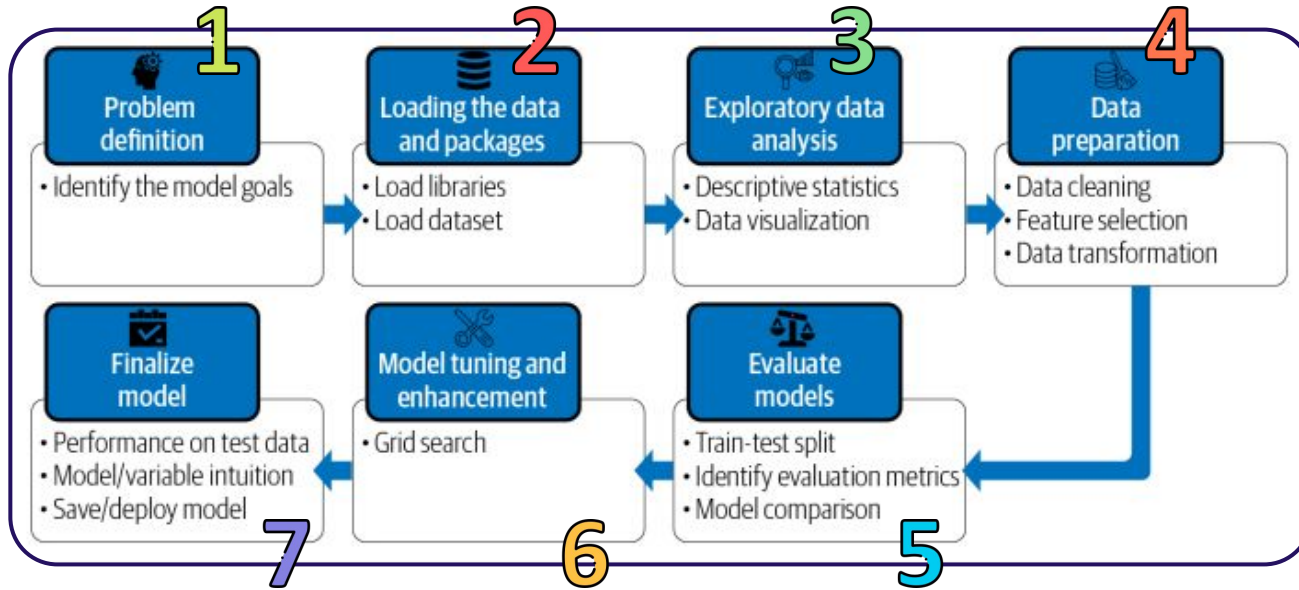


## DATA SCIENTIST JOB - REALITY



# Etapas de un proyecto de DS y ML

Ciencia de datos



Aprendizaje de Máquina



## Bloque B

“Toy example” y  
Ejemplo sencillo “end-to-end.”

# Librerías a utilizar



**NumPy**

`pip install numpy`



**Pandas**

`pip install pandas`



**Matplotlib**

`pip install matplotlib`



**Python 3.7+**

Anaconda  
JupyterLab  
Jupyter Notebook  
Google Colab



**Scikit-Learn**

`pip install -U scikit-learn`



**ML Repository**

`pip install ucimlrepo`



# Sencillo ejemplo “end-to-end”



## Infrared Thermography Temperature

External

Linked on 11/20/2023

The Infrared Thermography Temperature Dataset contains temperatures read from various locations of inferred images about patients, with the addition of oral temperatures measured for each individual. The 33 features consist of gender, age, ethnicity, ambient temperature, humidity, distance, and other temperature readings from the thermal images. The dataset is intended to be used in a regression task to predict the oral temperature using the environment information as well as the thermal image readings.



### Dataset Characteristics

Tabular

### Subject Area

Health and Medicine

### Associated Tasks

Regression

### Feature Type

Real, Categorical

### # Instances

1020

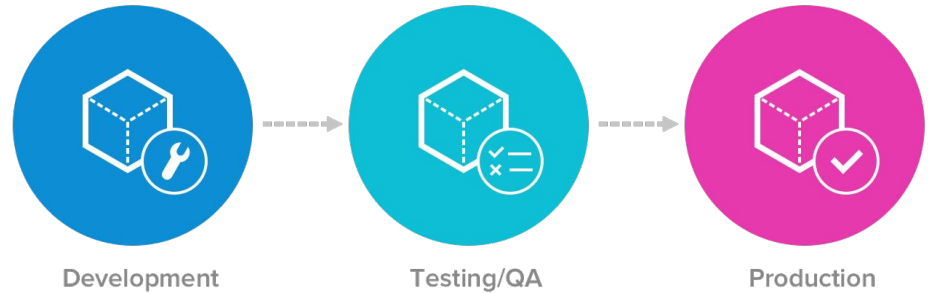
### # Features

33

[Enlace](#)



# Training y Test subsets



## Etapas de una solución



# Actividad recomendada

Explorar manualmente los datos de entrada para obtener un modelo más ligero con menor error.



# ¡Gracias por su atención!

**Datos de contacto:**

[rcanizales@uca.edu.sv](mailto:rcanizales@uca.edu.sv)

[rcanizal@colostate.edu](mailto:rcanizal@colostate.edu)

<https://www.cs.colostate.edu/~rcanizal/>

[www.linkedin.com/in/ronaldo-canizales/](https://www.linkedin.com/in/ronaldo-canizales/)

<https://x.com/ArmandoCodigos>

**CREDITS:** This presentation template was created by [Slidesgo](#), and includes icons by [Flaticon](#), and infographics & images by [Freepik](#)