



# Introducción al Aprendizaje Automático (Inteligencia Artificial)

Chaparro Amaro Oscar Roberto  
Martínez Felipe Miguel de Jesús  
Martínez Castro Jesús Alberto

23 de noviembre de 2022

Instituto Politécnico Nacional  
Centro de Investigación en Computación

# Agenda

Introducción

Aplicaciones

Referencias



## Consideraciones

Es importante mencionar que algunas **tareas** no se pueden definir correctamente.

Los humanos diseñan las máquinas, debido a los ambientes en que se desenvuelven.

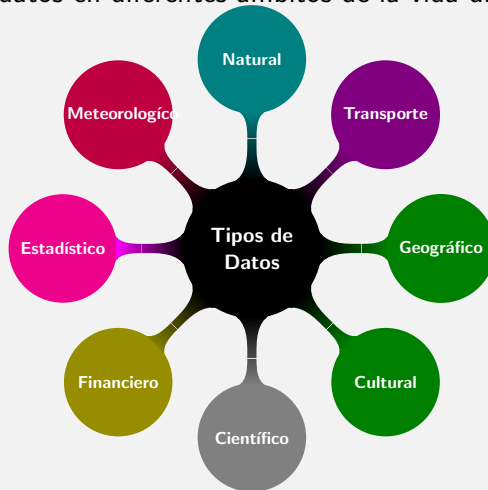
La **cantidad de conocimiento disponible** puede llegar a ser bastante grande.

Se requiere adecuar el **nuevo conocimiento** que el ser humano va descubriendo.

- **Nuevos eventos** en el mundo.
- **Rediseño** de los sistemas de los métodos que trabajan con este nuevo conocimiento.

# Datos

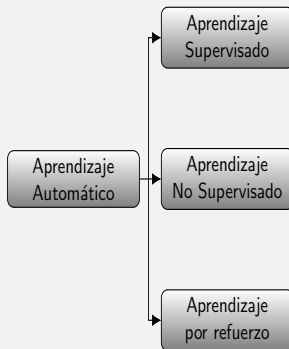
A través de los años el humano ha creado una tremenda cantidad de datos en diferentes ámbitos de la vida diaria:



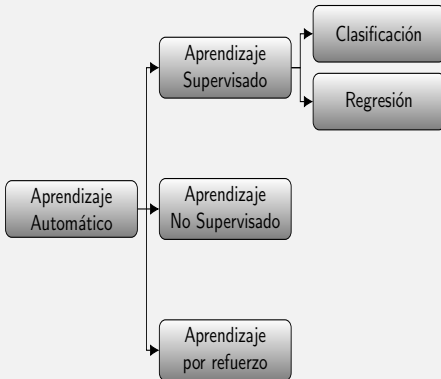
# Ramas principales del aprendizaje automático

Aprendizaje  
Automático

# Ramas principales del aprendizaje automático

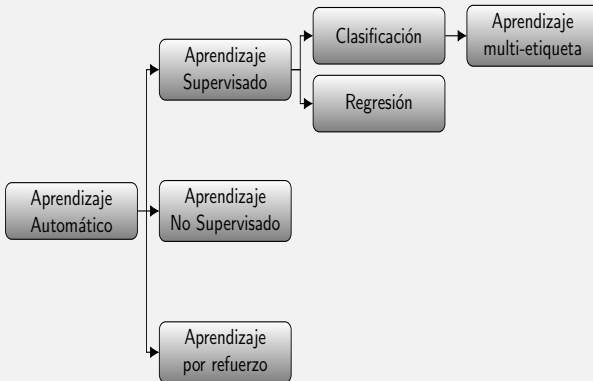


# Ramas principales del aprendizaje automático

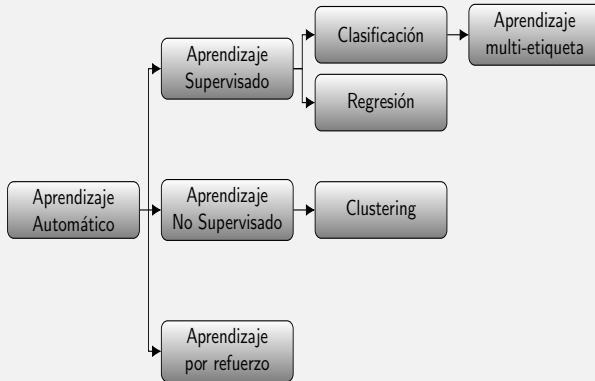




# Ramas principales del aprendizaje automático



# Ramas principales del aprendizaje automático



# Aprendizaje supervisado

# Aprendizaje supervisado

El nombre de aprendizaje supervisado proviene del hecho que es **necesaria** la **intervención** de un maestro o **supervisor** que **asigne** correctamente el valor de la **etiqueta** en casos previamente adquiridos.

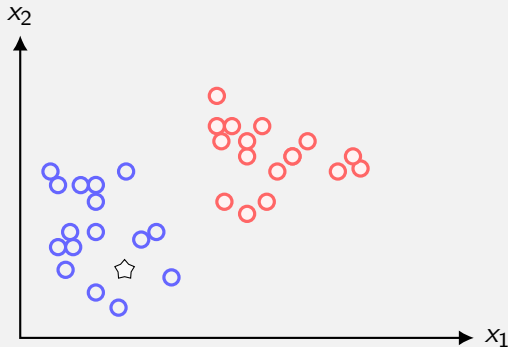
## Aprendizaje supervisado

El nombre de aprendizaje supervisado proviene del hecho que es **necesaria** la **intervención** de un maestro o **supervisor** que **asigne** correctamente el valor de la **etiqueta** en casos previamente adquiridos.

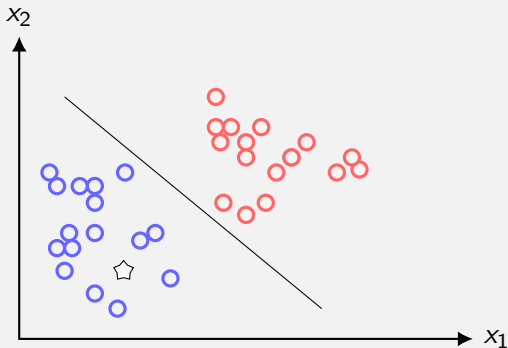
Desde un punto de vista formal, **dado** un conjunto de parejas **dato-etiqueta**, el aprendizaje supervisado **busca** un modelo que permita realizar una **correspondencia** de los datos a las etiquetas **de tal forma que** dado un **nuevo dato** que no se haya visto anteriormente **el algoritmo** de aprendizaje supervisado **permita predecir la etiqueta** que le correspondería.

Cuando la etiqueta puede tomar un **número finito** de valores estamos frente a un problema de **clasificación**.

Cuando la etiqueta puede tomar un **número finito** de valores estamos frente a un problema de **clasificación**.

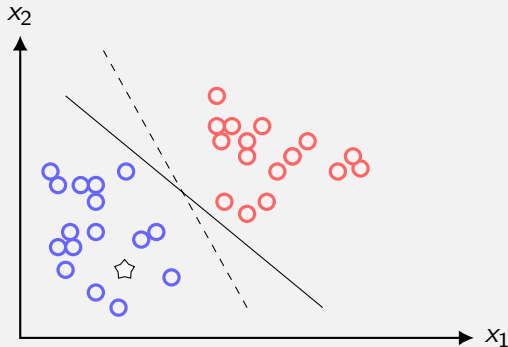


Cuando la etiqueta puede tomar un **número finito** de valores estamos frente a un problema de **clasificación**.



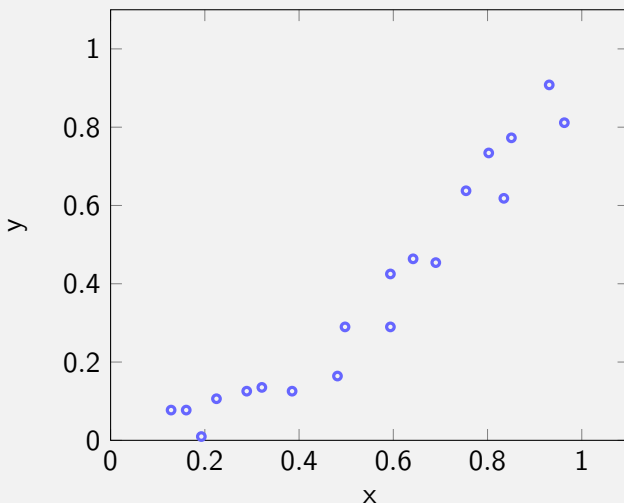


Cuando la etiqueta puede tomar un **número finito** de valores estamos frente a un problema de **clasificación**.

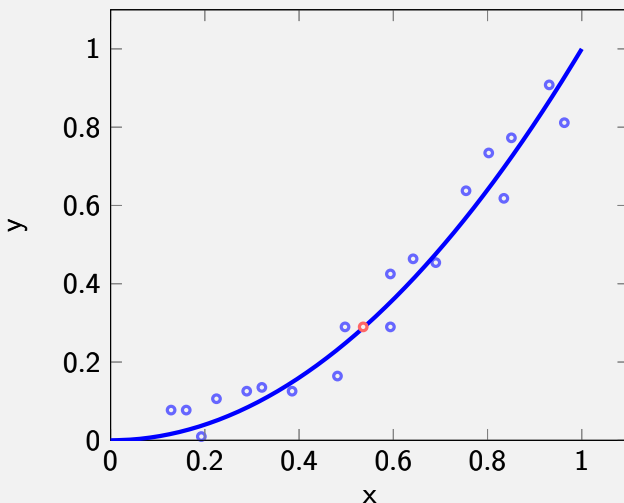


Si la etiqueta puede tomar un **número infinito** de valores, por ejemplo la predicción de un número real, llamamos a ese problema, **regresión**.

Si la etiqueta puede tomar un **número infinito** de valores, por ejemplo la predicción de un número real, llamamos a ese problema, **regresión**.



Si la etiqueta puede tomar un **número infinito** de valores, por ejemplo la predicción de un número real, llamamos a ese problema, **regresión**.



# Aprendizaje no supervisado

## Aprendizaje no supervisado

Este tipo de aprendizaje se distingue del supervisado por que **no existe etiqueta** que se quiera predecir, ni un tutor que anote previamente los datos e identifique cual es el resultado deseado.

## Aprendizaje no supervisado

Este tipo de aprendizaje se distingue del supervisado por que **no existe etiqueta** que se quiera predecir, ni un tutor que anote previamente los datos e identifique cual es el resultado deseado.

El aprendizaje no supervisado es importante cuando **objetivo** del aprendizaje es **explorar** y ganar más información sobre **los datos**.

## Aprendizaje no supervisado

Este tipo de aprendizaje se distingue del supervisado por que **no existe etiqueta** que se quiera predecir, ni un tutor que anote previamente los datos e identifique cual es el resultado deseado.

El aprendizaje no supervisado es importante cuando **objetivo** del aprendizaje es **explorar** y ganar más información sobre **los datos**.

Su componente clave es el concepto de **similitud entre los datos**.



## Aprendizaje no supervisado

Este tipo de aprendizaje se distingue del supervisado por que **no existe etiqueta** que se quiera predecir, ni un tutor que anote previamente los datos e identifique cual es el resultado deseado.

El aprendizaje no supervisado es importante cuando **objetivo** del aprendizaje es **explorar** y ganar más información sobre **los datos**.

Su componente clave es el concepto de **similitud entre los datos**.

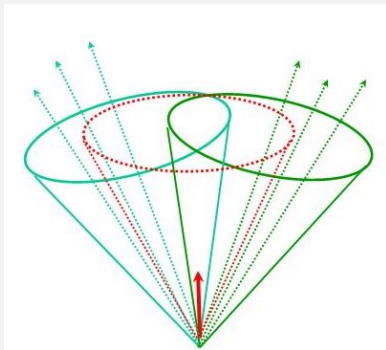
Se pueden distinguir **dos** grandes **sub-familias** dentro de este grupo:

## Métodos de agrupación o mixturas o clustering

Pretende dividir los datos en **grupos excluyentes** (un dato en un solo grupo) o **no excluyentes** (un dato puede pertenecer a diversos grupos con distinto índice de pertenencia).

## Métodos de agrupación o mixturas o clustering

Pretende dividir los datos en **grupos excluyentes** (un dato en un solo grupo) o **no excluyentes** (un dato puede pertenecer a diversos grupos con distinto índice de pertenencia).



**Figura:** Una partícula puede o no pertenecer a jet dependiendo de parámetros geométricos.

# Aprendizaje por refuerzo

# Aprendizaje por refuerzo

Este tipo de aprendizaje se basa en recompensas y penalizaciones.

## Aprendizaje por refuerzo

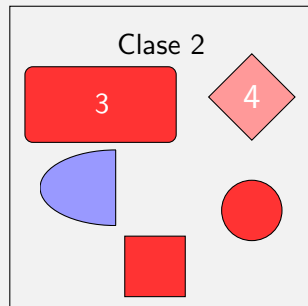
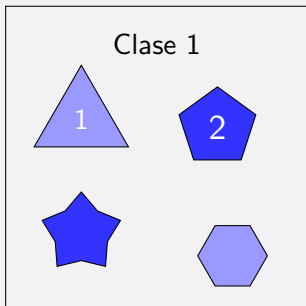
Este tipo de aprendizaje se basa en recompensas y penalizaciones.

En el aprendizaje por refuerzo no se especifica aquello que se busca como en el aprendizaje supervisado, sino que cuando el sistema de aprendizaje realiza una acción aceptable o correcta se le recompensa y se le penaliza en el caso contrario.

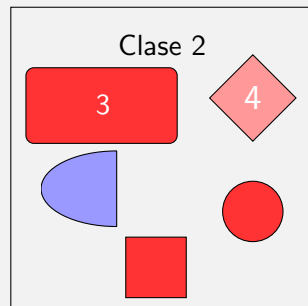
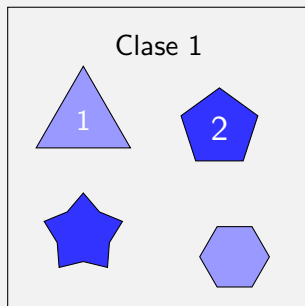
El **primer paso** necesario para el diseño y uso de los algoritmos de aprendizaje automático es la **representación del conocimiento** o de la experiencia previa.

Para ello se necesita, en general, una **representación numérica** de los **conceptos** y de las **abstracciones** de la vida real.

A este proceso se le conoce como **extracción de características** y consiste en la **descripción de rasgos** relevantes de los objetos de interés **en relación a la tarea** que debe llevarse a cabo y su posterior representación numérica.

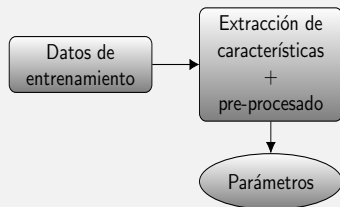


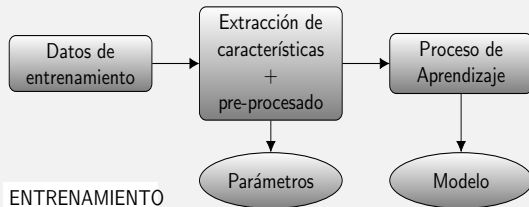


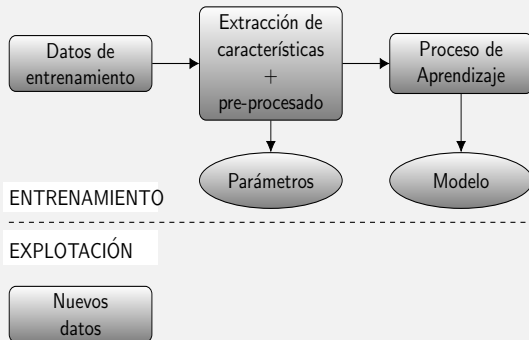


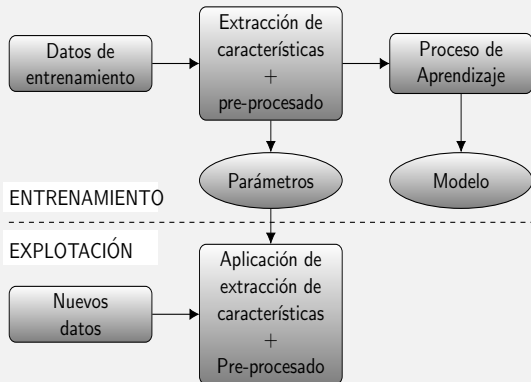
Elemento	Color	Número de vértices	Vértices redondos	Clase
1	Azul claro	3	no	1
2	Azul oscuro	5	no	1
3	Rojo oscuro	4	sí	2
4	Rojo claro	4	no	2

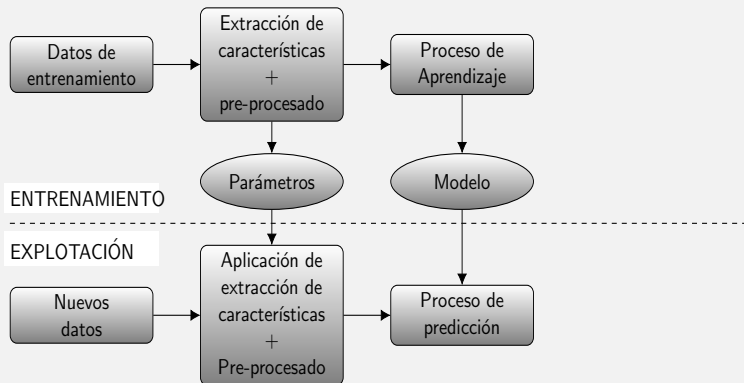
Datos de  
entrenamiento

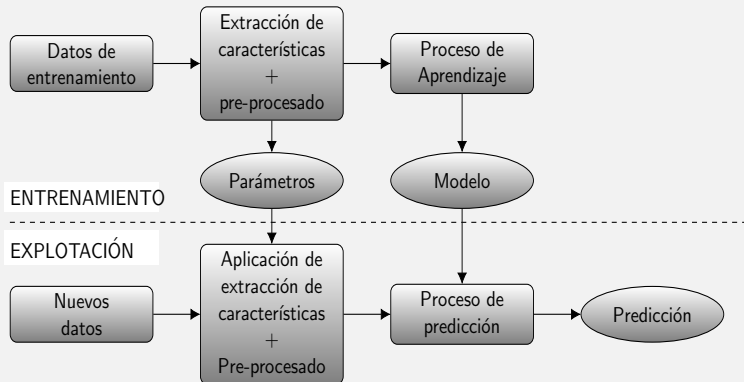














## Métodos usados en Aprendizaje Automático

### Regresión lineal[2]

Pros: Resultados fáciles de interpretar, computacionalmente barato.

Cons: No es muy bueno modelando información no lineal.

Usar con: Valores numéricos y nominales.

### Agrupamiento por k-medias (kNN) [2]

Pros: Fácil de implementar.

Cons: Puede converger a un mínimo local, lento en conjuntos grandes.

Usar con: Valores numéricos.

## Máquinas de vectores de soporte

Pros: Error de generalización bajo, computacionalmente barato y fácil de interpretar.

Cons: Sensible a ajustes en los parámetros y al kernel elegido. De forma nativa, solo trabaja con clasificación binaria.

Usar con: Valores numéricos y nominales.

## Regresión basada en árboles[2]

Pros: Se ajusta a información compleja y no lineal.

Cons: Los resultados son difíciles de interpretar.

Usar con: Valores numéricos y nominales.

# Redes neuronales

Es un sistema capaz de **aprender** las **mejores características** para un determinado problema, y es la **base de** la rama de aprendizaje automático conocida como **aprendizaje profundo**.

# Perceptrón

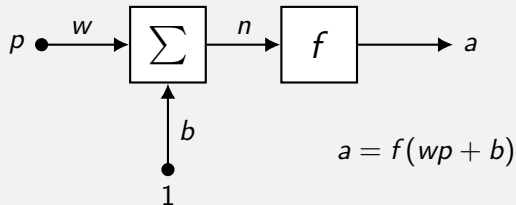


Figura: Una neurona.

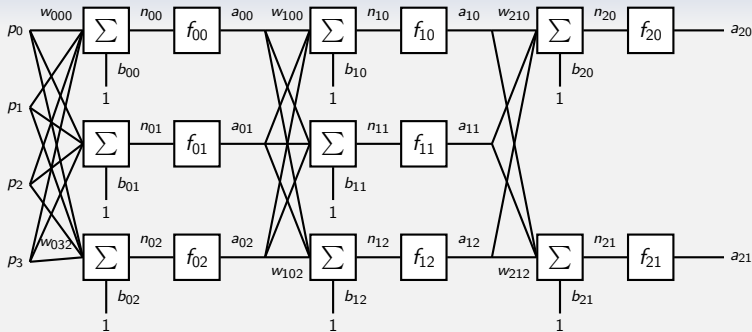


Figura: Perceptrón multicapa.

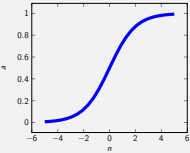
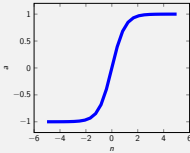
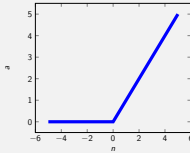
$$a_0 = f_0(W_0 p + b_0), \quad a_1 = f_1(W_1 a_0 + b_1), \quad a_2 = f_2(W_2 a_1 + b_2)$$

$$p = \begin{pmatrix} p_0 \\ \vdots \\ p_r \end{pmatrix}, \quad a_i = \begin{pmatrix} a_{i0} \\ \vdots \\ a_{i_r} \end{pmatrix}, \quad W_i = \begin{pmatrix} w_{i00} & \cdots & w_{i0c} \\ \vdots & \ddots & \vdots \\ w_{ir0} & \cdots & w_{irc} \end{pmatrix}, \quad b_i = \begin{pmatrix} b_{i0} \\ \vdots \\ b_{i_r} \end{pmatrix}$$

## Funciones de activación

¿Por qué? Incorporan la no linealidad que le da poder a una red neuronal.

¿Cuáles? Existen muchas, pero las que usaremos son:

Sigmoide	Tangente hiperbólica	ReLU
$1/(1 + e^{-n})$	$\tanh(n)$	$\max(0, n)$
		

## Más información en...

- Peter Harrington. *Machine Learning in Action*.  
Manning Publications Co., USA, 2012
- Howard B. Demuth, Mark H. Beale, Orlando de Jesús, and  
Martin T. Hagan. *Neural Network Design*.  
Martin Hagan, Stillwater, OK, USA, 2nd Edition, 2014

Los ejercicios contemplados son:

- Aprendizaje Supervisado Clasificación.
- Aprendizaje Supervisado Regresión.
- Aprendizaje No Supervisado.



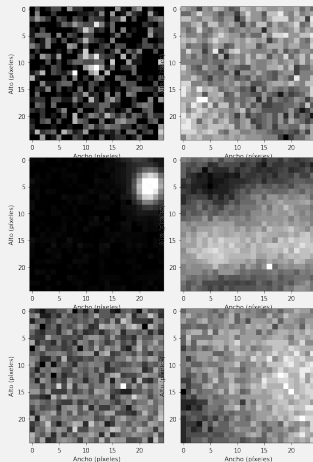
## Ejemplo 1

# Análisis de imágenes y Aprendizaje automático

¿Podrían usarse para clasificar imágenes?

# Análisis de imágenes y Aprendizaje automático

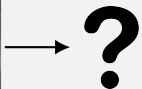
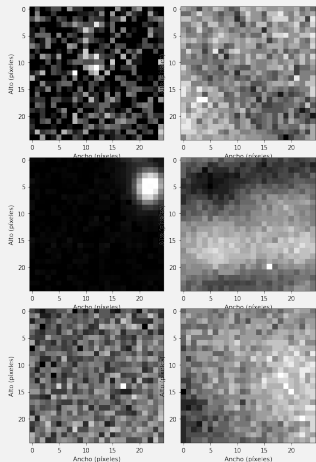
¿Podrían usarse para clasificar imágenes?



?

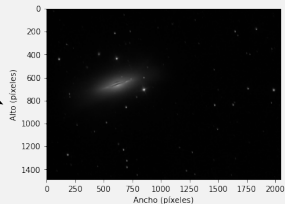
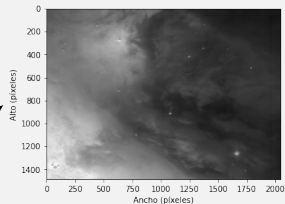
# Análisis de imágenes y Aprendizaje automático

¿Podrían usarse para clasificar imágenes?



Nebulosa

Galaxia



# Análisis de imágenes y Aprendizaje automático

Diagrama para clasificar fragmentos de imágenes que pertenecen a una nebulosa o a una galaxia

Adquisición de  
imágenes

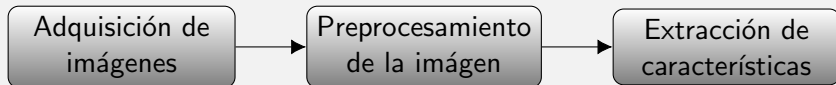
# Análisis de imágenes y Aprendizaje automático

Diagrama para clasificar fragmentos de imágenes que pertenecen a una nebulosa o a una galaxia



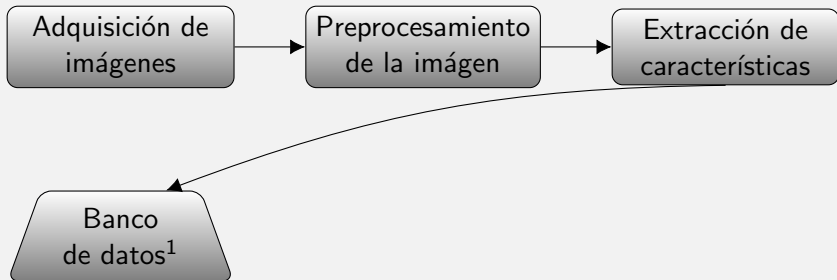
# Análisis de imágenes y Aprendizaje automático

Diagrama para clasificar fragmentos de imágenes que pertenecen a una nebulosa o a una galaxia



# Análisis de imágenes y Aprendizaje automático

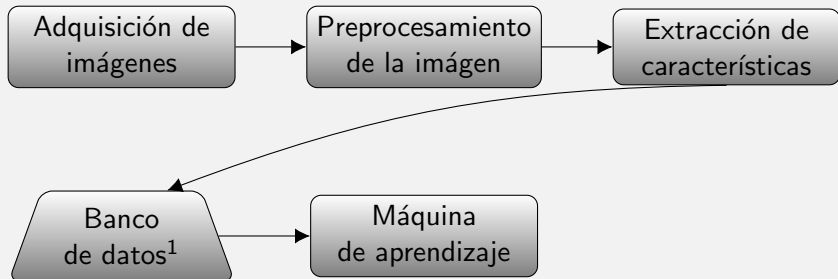
Diagrama para clasificar fragmentos de imágenes que pertenecen a una nebulosa o a una galaxia





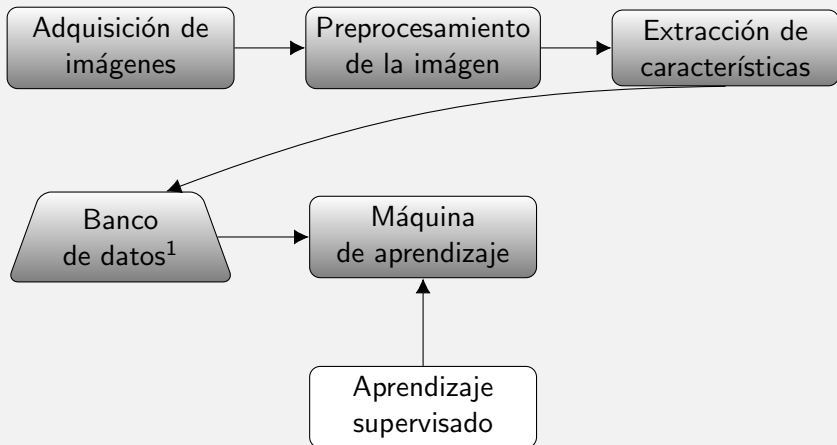
# Análisis de imágenes y Aprendizaje automático

Diagrama para clasificar fragmentos de imágenes que pertenecen a una nebulosa o a una galaxia



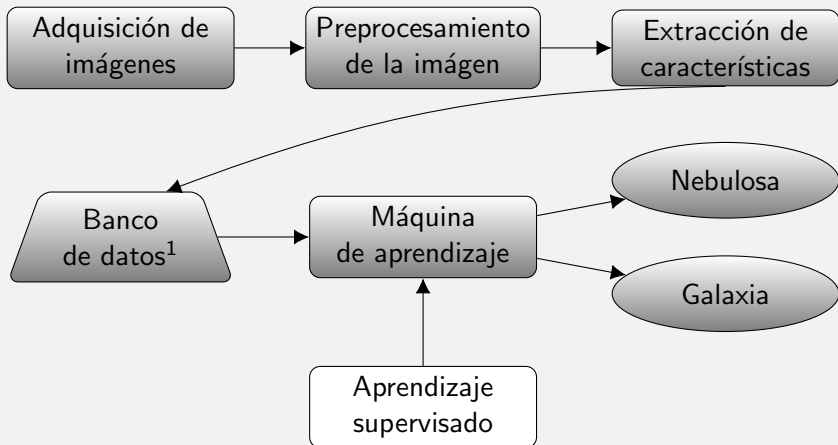
# Análisis de imágenes y Aprendizaje automático

Diagrama para clasificar fragmentos de imágenes que pertenecen a una nebulosa o a una galaxia



# Análisis de imágenes y Aprendizaje automático

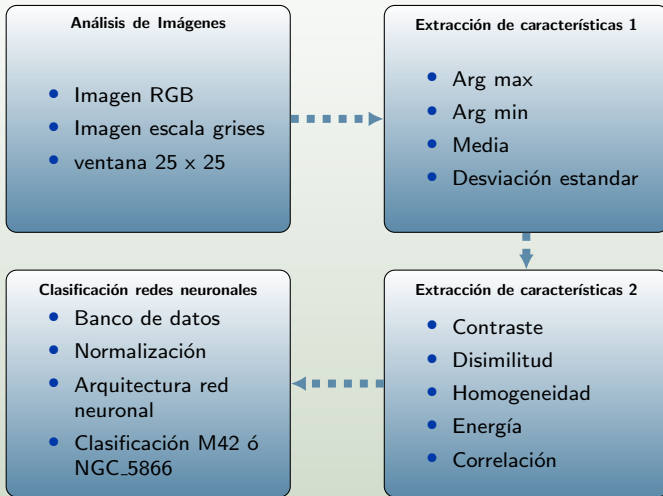
Diagrama para clasificar fragmentos de imágenes que pertenecen a una nebulosa o a una galaxia



<sup>1</sup>Holdout: criterio para separar banco de datos

# Análisis de imágenes y Aprendizaje automático

## Diagrama técnico para clasificación de imágenes usando aprendizaje supervisado



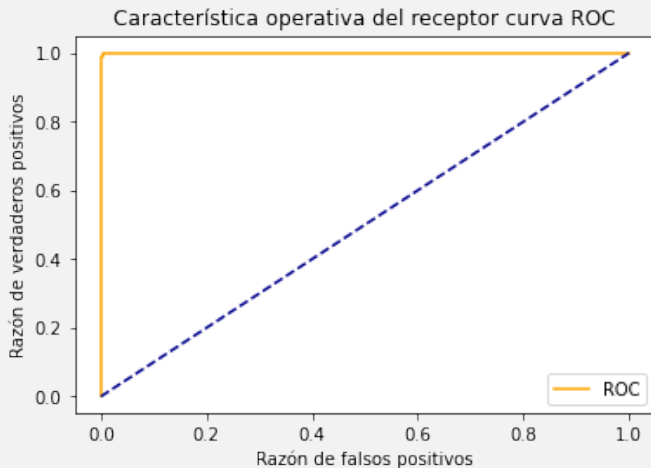
## Matriz de confusión

		salida del predictor	
		M42	NGC 5866
valor real	M42	756	4
	NGC 5866	0	716

**Exactitud:[3]**

$$\frac{(VP + VN)}{(VP + FP + FN + VN)} = \frac{(756 + 716)}{(756 + 0 + 4 + 716)} = 0.99$$

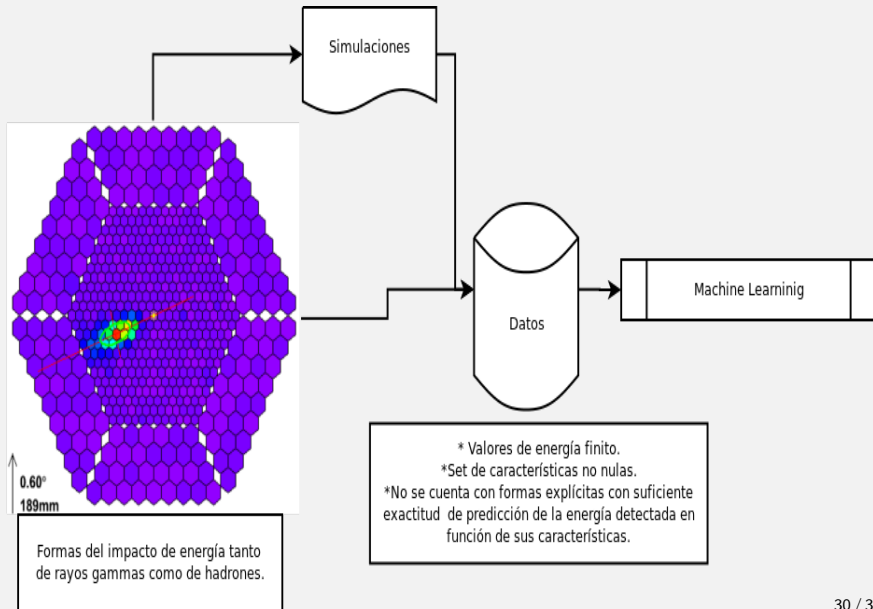
## Curva ROC <sup>1</sup>



<sup>1</sup>ROC:Receiver Operating Characteristic

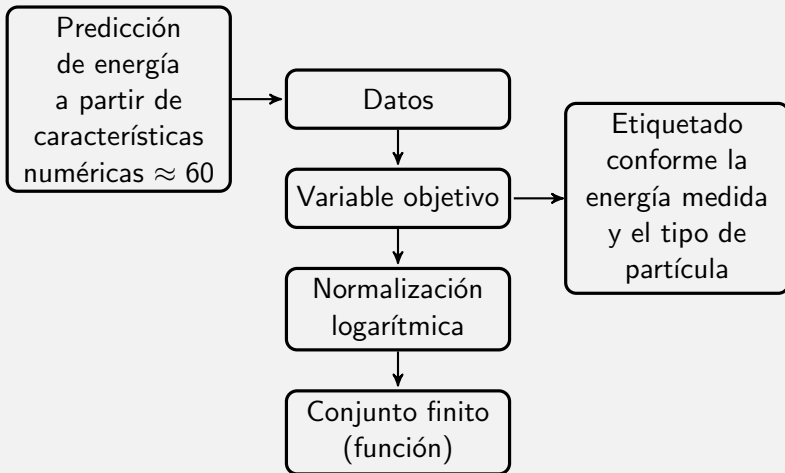
## Ejemplo 2

## Predicción de energía

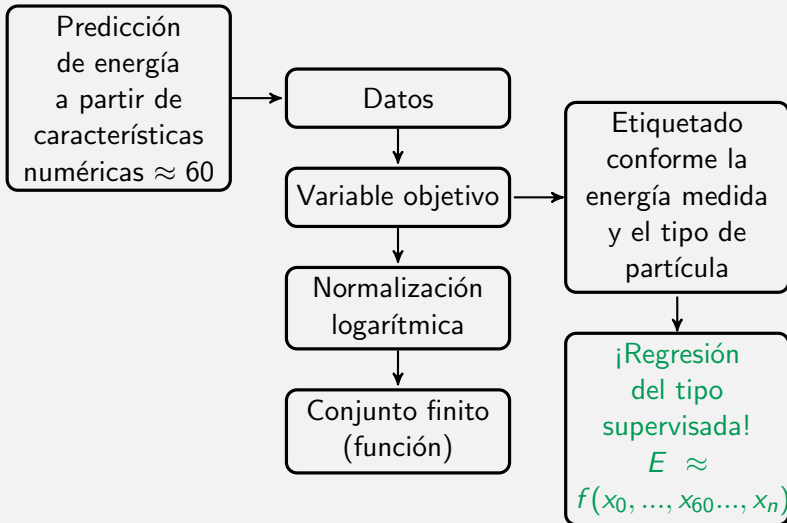




## Problema particular



## Problema particular



## Modelos de predicción y sus métricas

- Ciertas características son más importantes que otras. ¿Alguna característica se puede descartar?,  
¿ Valdría la pena?

## Modelos de predicción y sus métricas

- Ciertas características son más importantes que otras. ¿Alguna característica se puede descartar?,  
¿ Valdría la pena? **Sí (proceso de descubrimiento).**

## Modelos de predicción y sus métricas



- Ciertas características son más importantes que otras. ¿Alguna característica se puede descartar?,  
¿ Valdría la pena? **Sí (proceso de descubrimiento).**
- ¿La importancia de las características depende del algoritmo empleado?

## Modelos de predicción y sus métricas


- Ciertas características son más importantes que otras. ¿Alguna característica se puede descartar?,  
¿ Valdría la pena? **Sí (proceso de descubrimiento).**
- ¿La importancia de las características depende del algoritmo empleado?  
**Sí (depende de los patrones encontrados).**

## Modelos de predicción y sus métricas



- Ciertas características son más importantes que otras. ¿Alguna característica se puede descartar?,  
¿ Valdría la pena? **Sí (proceso de descubrimiento).**
- ¿La importancia de las características depende del algoritmo empleado?  
**Sí (depende de los patrones encontrados).**

- Factor de correlación   $\approx 1$    $\approx 0$ :

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

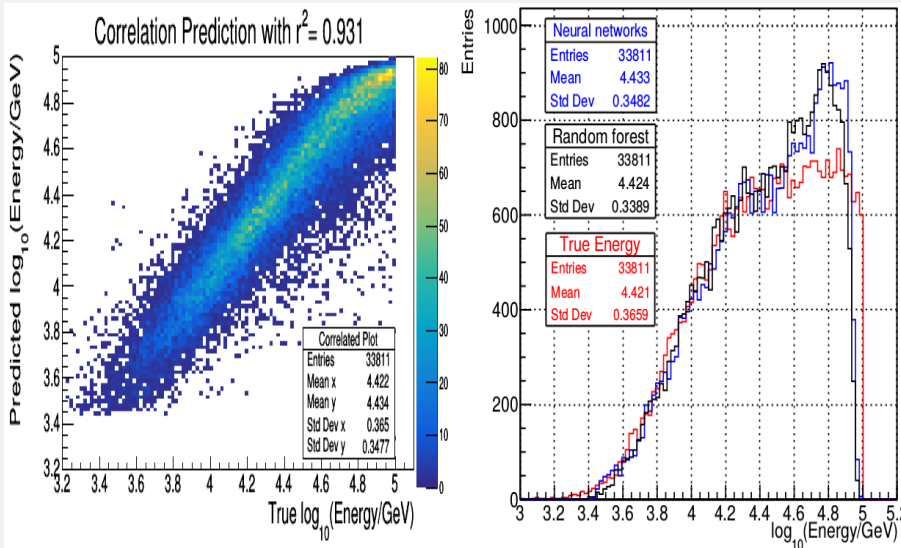
- Error absoluto medio   $\approx 0$    $\approx +/ - \infty$ :

$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n} = \frac{\sum_{i=1}^n |e_i|}{n} \quad (2)$$

- Bias, desviación o sesgo   $\approx 0$    $\approx +/ - \infty$ :

$$B \equiv \langle E_{Objetivo} - E_{predicho} \rangle \quad (3)$$

# ¿Qué tan bien predice y qué aprendimos?







Miguel de Jesús Martínez Felipe



Linkedin



mmartinezf2002@alumno.ipn.mx



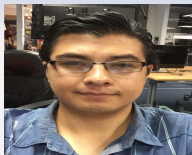
Orcid



mjmf2402@hotmail.com

**Curriculum Vitae resumido. Perfil:**

Ingeniero en informática egresado de **UPIICSA-IPN** con maestría en ciencias de la computación en **CIC-IPN** enfocado al área de análisis de imágenes, máquinas de aprendizaje e información de computación cuántica, cuenta con experiencia como desarrollador de cómputo en la nube en el tribunal federal de justicia administrativa, Actualmente estudia un doctorado en ciencias de la computación en el CIC-IPN, colaborando con el observatorio **HAWC** en un proyecto de aprendizaje automático.



## Oscar Roberto Chaparro Amaro



Linkedin



a170611@sagitario.cic.ipn.mx



Orcid



saga9211@hotmail.com

### Curriculum Vitae resumido. Perfil:

Ingeniero biomédico egresado de **UPIBI-IPN** con maestría en ciencias de la computación en **CIC-IPN** enfocado en proteómica y bioinformática con colaboración en un proyecto de rehabilitación de hueso en UPIBI, cuenta con experiencia como ingeniero de servicio en equipo médico, con colaboraciones con **FermiLab** en los proyectos de **GeantV** y **VecCore**, teniendo una estancia de investigación en el **ICN2** en nanopartículas magnéticas en aplicaciones biomédicas. Actualmente estudia un doctorado en ciencias de la computación en el CIC-IPN, colaborando con el observatorio **HAWC** en un proyecto de aprendizaje automático.

## Referencias bibliográficas I



Ian H. Witten, Eibe Frank, and Mark A. Hall.

*Data Mining: Practical Machine Learning Tools and Techniques.*

Morgan Kaufmann Series in Data Management Systems.

Morgan Kaufmann, Amsterdam, 3 edition, 2011.



Peter Harrington.

*Machine Learning in Action.*

Manning Publications Co., USA, 2012.



Atulya Nagar, Durga Prasad Mohapatra, and Nabendu Chaki.

*Proceedings of 3rd International Conference on Advanced Computing, Networking and Informatics: ICACNI 2015, Volume 1, volume 43.*

Springer, 2015.