

# Aprendizaje automático I

Especialización en Analítica y Ciencia de Datos

Universidad de Antioquia

Jorge Bedoya

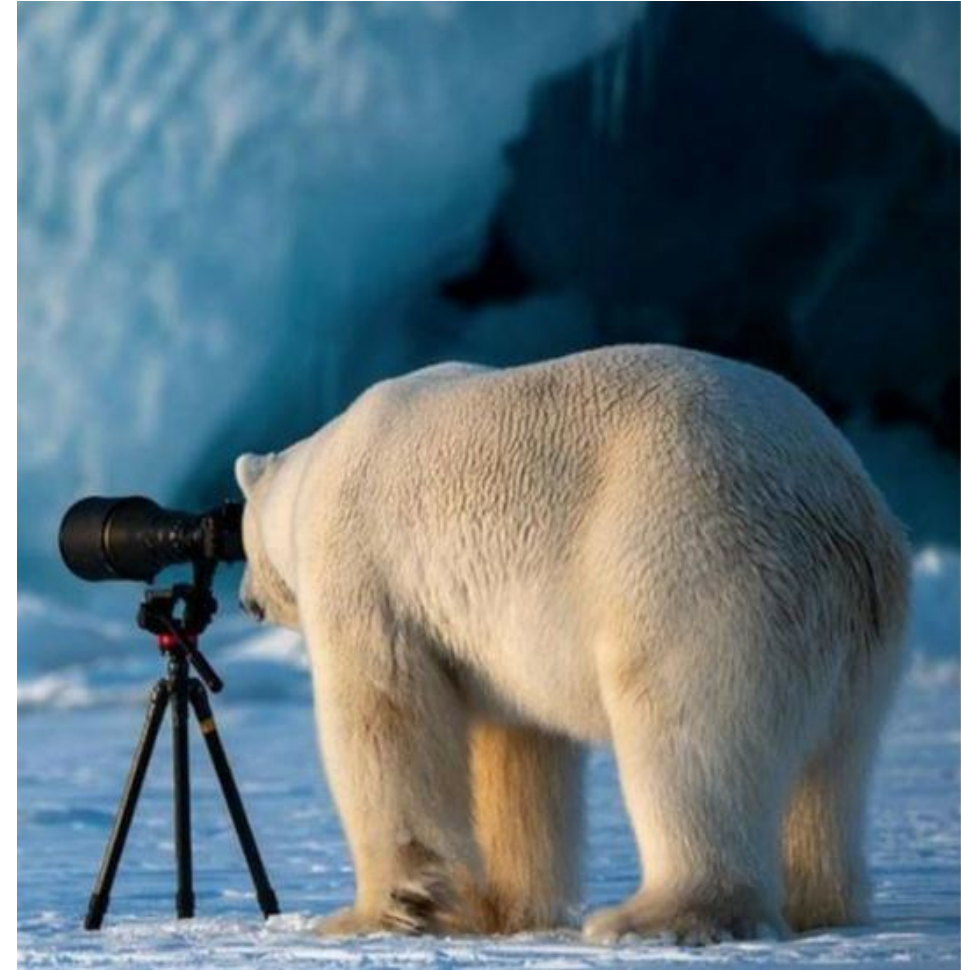
# Aprendizaje automático I

## Objetivo General:

Dotar al estudiante de los **elementos teóricos y las capacidades** para diseñar, implementar y evaluar sistemas basados en **aprendizaje automático** y usarlos para **solucionar problemas aplicados**.

## Objetivos Específicos:

- ✓ Desarrollar en el estudiante la capacidad para identificar y clasificar los tipos de problemas de aprendizaje y asociarlos con problemas de aplicación
- ✓ Dotar al estudiante con las capacidades para seleccionar el tipo de modelo más apropiado de acuerdo con las restricciones del problema a resolver.
- ✓ Proporcionar los elementos básicos para el diseño de un sistema basado en técnicas de aprendizaje de automático, desde el análisis de las características del problema, hasta solución de un modelo



<https://mag.elcomercio.pe/listas/vida-salvaje-animales-imagenes-compiten-foto-divertida-2018-noticia-561102-noticia/?foto=33>



# Unidades de aprendizaje

## Unidad 1: Introducción y fundamentos del aprendizaje automático

- Introducción, Definiciones, Taller Sklearn Script básico de una simulación en ML
- Regresión lineal y regresión logística + Taller
- Taller con dataset grande limpieza de datos + train/test con métrica de score básica para regresión y para clasificación

## Unidad 2: Clasificación y selección de modelos

- Paramétrico vs No paramétrico: K-nn vs Gaussian. Taller sobre los modelos, fronteras de decisión.
- Selección de modelos, overfitting y regularización.
- Taller con dataset real selección de modelos: k-fold, k-folds estratificado, k-fold por grupos, Bootstrapping.














## Unidad 3: Árboles de decisión y máquinas de vectores de soporte

- Árboles, Bagging + Random Forest.
- Máquinas de Vectores de Soporte, One vs All, All vs All
- Taller práctico comparación de modelos de la semana

## Unidad 4: Boosting y selección de características

- Boosting: Adaboost y Gradient Boosting
- Selección de características e importancia de variables
- Taller de aplicación de las técnicas de la semana

# Cronograma

	Tema	S1	S2	S3	S4	S5	S6	S7	S8
Unidad 1	Introducción								
	Definiciones, Sklearn Script básico de una simulación en ML								
	Regresión lineal y regresión logística								
	<b>Taller 0</b> con dataset grande limpieza de datos + train/test con métrica de score básica para regresión y para clasificación								
Unidad 2	Paramétrico vs No paramétrico: K-NN vs Gaussian. Taller sobre los modelos, fronteras de decisión.								
	Selección de modelos, overfitting y regularización.								
	<b>Taller 1</b>								
Unidad 3	Máquinas de Vectores de Soporte, One vs All, All vs All								
	Árboles, Bagging + Random Forest.								
	<b>Taller 2</b>								
Unidad 4	Boosting: Adaboost y Gradient Boosting								
	Selección de características e importancia de variables								
	<b>Taller 3</b>								

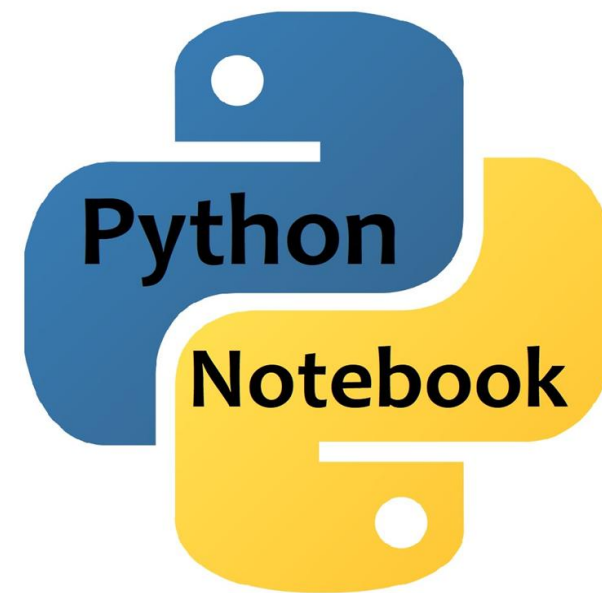


# Metodología

---

- Realización de ***clases magistrales*** para algunos conceptos y definiciones básicas.
- ***Clases tipo taller*** donde los estudiantes aplican técnicas y métodos, llevados por el docente, para el análisis de datos, creación de modelos e interpretación de resultados.
- Realización del ***talleres de clase y complemento en casa.***

# Herramientas



colab



# Evaluación

1

Hasta 4 notebooks comentados y ejecutados\* con:

1. Autores
2. Descripción corta del dataset
3. Objetivo a desarrollar
4. Resumen del proceso realizado
5. Desarrollo de experimentos
  - Preparación de datos
  - Creación de modelos
  - Evaluación de modelos
  - Etc.
6. Conclusiones
7. Referencias

2

Presentación del trabajo  
realizado (10 min)



\* Incluir dataset completo o muestra del dataset según el tamaño

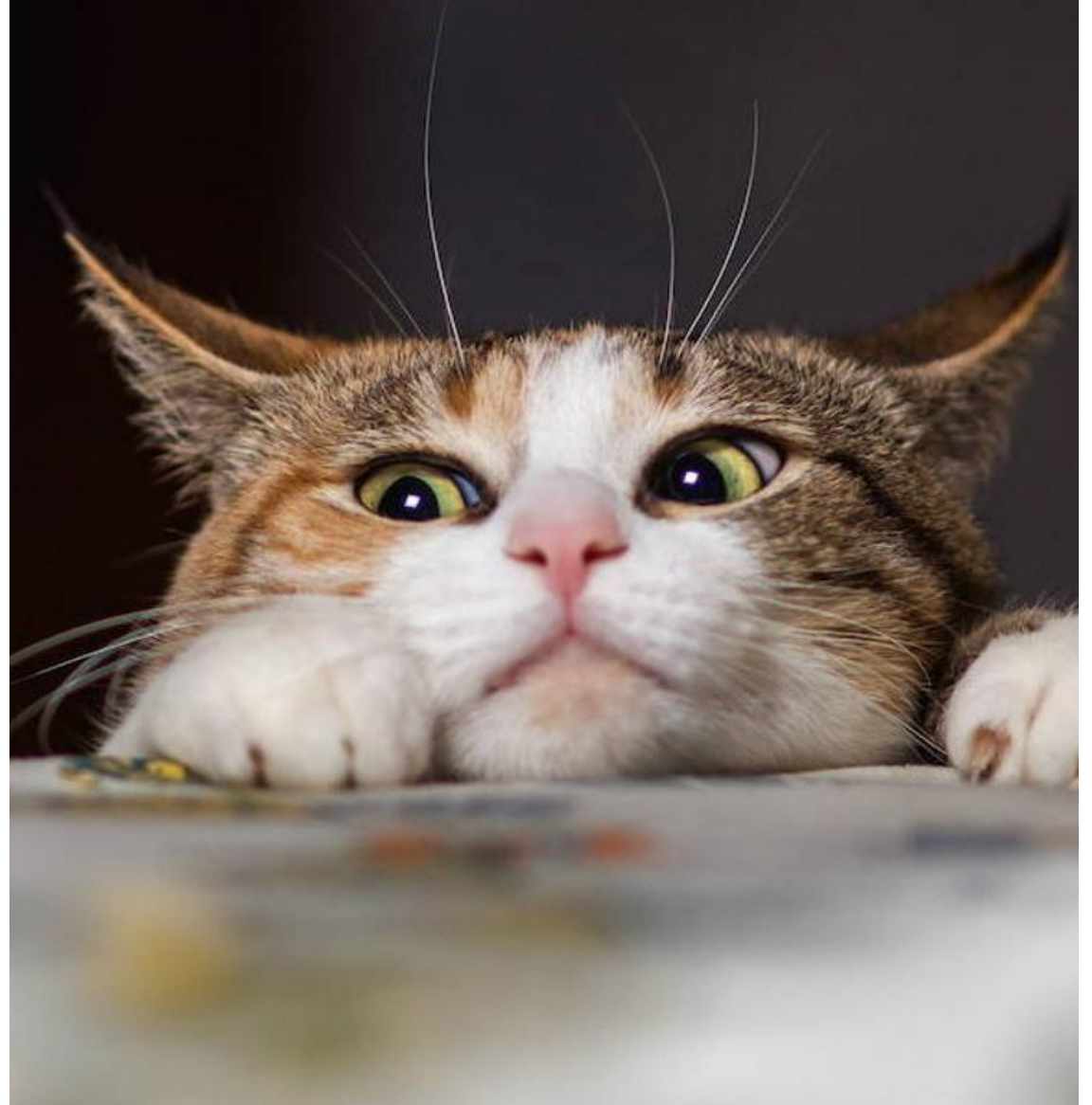
# Tipos de trabajos

---

1. 1 dataset y múltiples modelos de aprendizaje
2. Varios datasets y varios modelos de aprendizaje

- **Objetivos:**

- Identificar y ajustar el “mejor” modelo
- Identificar características de las técnicas de aprendizaje







# Evaluación

---

- Aspectos a valorar en el trabajo entregado:
  - Relación con lo visto en el curso
  - Nivel de profundización en los experimentos
  - Nivel de investigación
  - Cantidad de técnicas empleadas vs nivel de complejidad
  - **Claridad de las conclusiones, descripciones y comentarios**
  - Ortografía y redacción

# Calendario

## Clases

Dia(s)	Hora Inicial	Hora Final	Fecha inicial	Fecha final
Viernes	06:00 p. m.	10:00 p. m.	03/10/2025	25/10/2025
Sábado	08:00 a. m.	02:00 p. m.		

## Taller – Asesoría

13/10/2025 y 20/04/2025	Por definir horario (05:30 p.m. - 07:30 p.m.)
-------------------------	--



# Envío de talleres

Email: [jabedoyap79@gmail.com](mailto:jabedoyap79@gmail.com)

## Material del curso

[https://drive.google.com/drive/folders/1Bv1Ry4XNAy0a1jOL\\_NZ9SBIrq-Rrqf-?usp=sharing](https://drive.google.com/drive/folders/1Bv1Ry4XNAy0a1jOL_NZ9SBIrq-Rrqf-?usp=sharing)

# Envío de talleres

Email: [jabedoyap79@gmail.com](mailto:jabedoyap79@gmail.com)

## Información de contacto

- Jorge Bedoya
- Email: [mrbedoya@gmail.com](mailto:mrbedoya@gmail.com)
- Cel: 3105039131

