

# Aprendizaje automático I

Especialización en Analítica y Ciencia de Datos

Universidad de Antioquia

Jorge Bedoya

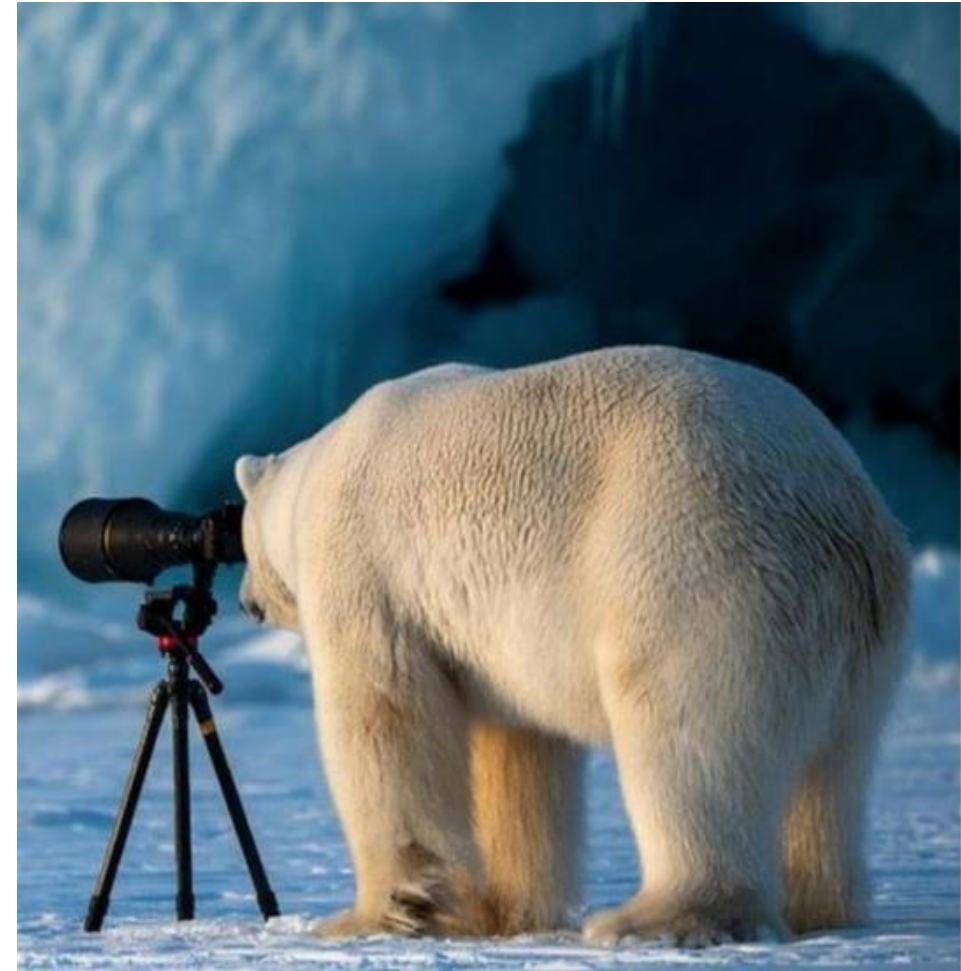
# Aprendizaje automático I

## 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 el despliegue 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								
	Taller 0 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.								
	Taller 1								
Unidad 3	Máquinas de Vectores de Soporte, One vs All, All vs All								
	Árboles, Bagging + Random Forest.								
	Taller 2								
Unidad 4	Boosting: Adaboost y Gradient Boosting								
	Selección de características e importancia de variables								
	Taller 3								



# 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

2 entregables (notebooks) que consisten en experimentos de sobre las técnicas vistas en el curso y 1 presentación

1. Notebooks comentados y ejecutados\* :
2. Presentación del trabajo realizado (10 min)

- Autor(es)
- Descripción
- Experimentos
- Conclusiones
- Referencias

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

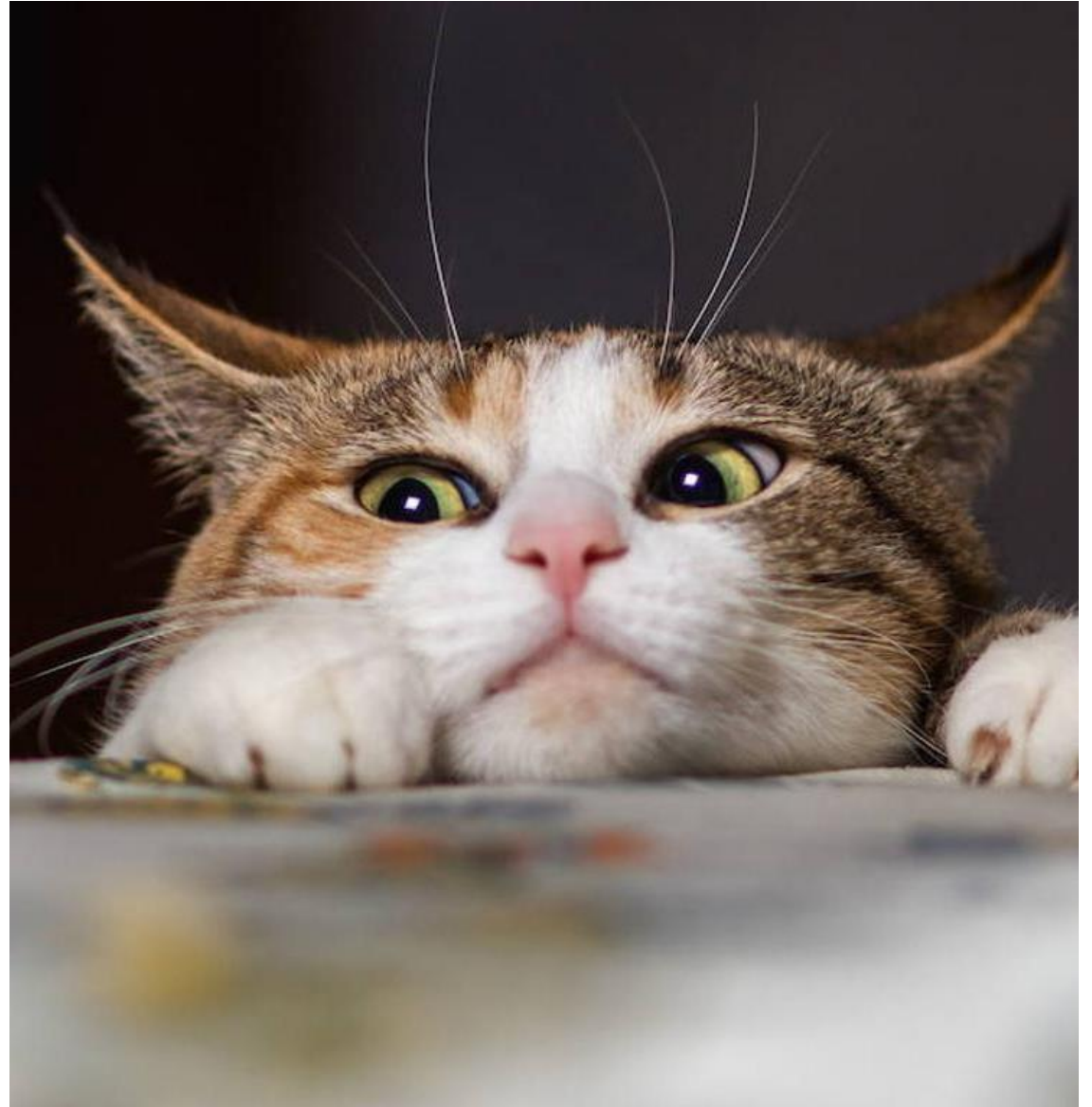
# Tipos de trabajos

---

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

<b>Dia(s)</b>	<b>Hora Inicial</b>	<b>Hora Final</b>	<b>Fecha inicial</b>	<b>Fecha final</b>
Viernes	6:00 PM	10:00 PM	19/05/2023	10/06/2023
Sábado	8:00 AM	02:00 PM		



# Envío de talleres

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

## Material del curso

<https://drive.google.com/drive/folders/1mKOfxwes55QXGLrjUlw4mRzHxtymWk2Y?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

