

2. Introducción

Ejercicio 2.1. Revisar y completar el notebook `notebook_1_herramientas.ipynb`.

Ejercicio 2.2. Describir para los siguientes problemas si se trata de aprendizaje supervisado o aprendizaje no supervisado. Especificar qué medida de performance y de un ejemplo de una base de datos que permita encarar el problema.

- (a) Detección de discurso de odio en tweets;
- (b) Segmentación de imágenes en los objetos que contiene;
- (c) Detección de fraude en tarjetas de crédito.
- (d) Agrupar voces según qué tan parecido suenan.

Ejercicio 2.3. Determinar para los siguientes problemas de aprendizaje supervisado si se trata de problemas de clasificación o de regresión. Para cada caso, indique los posibles valores que puedan tomar las etiquetas, especificando detalladamente **el tipo de datos (computacional)** que le corresponde. Tip 1: Imaginar la base de datos (es decir, qué le hubieran pedido a anotadores expertos que completen). **Warning: chatGPT no hace bien este ejercicio**

Ejemplo:

- **Problema:** Dado un tweet, determinar si habla a favor de un candidato presidencial.
 - **Etiquetas posibles:** Sí, No.
 - **Tipo etiquetas:** bool
 - **Respuesta:** Clasificación
- (a) Predecir cuánto gastará una empresa en luz el próximo semestre.
 - (b) Dado un tweet, predecir la probabilidad de que hable en contra o a favor de un candidato.
 - (c) Predecir a qué distancia de la facultad vive una persona.
 - (d) Predecir si se gastará más o menos que \$50.000 por mes de luz el próximo semestre.
 - (e) Predecir la probabilidad de que se gaste más o menos que \$50.000 por mes de luz el próximo semestre.
 - (f) Predecir la nota que tendrá un alumno en un examen cuya nota puede ser $0, 1, 2, \dots, 10$
 - (g) Predecir la nota que tendrá un alumno en un examen cuya nota puede ser "A", "R" o "T".
 - (h) Predecir dónde vive una persona.
 - (i) Predecir la próxima palabra a autocompletar dadas las oraciones anteriores.

Revisar el punto (d): ¿Qué responderían si en la base de datos tenemos etiquetas \mathbb{R} ? ¿Y si tuviéramos etiquetas binarias?

Ejercicio 2.4. Sea un problema de clasificación en el cual cada instancia tiene 2 atributos numéricos y pertenece a una de dos clases posibles (blanco o negro). Se tienen tres tipos de hipótesis ilustrados en Fig 1 que representan (a) rectas y la dirección en la cual se clasifica a una instancia como blanco, (b) dos rectas sin inclinación (ya sea horizontales o verticales, entre las cuales las instancias son clasificadas como blancos, (c) 3 elipses (que delimitan instancias blancas). Para cada uno de ellos, se pide:

- Describir el espacio de hipótesis H ;
- Identificar la cantidad de parámetros mínimos para describir una hipótesis de esta forma.

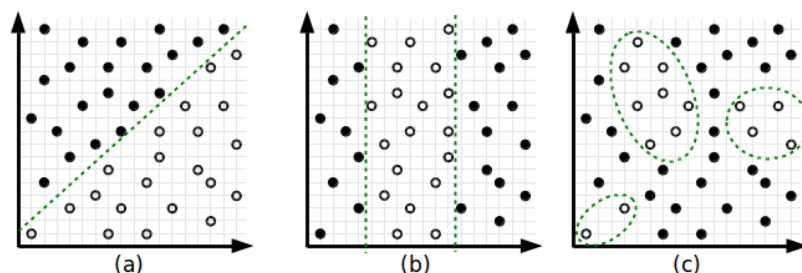


Figura 1: Tipos de Hipótesis

Ejercicio 2.5. Completar el notebook `notebook_2_titanic.ipynb`.