

Universidad Nacional Autónoma de México
Facultad de Ciencias



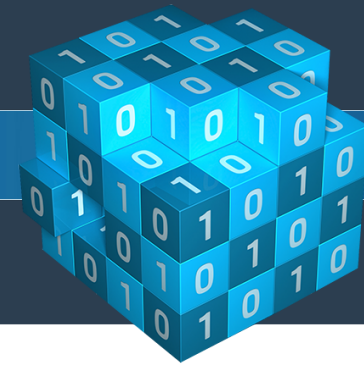
Profesores:

Dra. Amparo López Gaona
M. I. Gerardo Avilés Rosas

Laboratorista:

José Luis Vázquez Lázaró





1. Objetivo

Al finalizar esta práctica el alumno será capaz de:

- Identificar la diferencia entre un problema de aprendizaje supervisado y uno no supervisado.
- Conocer herramientas de construcción de árboles de decisión.
- Aplicar la construcción de árboles de decisión a un problema de predicción.

2. Marco teórico

Un **árbol de decisión** es un modelo de predicción utilizado en diversos ámbitos que van desde la inteligencia artificial hasta la **Minería de Datos**. Dado un conjunto de datos se fabrican diagramas de construcciones lógicas, muy similares a los sistemas de predicción basados en reglas, que sirven para representar y categorizar una serie de condiciones que ocurren de forma sucesiva, para la resolución de un problema.

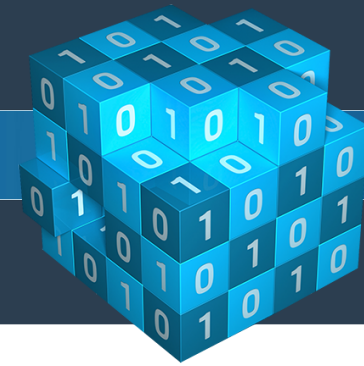
3. Instrucciones

- Utilizar las herramientas de minería de datos **RapidMiner** y **R** (la biblioteca **rattle**).
- Analizar el dataset **wine.data.csv** adjunto a esta práctica, así como el archivo **wine.names.txt** (este contiene la información del dataset).

4. Actividades

1. Responda las siguientes preguntas:
 - a) Liste algunas de las ventajas que ofrece un árbol de decisión y qué los hace tan populares con respecto a otros métodos de predicción.
 - b) En un árbol de decisión ¿un atributo puede ser revisado en más de una ocasión?, ¿existen atributos que, por el contrario, sean revisados una sola vez? (Justifique).





- c) ¿Cómo y por qué puede utilizarse un árbol de decisión como método de reducción de datos?
- d) Dado un conjunto de datos ¿el árbol de decisión construido a partir de este es único?
2. Realiza una exploración del conjunto de datos **wine** como en la práctica 9.
3. Elabora un clasificador para el conjunto de datos **wine** en **RapidMiner** utilizando un árbol de decisión. Explique brevemente la configuración y construcción de su proceso.
4. Elabora un clasificador para el conjunto de datos **wine** en **R** utilizando un árbol de decisión. Explique brevemente la configuración y construcción de su proceso.
5. Compara los dos árboles de decisión obtenidos y justifica cuál de los dos es “mejor”.

5. Entregables

Deberás enviar un archivo .zip, con nombre <número de cuenta>_practical12, que contenga lo siguiente:

- Un archivo .pdf con el desarrollo y resultado de las 5 actividades (para las actividades 3 y 4 incluye la captura del árbol de decisión correspondiente).
- Un archivo README.txt que contenga tu nombre completo, tu número de cuenta y tu correo.

a la dirección de correo luis_lazaro@ciencias.unam.mx con el asunto [A&MD2018-2]Practical12.

