

Universidad Nacional Autónoma de México
Facultad de Ciencias



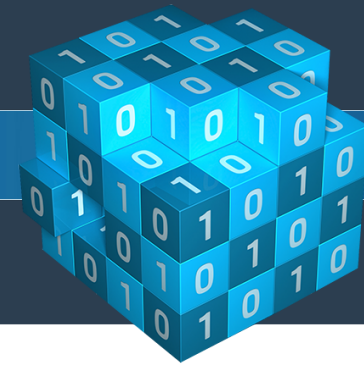
Profesores:

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

Laboratorista:

José Luis Vázquez Lázar





1. Objetivo

Al finalizar esta práctica el alumno será capaz de:

- Buscar patrones frecuentes dentro de un conjunto de datos e interpretar los resultados.
- Apreciar la relevancia del soporte y la confianza para reducir el número de patrones y facilitar su revisión.

2. Marco teórico

Las **Reglas de Asociación** es una técnica de **Minería de Datos** (propuesta por *Agrawal* in 1993) que trabaja sobre datos categóricos con la finalidad de revelar propiedades importantes e intrínsecas dentro de un conjunto de datos (*dataset*). El ejemplo típico de uso de esta técnica es el análisis de canasta de compra (en el contexto de los supermercados), donde la información obtenida se puede utilizar como base para tomar decisiones sobre marketing como precios promocionales para ciertos productos o dónde ubicar éstos dentro del supermercado.

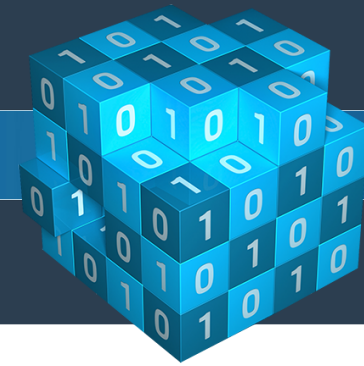
3. Instrucciones

- Utilizar el software **RStudio**
- Analizar el dataset **bank-additional-full.csv** adjunto a esta práctica, este será el dataset con el que se trabajará. La información de este dataset la puedes consultar en <https://archive.ics.uci.edu/ml/datasets/bank+marketing>

4. Actividades

1. Instale los paquetes **grid** y **arules**.
2. Ocupe la función **apriori** para generar las reglas de asociación con la configuración por defecto. Observe el resultado, ¿encuentras reglas útiles y significativas?. Justifique.





3. Ejecute de nuevo la función **apriori** (cinco veces al menos) pero ahora variando el soporte. Observe el resultado, ¿observa reglas útiles y significativas?, ¿qué cambio con respecto al soporte por defecto?, ¿en qué ayudan este tipo de cambios (del soporte)? Seleccione el soporte que más le parezca apropiado (justifique su respuesta).
4. Repite el paso anterior pero ahora modificando la confianza. Responda las mismas preguntas que el ejercicio anterior pero ahora con respecto al parámetro actual.
5. Tras las iteraciones anteriores, arroje el conjunto que, a su juicio, es el más significativo, útil y novedoso. No olvide indicar con que parámetros los obtuvo.

5. Entregables

Deberás enviar un archivo .zip, con nombre <número de cuenta>_practical1, que contenga lo siguiente:

- El proyecto RStudio con las 5 actividades solicitadas (debe incluir las respuestas a las preguntas).
- Un archivo .pdf que resuma los resultados obtenidos en las últimas 4 actividades.
- 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]Practical1.

