



**CURSO: MINERÍA DE DATOS**  
**MAESTRÍA EN EXPLOTACIÓN DE DATOS Y DESCUBRIMIENTO DE CONOCIMIENTO**

**TRABAJO PRÁCTICO ENTREGABLE II: Reglas de Asociación**

## **INTRODUCCIÓN**

En este Trabajo Práctico, se incorpora la técnica de *reglas de asociación* para abordar un problema puntual y tratar de entenderlo a partir de observar la coocurrencia de factores.

Para la exploración de este tema, se utilizará el IDE R-Studio del lenguaje de programación R con el objetivo de ejercitar los conceptos abordados en las clases teóricas.

## **OBJETIVO GENERAL**

El objetivo general de este trabajo es realizar un *análisis exploratorio* del dataset<sup>1</sup> y el posterior *preprocesamiento*, de acuerdo a las técnicas vistas en clase, para luego lograr extraer un conjunto de *reglas de asociación* que determine la coocurrencia de hechos.

El dataset contiene un relevamiento de ataques de tiburones a personas desde el año 1845 hasta 2016 con un conjunto de características y el objetivo del trabajo es encontrar asociaciones en los datos para, entre otras cosas, determinar qué características de esos ataques hacen que sea mortal para la persona.

*Si bien el trabajo tiene un carácter netamente exploratorio y se definen las consignas de manera abierta, se evaluará la aplicación de las técnicas vistas en clase así como también el carácter innovador de la solución propuesta.*

---

<sup>1</sup> El dataset *Shark\_Attack\_Data.csv* contiene un relevamiento con los ataques de tiburones a personas entre 1845-2016.



CURSO: MINERÍA DE DATOS

MAESTRÍA EN EXPLOTACIÓN DE DATOS Y DESCUBRIMIENTO DE CONOCIMIENTO

## CONSIGNAS PROPUESTAS<sup>2</sup>

### 1. ANÁLISIS EXPLORATORIO DE DATOS

- a. A partir del dataset *Shark\_Attack\_Data.csv*, explore sus características y detalle que *features* posee.
- b. De acuerdo a las técnicas abordadas en las prácticas de laboratorio, realice un análisis exploratorio para identificar cual es la distribución de sus variables y si existe relación entre las mismas<sup>3</sup>.

2. **PREPROCESAMIENTO:** A partir de las técnicas de preprocesamiento vistas en clase, y las *features* del dataset, realice el preprocesamiento correspondiente con la eventual transformación y construcción<sup>4</sup> de nuevas variables para la posterior construcción de reglas de asociación. Deberá documentar paso a paso las operaciones realizadas y fundamentarlas.

### 3. REGLAS DE ASOCIACIÓN

- a. Genere reglas de decisión que expliquen los factores más asociados en cuanto a ataques de tiburones. Explique cuáles son las reglas más robustas, justifique y explique el conocimiento que estas aportan.
- b. Indague a través del análisis de las reglas sobre cuáles son las características que hacen que un ataque sea fatal.
- c. Documente las iteraciones realizadas entre preprocesamiento y la construcción de las reglas y haga referencia a cuales son las transformaciones con las que obtuvo una mejor configuración.
- d. Busque y experimente alguna forma de visualización<sup>5</sup> para el análisis de las reglas de asociación generadas.

---

<sup>2</sup> En todas las consignas -principalmente en las relacionadas con análisis exploratorio y preprocesamiento- deberá documentar las actividades realizadas y argumentar sus decisiones en función de las técnicas (analíticas y gráficas) abordadas en clase.

<sup>3</sup> Deberán buscarse relaciones entre las variables numéricas y también entre variables categóricas.

<sup>4</sup> Se ponderará la construcción de nuevas variables a partir de las existentes.

<sup>5</sup> Se sugiere trabajar con la librería **arulesViz** de R.