

# Evaluación Electivo MDD

Cristóbal Gallardo Cubillos

Departamento de Matemática y Ciencia de la Computación,  
Universidad de Santiago de Chile, Santiago, Chile

20 de Mayo de 2025

## Etapa 0:

- **Área de estudio:** Nutrición y comida rápida
- **Problemática:** El consumo excesivo de alimentos de comida rápida se ha asociado con enfermedades como obesidad, hipertensión y diabetes. Sin embargo, no todos los productos ofrecidos por estas cadenas son igualmente dañinos.
- **Solución propuesta:** Utilizar minería de datos para clasificar alimentos según su contenido nutricional, identificando patrones que permitan distinguir entre productos más y menos saludables entre distintas cadenas de comida rápida.

## Etapa 1:

Se utiliza el conjunto de datos "fastfood\_calories.csv", el cual contiene aproximadamente 8000 registros con las propiedades nutricionales de diversos productos de comida rápida. Las columnas incluyen información como calorías, grasas, colesterol, sodio, carbohidratos, proteínas y otros nutrientes esenciales.

## Etapa 2:

Se realizó una revisión de los datos para detectar valores nulos, inconsistencias o datos irrelevantes. Se eliminaron las columnas 'salad', 'cal\_fat', 'vit\_a', 'vit\_c' y 'calcium' por ser índices redundantes o que no aportan al estudio. También se verificó que los datos nutricionales estén en formato numérico correcto y se transformaron algunas columnas categóricas para facilitar el análisis. Además, se eliminaron las filas que contenían datos nulos o NA.

## Etapa 3:

Se realizaron transformaciones como:

- Se crea una columna para el Porcentaje de calorías que provienen de la grasa.
- Se crea una columna para una métrica compuesta que favorece el contenido de fibra y proteínas, y penaliza la grasa saturada y el azúcar.
- Se crea una columna para la relación entre el contenido de sodio y proteína.
- Normalización de algunas variables numéricas.
- Agrupación por cadena de restaurante para análisis comparativo.
- Clasificación de los productos en categorías según niveles de calorías y grasas. Para saber si es o no Saludable.

Estas transformaciones permitirán aplicar técnicas de minería de datos en las siguientes etapas.