

## Contexto:

Tierra del Fuego presenta desafíos particulares en su sistema educativo debido a su geografía, distribución poblacional y condiciones sociales. Detectar a tiempo los factores que inciden en el abandono puede permitir acciones preventivas por parte de los equipos educativos y de gestión, mejorando la trayectoria escolar de los estudiantes.

Este proyecto tiene como objetivo predecir la probabilidad de abandono escolar en el nivel secundario en la provincia de Tierra del Fuego, utilizando técnicas de [Aprendizaje Automático](#). Se parte de datos oficiales que incluyen tasas de abandono interanual, promoción efectiva, sobreedad, repitencia y matrícula. El modelo busca identificar patrones que permitan anticipar el abandono, especialmente en sectores vulnerables.

## Objetivo:

- Desarrollar un modelo de clasificación que prediga la probabilidad de abandono escolar en el nivel secundario, en función de variables como promoción efectiva, sobreedad, repitencia, tipo de gestión educativa y matrícula.
- Identificar variables que determinen poblaciones de riesgo.

## Datasets utilizados:

***Dataset de Repitentes (04-1-10-Secundario-comun.-Repitentes-por-sector-de-gestion-y-ano-de-estudio.-2008-2021.xlsx)***

Cantidad de instancias: 28

Cantidad de columnas: 27

Variables principales:

- Año
- Sector de gestión (Estatal, Privado)
- Año de estudio (1° a 6°)
- Departamento (Ushuaia, Río Grande)

Tipos de datos:

- Numéricos: principalmente enteros (cantidad de repitentes, porcentajes).
- Categóricos: año, sector de gestión, año de estudio, departamento.

***Dataset de Promoción, sobreedad y abandono interanual (04\_01\_06-Secundario-comun.-Promocion-efectiva-sobreedad-y-abandono-interanual-por-ano-de-estudio.-TDF.xlsx)***

Cantidad de instancias: 9

Cantidad de columnas: 1

Variables principales descritas:

- Tasa de promoción efectiva
- Tasa de sobreedad

- Tasa de abandono interanual

Tipos de datos:

- Las variables son categóricas/descriptivas (indicadores sobre tasas específicas).  
Requiere estructuración y extracción manual de datos numéricos para análisis.

***Dataset de Matrícula (04-1-09-Secundario-comun.-Matricula-por-sector-de-gestion-ano-de-estudio-y-departamento.-TDF-USH-RGA-y-AA.xlsx)***

Cantidad de instancias: 38

Cantidad de columnas: 18

Variables principales:

- Año
- Sector de gestión (Estatal, Privado)
- Año de estudio (1° a 7°)
- Departamento (Ushuaia, Río Grande, Tolhuin, Antártida Argentina)

**Tipos de datos:**

- Numéricos: principalmente matrícula por año y por sector.
- Categóricos: año, sector de gestión, año de estudio, departamento.

Fuente: <https://ipiec.tierradelfuego.gob.ar/estadisticas-sociales-y-demograficas/>

No descarto incorporar nuevos data set que permitan ampliar información sobre las características poblacionales.