



# Prueba Analitica

Modelo de Predicción de Estado de Impago Financiero

# Objetivo de la prueba

- Medir el pensamiento crítico y creatividad analítica del participante mediante el desarrollo de un modelo estadístico computacional que logre los siguientes resultados:
  1. Desarrollar un Modelo para calcular:
    - la Probabilidad de Impago (Default) de un cliente [Variable Continua: 0.0 – 1.0]
    - El estatus de default [variable Binaria: Si=1, o No=0]
  2. Estimar el monto total (en unidad de dinero) de la cartera en perdida.

# Dataset

## 1. Nombre: *default\_dataset.csv*

- Descripción:

- El dataset presentado contiene información de clientes de algún banco extranjero en el que se mide diferentes variables y el estatus de impago, o también conocido como default, del cliente.

- Variable Objetivo: default **[Binaria (Si=1, No=0)]**

- Variables Explicativas: 24 **[Mixtas (Continuas y Categóricas)]**

## 2. Nombre: *validation\_dataset.csv*

- Descripción:

- Contiene el set de datos de validación para calificación

- Variable Objetivo: default **[Binaria (Si=1, No=0)]**

- Variables Explicativas: 24 **[Mixtas (Continuas y Categóricas)]**

# Criterios de Evaluación

- Lograr una métrica de rendimiento generalización por encima del 98% sin overfitting en el dataset de evaluación (validation\_dataset.csv)
  - Si bien existen varios tipos de métricas de clasificación, solo un subgrupo de ellas son aptas para el problema presentado en esta prueba. Escoger una y sustentar su escogencia.
- Documentación de la metodología de desarrollo que contemple (En caso de aplicar):
  - Metodología Limpieza de datos
  - Metodología de Selección de variables
  - Metodología de Imputación
  - Metodología de Escogencia de algoritmo de modelación
  - Metodología de Optimización de hiper parámetros
  - Metodología de evaluación
  - Metodología de estimación de la perdida esperada, en unidad de dinero, predicha por el modelo.
  - Etc.
- Adicionalmente, el dataset cuenta con problemáticas que surgen por la naturaleza del problema y estas dictan las decisiones de modelación. El participante debe documentar y sustentar sus decisiones.

# Criterios de Evaluación (Cont.)

- El participante deberá utilizar el dataset llamado ***validation\_dataset.csv*** para predecir a los clientes cuya variable objetivo *default [Binaria (Si=1, No=0)]* debe ser predicha.
- El participante cuenta con 5 días hábiles a partir del recibo de la prueba para entregar los resultados de la prueba.

# Requisitos de entrega

- Lenguaje de Desarrollo: Python
- Método de presentación de la respuesta: **Jupyter notebook**
  - Los sustentos textuales deben estar presentes como comentarios dentro del Jupyter Notebook.
- Emails de entrega:
  - Asunto: Resultados\_<apellido del participante>
  - [eduardo.perez@banistmo.com](mailto:eduardo.perez@banistmo.com)
  - [guillermo.marin@banistmo.com](mailto:guillermo.marin@banistmo.com)
- Banistmo tomará 3 días hábiles en calificar la prueba.
- Una vez calificada se contactará al participante el resultado obtenido y de ser exitoso pasar a una segunda etapa de entrevista técnica en persona.