



<DATA SCIENCE>

# ¡Bienvenidos/as

MÓDULO 4



→ soyhenry.com



¡Hola! Soy  
**Mariano**

Data Science Instructor



→ [soyhenry.com](http://soyhenry.com)



.01

# ¿Dónde estamos?





## BOOTCAMP

M1 | FUNDAMENTOS

M2 | BASE DE DATOS  
Y ARQUITECTURA

M3 | BUSINESS  
INTELLIGENCE

M4 | MACHINE  
LEARNING

M5 | FUNDAMENTOS  
DE NUBE





# BOOTCAMP

## M4 | MACHINE LEARNING

### TEMAS

Introducción al Machine Learning

Aprendizaje supervisado I: Regresión

Aprendizaje supervisado II: Regresión logística

Aprendizaje supervisado II: Modelos clasificación y métricas de evaluación

Modelos de ensamble

Optimización de modelos

Aprendizaje no supervisado I

Aprendizaje no supervisado II

Análisis de series temporales

Introducción al Deep Learning.



<FULLTIME/>

**SEMANAS  
BOOTCAMP**

01 | 02 | 03 | 04 | 05 | 06 | 07 | 08 | 09 | 10 | 11 | 12 | 13 | 14 | 15 | 16

**MÓDULO 1**



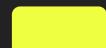
Total: 3 semanas

**MÓDULO 2**



Total: 3 semanas

**MÓDULO 3**



Total: 2 semanas

**MÓDULO 4**



Total: 3 semanas

**MÓDULO 5**



Total: 3 semanas



Cursada



Presentación y defensa de último PI



→ soyhenry.com



.02

# Proyecto integrador





<Proyecto Integrador>

# Detalle de Avances



~~HENRY~~

LET'S GO!/>

# Predicción de Fuga de clientes en el sector Financiero

PIM4 - Data Science



→ soyhenry.com



# Proyecto Integrador

## Contexto del proyecto y rol del estudiante:

**FinanceGuard es un banco digital que ha experimentado un crecimiento exponencial en los últimos años, alcanzando 10000 clientes activos.** Sin embargo, la dirección ha detectado un incremento preocupante en la tasa de abandono de clientes (churn), que actualmente alcanza el 20% anual, resultando en pérdidas millonarias.

Como Científico de Datos Junior, **tu misión es desarrollar un modelo de Machine Learning completo que prediga qué clientes tienen mayor probabilidad de abandonar el banco.** Este modelo permitirá al equipo de retención implementar estrategias personalizadas para retener a los clientes en riesgo, con el objetivo de reducir el churn al 15%.





# Objetivos del PI

- Implementar y comparar múltiples algoritmos de aprendizaje supervisado
- Optimizar modelos mediante validación cruzada y búsqueda de hiperparámetros
- Aplicar técnicas de ensamble para mejorar el rendimiento predictivo
- Manejar datos desbalanceados típicos en problemas de churn
- Implementar algoritmos básicos de aprendizaje no supervisado
- Entender e implementar los principios de aprendizaje profundo





<Proyecto Integrador>

# Entregable

## Final





# Proyecto Integrador

El entregable debe ser una carpeta de Drive (pública) en Drive con la siguiente estructura:  
ProyectoM4\_NombreEstudiante/

```
ProyectoM4_NombreEstudiante/
    └── Notebooks/
        ├── 1_EDA_RegresionLogistica.ipynb
        ├── 2_GradientBoosting_Optimizacion.ipynb
        ├── 3_AprendizajeNoSupervisado.ipynb
        └── 4_Extra_credit.ipynb
    └── Documentacion/
        └── Reporte_Modelos.pdf      # Comparación de modelos
```



**Se sugiere realizar los entregables en Google Colab, una herramientas gratuita en la nube para ejecutar código Python.**





# somos project based

Clase 3 →

## PRIMER AVANCE

Regresión Logística

Clase 6 →

## SEGUNDO AVANCE

Gradient Boosting y  
Stacking

Clase 8 →

## TERCER AVANCE

Aprendizaje No  
Supervisado

Clase 9 →

## CUARTO AVANCE

Comparaciones  
resultados de modelos  
supervisados vs no  
supervisados





# SOMOS project based

- ✓ Te recomendamos realizar los avances en las clases indicadas. ⌘
- ✓ Los avances incluyen conocimientos y habilidades aprendidas en clases anteriores. Son acumulativos. 💪
- ✓ Los avances son iterativos, apuntando siempre a la mejora u optimización de capacidades. 🔄
- ✓ No son entregables, solo se entrega el PI completo al finalizar el módulo. 🚀
- ✓ Es fundamental que realices los avances antes del Hands On con tu instructor para aprovechar al máximo esta instancia formativa, evacuar dudas y refinar tus tareas. 😊



~~HENRY~~

**¡Muchas gracias!**



→ soyhenry.com