

# Titanic

## Machine Learning from Disaster

Cepeda Villanueva Andrés M, Carvajal Garnica Julián, Moreno Barragán

Jhonatan D, Ramos Rojas Andrés C. School of Engineering



UNIVERSIDAD DISTRITAL  
FRANCISCO JOSÉ DE CALDAS

## Introduction

The Titanic Kaggle competition aims to **build a predictive system** to determine passenger survival using real historical data from 1912.

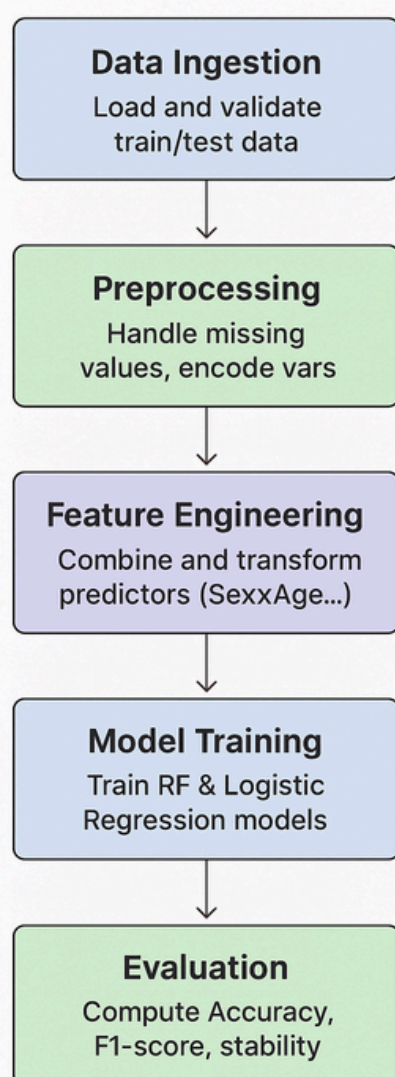
By applying Systems Analysis principles and machine learning algorithms, this project focuses on designing a **robust** and **modular architecture** that improves prediction accuracy and system reliability. It emphasizes understanding **data sensitivity**, **feedback**, and **chaotic variability** as key factors influencing model performance.

## Goal

The **goal** is to **design a modular predictive architecture** that improves the accuracy and stability of the Titanic survival model.

The **challenge** is the instability of existing models caused by **missing data**, **biases**, and **complex relationships**. The proposal integrates **systems analysis** and **machine learning** to achieve a structured and coherent system.

## Proposed Solution



A **modular architecture** with **five** stages is proposed: ingestion, preprocessing, feature engineering, training, and evaluation.

Each module improves the **model's stability** and **accuracy**, while a feedback loop continuously adjusts the system to ensure **reproducible results**.

The suggested implementation is based on **Python**, using Pandas, NumPy, and Scikit-learn under the principles of **modularity**, **sensitivity control**, and **reproducibility**.

In progress...