Workshops of the System Analysis project

Julián Carvajal Garnica

20242020024

Andrés Mauricio Cepeda Villanueva

20242020010

Jhonatan David Moreno Barragan

Andrés Camilo Ramos Rojas

20242020005

Faculty of Engineering, Universidad Distrital Francisco José de Caldas

System Analysis

Carlos Andrés Sierra

Bogotá D.C.

2025

Part 1 of the System Analysis project

1. Competition Overview

The sinking of the Titanic on April 15, 1912, during its maiden voyage, is one of history's most infamous shipwrecks. After hitting an iceberg, the lack of sufficient lifeboats led to 1,502 deaths out of 2,224 people onboard. While luck played a role in survival, some groups had higher chances of surviving than others.

The Kaggle competition *"Titanic: Machine Learning from Disaster"* challenges participants to predict which passengers survived the sinking of the Titanic.

Objectives:

- The main objective of this project is to develop a predictive model capable of determining, based on the provided passenger information, whether an individual aboard the Titanic would survive or not. The prediction is formulated as a binary classification problem, where the output variable takes the value 0 for passengers who did not survive and 1 for those who did.
- To understand how machine learning algorithms operate to derive a concrete and reliable solution to the survival prediction problem.
- To improve the predictive system after obtaining an initial solution by incorporating additional factors and refinements, thereby increasing the overall accuracy and robustness of the model.

Data structure:

- train.csv → passenger features + target Survived.
- test.csv → passenger features without target.
- gender_submission.csv → example of the expected submission format.

Features: Passenger class, sex, age, fare, cabin, and embarkation port, etc.

Target variable: Survived.

Metric: Accuracy (percentage of correct predictions).

Constraints: Submissions must include only PassengerId and Survived.

Example:

PassengerId,Survived
892,0
893,1
894,0
Etc.

2. **System Analysis Report:**

Systemic Analysis: