

Project: Titanic

Introduction. On April 15, 1912, the largest passenger liner ever built collided with an iceberg during its maiden voyage. When the Titanic sank, 1502 of the 2224 passengers and crew were killed. This sensational tragedy shocked the international community and led to improved safety regulations for ships. One of the reasons that the shipwreck resulted in such a loss of life was that there were not enough lifeboats for the passengers and crew. Although there was an element of luck involved in surviving the sinking, some groups of people were more likely to survive than others.

Data. The `titanic.csv` file in the directory `data` contains data for 887 of the real Titanic passengers. Each row represents one passenger. The columns describe attributes about the passengers, including whether they survived (S), their passenger class (C), their gender (G), their age (A), and the fare they paid (F).

Column S encodes survival as 1 and death as 0; column G encodes male as 0 and female as 1. Passenger classes C are 1 (top), 2 (middle), and 3 (bottom).

Task. Write a 1-nearest neighbor classifier using NumPy to predict whether a Titanic passenger survived or not. Follow the instructions in the notebook `titanic.ipynb`.