
ICS 352 Machine Learning

Homework Assignment 1

Due: Monday March 24th, 2025 at 11:30 PM

Upload the lab to D2L into homework assignment, Assignment 1 dropbox

Submission: Screen shots of code working, your code in *.txt files and *.py files (yes both are needed as *.py files do not open under D2L). No zip files.

Titanic Dataset contains information about the passengers aboard the RMS Titanic, which sank after hitting an iceberg on its maiden voyage. The dataset includes various features such as passenger id, passenger class, name, gender, ...

1. Data Exploration:
 - a. Load Titanic dataset into a Pandas DataFrame.
 - b. Show the first rows of the dataset and its summary statistics.
 - c. Identify and handle any missing values in the dataset.
2. Data Visualization:
 - a. Create visualization to show the distribution of key variables (age, passenger class, gender, and fare)
 - b. Generate a pie chart to illustrate the age distribution of survivors to non-survivors.
 - c. Use a histogram to visualize the age distribution of passengers.
3. Data Analysis and Interpretation: (Bayes Theorem)
 - a. Analyze the correlation between passenger class and survival rate.
 - b. Investigate the impact of gender on the survival rate.
 - c. Determine the effect of the number of sibling/spouses and parents/children on the survival rate.
 - d. Explore the relationship between fare paid and survival.