

Q1. This file below contains hiring statistics for a firm such as the candidate's age, experience, gender. Based on these 3 factors, HR will decide the salary. Given this data, you need to build a machine learning model for the HR department that can help them decide salaries for future candidates.

Also build a ML model that predicts candidates to greater than or less than the median salary given their age, gender and experience.

Dataset

https://drive.google.com/file/d/1SLFDYW559yHt8L4j1_EecV44IEU4hlQ8/view?usp=sharing

Q2. Assess the Titanic Disaster Survival on the dataset given below.

Import Data & Python Packages

Assess Data Quality & Missing Values

2.1 Age - Missing Values

2.2 Cabin - Missing Values

2.3 Embarked - Missing Values

2.4 Final Adjustments to Data

2.4.1 Additional Variables

Exploratory Data Analysis

Logistic Regression

4.1 Hold-Out Testing & Logistic Model Assessment

4.2 Kaggle "Test" Dataset

4.3 Re-run Logistic Regression w/ 80-20 Split

4.4 Out-of-sample test results

4.5 Logistic Regression Conclusions

Dataset folder:

https://drive.google.com/drive/folders/12IB_IgxkkiSPHcFudnY1x2xAxMPGjyCc?usp=sharing