|  |  |
| --- | --- |
| Name: Rahul Thambi | Roll Number: A059 |
| Subject: DWM | Project: #1 |

**Problem Description**: To create a model that predicts which passengers survived the Titanic shipwreck. The sinking of the Titanic is one of the most infamous shipwrecks in history. On April 15, 1912, during her maiden voyage, the widely considered “unsinkable” RMS Titanic sank after colliding with an iceberg. Unfortunately, there weren’t enough lifeboats for everyone onboard, resulting in the death of 1502 out of 2224 passengers and crew. While there was some element of luck involved in surviving, it seems some groups of people were more likely to survive than others. In this challenge, we ask you to build a predictive model that answers the question: “what sorts of people were more likely to survive?” using passenger data (ie name, age, gender, socio-economic class, etc).

**DataSet:**

Train.csv will contain the details of a subset of the passengers on board (891 to be exact) and importantly, will reveal whether they survived or not, also known as the “ground truth”.

The test.csv dataset contains similar information but does not disclose the “ground truth” for each passenger.Using the patterns we find in the train.csv data, predict whether the other 418 passengers on board (found in test.csv) survived.

|  |  |  |
| --- | --- | --- |
|  | Attribute | DataType |
| 1 | Passenger | Int |
| 2 | Survived | Int |
| 3 | Pclass | Int |
| 4 | Name | String |
| 5 | Sex | String |
| 6 | Age | Int |
| 7 | SibSp | Int |
| 8 | Parch | Int |
| 9 | Ticket | String |
| 10 | Fare | Float |
| 11 | Cabin | String |
| 12 | Embarked | Char |