Goal

It is your job to predict if a person won the lottery or not.   
For each PassengerId in the test set, you must predict a 0 or 1 value for the Survived variable.

Metric

Your score is the percentage of passengers you correctly predict.

### Overview

The data has been split into two groups:

* training set (train.csv)
* test set (test.csv)

The training set should be used to build your machine learning models. For the training set, we provide the outcome (also known as the “ground truth”) for each passenger. Your model will be based on “features” like passengers’ gender and class.

The test set should be used to see how well your model performs on unseen data. For the test set, we do not provide the ground truth for each passenger. It is your job to predict these outcomes. For each passenger in the test set, use the model you trained to predict whether or not they won the lottery.

**Data Dictionary**

|  |  |  |
| --- | --- | --- |
| Variable | Definition | Key |
| survival | Lottery Win | 0 = No, 1 = Yes |
| pclass | Ticket class | 1 = 1st, 2 = 2nd, 3 = 3rd |
| sex | Sex |  |
| Age | Age in years |  |
| sibsp | # of siblings / spouses aboard the Titanic |  |
| parch | # of parents / children aboard the Titanic |  |
| ticket | Ticket number |  |
| fare | Passenger fare |  |
| cabin | Cabin number |  |
| embarked | City | C = Cherbourg, Q = Queenstown, S = Southampton |