# Titanic: Machine Learning from Disaster

The sinking of the RMS Titanic is one of the most infamous shipwrecks in history.  On April 15, 1912, during her maiden voyage, the Titanic sank after colliding with an iceberg, killing 1502 out of 2224 passengers and crew. This sensational tragedy shocked the international community and led to better safety regulations for ships.

One of the reasons that the shipwreck led to such loss of life was that there were not enough lifeboats for the passengers and crew. Although there was some element of luck involved in surviving the sinking, some groups of people were more likely to survive than others, such as women, children, and the upper-class.

In this challenge, we ask you to complete the analysis of what sorts of people were likely to survive. In particular, we ask you to apply the tools of machine learning to predict which passengers survived the tragedy.

**Overview**

The sinking of the Titanic is one of the most infamous shipwrecks in history. On April 15, 1912, during her maiden voyage, the Titanic sank after colliding with an iceberg, killing 1502 out of 2224 passengers and crew. That's why the name ****DieTanic****. This is a very unforgetable disaster that no one in the world can forget.

It took about $7.5 million to build the Titanic and it sunk under the ocean due to collision. The Titanic Dataset is a very good dataset for begineers to start a journey in data science and participate in competitions in Kaggle.

The Objective of this notebook is to give an ****idea how is the workflow in any predictive modeling problem****. How do we check features, how do we add new features and some Machine Learning Concepts. I have tried to keep the notebook as basic as possible so that even newbies can understand every phase of it.

If You Like the notebook and think that it helped you..****PLEASE UPVOTE****. It will keep me motivated.

## Contents of the Notebook:

#### Part1: Exploratory Data Analysis(EDA):

1)Analysis of the features.

2)Finding any relations or trends considering multiple features.

#### Part2: Feature Engineering and Data Cleaning:

1)Adding any few features.

2)Removing redundant features.

3)Converting features into suitable form for modeling.

#### Part3: Predictive Modeling

1)Running Basic Algorithms.

2)Cross Validation.

3)Ensembling.

4)Important Features Extraction.

# Predictive Modeling[¶](https://www.kaggleusercontent.com/kf/2739619/eyJhbGciOiJkaXIiLCJlbmMiOiJBMTI4Q0JDLUhTMjU2In0..7nAM4fNow1v08x1zwReRTQ.JOR6091ctQ2JWXhJwaUmleEQjdPcidwQSsEdbNHUsLM65WpgRGtsYOgeDWctomdHHVC1sPRZimE6z4iL3e66fNHnjk_cIVY1h3rHiSd4sH2OuYMLWFvIM7ICS-zAgYkMpfSTBB9voNN3cTNfwZl0BQ.p2_75Dh_QplCBR8MmqKYzw/__results__.html" \l "Part3:-Predictive-Modeling" \t "https://www.kaggleusercontent.com/kf/2739619/eyJhbGciOiJkaXIiLCJlbmMiOiJBMTI4Q0JDLUhTMjU2In0..7nAM4fNow1v08x1zwReRTQ.JOR6091ctQ2JWXhJwaUmleEQjdPcidwQSsEdbNHUsLM65WpgRGtsYOgeDWctomdHHVC1sPRZimE6z4iL3e66fNHnjk_cIVY1h3rHiSd4sH2OuYMLWFvIM7ICS-zAgYkMpfSTBB9voNN3cTNfwZl0BQ.p2_75Dh_QplCBR8MmqKYzw/_self)

After gaining insights from the EDA part. But with that, we cannot accurately predict or tell whether a passenger will survive or die. So now we will predict the whether the Passenger will survive or not using some great Classification Algorithms.Following are the algorithms I will use to make the model:

1)Logistic Regression

2)Support Vector Machines(Linear and radial)

3)Random Forest

4)K-Nearest Neighbours

5)Naive Bayes

6)Decision Tree

7)Logistic Regression