**题目：Titanic生存率预测项目**

1. 背景介绍

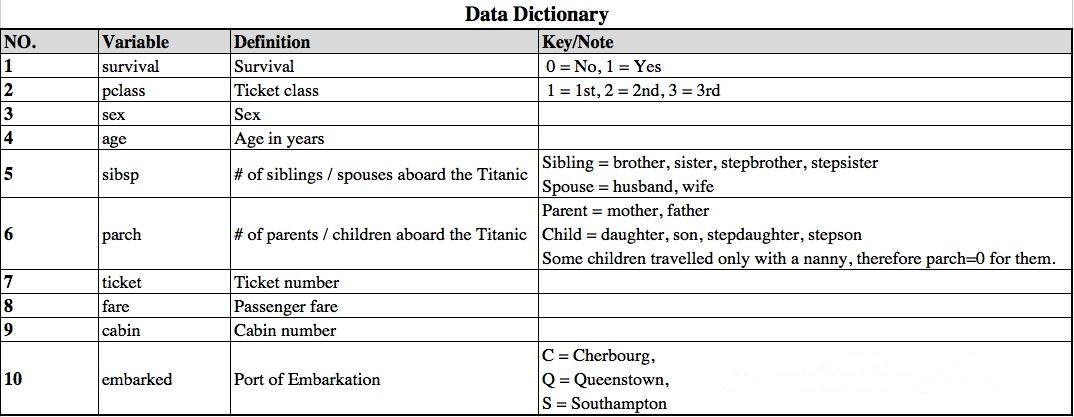
泰坦尼克号于1909年3月31日在爱尔兰动工建造，1911年5月31日下水，次年4月2日完工试航。她是当时世界上体积最庞大、内部设施最豪华的客运轮船，有“永不沉没”的美誉。然而讽刺的是，泰坦尼克号首航便遭遇厄运：1912年4月10日她从英国南安普顿出发，途径法国瑟堡和爱尔兰昆士敦，驶向美国纽约。在14日晚23时40分左右，泰坦尼克号与一座冰山相撞，导致船体裂缝进水。次日凌晨2时20分左右，泰坦尼克号断为两截后沉入大西洋，其搭载的2224名船员及乘客，在本次海难中逾1500人丧生。

在该案例中，我们将探究什么样的人在此次海难中幸存的几率更高，并通过构建预测模型来预测乘客生存率。

1. 数据

数据来源：根据blackboard上面提供的train.csv和test.csv文件。

字段说明：



1. 实验要求
2. 对缺失值和异常值进行处理；
3. 用Gradient Boosting Classifier和Logistic Regression模型进行建模；
4. 对比分析实验结果（评价指标包含Accuracy、Recall以及F1-score等）。
5. 实验内容及结果
6. 实验分析与总结