

## 实验四 集成学习二

### 一. 简单介绍 GBDT 和 XGBoost 算法的原理

### 二. Titanic 数据实验

1.对 Titanic 数据进行探索性分析

2.数据预处理：缺失值处理、哑变量处理、分箱操作等

3.特征工程：特征提取、特征生成等

4.以 Precision、Recall、F1 和 Auc 等指标为评价标准，利用网格搜索等方法找到合适的模型参数，探究并比较 AdaBoost、GBDT、XGBoost 和 Random Forest 集成算法的性能。