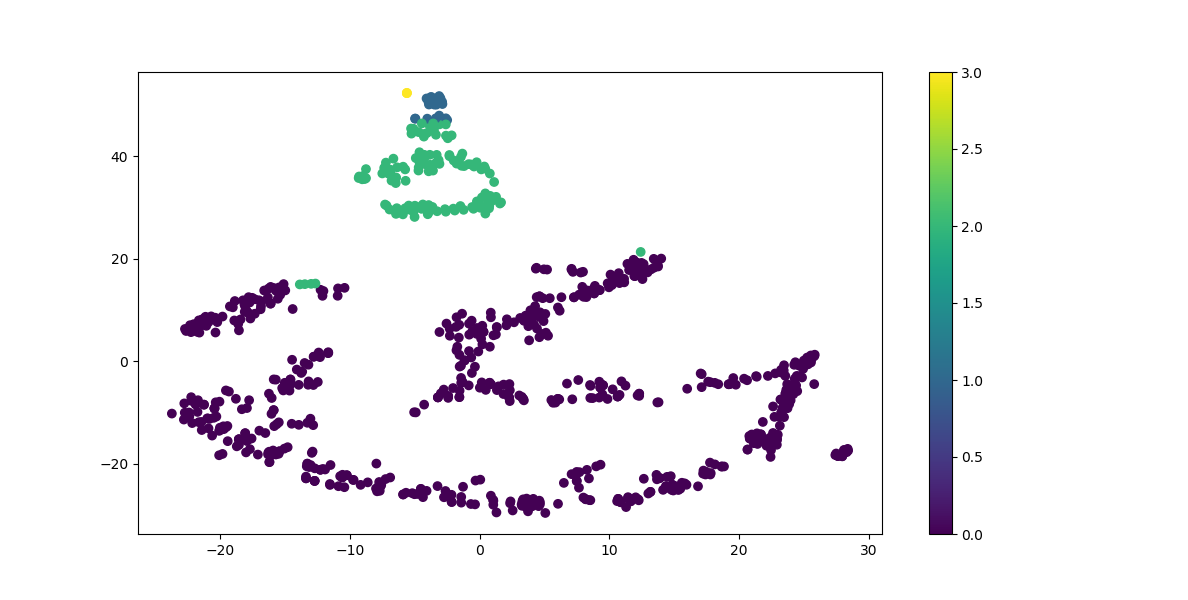
数据挖掘结果分析作业三

王钰

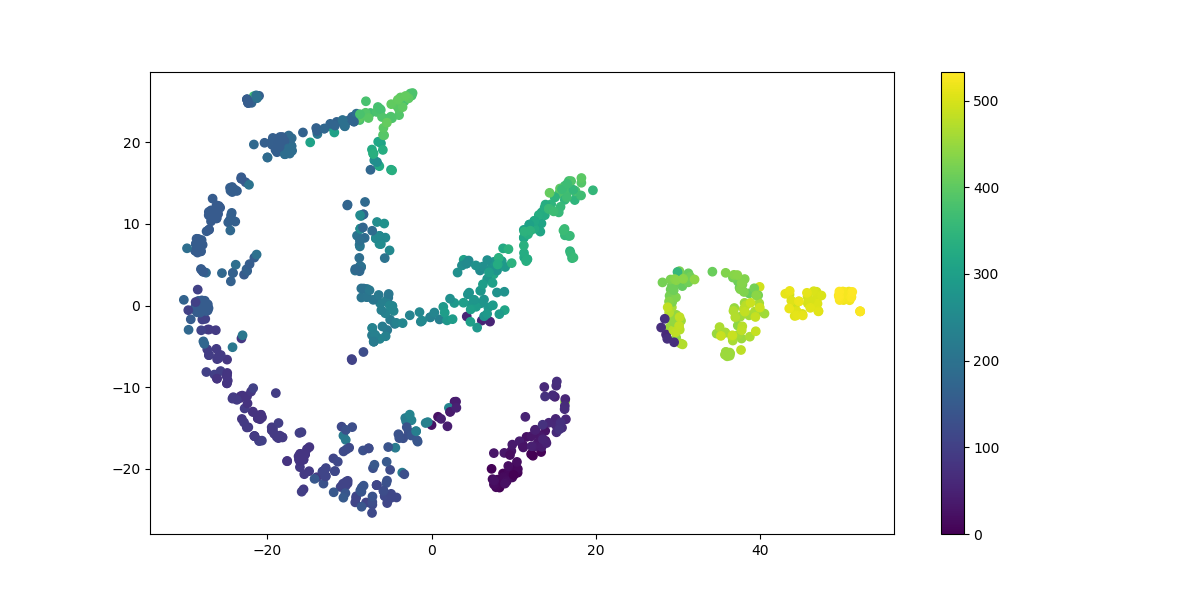
2120171076

1. **聚类结果**

**（1）k-means算法**

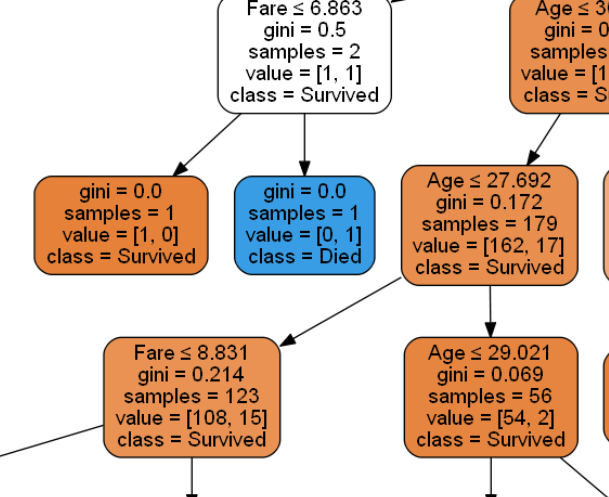
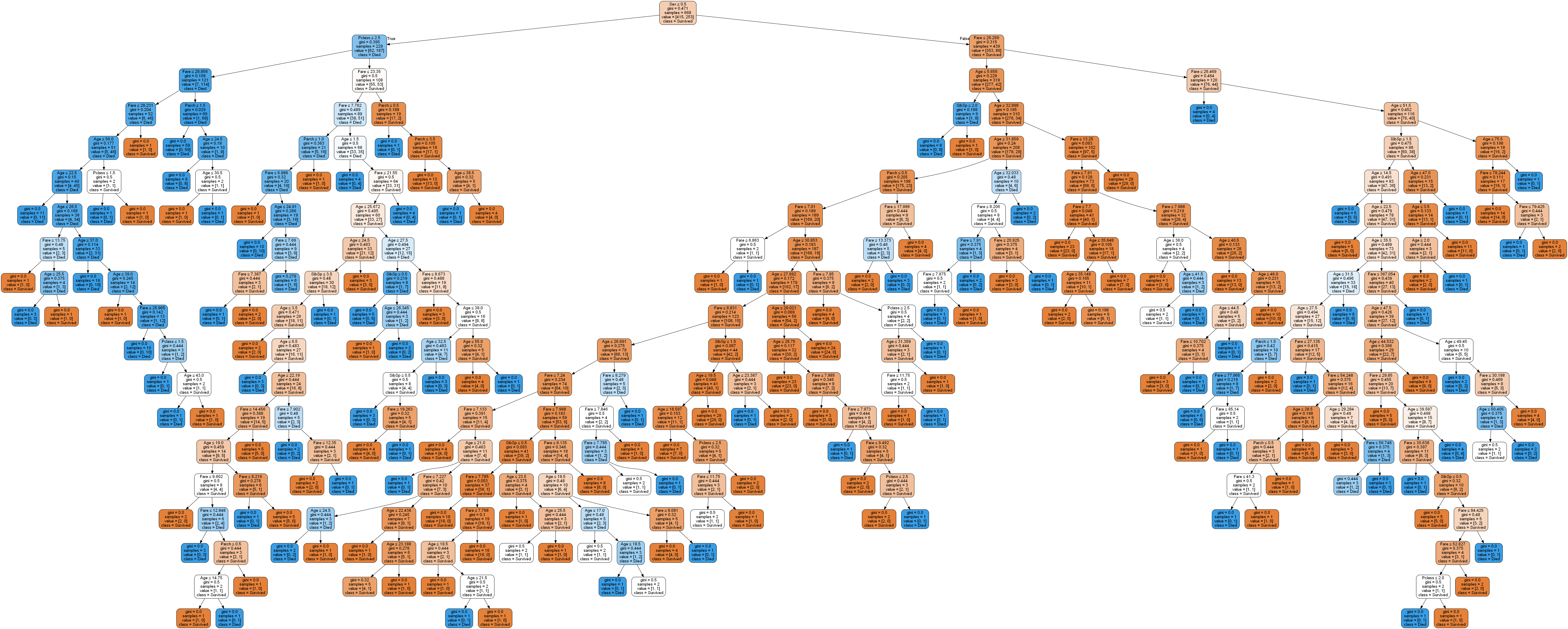


**（2）birch算法**



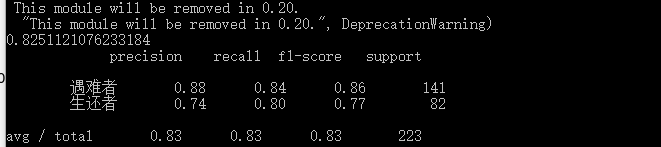
1. 分类模型

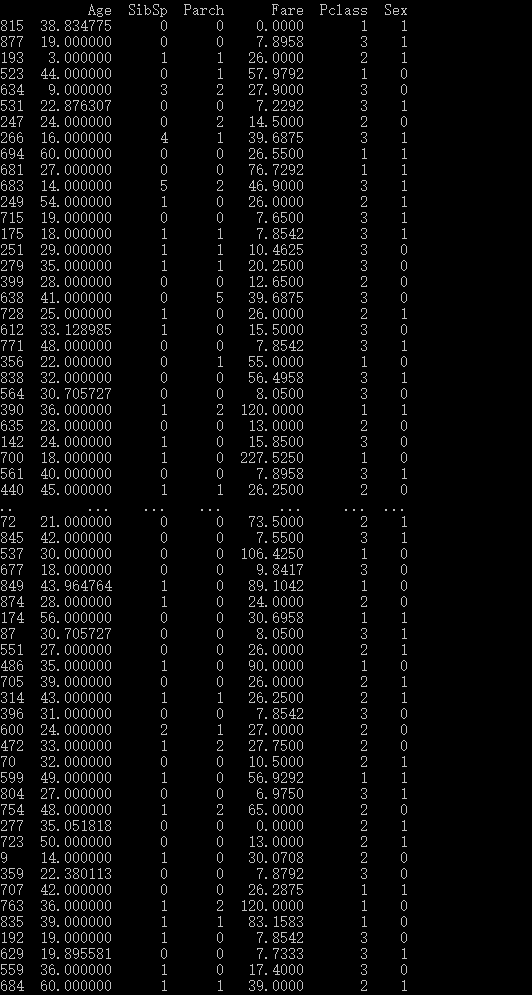
将决策树可视化，保存为png格式：

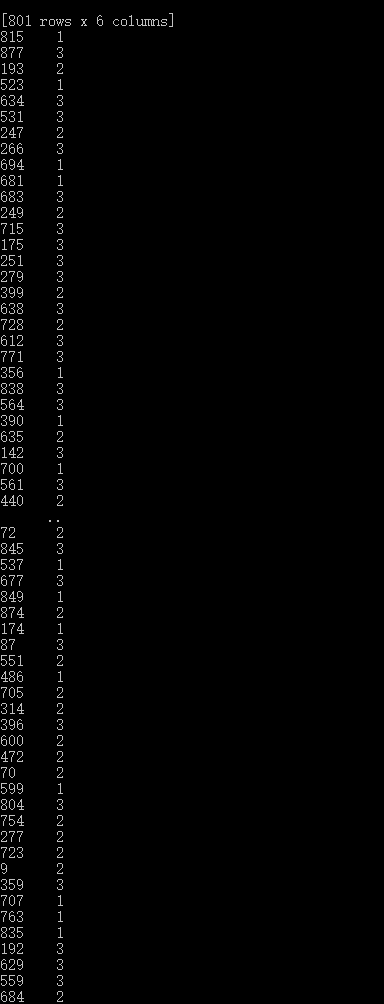


黄色系方块表示存活者，蓝色系方块表示遇难模块，颜色深浅取决于样本个数多少。

1. 结果分析

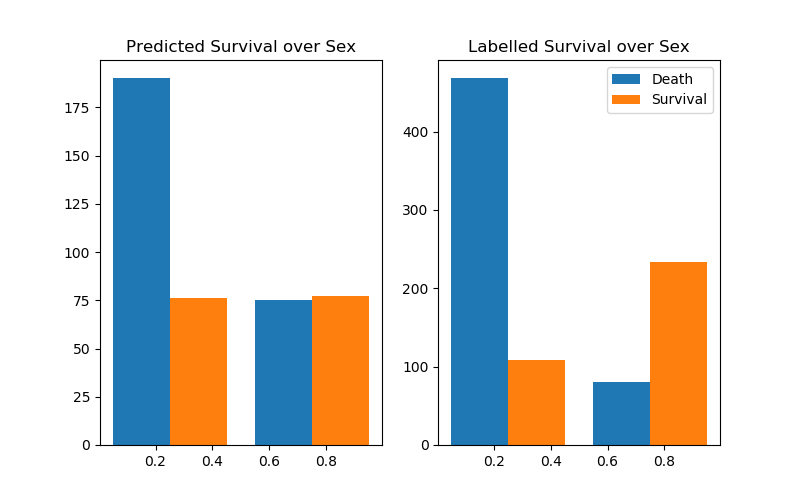


****

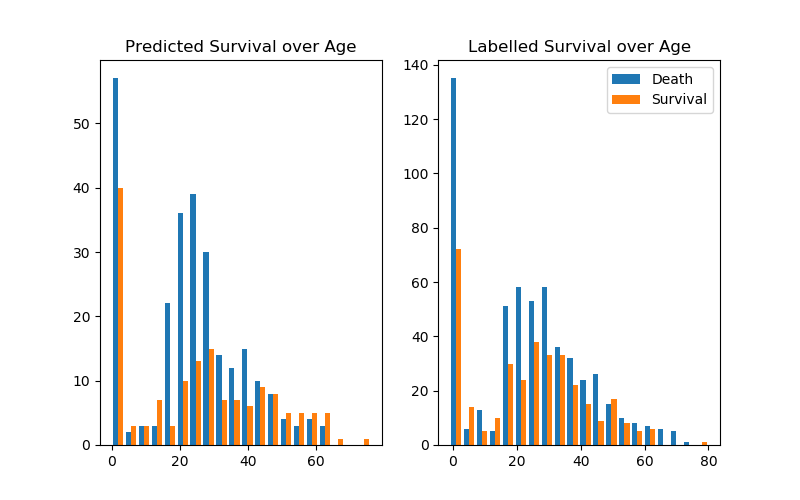
****

从各种聚类和分类模型中可以看出：

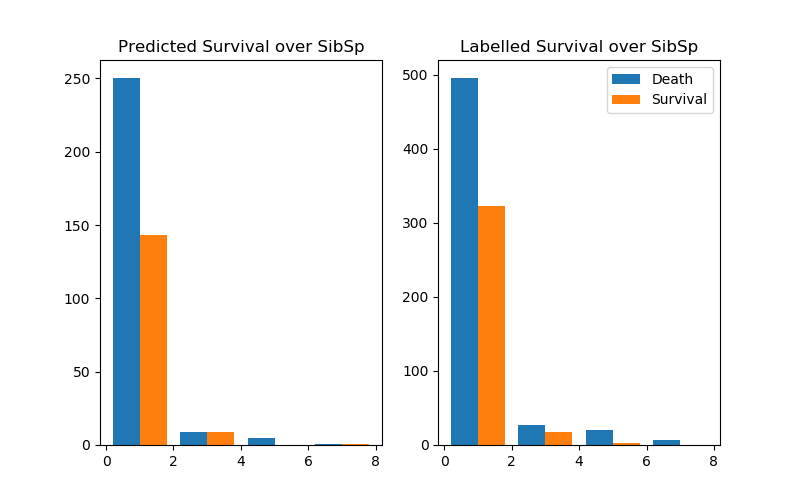
1. Sex女性的幸存率远远大于男性



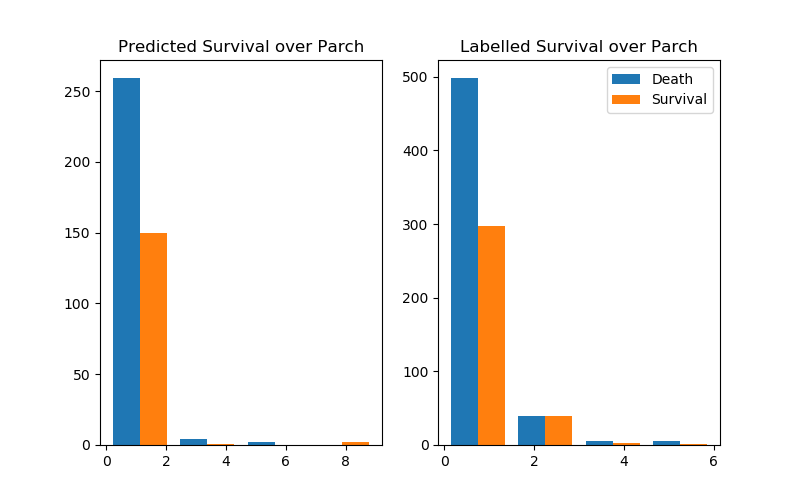
1. Age未成年人的幸存率要远远高于青壮年



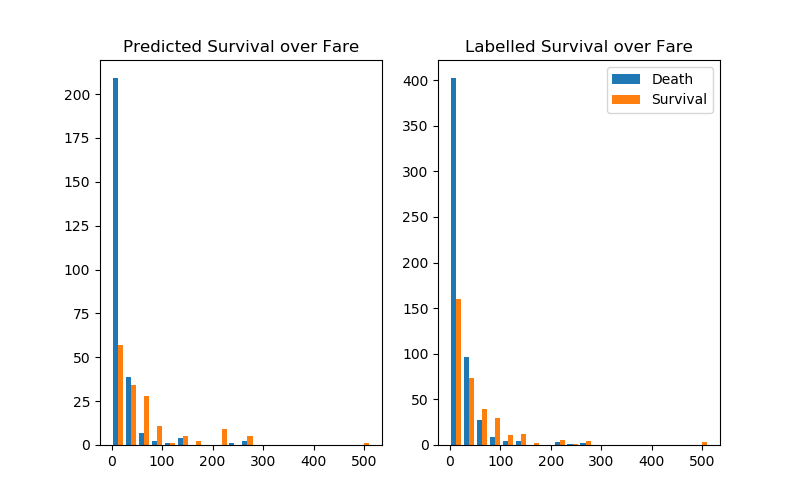
1. SibSp为1或2的乘客幸存率最高



1. Parch为1到3的乘客幸存率最高



1. Fare船票越贵的乘客的存活率越高



1. Pclass: 购买一级船票的存活率最高，购买三级船票的死亡率最高

