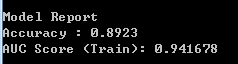
Age：

缺失数177，未缺失714

Estimator = 37



XGBClassifier(reg\_alpha=1e-5,learning\_rate = 0.05,n\_estimators=73,max\_depth =3, min\_child\_weight=1,gamma=0.4,subsample=0.8,colsample\_bytree=0.8,

objective= 'binary:logistic',nthread=4,scale\_pos\_weight=1,seed=27)

将特征重要性等于0的去除，xgboost尽量选择小的学习率，即使无用的特征也可能对分数起作用

采样未标签0.55 0.8022

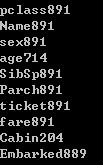
0.44 0.8035

加入半监督学习，即采样0.44的测试集预测后再合并训练预测测试集，有一点提高

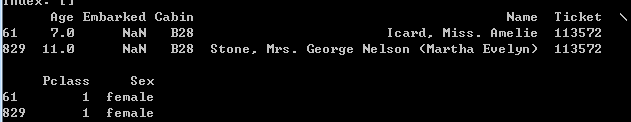
Feature去除无用特征分数提高

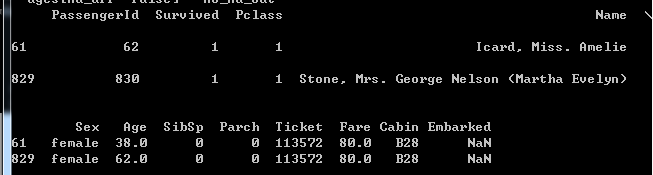
5个基模型，xgboost第二层5个特征，采用extraclassifer ,0.78947

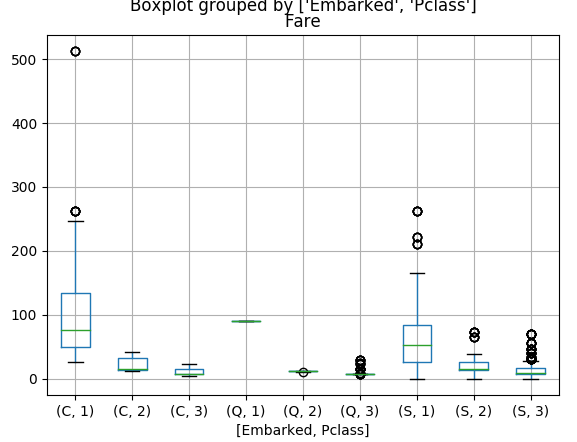
同上的模型，第二层采用原始特征加5个特征，0.79425



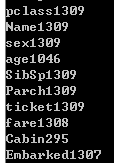
Embark缺失







填充C



Svm 0.784，归一化，填充一些缺失值