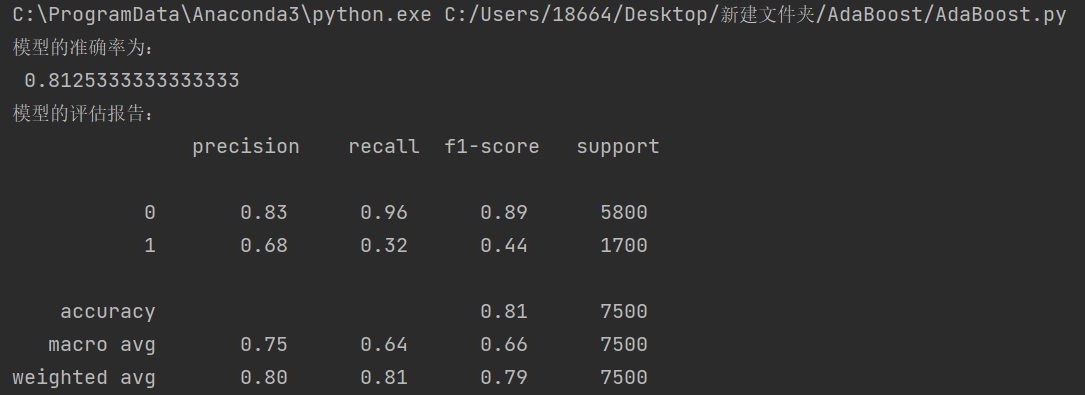
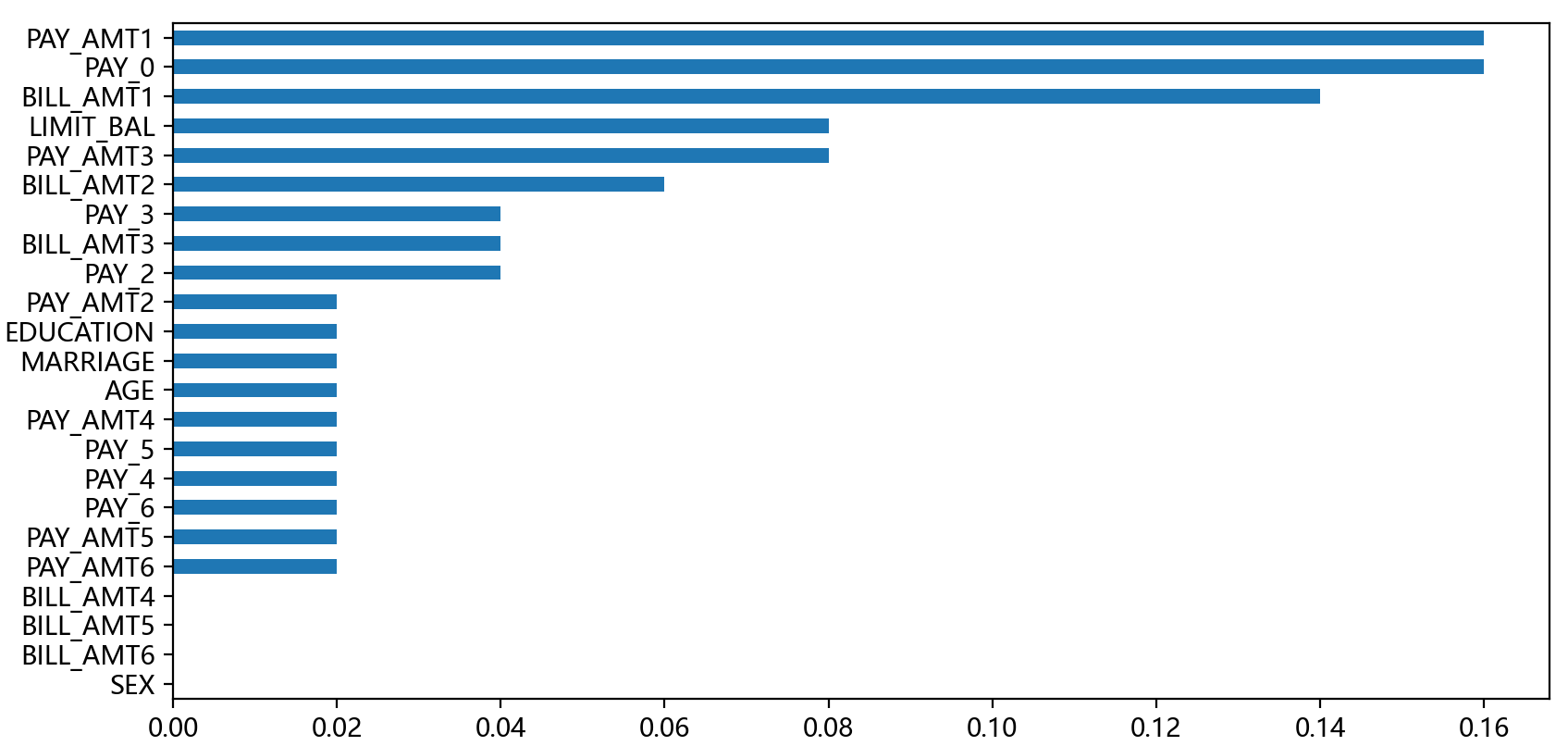
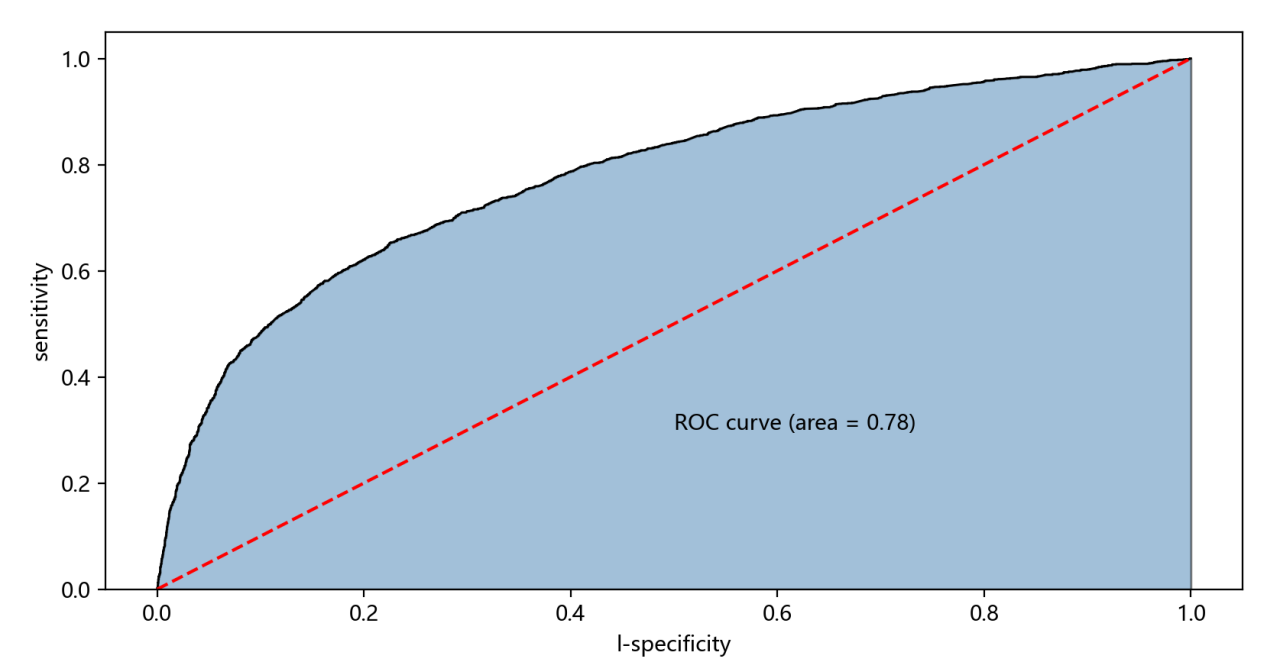
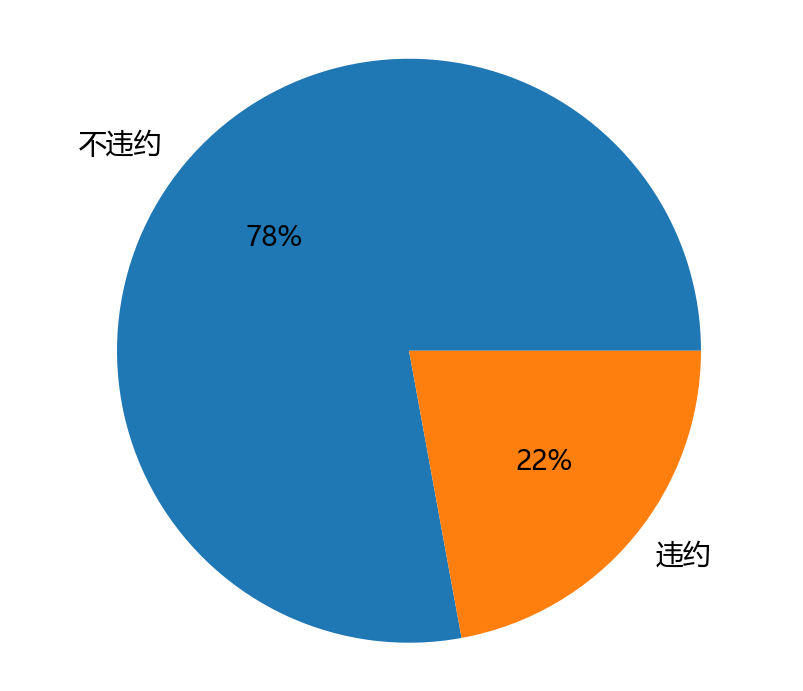
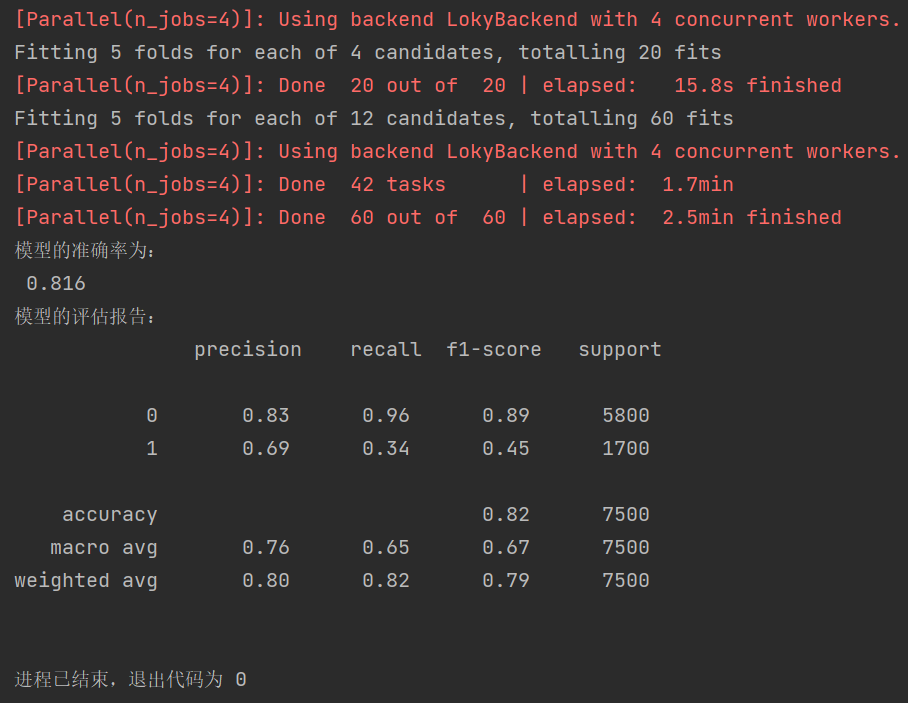
运行结果：



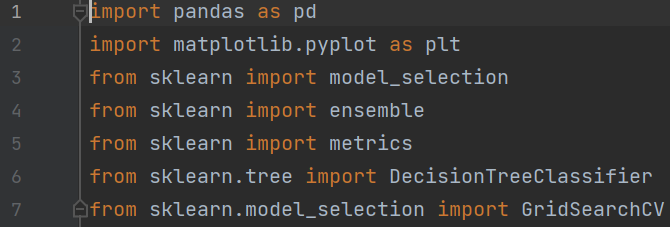




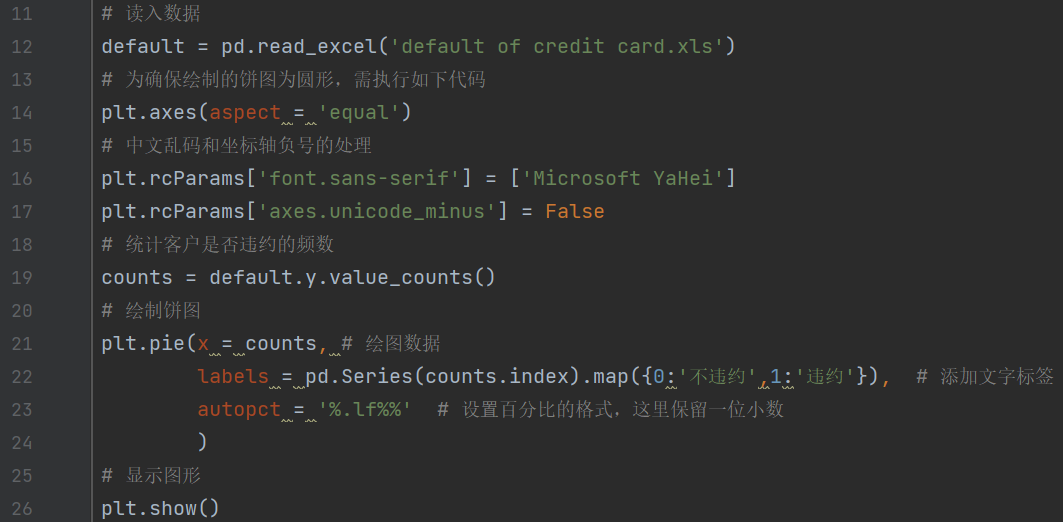
以信用卡违约数据为例，该数据集来源于UCI网站，一共包30 000条记录和25个变量，其中自变量包含客户的性别、受教育水平、年龄、婚姻状况、信 用额度、6个月的历史还款状态、账单金额以及还款金额，因变量y表示用户在下个月的信用卡还款中是否存在违约的情况（1表示违约，0表示不违约）。

Python代码如下：

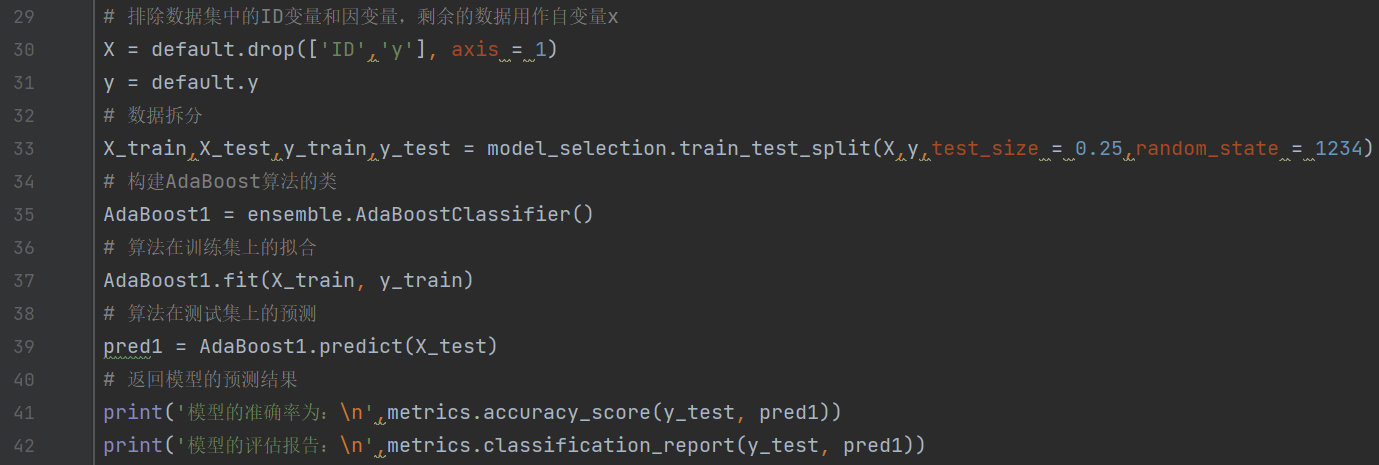
导入python自带包以及第三方包



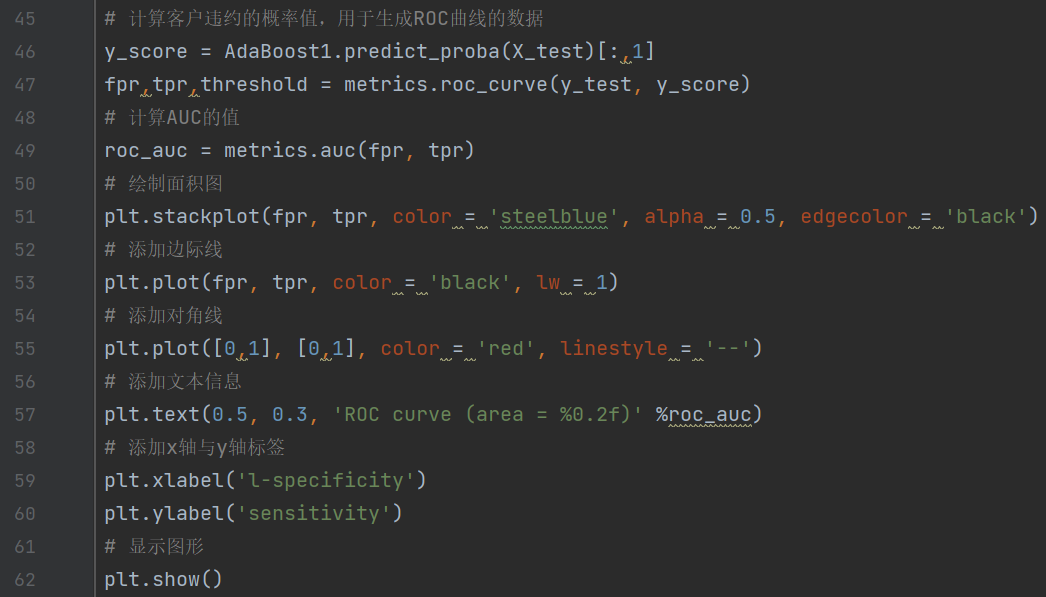
绘图查看训练集标签类别是否不平衡



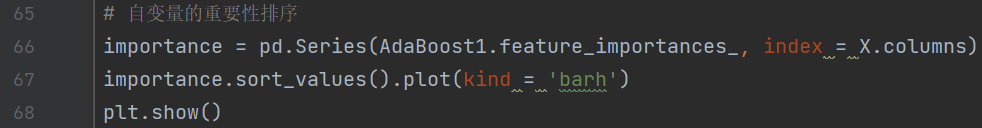
利用AdaBoost模型进行预测，输出模型评估报告



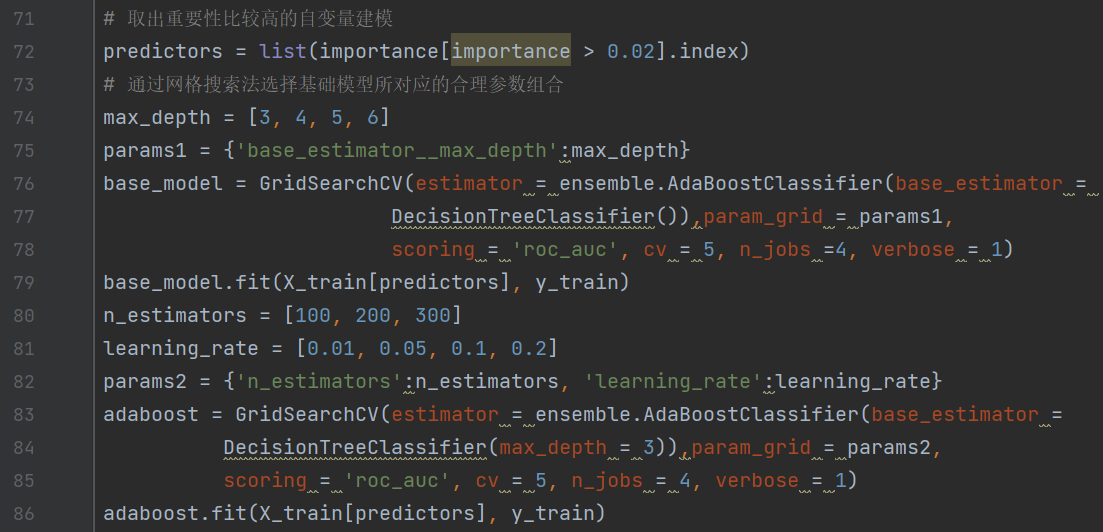
利用ROC曲线对模型进行评估



查看自变量的重要性排序



利用网格搜索寻找AdaBoost模型最优参数



利用优化后的AdaBoost模型进行预测，输出模型评估报告

