# 基于逻辑回归模型的心脏病预测

1. **设计背景**

在机器学习算法中，二分类问题一直是一个经典的研究课题。因此，本次作业中，使用了kaggle网站上的Heart disease数据集进行是否患心脏病的分类预测。Heart disease是一个比较经典的数据集，常出现在一些论文和实验验证中。本次使用到的数据集较小，只有304条数据，每条数据一共有13个特征值，包括性别，年龄等等。在程序将使用前12个特征值进行逻辑回归模型的训练，来帮助我们预测是否患心脏病。

**二**．**环境搭建**

1)操作系统: windows 10

2)编程语言: python 3.8.3

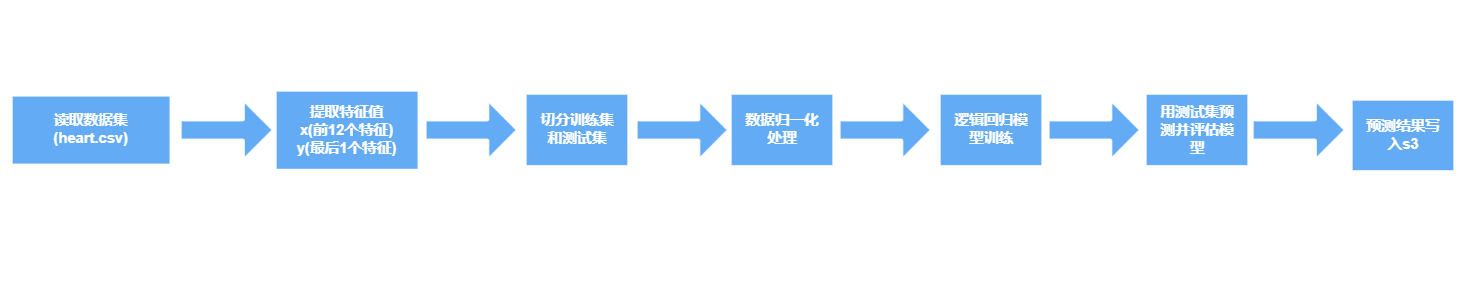
3)集成开发环境: spyder4(anaconda3)

4)服务器客户端接口: S3 Browser

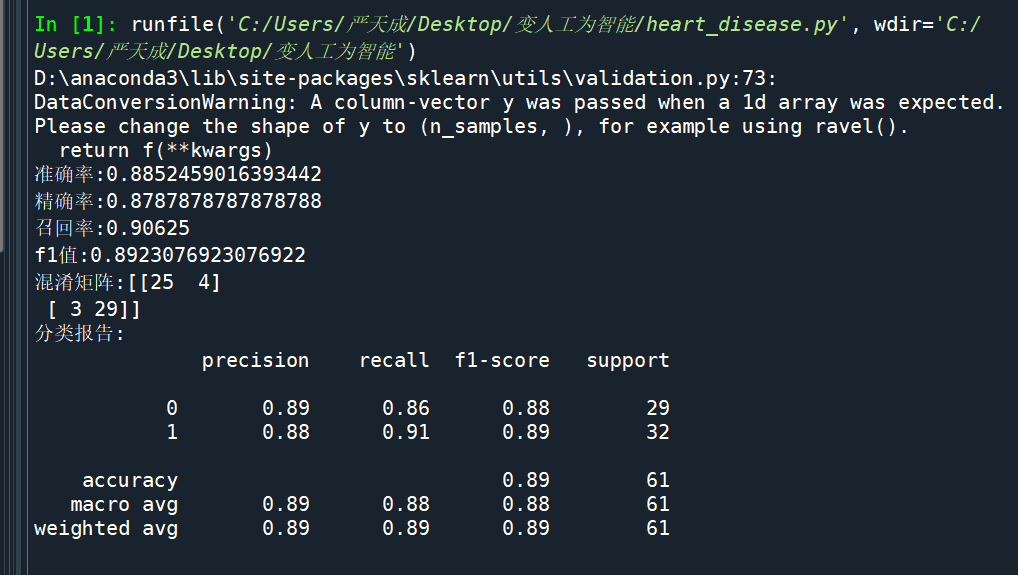
5)python第三方库:sklearn，boto， pandas

6)数据集:Kaggel网站 heart disease数据集

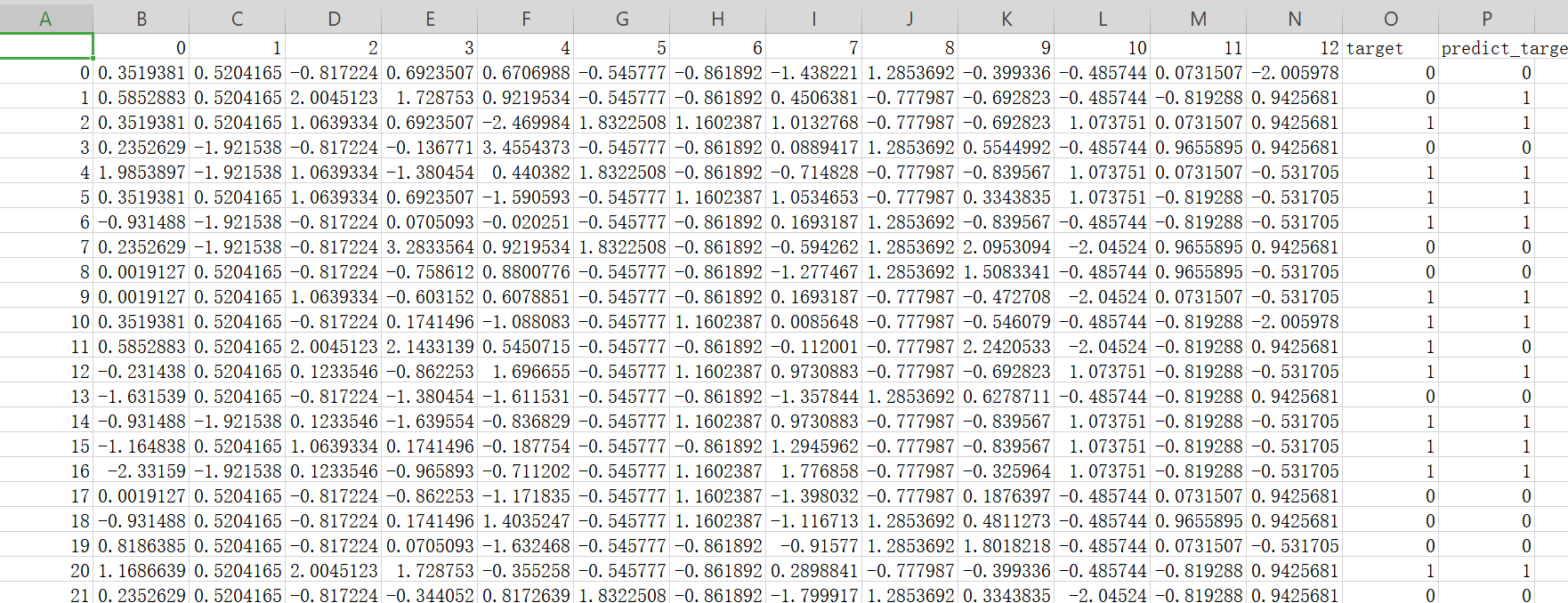
**三．设计流程图**



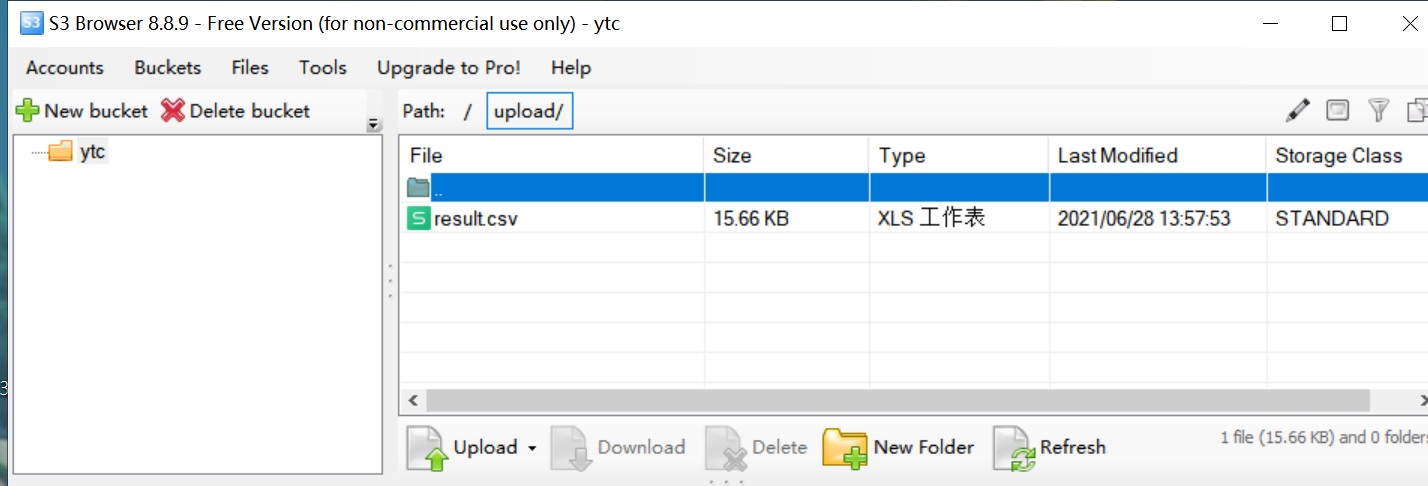
1. **程序使用**
2. 安装anaconda3，下载需要的依赖库
3. 使用spyder4打开heart\_disease.py文件并运行
4. 查看s3桶上/upload/result.csv
5. **运行结果**



训练好的模型在测试集上做评估



查看本地保存的预测结果



在s3上查看已写入的预测结果