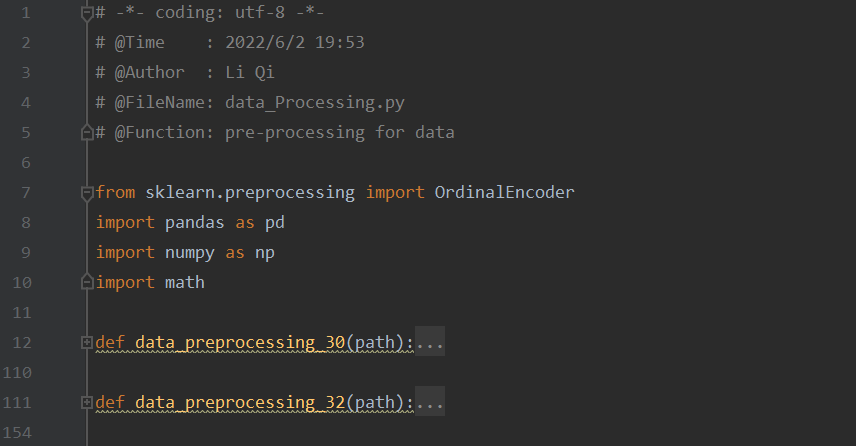
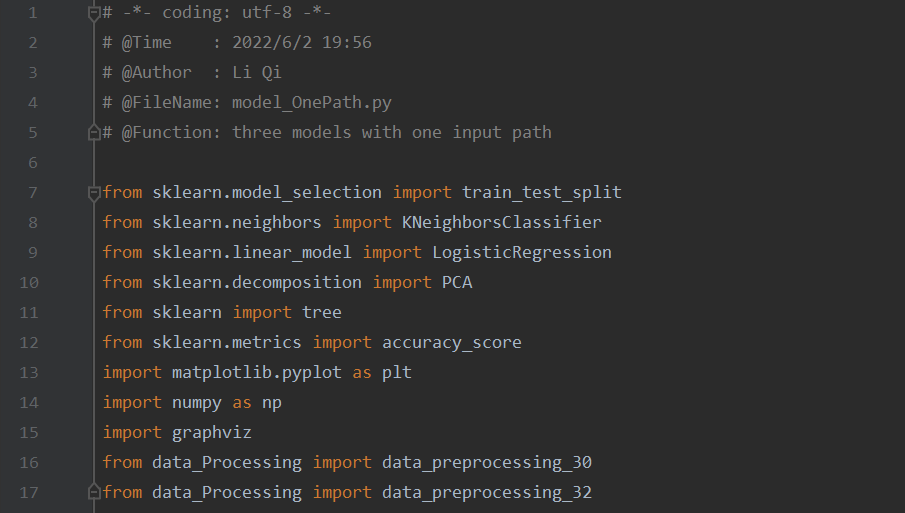
本次工程2在python下实现了knn，逻辑回归和决策树，并且将数据集进行预处理，在30个特征和32个特征的条件下分别用三个算法进行验证，得到了实验结果。总代码量一千行左右。

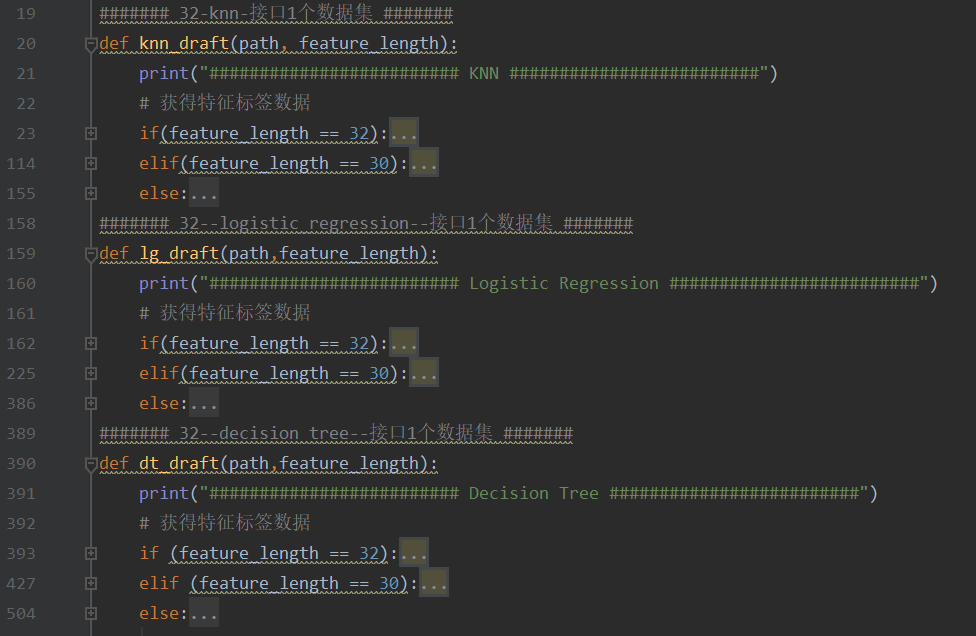
**1）data\_Processing.py：**



在该文件中实现了数据的预处理，具体来说，先分别提取数据集中的特征和标签，再将数据根据标签的大小排序（data\_preprocessing\_32()中根据G3排序，data\_preprocessing\_30()中分别根据G1、G2、G3排序），之后将数据按成绩排序的10%-20%-40%-20%-10%划分为5类，根据类别对G1、G2、G3重新赋值。

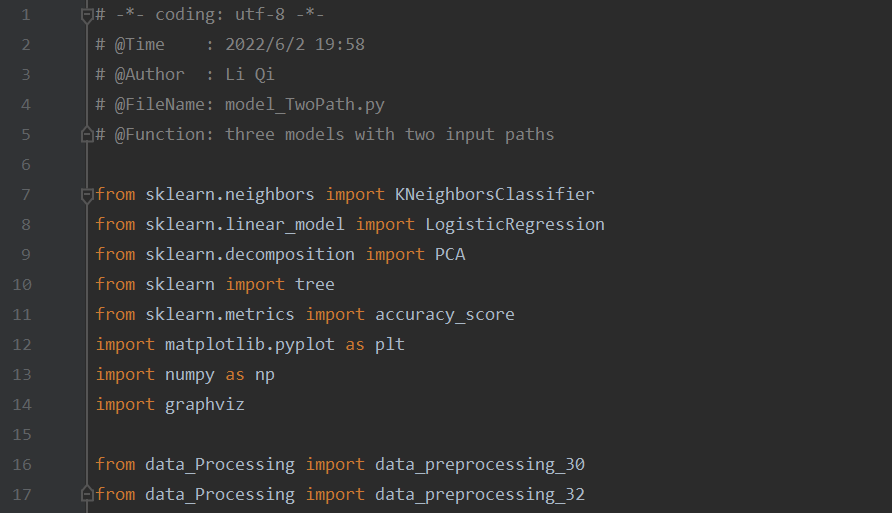
**2)model\_OnePath.py:**

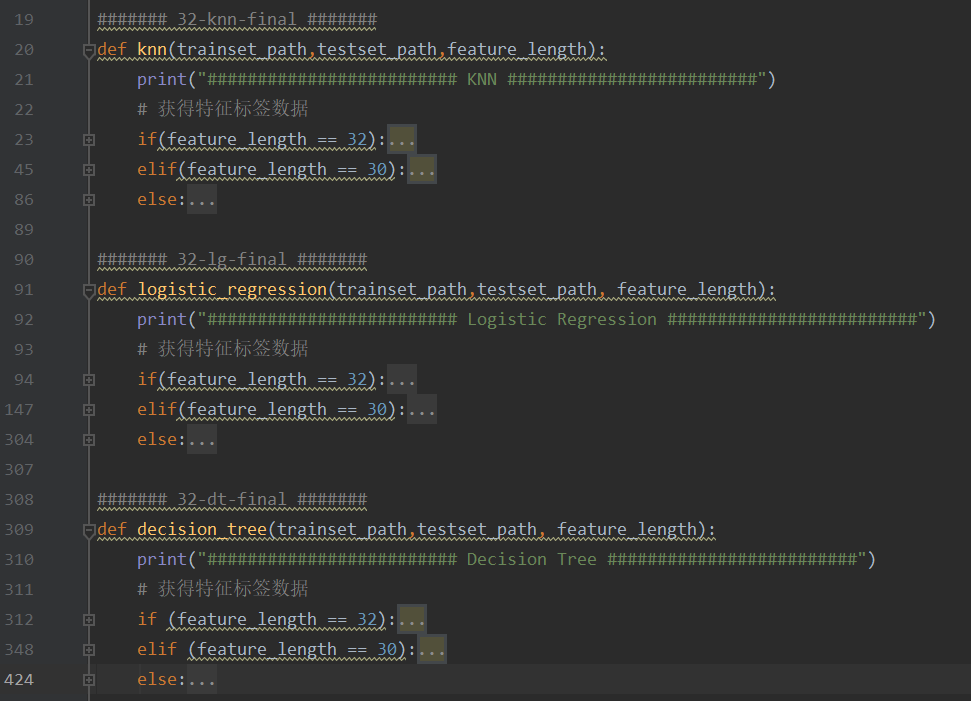




在该文件中实现了接口为1个数据集的三个模型，三个模型分别有两个参数，path用来传输数据集的存放路径，feature\_length用来传输数据的特征长度（除30和32外的其他值均非法）。

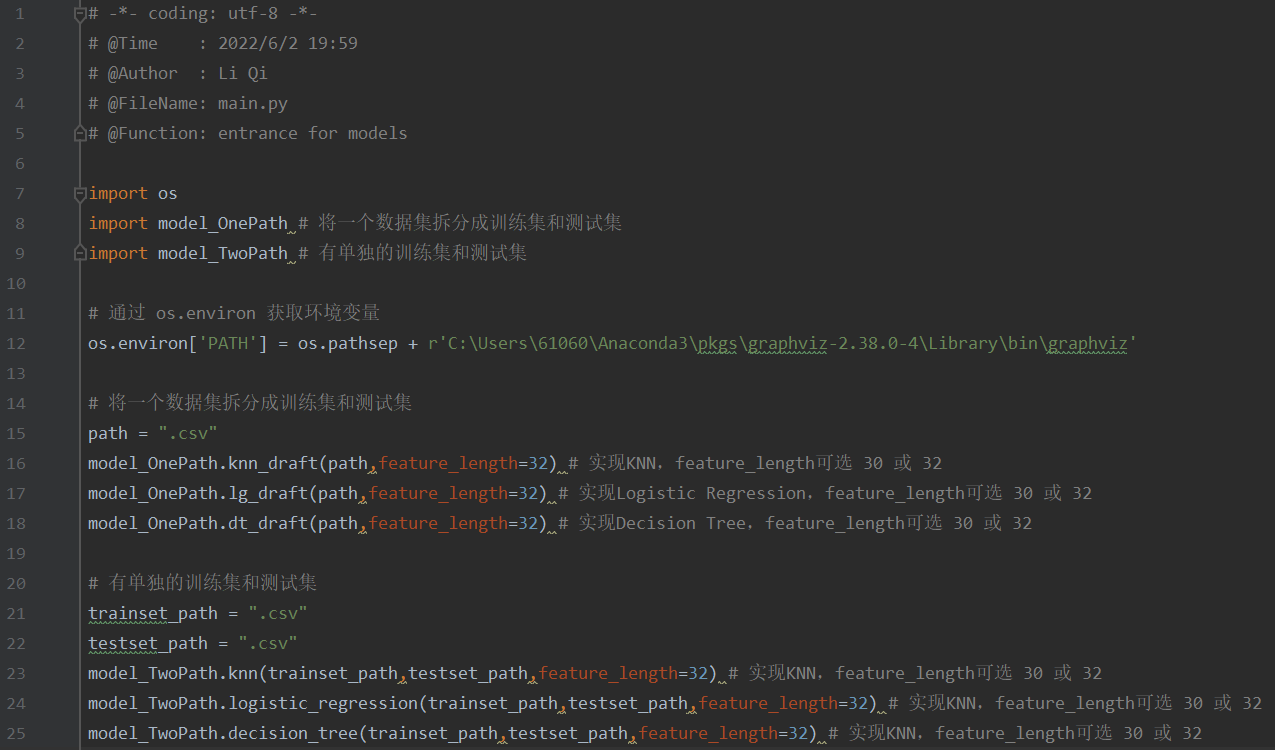
**3)model\_TwoPath.py:**





在该文件中实现了接口为2个数据集的三个模型，三个模型分别有三个参数，trainset\_path用来传输训练集的存放路径，testset\_path用来传输测试集的存放路径，feature\_length用来传输数据的特征长度（除30和32外的其他值均非法）。

**4)main.py:**



模型训练的入口，在这里导入自己编写的model\_OnePath、model\_TwoPath，调用其中的函数对模型进行训练，可以直接运行。

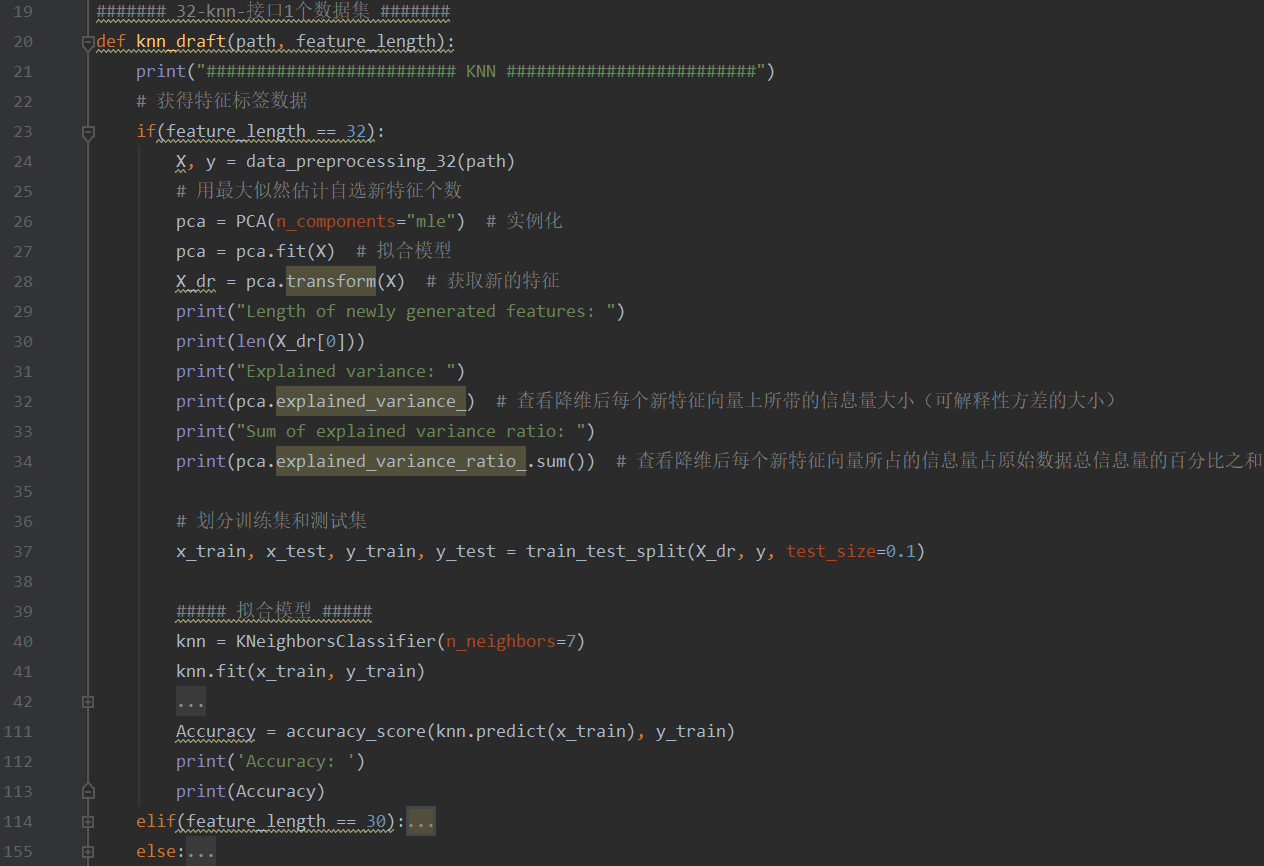
**5)whole\_work\_combine.py:**

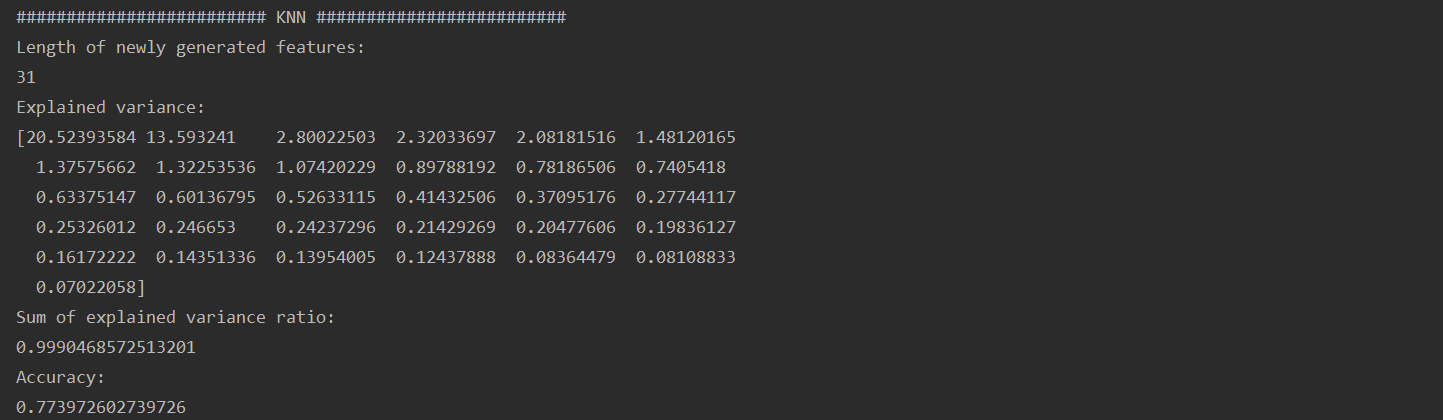
将以上所有内容整合到一个文件里，可以直接运行。

**具体实现及运行结果如下：**

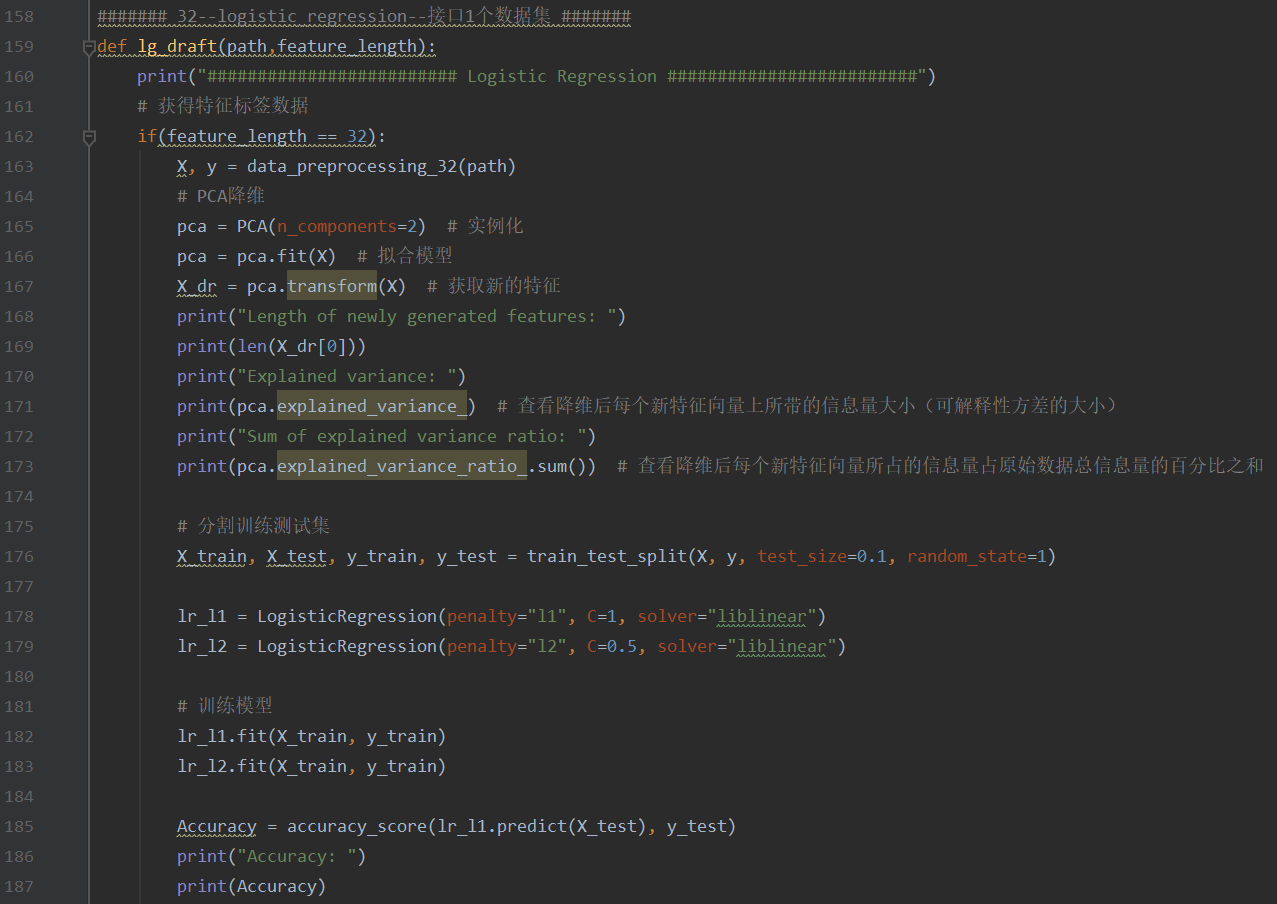
**1）32特征预测G3：**

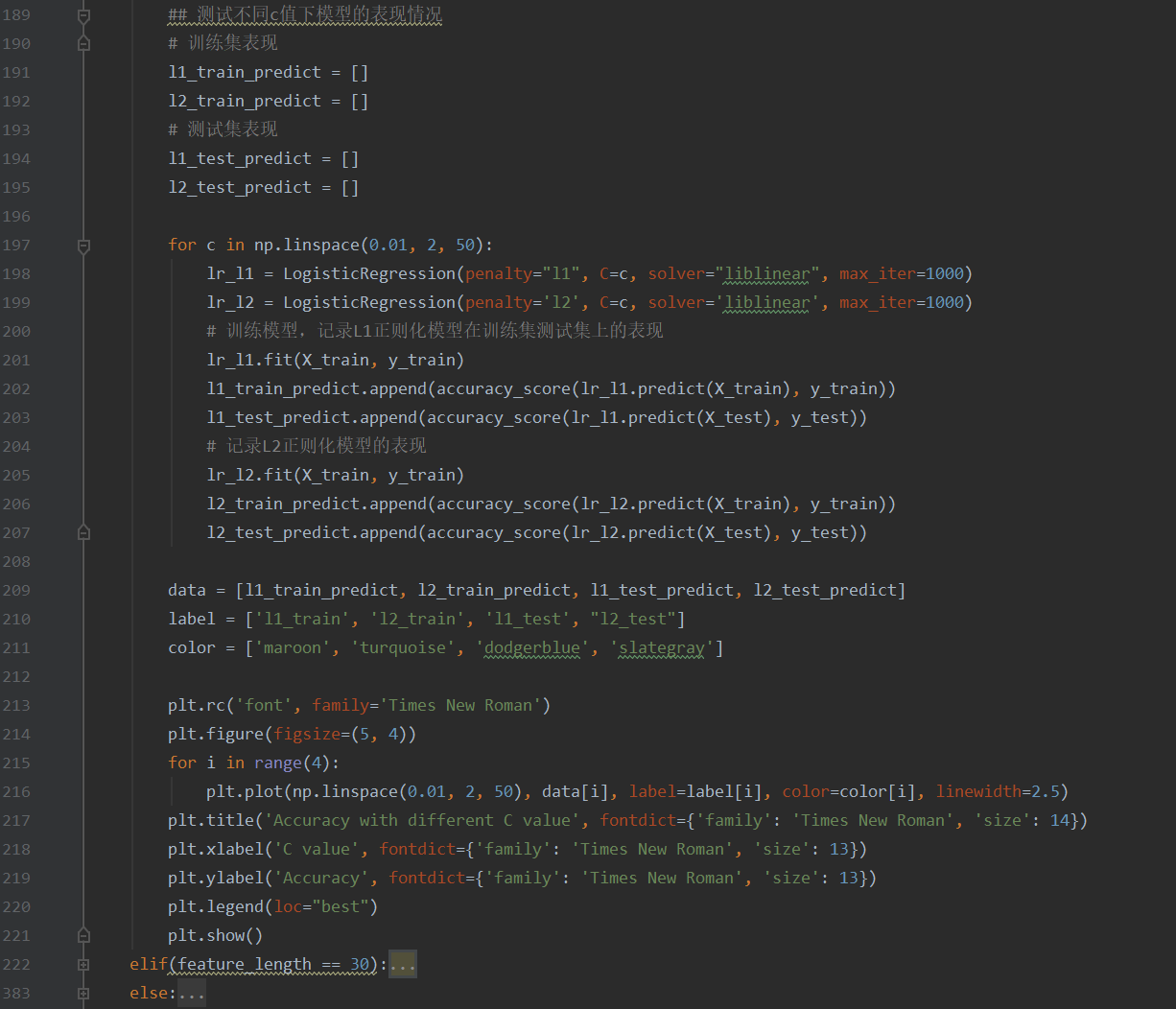
**KNN：**

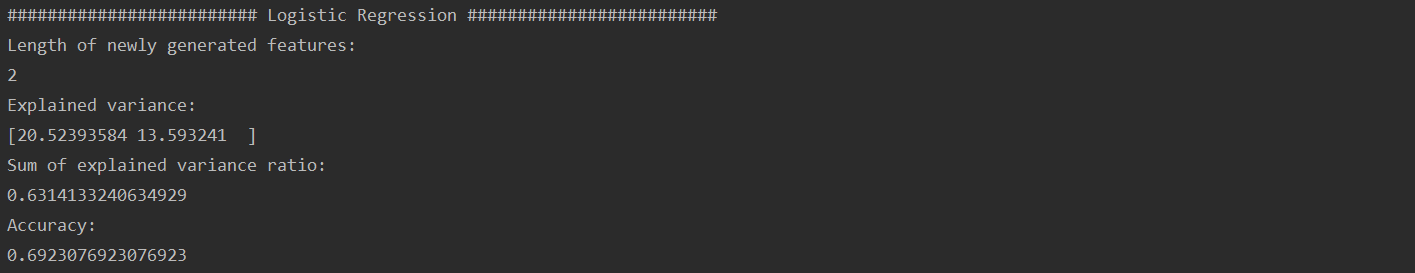


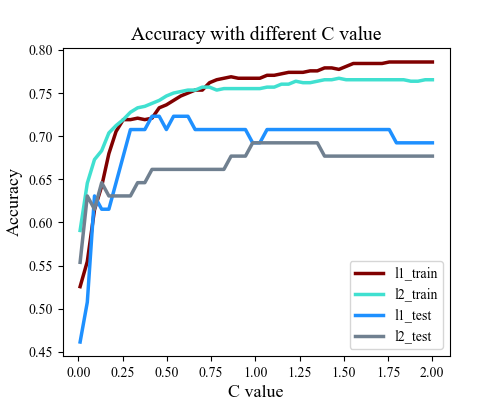


**Logistic Regression：**

****

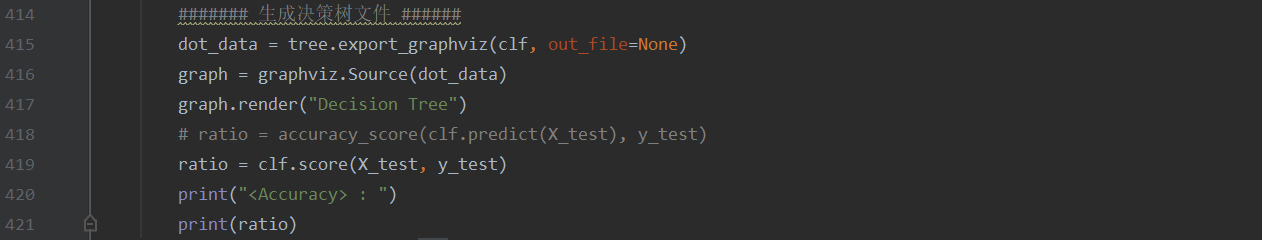
****

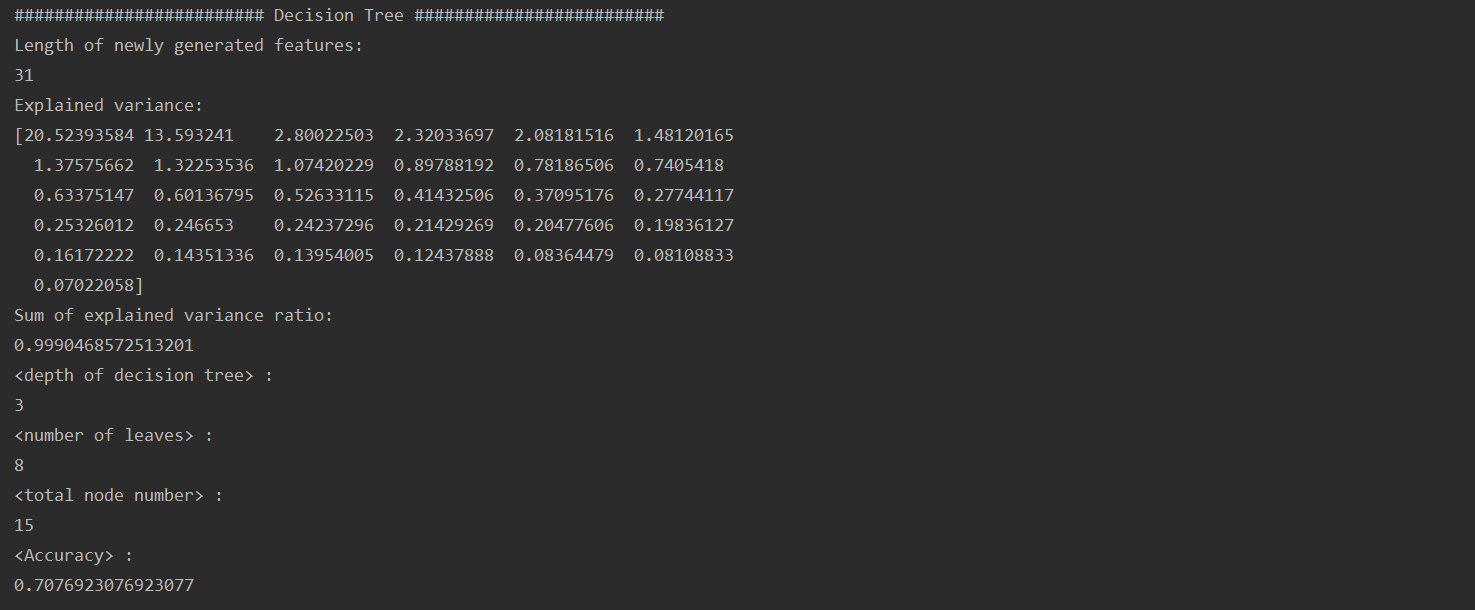
****

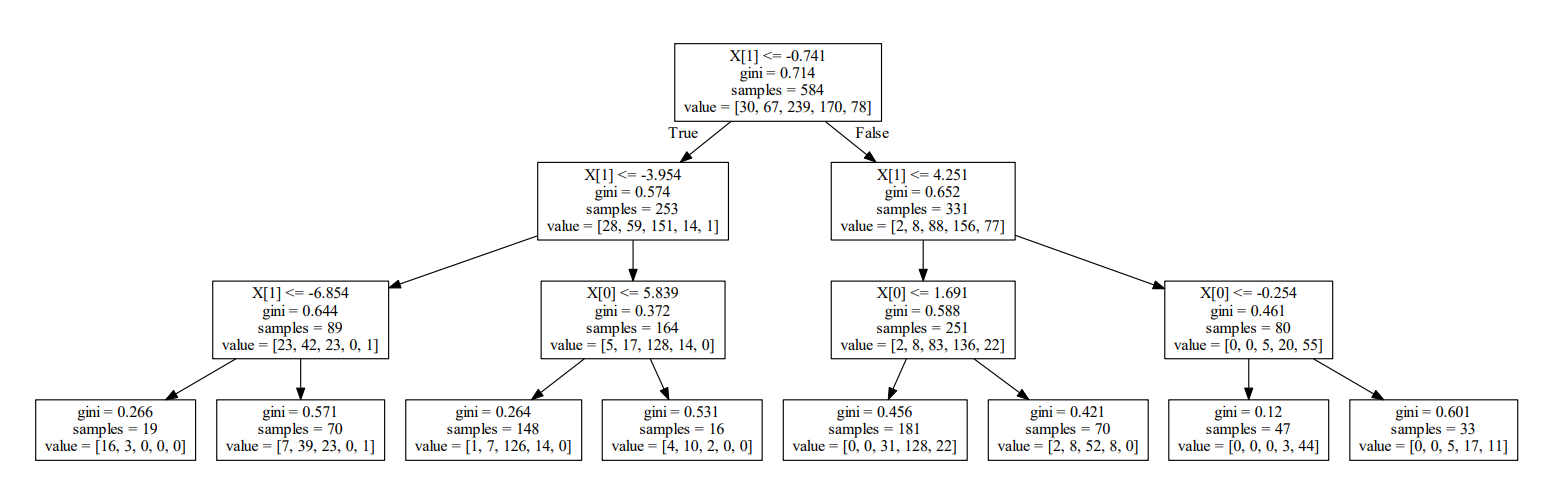


**Decision Tree：**

****

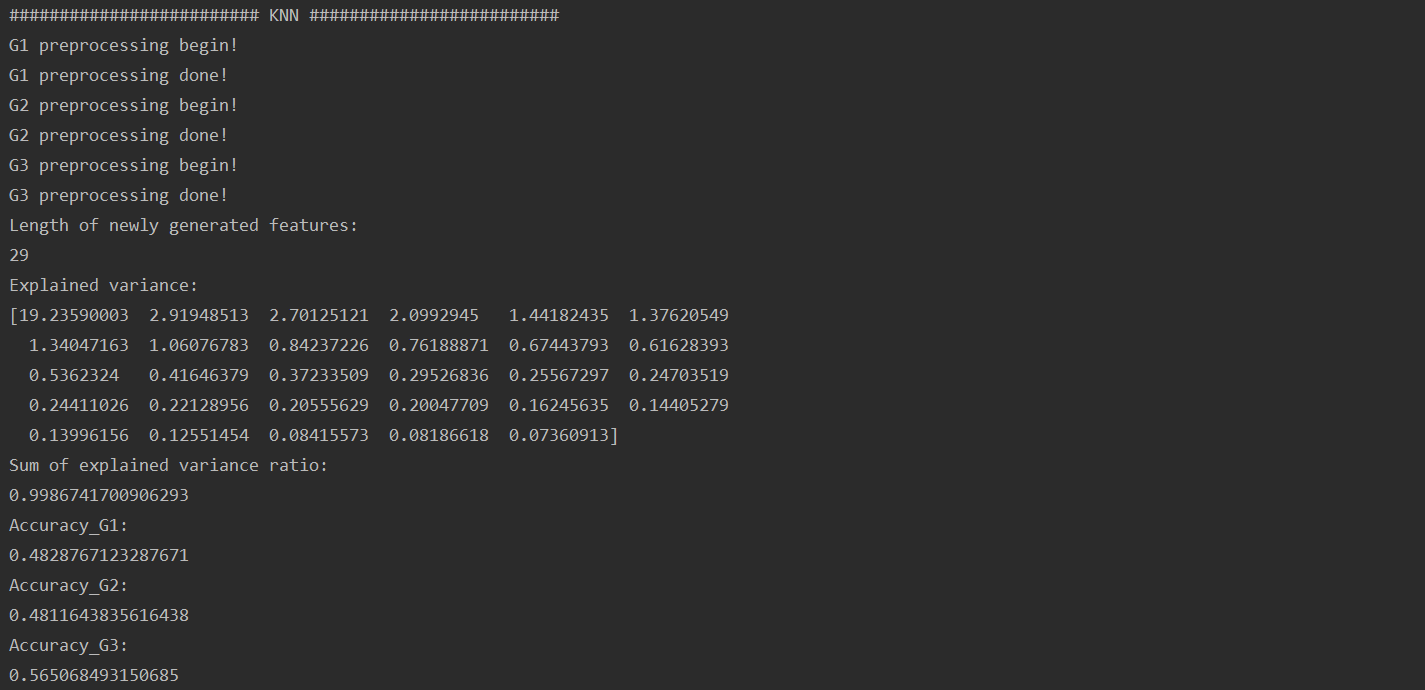
****

****

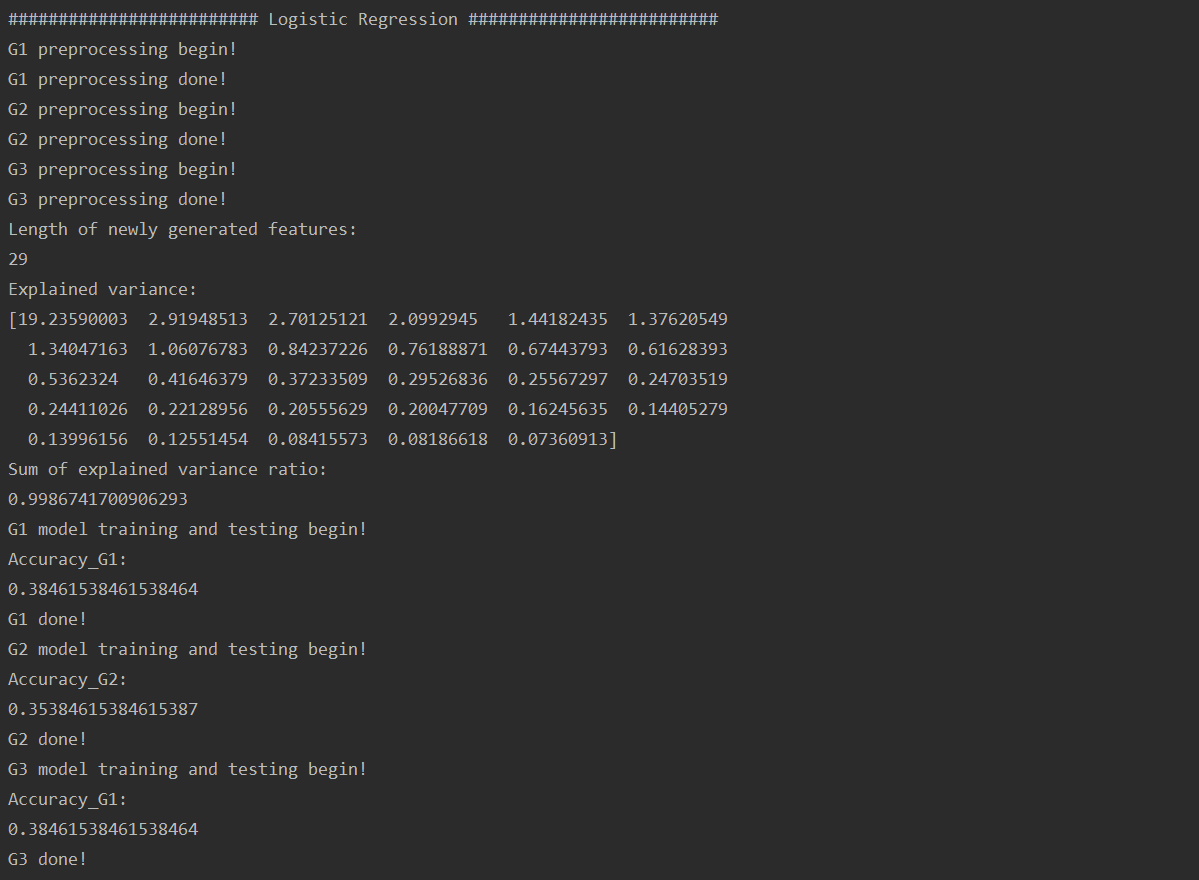
****

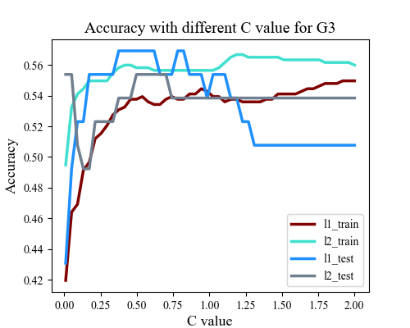
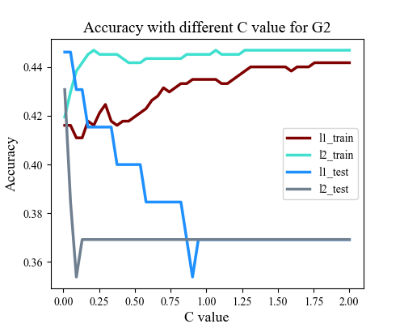
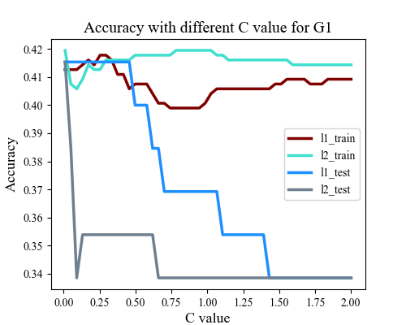
**2）30特征预测G1、G2、G3：**

**KNN：**

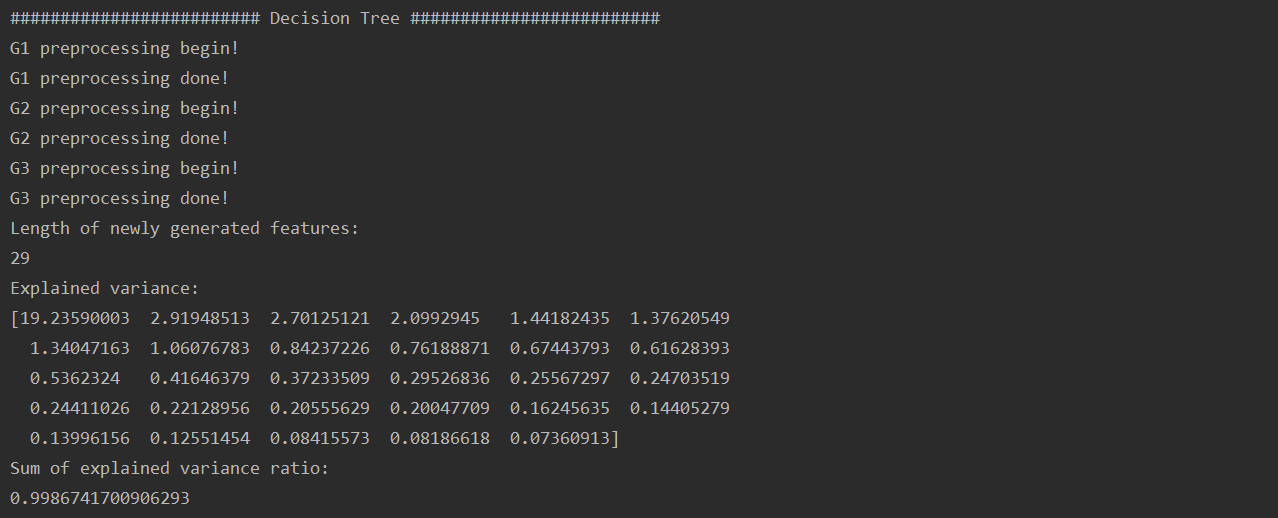


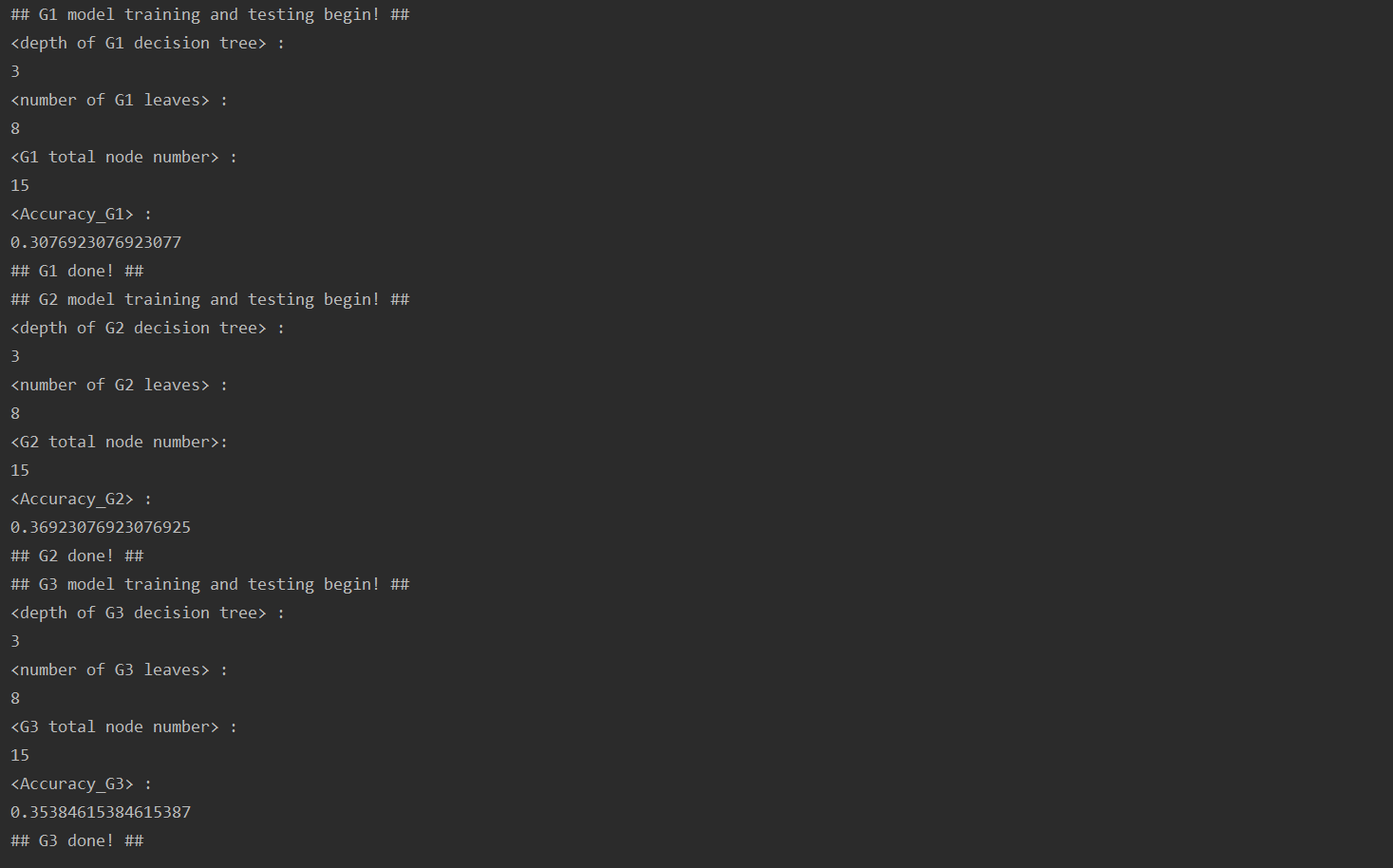
**Logistic Regression：**

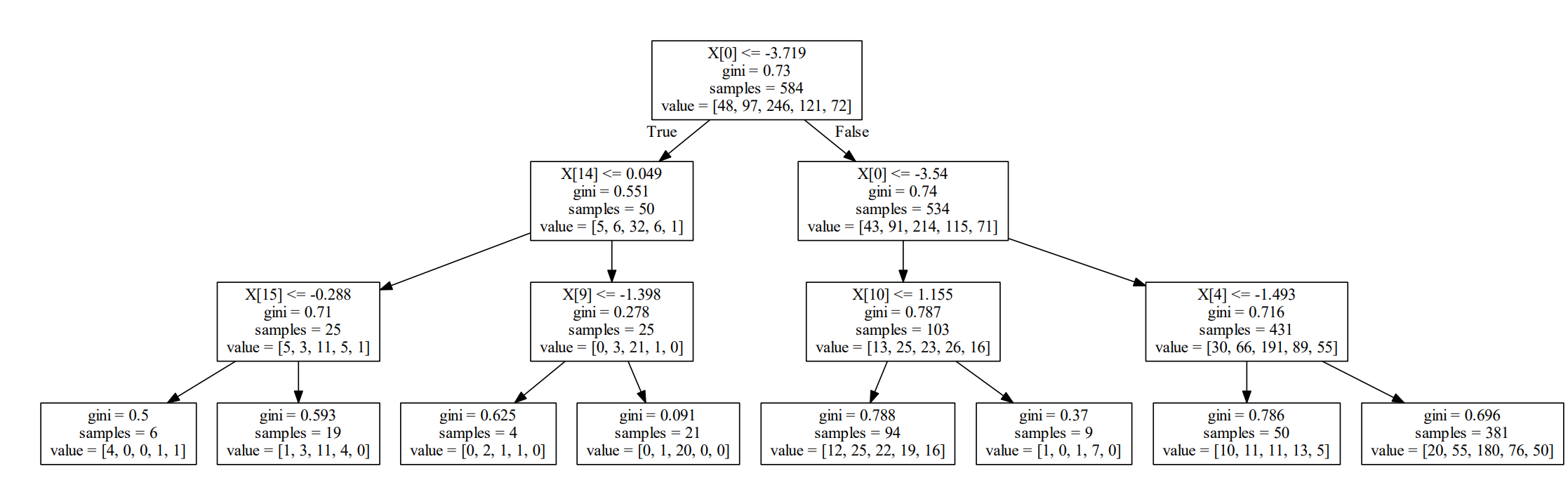
****

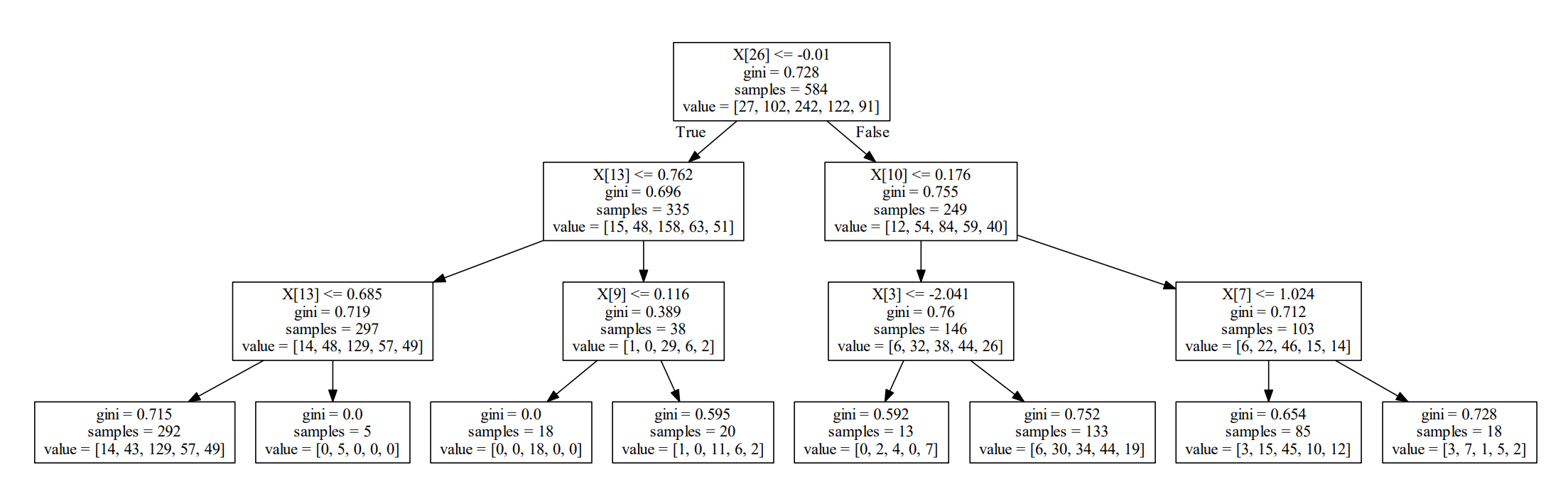


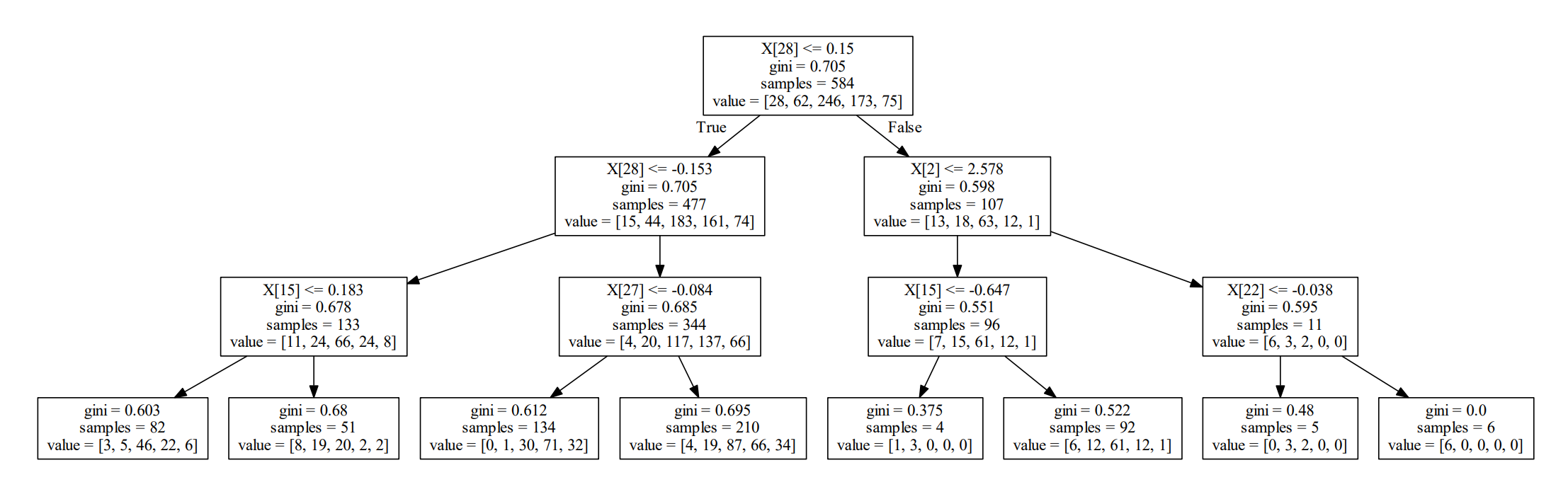
**Decision Tree：**

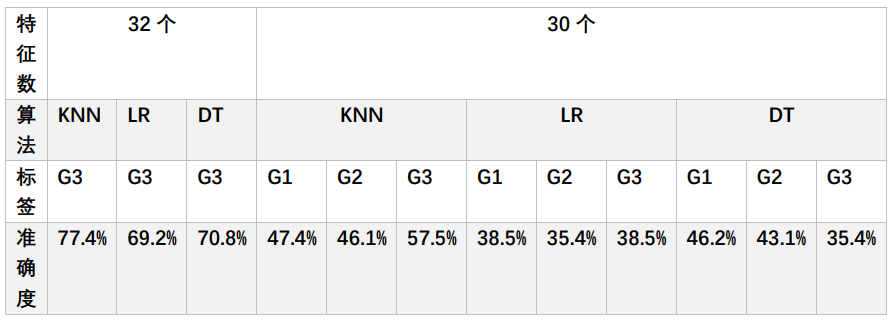
****

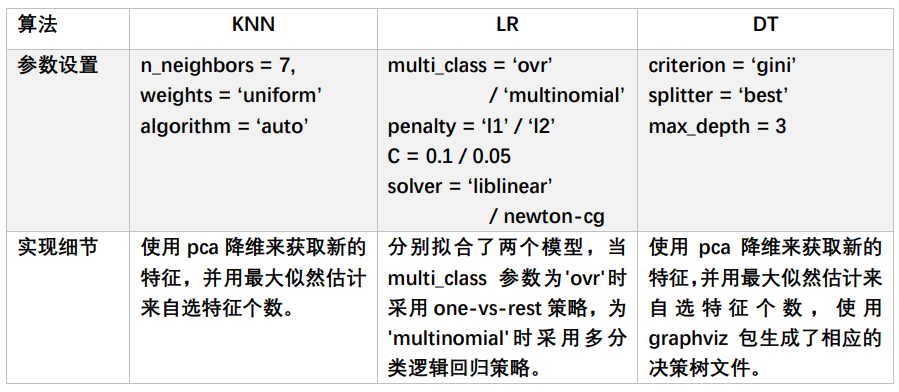
****

****

****

****

****

****