**本 科 实 验 报 告**

实验名称： 多分类任务补充实验

学 员： 孙源辰 学 号： 201802001067

培养类型： 无军籍本科学员 年 级： 大三

专 业：人工智能与大数据 所属学院： 计算机学院

指导教员： 刘新旺 职 称： 副教授

实 验 室： 305-205 实验日期： 2021春季学期

国防科学技术大学训练部制

《本科实验报告》填写说明

1．学员完成人才培养方案和课程标准要所要求的每个实验后，均须提交实验报告。

2．实验报告封面必须打印，报告内容可以手写或打印。

3．实验报告内容编排及打印应符合以下要求：

（1）采用A4（21cm×29.7cm）白色复印纸，单面黑字打印。上下左右各侧的页边距均为3cm；缺省文档网格：字号为小4号，中文为宋体，英文和阿拉伯数字为Times New Roman，每页30行，每行36字；页脚距边界为2.5cm，页码置于页脚、居中，采用小5号阿拉伯数字从1开始连续编排，封面不编页码。

（2）报告正文最多可设四级标题，字体均为黑体，第一级标题字号为4号，其余各级标题为小4号；标题序号第一级用“一、”、“二、”……，第二级用“（一）”、“（二）” ……，第三级用“1.”、“2.” ……，第四级用“（1）”、“（2）” ……，分别按序连续编排。

（3）正文插图、表格中的文字字号均为5号。

一、实验目的和要求

通过编程实现多分类任务ECOC模型，理解并掌握多对多ECOC原理以及不平衡数据处理办法。

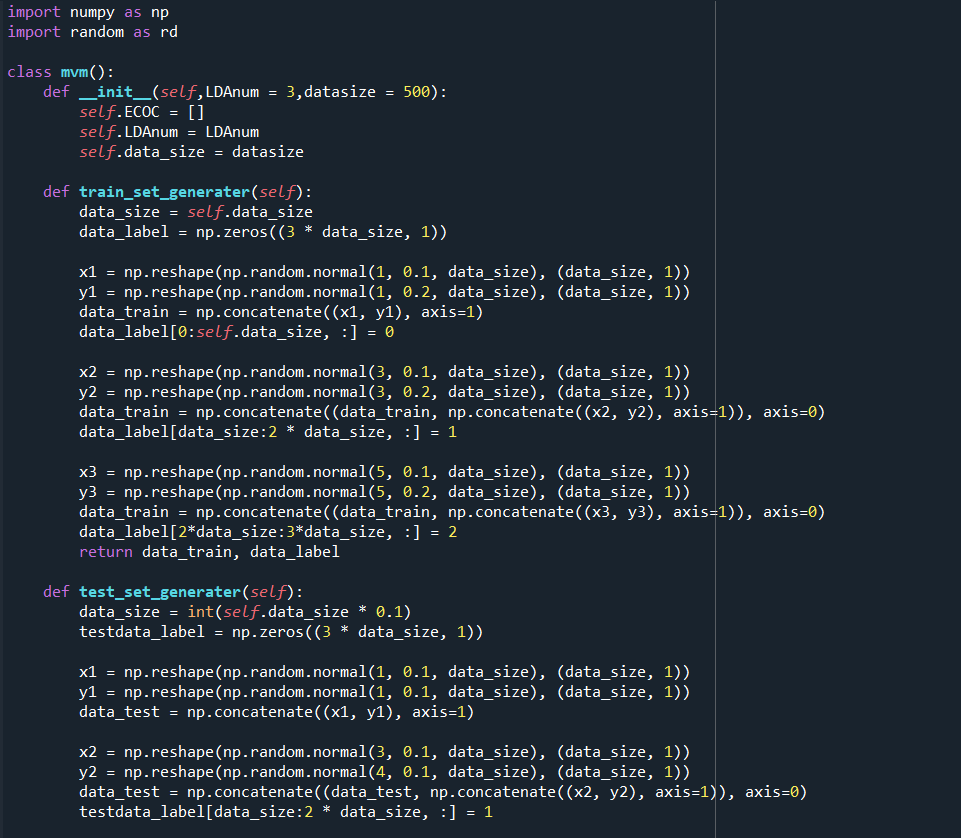
二、实验内容和原理

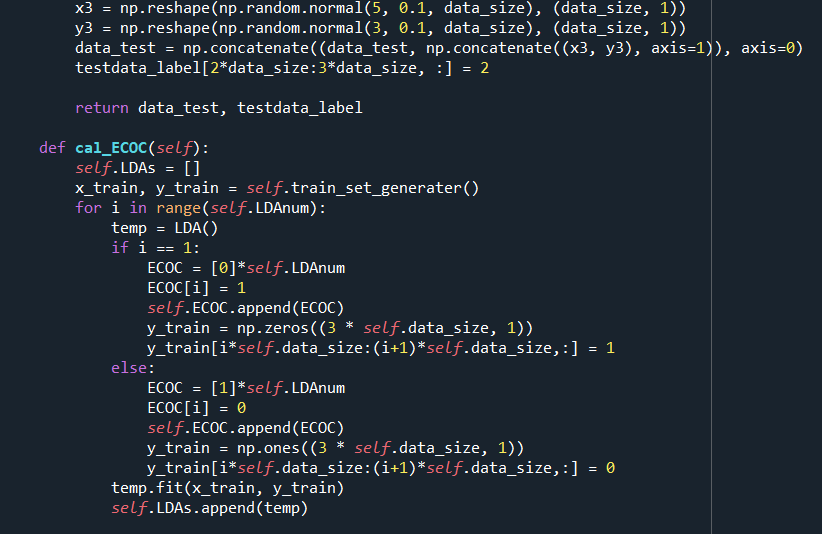
用python实现ECOC和不平衡数据处理。

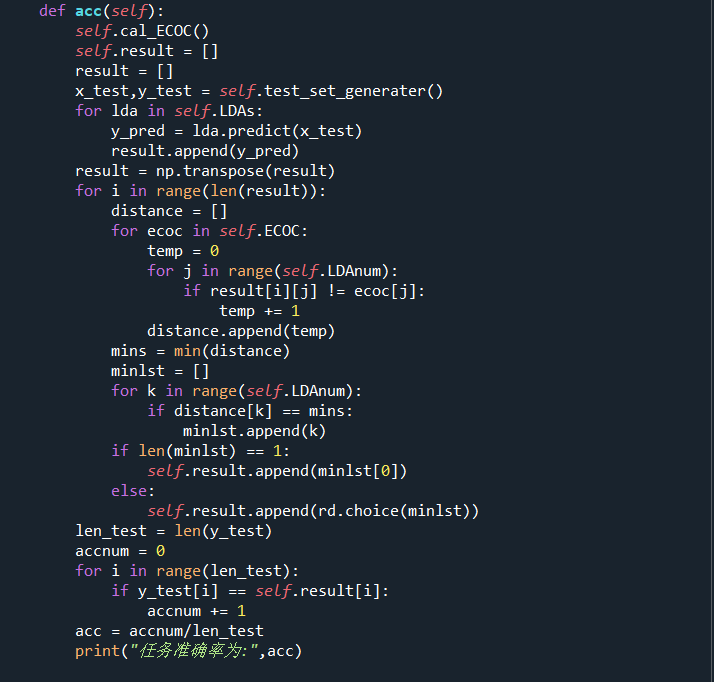
三、操作方法与实验步骤

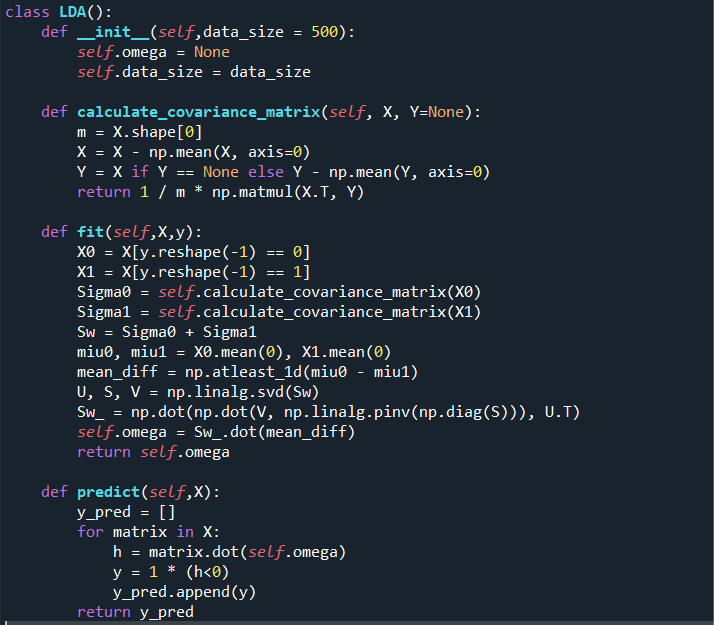
(一).编程思路

1.ECOC



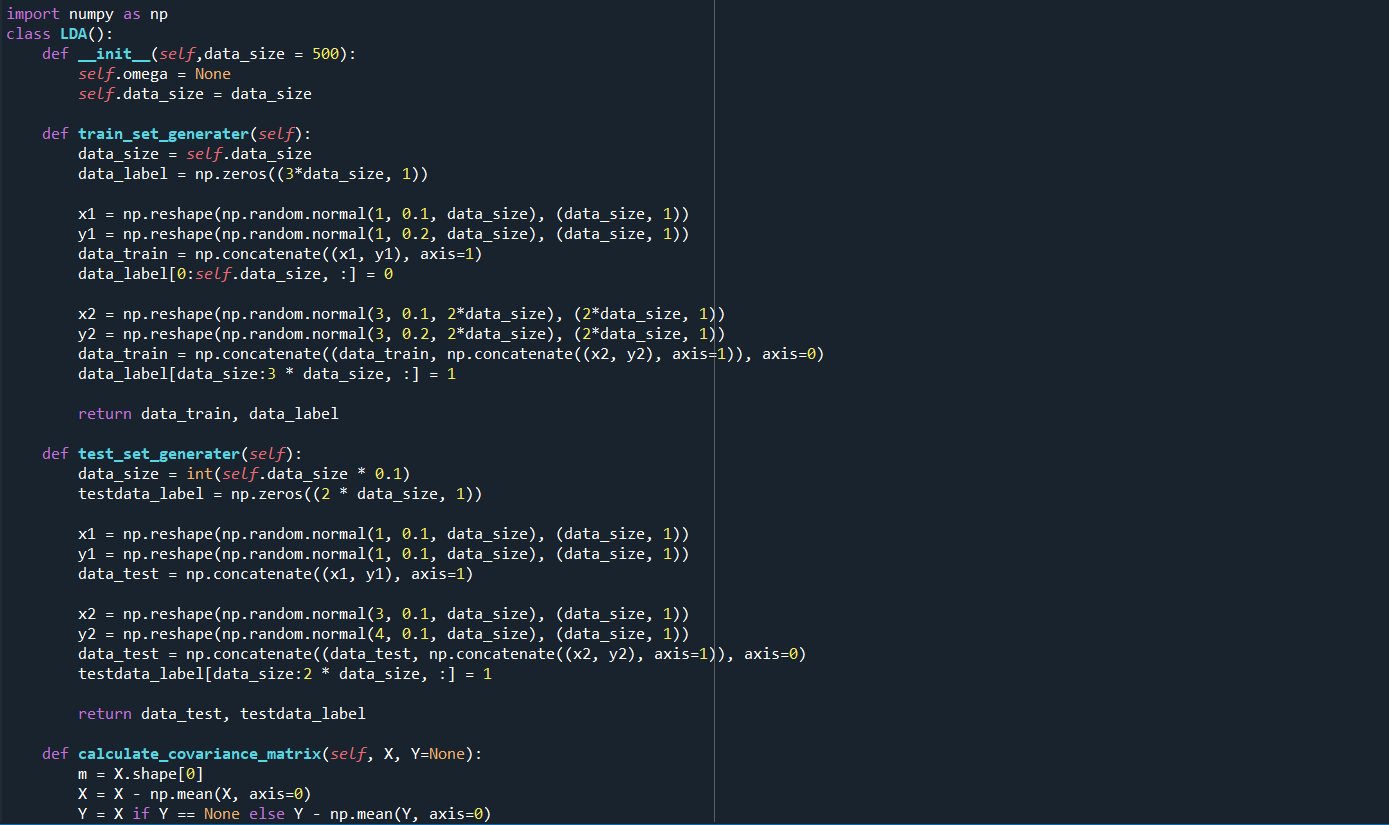






这里选择了最简单的多分类任务，三个线性LDA组成ECOC的三分类任务。首先是初始化部分，相同数量的三种数据混合，之后直接定义了ECOC码的计算函数，训练分类器。之后是对同样三类的数据集进行预测分析，计算海明距离分类，对距离相同的随机选择。

2.不平衡





这里将一类变成了另一类的二倍，框架依然是二分类LDA。

四、实验结果及分析

(一)ECOC

多分类result

可以看到，几个LDA综合分类结果准确性很高，几个LDA不同分类却集成出了很好的效果。

(二)不平衡

不平衡result1

上面的是处理前的LDA，可以看到对比原LDA的90以上分类准确率有较大下降。

不平衡result2

经过处理后，判别标准变为正例除以反例，准确率回归。