第3章监督学习作业

**题目：**

选择sklearn库中的葡萄酒数据集（同学们可以自己找数据集），采用三种以上的监督学习算法进行分类，选择你认为合适的模型评估参数（错误率、查准率、查全率、F1等）对这些算法进行评估对比，选出最优模型。

**数据集**：

葡萄酒数据集：sklearn.datasets.load\_wine

自选数据集

**提交材料：程序代码，实验报告和ppt，全部文件打包压缩后以“学号\_姓名”命名上传课程平台。截止时间为11月6日**

**程序代码说明**：对应每个题目提交可执行的python代码，需包含相应注释，以“学号\_题目号”命名，例如：“20211012\_1.py”

**实验报告说明**：上述题目需包含三部分内容：算法简介、结果和分析。算法简介部分应给出对数据集的介绍，数据预处理、训练集测试集的划分，模型参数的选择等信息；结果部分应以图或者表的形式呈现结果，分析部分对结果进行简要的分析，如效果好或者不好的原因等。实验报告格式见下页。

**PPT说明**：主要包含本题内容，介绍数据集，数据集预处理以及划分情况，选择的几种模型，模型训练的结果，性能参数的计算，结果的对比分析等内容，将用于本章结束时课堂分享。

# 实验报告：监督学习

学号：

姓名：

1. k近邻算法
   1. 算法简介
   2. 结果

* 1. 分析

1. 决策树
   1. 算法简介
   2. 结果

* 1. 分析

1. 线性模型
   1. 算法简介
   2. 结果

* 1. 分析

1. SVM
   1. 算法简介
   2. 结果

* 1. 分析

1. 朴素贝叶斯
   1. 算法简介
   2. 结果

* 1. 分析

1. 模型对比
   1. 算法简介（需介绍选择的哪几种算法，对数据集的预处理，数据集的划分，对每个模型参数的选择，性能评估参数的选择等）
   2. 结果（以表的形式给出每种模型的结果）

* 1. 分析