**一、第二次大作业为集成学习综合对比分析，具体任务主要有两个：**

1. 实现AdaBoost算法
2. 实现Bagging集成方法

大作业为组队形式完成，最多三人一组，提交实验报告（模板见附件），在交作业的时候写上组内每个人的姓名和学号。

**二、实现AdaBoost算法**

首先需要同学们自行搜集分类任务的数据集（分类任务），然后自己实现AdaBoost算法，达到以下要求：

1. 基学习器自行选择（可以使用Scikit-learn现成的分类器），使用不同的基学习器实现2种以上的AdaBoost。
2. 运用Accuracy, Precision, Recall, F1四个指标进行对比评测，进行十折交叉验证，把结果绘制成表格
3. 与Scikit-learn 的AdaBoostClassifier得到的结果进行对比（基学习器和你自己实现的AdaBoost相同）
4. 详细分析不同的基学习器的AdaBoost的表现差异以及可能的原因。

**三、实现Bagging集成方法**

使用和上节AdaBoost一样的数据集，然后自己实现Bagging算法，达到以下要求：

1. 基学习器自行选择（可以使用Scikit-learn现成的分类器），使用不同的基学习器实现2种以上的Bagging。
2. 运用Accuracy, Precision, Recall, F1四个指标进行对比评测，进行十折交叉验证，把结果绘制成表格
3. 与Scikit-learn 的BaggingClassifier得到的结果进行对比（基学习器和你自己实现的Bagging相同）
4. 详细分析不同的基学习器的Bagging的表现差异以及可能的原因。
5. 与AdaBoost的结果进行对比，分析两种集成方法的优劣

**四、辅助材料**

1. UCI数据集网址：<https://archive.ics.uci.edu/ml/datasets.php>
2. Scikit-learn 数据集模块文档：[https://scikit-learn.org/0.16/modules/classes.html#module-sklearn.datasets](https://scikit-learn.org/0.16/modules/classes.html" \l "module-sklearn.datasets)
3. Scikit-learn AdaBoost模块文档：<https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.AdaBoostClassifier.html#sklearn.ensemble.AdaBoostClassifier>
4. Scikit-learn Bagging模块文档：<https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.BaggingClassifier.html#sklearn.ensemble.BaggingClassifier>
5. AdaBoost教程，仅供参考，严禁抄袭<https://xavierbourretsicotte.github.io/AdaBoost.html>
6. Bagging教程，仅供参考，严禁抄袭<https://machinelearningmastery.com/implement-bagging-scratch-python/>