

完成以下三个任务，提交源代码和实验报告。请在实验报告中记录你的实验过程、实验结果和思考。

任务1 回归（对应文件：data_akbilgic.csv）

任务描述：在给定的数据文件中，每一行代表一个开盘日中的股指交易涨跌值，第一列记录具体日期，其后每一列代表一项股指数据，共九列，依次为：ISE(TL-based), ISE(usd), SP, DAX, FTSE, NIKKEI, BOVESPA, EU, EM。回归任务是通过后八项股指来对第一项股指（ISE(TL-based)）的数值进行预测。请完成下列工作：

1. 使用可视化的方法观察数据之间的关联，推测该数据是否适合进行回归分析/线性回归分析。
2. 使用回归分析的方法（如线性回归）进行回归分析，并与你的推测结果进行对比和思考。（实验过程中请注意评价指标、训练误差、泛化误差、测试数据划分等内容，并记录在实验报告中。）
3. （附加题）尝试使用降维前后的数据表示分别进行回归，并比较回归的结果，思考降维对该回归任务的影响。

数据来源于 <https://archive.ics.uci.edu/ml/datasets/ISTANBUL+STOCK+EXCHANGE>

任务2 二分类（对应文件：wdbc.csv）

任务描述：该数据为乳腺癌检查的医疗检测数据，每行对应一个案例(以逗号分隔每个数据)，其中第一列为案例编号，第二列为诊断结果(M, malignant; B, benign)，其后十列为各项检查数据。分类任务是通过后十项的检查数据来预测诊断结果。请完成下列工作：

1. 对检查数据进行处理并使用降维方法（如PCA）进行降维（2维或3维）。通过可视化观察降维结果，并推测该数据是否适合进行分类学习。
2. 使用分类方法（如logistic regression）对上述问题进行分类学习，并与你的推测结果进行对比和思考。（实验过程中请注意评价指标、训练误差、泛化误差、测试数据划分等内容，并记录在实验报告中。）
3. （附加题）尝试使用降维前后的数据表示分别进行分类，并比较分类的结果，思考降维对该分类任务的影响。

数据来源于 <https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29>

任务3 多分类（对应文件：dataset.csv）

任务描述：有1500个样例数据（每一行为一个样例，每个样例中有十一列数据，其中第一列为样例的类别，共四种A、B、C、D，其后十列为样例的输入特征）。分类任务是通过输入特征来预测样例的类别。请完成下列工作：

1. 用降维算法将数据降为2维或者3维，并以不同的颜色表示各类别进行可视化。
2. 尝试比较不同的降维算法（如自己实现的PCA算法、sklearn中的PCA方法，以及其他可能的降维方法等）的结果差异，如通过可视化结果进行比较等。
3. （附加题）尝试对此数据进行分类（分类方法、评估方法自行选择），并报告分类结果。