

2021 新人任务第一弹

大一（一周任务）：

一、数据处理

1: 缺失值处理

包括但不限于，knn填补，众数，均值填补，补零

2: 数据标准化，归一化

3: 对于离散型变量的OneHotEncoder

二、逻辑回归

实现

防止过拟合的算法

尝试多种优化方式，包括但不限于，sgd，adam

数据集：

https://uniquestudio.feishu.cn/drive/folder/fldcnV0PzAB5J8ZaoMp8WXho8if?from=from_copylink

三、keams

实现

尝试多种优化初始点的选择

四、其他

1. 学习使用Github
2. 学习Jupyter
3. 学习Pycharm或VSCode
4. 了解Linux

PS：验收要求

组会现场推导逻辑回归

检查代码的真实性，和运行结果，准确率等

大二（两周任务）

一、论文阅读

每周阅读两篇论文，并且书写论文笔记，上传语雀

二、论文复现

复现faster-rcnn论文

数据集用voc2007

https://uniquestudio.feishu.cn/drive/folder/fldcnV0PzAB5J8ZaoMp8WXho8if?from=from_copylink

Pascal VOC Dataset Mirror

Here is a mirror for the Pascal VOC files in case, you know, you want to download them at a somewhat decent rate.

<https://pjreddie.com/projects/pascal-voc-dataset-mirror/>

只能在jupyter复现所有代码在一个文件中，两周内完成，第一次组会要把源码看明白并且讲述实现细节

可以通过colab等平台训练

验收要求

输入的图片经过faster-rcnn有初步的检查效果，不需要完全训练