Projects

Due 2020/11/30

Topic 2: 天体光谱智能识别分类

背景：

天文望远镜每晚能采集万余条光谱，使得传统的人工或半人工的利用模板匹配的方式不能很好应对，需要高效而准确的天体光谱智能识别分类算法。

任务描述：

给出18万个天体的光谱数据，要求根据光谱数据，把未知天体分成行星（star），星系（galaxy）和类星体（qso）三类。评价指标即预测准确率。

数据描述

数据为csv文件，每一行代表一个天体，每一列代表这个天体在不同波长上的值。

训练集带有数据标签，即每个天体的分类。star代表恒星，galaxy代表星系，qso代表类星体。测试集要求预测天体的标签。

训练集：

训练集数据集为csv文件，数据包括14万个天体的信息。其中 id列 为天体 ID，FE0列,FE1列,FE2列,…,FE2599列,共计2600列为波长信息，作为数据的输入特征。天体的类别在对应label文件中。

验证集：

验证集数据有4万行，格式和训练集格式相同。（需要注意，由于存入一个文件太大，每2万数据为一个文件存储，共9个文件，同时只有第一个文件有表头信息，训练集是sets\_v1\_0至sets\_v1\_6，验证集是sets\_v1\_7和sets\_v1\_8）

下载地址：

地址：https://pan.baidu.com/s/1Di5VMzz0HC0qc--W5NNrig

提取码：muls

最后验收请提交实验报告来阐述你的实验方案、方法和性能分析。同时，现场会提供测试集文件和label文件，同学需要当场运行程序进行预测并计算准确率。

Topic 3: Foreign Object Detection on Roadways

*Task description*

Given an input image, an algorithm is expected to produce a set of tight boxes around objects with classification labels. You can exploit various state-of-the-art methods or propose a novel method.

*Data introduction*

We provide 666 images, taken on the road at the Zhongguancun Campus. The foreign objects in the images include yellow/black clippers, screws and small wrenches for screwing. For example,



* Training dataset: 666 images with annotation files
* Testing dataset: 100 images with annotation files (will be provided by the end of this course)

*Evaluation method*

Three indicators will be used in the evaluation: recall, precision and average precision (ap). TP indicates the number of positive samples detected as positive samples; FP indicates the number of negative samples detected as positive samples; FN indicates the number of positive samples detected as negative samples. Recall and precision are defined as follows:



Average precision (ap) is calculated using VOC07 method.

下载地址：

链接：https://pan.baidu.com/s/1xZCU6T4AS4csyrSxWUZq2w

提取码：ppla

最后验收请提交实验报告来阐述你的实验方案、方法和性能分析。同时，现场会提供测试集文件和label文件，同学需要当场运行程序进行预测并计算准确率。