google facets是google一个开源的工具，它包括overview和dive这两个可视化的工具，用来了解和分析机器学习的数据集。

overview提供了关于一个或多个数据集的高级视角，提供了特征对特征的分析视角。工具可以处理数字和字符特征。

overview能对数据集有以下帮助：

\*意外的特征值

\*缺少大量示例的特征值

\*训练/服务偏差

\*训练/测试/验证设置偏差

可视化的关键方面是跨多个数据集的异常值检测和分布比较。红色显示感兴趣的值（例如，高比例缺失数据、跨多个数据集的特征分布）。

特征可以按照感兴趣值排序，例如缺少值的数量或不同数据集之间的偏差。

DIVE是一种交互式探索多达数万个数据点的工具，允许用户在高级概述和低级细节之间进行无缝切换。

每个示例在可视化中被表示为单个项目，并且可以通过其特征值来定位点。结合平滑动画缩放、细分、过滤，dive能够轻松的在复杂数据集中发现模式和异常值。

facets介绍

<https://pair-code.github.io/facets/>

下载方法：

如果你的电脑可以直接联网，使用下面的git clone可以直接从github上下载。

git clone https://github.com/PAIR-code/facets

如果你的电脑可以不能联网， 找一个可以上网的电脑从<https://github.com/PAIR-code/facets>上下载即可。

相关知识点：

1. google项目构建工具bazel，以及对应的BUILD语法。
2. html
3. JavaScript
4. typescript
5. jupyter notebook
6. python