1.题目

**TF-IDF练习**

2.数据集

参考任务描述

3.任务描述

（1）自行准备数据集：

比如doc1: Olympic games end

Doc2: football games begin

Doc3: basketball games ongoing

Doc4: volleyball games draw

打开网址<https://remykarem.github.io/tfidf-demo/>；

输入4个文档，可以参考上述文档，也可以自行准备4个文档；

生成T-IDF；

针对其中两个词项，分析其TF-IDF的计算过程（自己用纸笔手写计算公式和计算过程，拍照，以图片形式嵌入到ipynb文件中(Markdown Cell)，图片和ipynb文件打包提交）。

（2）参考“run\_TF\_IDF.ipynb”示例代码。

使用Jupyter notebook运行本目录下run\_TF\_IDF.ipynb文件；

增加cell，对示例代码进行注释；

针对其中两个词项，分析其TF-IDF的计算过程（自己用纸笔手写计算公式和计算过程，拍照，以图片形式嵌入到ipynb文件中(Markdown Cell)，图片和ipynb文件打包提交）。

4.成果提交

以ipynb形式提交。

5.参考文献