第三次大数据分析作业

1. 工作流程、思路

首先使用百度AIP的ocr技术读取图片中矩阵信息，然后更据pagerank1迭代计算公式计算每次迭代（即将自己的pagerank1修正为含有入链的pagerank）

1. 技术选型

使用Python语言作为编程语言，使用baidu AIP开源sdk作为图像识别工具，然后使用numpy的矩阵模型存放转移矩阵以及pagerank的值并迭代计算，最终不断迭代产生修正的pagerank

1. 十次迭代结果

迭代之前：

A:

0.25

B:

0.25

C:

0.25

D:

0.25

第1次迭代：

A:

0.25

B:

0.20833333333333331

C:

0.20833333333333331

D:

0.3333333333333333

第2次迭代：

A:

0.2708333333333333

B:

0.25

C:

0.1875

D:

0.29166666666666663

第3次迭代：

A:

0.2708333333333333

B:

0.23611111111111108

C:

0.21527777777777776

D:

0.2777777777777778

第4次迭代：

A:

0.2569444444444444

B:

0.22916666666666666

C:

0.20833333333333331

D:

0.3055555555555555

第5次迭代：

A:

0.2673611111111111

B:

0.2384259259259259

C:

0.20023148148148145

D:

0.29398148148148145

第6次迭代：

A:

0.26620370370370366

B:

0.2361111111111111

C:

0.20833333333333331

D:

0.2893518518518518

第7次迭代：

A:

0.26273148148148145

B:

0.23341049382716045

C:

0.2067901234567901

D:

0.29706790123456783

第8次迭代：

A:

0.2652391975308641

B:

0.23611111111111105

C:

0.20428240740740738

D:

0.29436728395061723

第9次迭代：

A:

0.2652391975308641

B:

0.23559670781893

C:

0.20646862139917688

D:

0.29269547325102874

第10次迭代：

A:

0.2641460905349794

B:

0.23476080246913572

C:

0.20621141975308638

D:

0.29488168724279823