

数据挖掘第一次作业实验数据说明文档

学号：1452822

姓名：洪嘉勇

文件列表

- result_lsh.json
- result_knn3.json
- result_knn4.json
- result_knn5.json

详情

result_lsh.json

 该文件为lsh的结果文件

数据结构

```

{
  'hash_size1':[
    '查询的路径id':{
      // 路径列表
      [
        // 点列表
        [
          ['经度', '维度']
          ['经度', '维度']
          ['经度', '维度']
          ...
        ]
      ]
    },
    ],
  hash_size2':[
    '查询的路径id1':{
      // 路径列表
      [
        // 点列表
        [
          ['经度', '维度']
          ['经度', '维度']
          ['经度', '维度']
          ...
        ]
      ]
    },
    '查询的路径id2':{
      // 路径列表
      [
        // 点列表
        [
          ['经度', '维度']
          ['经度', '维度']
          ['经度', '维度']
          ...
        ]
      ]
    },
    ...
  ],
  ...
}

```

result_knnX.json

这些文件为 $k = x$ 对应的结果文件

数据结构

```
{
  [
    {
      '查询的路径id1': [
        [ '经度', '维度' ]
        [ '经度', '维度' ]
        [ '经度', '维度' ]
        ...
      ],
      '查询的路径id2': [
        [ '经度', '维度' ]
        [ '经度', '维度' ]
        [ '经度', '维度' ]
        ...
      ],
      '查询的路径id3': [
        [ '经度', '维度' ]
        [ '经度', '维度' ]
        [ '经度', '维度' ]
        ...
      ],
      ...
    }
  ]
}
```

注意事项

在提交的作业中的实验结果文件仅仅是由代码跑出来的众多结果中的一组，因为局部敏感哈希得到的结果每次都不一样，所以我只选取了一组。

在[网页](#)中显示的结果也仅为实验中的一组，见谅。

谢谢~

