1. **算法说明：**

可推荐出国留学课程人数&意向用户的校区分布

1. **部署方式：**

****

**定时任务：**

1. 拷贝predictenrollnum.jar文件至指定文件夹。
2. 修改predictenrollnum.sh 最后一行predictenrollnum.jar位置，如果大数据环境变量有修改，做相应修改。
3. 拷贝predictenrollnum.sh至指定文件脚本文件夹，并使用chmod 修改为可执行文件。

chmod +x predictenrollnum.sh

1. 添加定时任务，时间应该定于上一年数据全部入库后的某个凌晨。

【示例中使用1月2日凌晨2点; /zhouzhou/predictenrollnum.sh 部分需要根据3）进行修改。】

示例：

# crontab –e

0 2 2 1 \* /zhouzhou/predictenrollnum.sh

**手动运行方式：**

./spark-submit --class "cn.xdf.www.analysis.etl.predictenrollnum" predictenrollnum.jar

1. **查询方式：**

**表名为**dw. dw\_predecit**\_预测年份**

**可推荐出国留学课程人数：**

* + 1. **查询语句**：

select a.nschoolid,a.month, a.coursecode, a.coursename,sum(predict) as predict

from dw. dw\_predecit\_2018 a

where nschoolid=1 and a.coursecode=20

group by a.nschoolid, a.month, a.coursecode, a.coursename

* + 1. **参数**

Nschoolid：学校

Coursecode：课程分类代码，中文表中coursename

* + 1. **结果**

|  |  |
| --- | --- |
| nschoolid | 学校id |
| month | 预测月份 |
| coursecode | 课程代码 |
| coursename | 课程名称 |
| predict | 人数 |

**意向用户的校区分布：**

1. **查询语句**：

select a.nschoolid,a.month, a.sareacode ,sum(predict) as predict

from dw. dw\_predecit\_2018 a

where nschoolid=1 and a.coursecode=20

group by a.nschoolid, a.month, a.coursecode, a.coursename, a.sareacode

1. **参数：**

Nschoolid：学校

Coursecode：课程分类代码，中文表中coursename

1. **结果：**

|  |  |
| --- | --- |
| nschoolid | 学校id |
| month | 预测月份 |
| sareacode | 地区代码 |
| predict | 人数 |