**作业12**

**在Spark MLlib中使用ALS算法序**

**备注：参考厦门大学数据库实验室博客《Spark 2.1.0 入门：协同过滤算法》http://dblab.xmu.edu.cn/blog/1461-2/**

1. **编程**

获取MovieLens数据集，使用Spark MLlib自带的ALS算法来构建推荐模型，并根据模型对评分预测的均方根误差来对模型进行评估。

**二、实验报告**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 实验报告 | | | | |
| 题目 | 在Spark MLlib中使用ALS算法序 | 姓名 | 吴雨娟 | 日期：  2021.11.20 |
| 实验环境：VMware workstation 16 pro，Ubuntu-20.04操作系统, JDK1.8,scala-2.12.10 | | | | |
| 实验内容与完成情况：  一、编程  获取MovieLens数据集，使用Spark MLlib自带的ALS算法来构建推荐模型，并根据模型对评分预测的均方根误差来对模型进行评估。  1.导入需要的包    2. 根据数据结构创建读取规范      3. 读取数据      4. 构建模型        5. 模型预测        6. 模型评估 | | | | |
| 出现的问题：  1.问题一：没有开启spark-shell，而是开启scala，发现包导入失败。 | | | | |
| 解决方案（列出遇到的问题和解决办法，列出没有解决的问题）：  1.问题一的解决方案：发现是因为没有开启spark-shell，开启spark-shell后再运行代码，可以得到正确结果。 | | | | |

（备注：相关实验过程可以截图放入实验报告）