**中北大学软件学院毕业设计中期报告**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 专业 | 软件工程 | 方向 | 云计算与大数据分析方向 | | 班级 | 14140Y01 |
| 姓名 | 张丹 | 学号 | 1414010513 | | 指导教师 | 曹辉 章永来 |
| 题目 | 大学生一卡通数据分析系统的设计与实现 | | | | | |
| 本人在该设计中具体  应完成的工作 | | | | (1)开发一卡通管理系统小程序，模拟用户的行为数据  (2)开发数据生成系统，根据样本数据生成更大数据集  (3)开发数据采集系统，从数据生成源采集数据到HDFS存储  (4)开发数据处理系统，对海量数据集进行处理、分析  (5)开发报表展现系统对分析结果进行可视化展现 | | |
| 1. 简述毕业设计开始以来所做的具体工作和取得的进展 2. 集群搭建 3. 搭建3台Linux服务器，分别为hdp1，hdp2，hdp3。 4. 集群内安装部署Hadoop，HDFS，HBase，Hive，Mysql，Sqoop，Flume，Kafka，Zookeeper，Spark，Scala，Mongodb，Redis。 5. 数据处理 6. 将已有数据上传至HDFS进行存储。 7. Hive建数据表，加载hdfs数据到Hive表中。 8. Mysql建有效字段的数据表，利用Sqoop将Hive表中数据导出到Mysql中。 9. 数据分析 10. 对学校管理员来说： 11. 应征入伍：分析学校12级应征入伍学生总人数；每个学院入伍人数；学院之间比较每一种消费类型；每一个学院入伍学生的人均消费。 12. 留级学生：分析学校13级留级学生总人数；每个学院留级人数；学院之间比较每一种消费类型；每一个学院留级学生的人均消费。 13. 普通学生：分析学校14级普通学生总人数；每个学院人数；学院之间比较每一种消费类型；每个学院每一专业人数；学院每个班级人均消费；比较每个班级每个人每一种消费类型，进而比较消费质量。 14. 对每类学生饮食金额、时间点做分析，进而分析留级学生留级原因。 15. 对消费类型来说：【普通学生】 16. 餐厅：分析每一pos机收益情况，进而去分析收益低的pos机对应的窗口收益低的原因，比如卫生、饭菜质量；分析学生消费时间点，进而调整开放窗口时间；分析每月消费次数，进而改变饭量等；分析一个季度每天消费次数，进而改变某天的饭量，比如月底，周末学生会外出或回家；分析学生每次消费金额，进而可以推广套餐等；分析每一pos机收益金额和刷卡次数，进而确定消费质量或贫困生评定； 17. 淋浴中心：分析每一pos机收益情况，进而去分析收益低的pos机对应的窗口收益低的原因，比如淋浴器质量；分析学生消费时间点，进而调整开放澡堂时间；分析每周每天消费次数，进而可以延长某天开放时间，避免拥挤；分析每月消费次数，进而改变供水量等；分析每一pos机收益金额和刷卡次数，进而确定消费质量； 18. 热水：分析每一pos机收益情况，进而去分析收益低的pos机对应的窗口收益低的原因，比如热水器质量；分析学生消费时间点的刷卡次数，进而调整热水器功率；分析每周每天消费次数，进而可以延长某天开放时间，避免拥挤；分析每月消费次数，进而改变供水量等；分析每一pos机收益金额和刷卡次数，进而确定消费质量； 19. 对学生来说：   分析三大消费类型的占比；分析每个月的消费支出；分析每日消费最低最高均值总价；分析每天消费时间点频率，进而得知自己的消费习惯和早起习惯；分析在每个pos机消费情况，进而得知饮食喜好，消费质量；   1. 数据展示   (1)开发大学生一卡通数据分析系统。  (2)系统管理员：可对学校管理员、消费区管理员、学生做增删改查；修改密码；退出系统。  (3)消费区域管理员：按照以上分析利用Echarts做可视化；修改密码；退出系统。  (4)学校管理员：按照以上分析利用Echarts做可视化；修改密码；退出系统。  (5)学生：按照以上分析利用Echarts做可视化；修改密码；退出系统。 | | | | | | |
| 1. 目前存在问题，下一步的主要设计任务，具体设想与安排   问题：   1. 感觉分析出来的内容很多，但不知道哪些是有效分析。 2. 感觉数据很多，不知道还能分析出些什么。 3. 没有用很高级很新的技术，但是也分析出来了，技术选择哪个重要吗？ 4. 没有学院学生学号、宿舍门禁记录、图书馆门禁记录进而分析不出哪些学生哪些时间段没有消费记录、夜不归宿情况。这是额外的事情。   下一步任务：   1. 开发微信小程序，模拟用户消费行为，并显示消费分析情况。 2. 优化完善大学生一卡通数据分析系统。 | | | | | | |
| 1. 指导教师对该学生前期设计工作的评价（是否同意继续设计工作）   该同学在前期设计工作中能够严格按照任务书要求设计，在此过程中合理应用网络平台的资料，借鉴可用部分作为自己开发设计需要，对大学生一卡通数据分析系统设计整个思路清晰，流程也较为完善，对于存在的问题，解决思路如下：   1. 分析数据过多问题，可适当减少无效分析，以实际的指导性为标准进行选择， 2. 如果对目前分析的结果不满意，还可以继续挖掘对实际工作有指导意义的分析， 3. 以结果为导向，运用成熟技术，新型技术也面临着不稳定的风险， 4. 对于数据质量问题，依据现有数据进行分析，如需扩展，另行模拟相关数据即可   ，在后期设计过程中有问题及时沟通。  该同学前期工作基本符合毕业设计的要求，同意继续设计工作。  指导教师亲笔签字：曹辉 2018年 4 月 7 日 | | | | | | |
| 备注：1、本表由学生填写，指导教师亲笔签署意见。  2、以上各项间距可以根据实际需要调整。 | | | | | | |
|  | | | | | | |