数据挖掘可视化系统-进度文档V1.0

作 者：    许继元 程络 黄倬熙 张平路

               刘浩斌 邓志聪

学   历：       本科

指导老师 ：           谢光强

# 项目基本情况介绍

## 1.1 基本工作情况

本项目为 QG 工作室中期考核项目。旨在开发一个面向数据挖掘新手的无编程数据挖掘可视化系统，提供原始数据的图表可视化，以及基于多元线性回归、决策树、SVM、朴素贝叶斯等经典机器学习算法进行的数据挖掘结果的可视化。用户可以根据需求选择数据集和算法，并通过该数据挖掘可视化系统轻松理解数据挖掘所得信息，发现其中的隐藏特征，特征相关度，算法模型和数据发展趋势等。

开发时间为 2020.7.25 星期六到 2020.8.1 星期六。

## 1.2 项目计划

### 1.2.1 项目阶段

|  |  |
| --- | --- |
| **阶段** | 要求 |
| 构思阶段 | 完成产品概念构思，进行分工并搭建基本框架 |
| 完善阶段 | 对产品具体功能进行实现，完成基本任务 |
| 优化阶段 | 对产品性能及外观进行优化并实现额外功能 |

### 1.2.2 项目参与人员

|  |  |  |
| --- | --- | --- |
| 姓名 | 组别 | 学历 |
| 许继元 | 数据挖掘组 | 本科 |
| 程络 | 数据挖掘组 | 本科 |
| 黄倬熙 | 数据挖掘组 | 本科 |
| 张平路 | 数据挖掘组 | 本科 |
| 刘浩斌 | 前端组 | 本科 |
| 邓志聪 | 设计组 | 本科 |

### 1.2.3 项目进度

|  |  |  |  |
| --- | --- | --- | --- |
| **项目模块** | **模块描述** | **时间安排** | **参与人员** |
| **Git学习** | 学习Git的多人协作功能 | 7.26号上午至中午 | 全体人员 |
| 服务器后端接口 | 与前端进行沟通的中间模块，为数据的吞吐进行调度 | 7.26号下午至7.28号上午 | 程络，刘浩斌 |
| 网页视觉设计 | 设计前端网页的美化方法 | 7.26号上午至7.28号下午 | 邓志聪 |
| 数据挖掘算法及算法可视化实现 | 使用Python实现算法并对决策过程可视化 | 7.26号下午至7.29号下午 | 黄倬熙 |
| 数据挖掘算法及可视化实现 | 实现算法并将高维数据可视化 | 7.26下午至7.29号下午 | 张平路 |
| 数据挖掘算法及其结果可视化 | 实现算法并可视化数据挖掘结果 | 7.26下午至7.29号下午 | 许继元 |
| 辅助小组成员完成答辩相关文件 | 完成答辩所需文档，演示文稿和演示视频 | 7.31号上午至8.1号下午 | 全体人员 |

# 项目进度日结

## 2.1 第一天

|  |  |
| --- | --- |
| **内容** | **参与人员** |
| 部署服务器，前端后端接口测试 | 程络，刘浩斌 |
| 完善支持向量机算法 | 黄倬熙 |
| 设计前端网页以及美化方法 | 邓志聪 |
| 测试KNN算法 | 张平路 |
| 完成对数几率回归算法代码实现 | 许继元 |

## 2.2 第二天

|  |  |
| --- | --- |
| **内容** | **参与人员** |
| 重构API，添加Python的DataFrame与JavaScript的数组格式转换的中间形式 | 程络 |
| 实现支持向量机算法可视化 | 黄倬熙 |
| 设计前端网页以及美化方法 | 邓志聪 |
| KNN算法可视化实现 | 张平路 |
| 完成对数几率回归算法结果可视化 | 许继元 |

## 2.3 第三天

|  |  |
| --- | --- |
| **内容** | **参与人员** |
| 完成后台数据集与模型的调度系统，部署朴素贝叶斯模型，修补bug | 程络 |
| 实现决策树算法可视化 | 黄倬熙 |
| 设计前端网页以及美化方法 | 邓志聪 |
| 制作网页 | 刘浩斌 |
| 提出高维数据可视化方法 | 张平路 |
| 研究连续值变量的高维可视化方法以及多元线性回归算法的提交 | 许继元 |

## 2.4 第四天

|  |  |
| --- | --- |
| **内容** | **参与人员** |
| 完成radViz、Andrew\_Curve、平行坐标图高维化图像的显示模块，修补bug | 程络，刘浩斌 |
| 选择适当的数据集训练算法，并保存其可视化图片 | 黄倬熙 |
| 与前端协同合作 | 邓志聪 |
| 完成iris数据集的高维可视化以及housing数据集的可视化 | 许继元 |

## 2.5 第五天

|  |  |
| --- | --- |
| **内容** | **参与人员** |
| 完成对自己所选的数据集简单的可视化 | 黄倬熙 |
| 与前端协同合作 | 邓志聪 |
| 协作前端改进与了解数据挖掘过程，为可视化提供接口 | 程络，刘浩斌 |
| 修改已上传的代码 | 张平路 |
| 根据后端接口修改代码架构 | 许继元 |

## 2.6 第六天

|  |  |
| --- | --- |
| **内容** | **参与人员** |
| 改进已完成的任务，对项目雏形提出建议 | 黄倬熙、张平路、许继元 |
| 完成网页制作并添加功能 | 刘浩斌、邓志聪 |
| 协作前端改进与了解数据挖掘过程，修补后台bug，完成朴素贝叶斯的可视化接口 | 程络、刘浩斌 |

## 2.7 第七天

|  |  |
| --- | --- |
| **内容** | **参与人员** |
| 修补bug，加强前端网页和后台的鲁棒性，完全部署前端网页 | 程络，刘浩斌 |
| 辅助小组成员完成答辩相关文件 | 黄倬熙、邓志聪、张平路、许继元 |
| 制作项目演示视频 | 许继元 |