

软件需求规格说明1.0.0

软件需求规格说明1.0.0

- 版本说明
- 最开始提出的需求
 - 研究背景
 - 主要功能
- 驳回的需求
- 参与者
- 用例
 - 用户
 - 普通用户
 - 超级管理员
- 用例图
- 细化用例

版本说明

此版本为最初版本

最开始提出的需求

研究背景

大数据分析工具平台作为大数据分析挖掘的基础工具组件

基于 Hadoop 体系架构构建

其主要功能是为各类大数据分析提供分析所需的

数据集成与选择、数据预处理、数学建模、参数配置、结果分析与展现等功能

满足大数据分析挖掘的需求

主要功能

1. 项目管理配置功能

分析项目配置功能主要完成对项目的创建、维护与管理

包括：新建项目、打开项目、修改项目、删除项目、保存项目与退出功能

2. 数据配置功能

平台中分析数据主要来源于数据中心的各类数据

跨部门数据以及统计年鉴数据等

因此，用于分析的数据可以从数据库（Oracle，SQL Server，MySQL，Hdfs，Hbase）导入

或从文档（Excel，Txt）中导入以及手工录入等方式

用于分析的数据一般有：全部数据，部分数据或增量数据

用于分析的变量可能是一个或多个

因此，在录入时应考虑为多维数据分析情形

在分析数据配置时应确定是单变量还是多变量回归分析问题

如果是多变量回归分析问题

还应该确定哪个是自变量，哪个是因变量

在分析数据配置中，还应该对分析结果保存进行配置

如保存在何处（数据库、文档保存目录）

以何种方式进行保存，是数据库还是文档，是追加还是覆盖，亦可是新建等

3. 分析模型配置功能

分析模型配置功能主要为各个分析主题进行分析模型的配置

配置方式分为：自动配置，手动配置两种方式

1. 自动配置

自动配置分析模型主要是根据各个分析主题关联变量的多少

结合各个分析模型的特点

自动地为每个分析主题配置相应的分析模型

各个分析模型的参数为模型的初始参数

在分析模型参数设置中设置

2. 手动配置

手动配置则是根据选定的分析主题

为每个分析主题配置合适的分析模型

并对每个分析模型设定相应的参数

4. 预测分析配置功能

预测分析功能菜单主要包括：

数据校验、批量预测、主题预测、预测设置与详细分析

5. 系统管理

1. 用户管理，包括权限、密码管理与设置

2. 退出

3. 帮助

驳回的需求

具体需求：

项目管理配置功能 —— 保存项目

驳回原因：

在设计该系统时，每一个项目内的主体功能包括四大块：数据集模块（包括导入数据及管理数据集）、模型模块、训练模块、预测分析模块，这四个模块的功能要么完成、要么没完成，不会出现中间需要保存的地方。比如数据导入功能，导入成功了系统就会有记录，导入结果被丢弃或者导入失败两种情况系统内都不会留有任何记录。再比如模型训练功能，当模型训练成功时就会得到一个模型并保存下来，当训练因某些原因失败或终止则系统内不会留有记录。另外像数据集的展示以及模型的展示都是属于一种已经存在的状态，也没有保存的需要。

参与者

用户、普通用户、超级管理员

【说明】该系统的主要用户为普通用户，普通用户在默认情况下可以使用平台提供的全部功能。超级管理员在普通用户的基础上拥有用户管理权限，能够添加用户、删除用户、更改普通用户权限。系统不提供注册功能，所有账号均由超级管理员开通。因为所有用户（包括普通用户和超级管理员）都拥有登录、退出等权限，因此把公有的这些功能抽象为用户这一参与者

用例

用户

【设计思路】

所有用户都可以在系统进行登录、退出，同时可以对个人账号进行设置，也可以查看帮助

【用例识别】

1. 登录
2. 退出
3. 设置
4. 查看帮助

普通用户

【设计思路】

普通用户可以在工作台新建任意数量的项目，每一个项目都拥有各自的项目空间。在每一个项目中，用户都可以对数据集和模型进行管理。

通常在新建一个项目后，用户需要先从数据源导入数据，得到数据集。一个项目可拥有多个数据集，并且允许用户对其中的数据进行添加、修改、删除。用户也可以对数据集进行重命名、删除。

新建的项目默认是没有模型的，用户必须通过自己的数据集新建一次训练以得到模型。在开始训练之前，首先要进行数据配置工作，即用户要选定自变量和因变量。之后用户可以选择系统自动配置模型，或者自己手动配置，但即使是选择系统自动配置，用户仍可以自由变更各类参数。

训练完成后将得到一个模型，用户必须对其选择保留或者丢弃，保留下来的模型可供日后使用。用户可以对对自己的模型进行重命名或者删除。

拥有模型后，用户可以新建一次预测分析。通过选定模型以及数据集中的数据范围，系统将给出分析结果。

【用例识别】

1. 管理项目
2. 导入数据
3. 编辑数据
4. 管理数据集
5. 训练模型
6. 配置数据

- 7. 配置模型
- 8. 管理模型
- 9. 预测分析

超级管理员

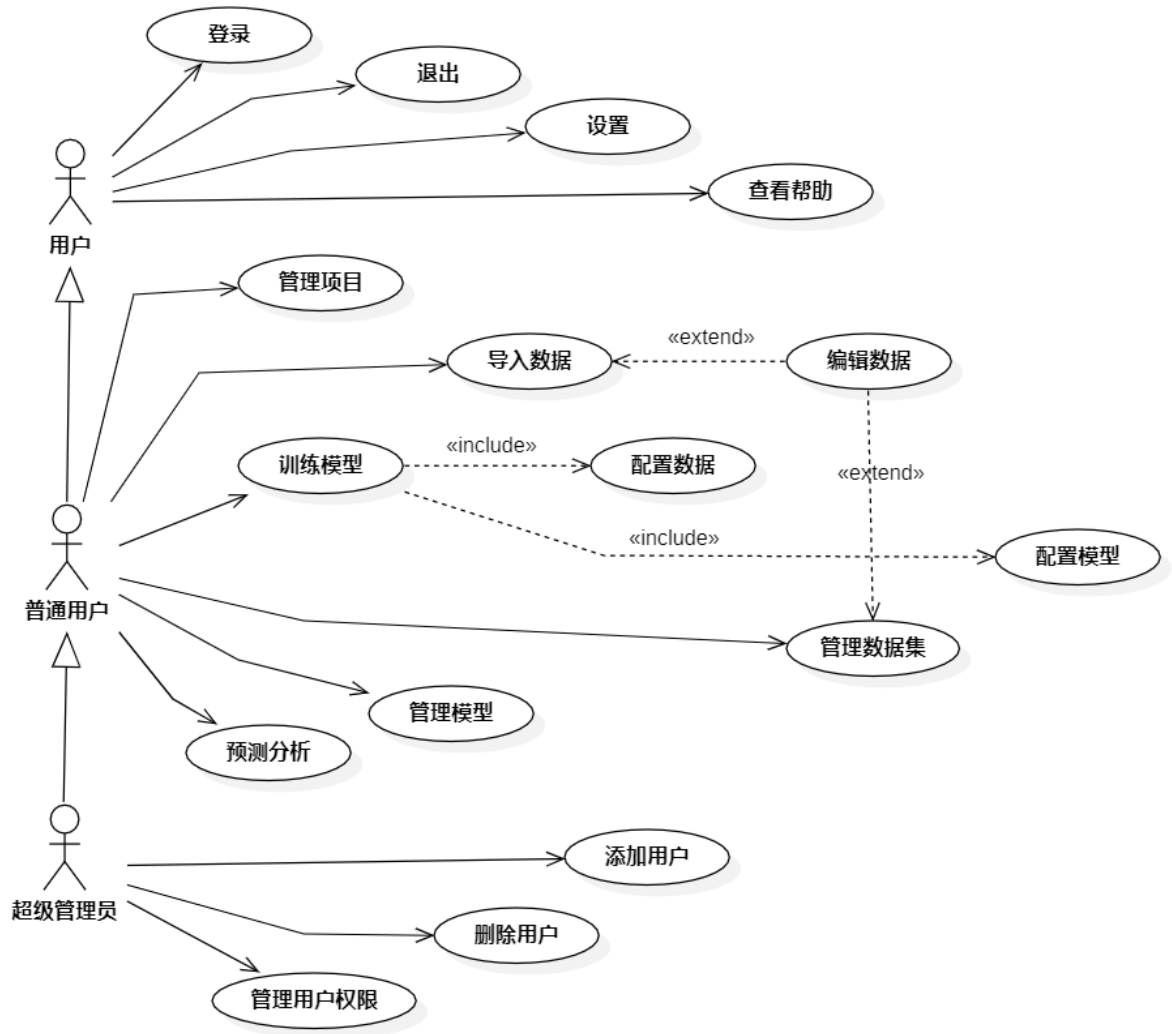
【设计思路】

系统不提供注册功能，所有账户均由超级管理员开通，因此超级管理员将拥有添加及删除用户的功能。另外，超级管理员可以对普通用户的权限进行限制，即可以禁用某一用户的某一功能模块。

【用例识别】

- 1. 添加用户
- 2. 删除用户
- 3. 管理用户权限

用例图



细化用例

- 1. 登录

描述：用户输入账号密码登录平台。

参与者：用户

2. 退出

描述：用户点击退出按钮后退出账号，系统回到登录界面。

参与者：用户

3. 设置

描述：用户可在设置里查看账号信息并修改密码。

参与者：用户

4. 查看帮助

描述：帮助信息用于介绍系统功能，旨在帮助用户快速上手。

参与者：用户

5. 管理项目

描述：

普通用户可以在工作台新建任意数量的项目，每一个项目都拥有各自的项目空间。在每一个项目中，用户都可以对数据集和模型进行管理。当需要进行项目时，用户可以打开指定项目，并在工作台进行操作。用户可以对项目进行重命名。当用户删除一个项目时，系统将自动把所有隶属该项目的数据（包括数据集及模型）删除。

参与者：普通用户

6. 导入数据

描述：

系统允许用户从多个数据源进行数据的导入（包括MySQL数据库以及Excel的.csv表格），系统将对导入的数据进行展示，此时用户可对数据进行编辑。若用户对导入的数据不满意可以进行舍弃，否则系统将要求用户对数据集进行命名并保存。

参与者：普通用户

7. 编辑数据

描述：系统允许用户对数据集中的数据进行修改，包括增加和删除操作。

参与者：普通用户

拓展：导入数据、管理数据集

拓展点：当用户希望更改数据集中的数据时

8. 管理数据集

描述：用户可以对数据集进行重命名或者删除，也可以对数据集内的数据进行编辑。

参与者：普通用户

9. 训练模型

描述：在用户拥有数据集后，用户可以新建一次训练并选定数据范围，在进行完数据配置以及模型配置后，系统将自动完成对模型的训练，并询问用户是否保存模型。

参与者：普通用户

包含：配置数据、配置模型

10. 配置数据

描述：在一次训练中，用户需指定自变量和因变量。

参与者：普通用户

11. 配置模型

描述：

这里的模型是指机器学习模型。用户可以选择系统自动配置，此时系统会根据问题类型和数据规模大小自动推荐算法，并且各类参数也会默认填写好，但在正式开始训练前，用户仍拥有修改算法和参数的能力；用户也可以进行手动配置，即自行选择算法并填写参数。

参与者：普通用户

12. 管理模型

描述：用户可以查看项目里的模型并对其进行重命名或者删除。

参与者：普通用户

13. 预测分析

描述：

在拥有模型后，用户可以新建一次预测分析，选择指定的模型、数据范围，并填写合适的自变量与因变量，系统就会自动进行预测分析。得到的结果会与数据集中的数据进行比对，从而得到准确率，同时系统也会以图形化的方式对结果进行展示。

参与者：普通用户

14. 添加用户

描述：

系统不提供注册功能，所有账户均由超级管理员开通。添加用户时，超级管理员需要指定新的用户名及密码。默认状态下新开通账户可以使用所有功能模块，超级管理员也可以对默认权限进行修改。

参与者：超级管理员

15. 删除用户

描述：超级管理员可以删除用户，届时，隶属于该用户的所有项目均会被删除。

参与者：超级管理员

16. 管理用户权限

描述：

普通用户可以使用的主体功能模块总共有四大块：数据集模块（包括导入数据及管理数据集）、模型模块、训练模块、预测分析模块，这四大块内容在默认情况下都是对普通用户开放的，但是超级管理员可以选择禁用某一用户的某一模块。另外，超级管理员不能对用户的基本权限进行限制，即登录、退出、修改密码等基本功能。

参与者：超级管理员