

**教育平台线上课程用户行为数据分析与展示系统**

**需求规格说明书**

Wisdom Group开发小组

**2022年6月**

**目录**

[1. 引言 3](#_Toc32494)

[1.1 编写目的 3](#_Toc12655)

[1.2 范围 3](#_Toc27212)

[1.3 预期读者和阅读建议 3](#_Toc4042)

[1.4 术语和缩略语 4](#_Toc11139)

[1.5 文档约定 4](#_Toc32365)

[1.6 参考资料 4](#_Toc29530)

[2. 项目概述 5](#_Toc27589)

[2.1 项目基本描述 5](#_Toc23647)

[2.1.1 项目背景 5](#_Toc7852)

[2.1.2 项目目标 5](#_Toc30313)

[2.1.3 系统功能概述 5](#_Toc17019)

[2.2 用户特点 5](#_Toc5847)

[2.3 假定条件和约束限制 6](#_Toc6333)

[2.4 运行环境 6](#_Toc23568)

[2.4.1 软件环境 6](#_Toc9524)

[2.4.2 硬件环境 6](#_Toc1501)

[3. 业务分析 6](#_Toc3926)

[3.1 系统构成图 6](#_Toc4481)

[3.2 系统业务简要分析 7](#_Toc3182)

[4. 功能描述 7](#_Toc16329)

[4.1 功能划分 7](#_Toc2)

[4.2 功能需求总述 8](#_Toc7273)

[4.2.1 功能需求总表 8](#_Toc11311)

[4.2.2 角色、权限需求 8](#_Toc6490)

[4.3 数据集获取 8](#_Toc23063)

[4.3.1 获取数据集 8](#_Toc9752)

[4.4 数据集清洗 9](#_Toc25241)

[4.4.1 清洗获取到的数据集 9](#_Toc10743)

[4.5 数据集分析 9](#_Toc488)

[4.5.1 数据集深度分析 9](#_Toc16866)

[4.6 查阅用户信息概览 9](#_Toc29887)

[4.6.1 查看用户信息概览功能 9](#_Toc25915)

[4.6.2 查找各个省份数据 10](#_Toc5181)

[4.7学习完成度深入挖掘功能 10](#_Toc29381)

[4.7.1 查看完成度影响数据 10](#_Toc3511)

[4.7.2 筛选学习完成度数据 11](#_Toc28081)

[4.8查看活跃用户查看活跃用户并预测未来用户数量功能 11](#_Toc32256)

[4.8.1 查看活跃用户 11](#_Toc13388)

[4.8.2 预测用户活跃量功能 11](#_Toc13297)

[4.9 分析市场热度 12](#_Toc1244)

[4.9.1 查看市场热度概况 12](#_Toc15417)

[4.9.2 查看对应地区、院校热度详情 12](#_Toc2738)

[4.10 探索原始数据功能 12](#_Toc10438)

[5. 非功能需求 13](#_Toc15483)

[5.1 性能需求 13](#_Toc22589)

[5.2 安全保密需求 13](#_Toc24139)

[5.3 可拓展性需求 13](#_Toc26196)

[5.4 稳定性需求 13](#_Toc19722)

[6. 运行需求 14](#_Toc29729)

[7. 接口需求 14](#_Toc9760)

[8. 其他需求 14](#_Toc12170)

1. 引言
   1. 编写目的

本文档的目的是详细的介绍教育平台线上课程用户行为数据分析与展示系统所包含的需求，以便客户及潜在客户能够确认产品的确切需求以及开发人员能够根据需求设计编码，以下叙述将结合文字描述以及数据流图、ER图等图来描述教育平台线上课程用户行为数据分析与展示系统功能、性能、用户界面、运行环境、外部接口以及针对用户操作给出的各种响应。

* 1. 范围

具有一定的实用性，操作简单，设计界面简洁、易懂、直观。

* 1. 预期读者和阅读建议

本文档的预期读者为客户及潜在客户、项目负责人、Scrum敏捷教练、开发工程师、需求工程师以及其他可能和该项目相关的其他项目干系人。

预期读者的阅读建议如下：

1. 项目负责人：该项目说明书很好的举出了该系统的功能需求，以及关于开发该项目时所用到的专业术语缩略语，有助于用户及潜在用户理解该项目，但有一处不足是没有详细的写出所有的功能，希望后期能够改进。
2. 测试人员：该说明书列举出了该系统的功能需求，方便测试。
3. 文档编写人员：文档排版比较合理。
4. 开发人员：该说明书很好的体现了做该项目的各项需求。
5. 用户：该说明书的各项写的都非常详细，大多数不理解的地方都可以进行查阅。
   1. 术语和缩略语

爬虫：用户收集网络数据集的工具。

数据集：10000~10000000项的大量数据集合。

* 1. 文档约定

项目严格遵守软件项目开发相关国家标准。《需求规格说明书》明确了教育平台线上课程用户行为数据分析与展示系统的功能性需求和非功能性需求。该说明书必须使得系统开发人员与用户的需求达成统一的、无二义性的认识，所描述的内容，可以作为系统确认测试的依据。

* 1. 参考资料

《计算机软件文档编制规范》GB/T 8567-2006

《软件工程术语》GB/T11457—1995

《计算机软件质量保证计划规范》GB/T12504—1990

《计算机软件配置管理计划规范》GB/T12505—1990

《计算机软件分类与代码》GB/T13702—1992

《计算机软件产品开发文件编制指南》GB/T8567—1988

《计算机软件需求说明编制指南》GB/T9385—1988

《计算机软件测试文件编制规范》GB/T9386—4988

《软件维护指南》GB/T14079—1993

《软件文档管理指南》GB/T16680—1996

《软件支持环境》GB/T15853—1995

《项目章程》2022-6-13，周昱琪

《项目管理计划》 2022-6-14，王松涛

《软件工程方法与实践》2016，北京：机械工业出版社，窦万峰

1. 项目概述

2.1 项目基本描述

2.1.1 项目背景

教育平台线上课程用户行为数据分析与展示系统适用于网上教育平台数据分析，由Wisdom Group进行开发。市面上大部分数据分析可视化软件虽然能够有效的帮助到用户，但是同时数据易读性较差，用户友好型差，所以我们教育平台线上课程用户行为数据分析与展示系统，主打内容为数据分析与预测。

2.1.2 项目目标

教育平台线上课程用户行为数据分析与展示系统是一个专注于收集、清洗、分析预测数据的软件系统，为各种网上教育平台的大量数据提供分析预测服务。本产品强调分享、交流和学习。教育平台线上课程用户行为数据分析与展示系统具有高准确率、高分析正确率且对用户友好的特点。

2.1.3 系统功能概述

教育平台线上课程用户行为数据分析与展示系统具有数据获取、数据清洗、数据分析、数据处理、数据预测等功能。

2.2 用户特点

本项目最终产品为一个大数据采集分析处理系统，主要面向的对象为网上教育平台数据。用户主体所拥有数据量大、数据清洁度差、数据复杂等特点。

2.3 假定条件和约束限制

硬件设施约束：最低Inter i3级处理器，500MB以上运行内存以进行数据处理与计算。最低需求300MB存储空间用于数据存储

软件使用者素质要求：要求软件使用者最好了解大数据分析行业，尽量有数据分析基础，了解神经网络数据预测，对于计算机最基础的知识要了解。

约束条件：

1. 编程语言的约束
2. 工具约束
3. 性能约束
4. 终端响应时间，任务切换时间等
5. 特殊场景约束

2.4 运行环境

2.4.1 软件环境

1. 操作系统：Windos 7.0及以上系统，Liunx 3.0及以上内核。

2. Python语言编译环境，r语言编译环境。

3. google浏览器或其他主流浏览器

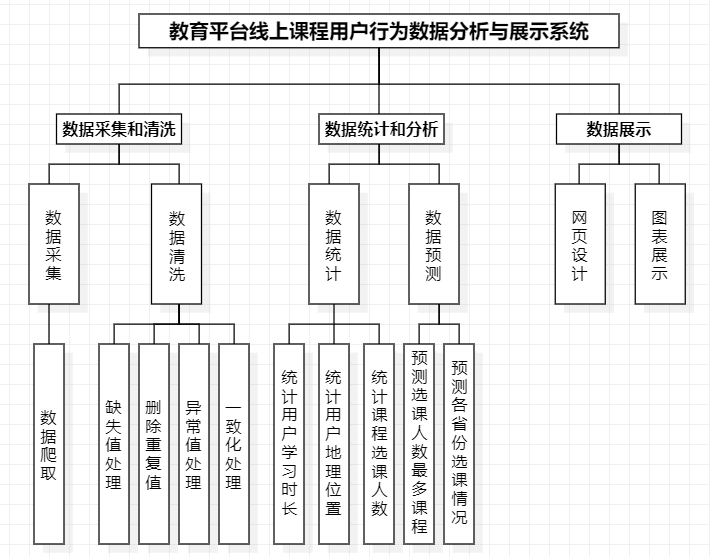
2.4.2 硬件环境

1.最低Inter i3级处理器，500MB以上运行内存以进行数据处理与计算

2.最低需求300MB存储空间用于数据存储

1. 业务分析

3.1 系统构成图



3.2 系统业务简要分析

我们不难看出，整个系统由多个部分构成，其中主要的部分为数据展示系统、数据采集和清洗系统、数据统计和分析系统的实现和交互，涉及到不同功能。我们会将大量的人力和时间用于这项工作。

1. 功能描述

4.1 功能划分

4.1.1 系统功能组成

系统大体功能如下：

1. 数据集获取
2. 数据集清洗
3. 数据集分析
4. 查看用户信息概述
5. 学习完成度深入挖掘
6. 用户活跃度预测
7. 分析热度

4.1.2 功能编号和优先级

1. 数据集获取——高

2. 数据集清洗——高

3. 数据集分析——高

4. 查看用户信息概述——高

5. 学习完成度深入挖掘——高

6. 用户活跃度预测——高

7. 分析热度——高

4.2 功能需求总述

4.2.1 功能需求总表

|  |  |  |
| --- | --- | --- |
| **编号** | **功能需求名称** | **优先级别** |
| SRS\_01 | 数据集获取 | 高 |
| SRS\_02 | 数据集清洗 | 高 |
| SRS\_03 | 数据集分析 | 高 |
| SRS\_04 | 查看用户信息概述 | 高 |
| SRS\_05 | 学习完成度深入挖掘 | 高 |
| SRS\_06 | 用户活跃度预测 | 高 |
| SRS-07 | 分析热度 | 高 |

4.2.2 角色、权限需求

|  |  |  |  |
| --- | --- | --- | --- |
| **序号** | **角色** | **功能** | **权限** |
| P1 | 用户 | SRS\_01——SRS\_07 | 数据处理并查阅 |

4.3 数据集获取

4.3.1 获取数据集

【功能概述】

从网络服务器爬取数据集。

【操作者】

用户

【展示界面】

CMD窗口。

【输入输出】

输入：命令

输出：爬取到的数据集

4.4 数据集清洗

4.4.1 清洗获取到的数据集

【功能概述】

对数据进行清洗并初步预览。

【操作者】

用户

【展示界面】

界面展示方式：CMD窗口。

【输入输出】

输入：命令

输出：初步处理后的数据集

4.5 数据集分析

4.5.1 数据集深度分析

【功能概述】

对数据集进行深度分析。

【操作者】

用户

【展示界面】

界面展示方式：可视化界面。

【输入输出】

输入：命令

输出：深度分析后的可视化界面

4.6 查阅用户信息概览

4.6.1 查看用户信息概览功能

【功能概述】

点击用户信息概览页面进行查看。

【操作者】

用户

【展示界面】

界面展示方式：图形、网格、文字信息、按钮。

【输入输出】

输入：无

输出：用户信息概览

4.6.2 查找各个省份数据

【功能概述】

根据系统提示再页面上方检索特定省份。

【操作者】

用户

【展示界面】

界面展示方式：图形、网格、按钮。

【输入输出】

输入：省份名称

输出：各省份数据

4.7学习完成度深入挖掘功能

4.7.1 查看完成度影响数据

【功能概述】

点击学习完成度关系挖掘页。

【操作者】

用户

【展示界面】

界面展示方式：图形、文字、按钮。

【输入输出】

输入：点击

输出：学习完成度图标

4.7.2 筛选学习完成度数据

【功能概述】

点击页面上方选项框进行学习完成度数据。

【操作者】

用户

【展示界面】

界面展示方式：图形、文字、按钮。

【输入输出】

输入：点击

输出：经筛选过的学习完成度数据

4.8查看活跃用户查看活跃用户并预测未来用户数量功能

4.8.1 查看活跃用户

【功能概述】

点击活跃用户数与预测页面，可以查看活跃用户数据随时间变化图并预测未来活跃用户数的变化。

【操作者】

用户

【展示界面】

界面展示方式：图形、文字、按钮。

【输入输出】

输入：点击

输出：用户活跃度分析数据

4.8.2 预测用户活跃量功能

【功能概述】

点击数据处理预测元素，会对未来数据进行预测。

【操作者】

用户

【展示界面】

界面展示方式：图形、文字、按钮。

【输入输出】

输入：点击网页元素

输出：预测未来数据

4.9 分析市场热度

4.9.1 查看市场热度概况

【功能概述】

点击进入热度分析页面，可以查看市场热度分析概述。

【操作者】

用户

【展示界面】

界面展示方式：图片、文字、按钮。

【输入输出】

输入：点击

输出：市场热度分析图表

4.9.2 查看对应地区、院校热度详情

【功能概述】

用户可以点击对应网页元素，查看相应地区、院校热度详情。

【操作者】

用户

【展示界面】

界面展示方式：图片、文字、按钮。

【输入输出】

输入：点击网页元素

输出：相应地区、院校热度详情

4.10 探索原始数据功能

4.10.1 探索原始数据

【功能概述】

用户对原始感兴趣时，可以单击进入原始数据探索界面对其具体查看。

【操作者】

用户

【展示界面】

界面展示方式：图片、文字、按钮，表格。

【输入输出】

输入：点击

输出：原始数据

1. 非功能需求

5.1 性能需求

数据处理能力：

支持的终端数量：1

处理数据的大小：100m

时间特性：

响应时间：小于1s

处理时间：小于100s

5.2 安全保密需求

本系统应具备以下的安全及保密功能：

防止数据集收到污染；

防止数据集遭到泄露；

一定程度上保证应用系统安全；

5.3 可拓展性需求

本系统在一定程度上具备一定的可拓展性：

1. 在不同数据集具有移植性。
2. 可以优化算法提高效率。
3. 可以简化系统。

5.4 稳定性需求

系统在一个运行周期内，在一定的压力下（数据处理速度200KB/s），系统的出错几率为万分之一，性能劣化趋势为一年内访问性能劣化千分之一以内。

1. 运行需求

6.1 用户界面

6.1.1 图形要求

本系统界面要求如下：

* 界面中的文字标识，含义准确、清楚、用通用词汇；
* 界面中基本信息部分尽量都采用下拉框选择方式。降低人为误操作错误；减少手工输入工作量；规范基本信息的一致性。
* 对于相互关联的内容，要有数据过滤。比如选择了省份后，单位下拉框中只列出该省内的单位，共操作者选择，降低操作员查找的不方便性。
* 数据量大时候，采用分页浏览的形式。

界面具体要求：

界面理解：界面中的文字标识，含义准确、清楚、用通用词汇，界面元素要提供充分而必要的提示，例如当鼠标移动到某个界面元素时，该界面元素图标旁边应出现功能提示。

界面风格：界面中，同类的界面元素要有相同的视觉感和操作方式，例如所有的按钮的形状、色彩以及对鼠标的响应方式都是一样的。

界面的合理布局：界面的布局应当整洁，界面元素应当在水平或垂直方向对其，行列的间距保持一致。窗体的尺寸要合适，各种按钮控件不能过分拥挤也不能过分宽松。

界面中基本信息部分采用点击选择方式。降低人为误操作错误；减少手工输入工作量；规范基本信息的一致性。

1. 接口需求

教育平台线上课程用户行为数据分析与展示系统。

1. 其他需求

8.1 质量属性

1. 可用性：用户可以安全使用

2. 可靠性：在给定时间内可以大致满足无错运行

3. 可维护性：数据处理结果与错误信息、写进日志

4. 安全性：对数据集进行保密