

文件编号: 奋斗的茄子-SWC2020-20200079

受控状态: ☒受控 ☐非受控

保密级别: ☐公司级 ☐部门级 ☒项目级 ☐普通级

采纳标准: CMMI DEV V1.2



Darwin's window

达尔文之窗

Darwin's window

项目开发文档

Version 3.6.0

2020.6.10

Written by 奋斗的茄子



All Rights Reserved

目录

1 项目概述	1
1.1 项目背景	1
1.1.1 技术背景	1
1.1.2 时代背景	1
1.1.3 万物互融	2
1.1.4 智适应	2
1.2 项目定位	2
1.2.1 应用场景	2
1.2.2 目标人群	2
1.3 项目方案	2
1.4 项目目标	3
1.5 项目价值	3
2 开发计划	3
2.1 最终呈现形式	3
2.2 主要功能描述	4
2.2.1 儿童端：图像信息处理	4
2.2.2 儿童端：知识体系建构	6
2.2.3 儿童端：个性推荐系统	6
2.2.4 家长端：儿童学习轨迹	7
2.2.5 家长端：亲子互动方式	7
2.2.6 家长端：商业价值拓展	8
2.2.7 快应用入口：负一屏	9
2.3 运行环境	9
2.4 验收标准	10
2.5 关键问题	10
2.6 进度安排	10
2.7 开发预算	11
3 可行性分析	11
3.1 技术可行性分析	11
3.2 资源可行性分析	13
3.3 市场可行性分析	13
4 需求分析	14
4.1 数据需求	14
4.1.1 静态数据	14
4.1.2 动态数据	15
4.1.3 数据词典	15
4.1.4 数据采集	16
4.2 功能需求	16
4.2.1 图片信息功能模块	17
4.2.2 知识体系建构功能模块	18

4.2.3 个性推荐功能模块.....	20
4.2.4 儿童学习轨迹功能模块.....	21
4.2.5 亲子互动功能模块.....	21
4.2.6 商品推荐功能模块.....	22
4.3 性能需求.....	22
4.3.1 时间特性.....	22
4.3.2 适应性.....	23
4.4 界面需求.....	23
4.5 接口需求.....	23
4.5.1 硬件接口.....	23
4.5.2 软件接口.....	23
4.6 其他需求.....	24
5 概要设计.....	24
5.1 处理流程.....	24
5.2 总体结构设计.....	26
5.3 功能设计.....	27
5.4 用户界面设计.....	28
5.4.1 UI 设计思路.....	28
5.4.2 儿童端.....	28
5.4.3 家长端.....	32
5.4.4 负一屏.....	34
5.5 数据结构设计.....	35
5.5.1 逻辑结构设计.....	35
5.5.2 物理结构设计.....	35
5.6 接口设计.....	36
5.6.1 外部接口.....	36
5.6.2 内部接口.....	37
5.7 错误/异常处理设计.....	38
5.7.1 错误/异常输出信息.....	38
5.7.2 错误/异常处理对策.....	39
5.8 系统配置策略.....	39
5.9 系统部署方案.....	40
5.10 其他相关技术与方案.....	41
5.10.1 跨域资源共享.....	41
5.10.2 错误日志.....	41
6 数据库设计.....	42
7 详细设计.....	43
7.1 智能语音处理功能模块.....	43
7.1.1 功能描述.....	43
7.1.2 性能描述.....	43
7.1.3 输入.....	44
7.1.4 输出.....	44
7.1.5 程序逻辑.....	44
7.1.6 限制条件.....	44

7.2 个性化推荐功能模块.....	44
7.2.1 功能描述.....	44
7.2.2 性能描述.....	45
7.2.3 输入.....	45
7.2.4 输出.....	45
7.2.5 程序逻辑.....	46
7.2.6 限制条件.....	46
7.3 图像识别模块.....	46
7.3.1 功能描述.....	46
7.3.2 性能描述.....	46
7.3.3 输入.....	47
7.3.4 输出.....	47
7.3.5 程序逻辑.....	47
7.3.6 限制条件.....	47
7.4 风格迁移模块.....	47
7.4.1 功能描述.....	47
7.4.2 性能描述.....	48
7.4.3 输入.....	48
7.4.4 输出.....	48
7.4.5 程序逻辑.....	49
7.4.6 限制条件.....	49
7.5 亲子互动推荐功能模块.....	49
7.5.1 功能描述.....	49
7.5.2 性能描述.....	49
7.5.3 输入.....	50
7.5.4 输出.....	50
7.5.5 程序逻辑.....	50
7.5.6 限制条件.....	51
7.6 商品推荐功能模块.....	51
7.6.1 功能描述.....	51
7.6.2 性能描述.....	51
7.6.3 输入.....	51
7.6.4 输出.....	51
7.6.5 程序逻辑.....	52
7.6.6 限制条件.....	52

记录更改历史

序号	更改原因	版本	作者	更改日期	备 注
1	创建项目概述模块	V1.0.0	队员 B	2019.11.13	完成项目概述部分 1.1-1.4
2	更新项目概述部分	V1.0.1	队员 B	2019.11.14	完成项目概述 1.5
3	创建开发计划模块	V1.1.0	队员 B	2019.11.19	完成开发计划模块 2.1-2.7
4	创建可行性分析模块	V1.2.0	队员 C	2019.11.29	完成可行性分析 3.1-3.3
5	完善可行性分析模块	V1.2.1	队员 A	2019.12.02	完善技术可行性分析模块
6	创建需求分析模块	V1.3.0	队员 B	2019.12.07	完成需求分析模块 4.1-4.3
7	完善需求分析模块	V1.3.1	队员 B	2019.12.09	完成需求分析模块 4.4-4.6
8	文档校验 1	V1.4.0	队员 D	2019.12.11	对整个文档进行检查、补缺、更新
9	更新文章逻辑	V1.4.1	队员 A	2019.12.14	修改部分段落描述
10	完善用户界面需求	V1.4.5	队员 A	2019.12.17	完善需求分析模块 4.4
11	文档校验 2	V1.5.0	队员 B	2019.12.22	对整个文档进行检查、补缺、更新
12	创建概要设计模块	V2.0.0	队员 A	2022.1.22	完成概要设计模块 5.1-5.3
13	完善概要设计模块	V2.0.1	队员 B	2020.1.28	完成概要设计模块 5.5、5.7-5.9
14	创建数据库设计模块	V2.1.0	队员 B	2020.1.28	完成 6 数据库设计
15	完善概要设计模块	V2.1.1	队员 D	2020.2.15	完成 5.4 用户界面设计
16	完善概要设计模块	V2.1.2	队员 C	2020.2.20	完成概要设计模块 5.6 和 5.10
17	文档校验 3	V2.2.0	队员 D	2020.3.5	对整个 5 概要设计行检查、补缺、更新
18	创建详细设计模块	V2.3.0	队员 A	2020.3.8	完成详细设计模块 7.1-7.2
19	完善详细设计模块	V2.3.1	队员 B	2020.3.10	完成详细设计模块 7.3-7.4

20	完善项目概述	V2.4.0	队员 A	2020.3.15	对 1 项目概述描述进行调整
20	完善开发计划	V2.4.1	队员 C	2020.3.18	完善 2.2 主要功能描述, 修改 2.6 进度安排
21	完善可行性分析	V2.4.2	队员 A	2020.3.20	完善 3 可行性分析
22	完善需求分析	V2.4.2	队员 D	2020.3.24	完善 4 需求分析
23	文档校验 4	V2.5.0	队员 B	2020.3.28	对 1-4 进行检查、补缺、更新
24	文档校验 5	V2.5.1	队员 D	2020.4.1	对 5-7 进行检查、补缺、更新
25	完善项目概述	V2.5.2	队员 A	2020.4.10	对项目概述部分进行调整
26	文档校验	V2.6.0	队员 B	2020.4.21	对整个文档进行检查、补缺、更新
27	完善详细设计模块	V3.0.0	队员 B	2020.4.31	完成详细设计模块 7.1——7.3
28	完善详细设计模块	V3.1.0	队员 C	2020.5.04	完成详细设计模块 7.4——7.6
29	文档校验 6	V3.2.0	队员 D	2020.5.05	对 1-4 进行检查、补缺、更新
30	文档校验 7	V3.2.1	队员 B	2020.5.07	对 5-6 进行检查、补缺、更新
31	完善可行性分析	V3.2.2	队员 A	2020.5.09	完善 1 可行性分析
32	完善项目概述	V3.3.0	队员 D	2020.5.11	对 1 项目概述描述进行调整
33	完善需求分析	V3.3.1	队员 D	2020.5.18	完善 3 需求分析
34	修改开发计划	V3.4.0	队员 C	2020.5.19	完善 2.3 主要功能描述, 修改 2.5 进度安排
35	更改文章逻辑	V3.5.0	队员 A	2020.5.21	修改文章的逻辑
36	文档校验	V3.5.1	队员 C	2020.5.26	校验整个文档的语句
37	检查排版	V3.6.0	队员 C	2020.6.10	检查文档的排版

1 项目概述

1.1 项目背景

1.1.1 技术背景

新时代背景下 5G+理念的提出,将使人工智能(Artificial Intelligence)、物联网(Internet of Things)、大数据(big data)、云计算(cloud computing)等各种技术互融,各种服务及设备互融,从而打破彼此的边界,实现用户、信息、场景的联接;同时,5G技术的出现使数据的传输与处理取得质的飞跃,为人工智能技术提供了更加丰富的应用场景,为用户提供了更快的反应速度和更智能的应用模式,使得“万物互融”成为可能。智能手机因其本身便携性和普适性成为用户主要的多媒体交互终端,必将成为万物互融的核心和枢纽。因此,在5G和Io技术的推动下,万物互融将会以手机为原点向外延伸,使用户在新的技术环境下有更顺畅的智能体验。

让·皮亚杰是儿童心理学、发生认识论的开创者,达尔文之窗基于皮亚杰的教育理论设计。皮亚杰的发生认识论既反对先天论,也反对学习被动论。他根据自己对儿童心理发展的研究,提出了认识的心理发生论。皮亚杰认为认识的心理发生既不是来自先天的遗传,也不是来自对客体的知觉,而是在动作中产生的:即鼓励儿童在探索中主动学习。这是我们灵感的来源,也是项目的科学立足点。

从技术的角度来看,人工智能和大数据领域的技术取得了非常大的发展,许多技术已经落地。智能推荐算法能够针对用户的兴趣不断反馈给用户喜欢的内容,在电子商务和新闻推送等领域得到了广泛应用。计算机视觉中的目标识别技术也已经发展的比较成熟,如“百度识物”、“形色”、“智能识图王”等。同时,语音和文本识别技术也在人机交互领域得到广泛应用。这些技术为实现个性化内容推荐、内容智能生成以及用户行为分析等提供了坚实的支撑。

借助于5G和人工智能相关技术,我们开发了一款基于手机的儿童智能学习应用软件。该软件能够准确地提供给用户个性化的学习内容,并对用户的学习行为和学习效率进行智能分析,很好地解决了个性化内容推荐质量较低、亲子互动性不高以及学习场景单调等现有学习软件存在的问题。

1.1.2 时代背景

随着智能手机的普及,儿童与智能手机的接触已经无法避免,而更多时候手机成为了游戏和观看短视频的机器,这给正处在黄金学习年龄儿童的成长带来了一定的负面影响,增加了家长和孩子间的矛盾进而产生沟通障碍,影响亲子关系。如果能够通过手机让儿童快乐地探索身边的奇妙世界,学习知识,增加亲子关系,那么将产生积极的影响。

本项目利用人工智能技术,将打造了一款**基于万物互融的亲子个性化知识探索快应用——达尔文之窗**,让手机相机变成孩子探索周围奇妙世界的窗口。达尔文之窗应用软件实现了**儿童与身边环境的互融**,构建知识体系**将物与物互联**,提供科学互动方式(如睡前故事、科普知识)**拉近亲子间的距离**,满足了启蒙学习个性化和亲子互动的核心需求。以儿童为中心实现了个性化学习、趣味学习,助力亲子互动,用科技的力量,给儿童一个更好的世界。

1.1.3 万物互融

达尔文之窗立足于皮亚杰理论,鼓励儿童在探索中主动学习,实现了儿童与**身边环境的互融**,构建知识体系将**物与物互联**,提供科学互动方式(如睡前故事、趣味问答,商品推荐)拉近亲子间的距离,实现**亲子互联**。满足了启蒙学习个性化和亲子互动的核心需求,以儿童为中心实现了个性化学习,趣味学习,助力亲子互动,用科技的力量,给儿童一个更好的世界,体现了万物互融,人本世界的思想。

1.1.4 智适应

通过知识图谱技术与推荐系统结合,尊重儿童的宝贵个性,根据儿童的学习轨迹为儿童智能推荐专属的学习内容,构建专属的**智适应学习体系**,为儿童的“因材施教”提供了更科学的可能。

1.2 项目定位

1.2.1 应用场景

我们团队针对儿童教育出现的新形式,将人工智能和大数据的相关技术应用到儿童教育领域。通过快应用平台,使之成为帮助孩子探索周围世界的奇妙窗口,让孩子享受到更加全面、生动、形象的知识体系,并且参与创作,乐在其中,快乐学习,快乐创作。

我们的软件主要针对儿童个性化的探索性学习场景所打造。在具体学习过程中,儿童可以通过手机快应用探索周围各个物品,让物品在儿童眼里具有生命力,有其丰富的内在价值。软件不仅提供了各种物体的基本信息说明、中英文解释、卡通动画、童话故事、语音朗读、科普故事、相关诗词、趣味问答等学习内容,还可以进行个性化内容推荐。通过建立立体的知识体系,帮助儿童丰富知识,培养兴趣,完善对世界的认知,成为儿童探索周围世界的好帮手。同时,家长可以通过软件实时了解孩子学习了什么,对哪些物品感兴趣,查看统计数据分析的儿童偏好,向朋友展示自己孩子的学习路程等,为建立健康的亲子关系边界提供新途径。

1.2.2 目标人群

达尔文之窗的主要的面向人群为学龄儿童,同时针对一些已经沉迷于手机娱乐的幼童。并提供家长端给家长,可查看儿童学习轨迹、实现偏好分析以及亲子互动。

1.3 项目方案

我们通过组建奋斗茄子团队,完成达尔文之窗这一项目的开发。我们团队旨在打造一款基于万物互融的亲子个性化知识探索快应用。该项目使用深度学习,机器学习和数据分析等技术,依托于快应用平台。项目预算约为13000元,项目时长约为5个月。最终呈现方式为可运行的快应用产品。

儿童端:

(1) 对于图像信息处理,该项目通过YOLO v3来实现图片识别,该项目通过快速风格

迁移模型来实现风格迁移, 增加趣味性。

- (2) 对于知识体系建构, 该项目通过 JRNN 事件抽取模型来实现知识图谱的构建, 该项目通过绘声绘色的童话、科普知识、知识汇集、诗词来填充知识体系
- (3) 对于个性推荐系统, 该项目通过知识图谱与协同过滤来实现针对用户的个性化推荐, 包括童话和物体推荐。

家长端:

- (1) 对于亲子互动方式, 该项目通过推荐睡前故事与趣味问答来实现多种亲子互动形式
- (2) 对于儿童学习轨迹, 该项目通过 python 来实现数据的可视化, 通过建立数学模型来实现对孩子数据的分析。
- (3) 对于商业价值拓展, 该项目通过协同过滤来实现智能商品推荐。

1.4 项目目标

达尔文之窗的目标是依据学龄段儿童的学习行为习惯, 为儿童提供一个以快应用为载体的学习平台, 利用 AI, 图像识别等技术识别物体, 联想童话、诗词、科普和相关知识, 建立知识体系, 并为家长提供多种亲子互动形式、通过数据分析了解孩子学习情况。为儿童周围的物品赋予生命并活灵活现地展现在儿童眼前正是万物互融的体现。本项目立足于皮亚杰教育理论, 提供更易于儿童接受的学习方式, 让学习不枯燥。童话和诗词联想, 辅以语音朗读, 精美趣味的插画, 可迁移风格的图像, 寓教于乐, 让孩子在探索中逐步认知世界, 记忆深刻。

1.5 项目价值

随着智能手机和快应用的普及, 孩子与智能手机的接触也越来越多。据相关研究表明, 儿童有强烈的探索周围未知世界的愿望。如何正确满足儿童的好奇心, 并将其向好的方面引导已成为家长关注的重要问题。达尔文之窗的价值在于: 为儿童提供了一个个性化的学习平台, 为家长提供了一个关注孩子成长的新途径。通过绘声绘色的卡通形象和童话故事帮助儿童探索周围世界, 激发儿童想象力, 实现儿童健康的学习, 同时根据儿童的主动点击为儿童建立一套最适合的知识体系; 此外, 通过数据分析可以了解每一个儿童对什么感兴趣, 适合什么, 这一点对家长、甚至社会关注儿童的成长有着重要的辅助意义。随着快应用的发展, 其呈现形态不断变化。它的简单、易用将会让儿童很容易上手使用, 并可以带来巨大的社会价值和商业价值。

2 开发计划

2.1 最终呈现形式

达尔文之窗旨在通过 AI 帮助儿童主动探索周围世界, 帮助儿童个性化学习、快乐学习, 同时让家长更了解儿童。因此需要一个适合儿童的互动性界面来完成一系列操作。我们选择快应用这个平台来实现我们的想法, 并且随着快应用的发展, 我们期望达尔文之窗可以运行在不同的平台之上。在快应用客户端, 我们拍照获取图片信息, 再通过后台服务器分析内容,

识别图片中的物体，联想童话，人声分割，风格迁移，智能推荐，为儿童构建知识体系。达尔文之窗通过快应用的形式封装内部复杂的接口，儿童用户能够拍照探索，进入个性化知识体系学习童话、科普、知识汇集、诗词等知识，并获取个性化推荐；家长用户可以获取儿童学习轨迹的分析，并获取与儿童的科学互动内容。达尔文之窗提供儿童端和家长端两款快应用。儿童端为主题，供儿童学习，建立知识体系。家长端为辅助，供家长了解儿童学习轨迹和提供亲子互动形式。

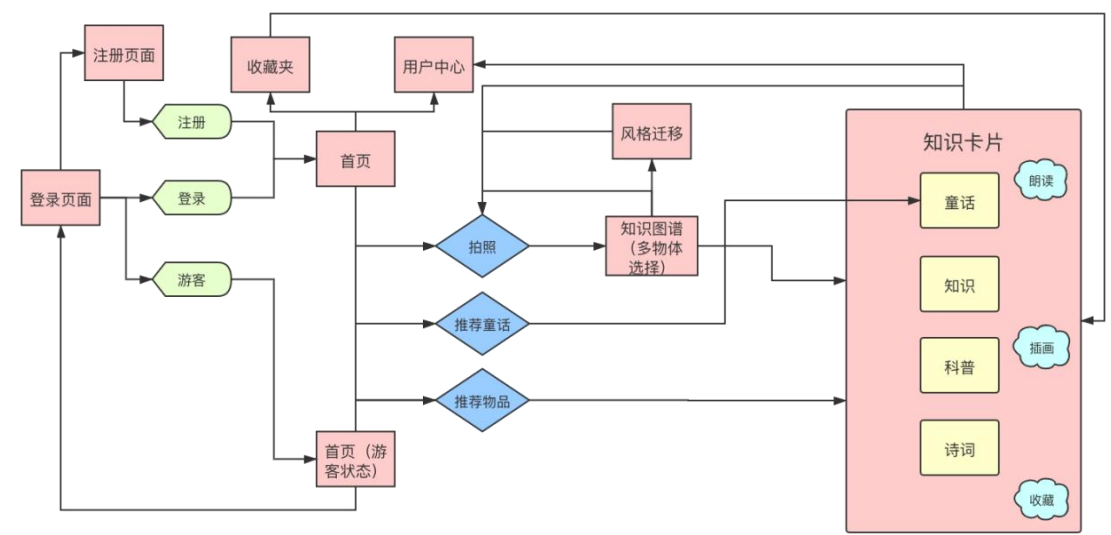


图 2-1 儿童端使用流程

达尔文之窗提供儿童端和家长端两款快应用。儿童端为主体，供儿童学习，建立知识体系。家长端为辅助，供家长了解儿童学习轨迹和提供亲子互动形式。

2.2 主要功能描述

2.2.1 儿童端：图像信息处理

图片识别：

根据儿童拍摄的图片，识别图中的单一或多个物体。孩子可以选择感兴趣的物体并进行下一步探索。

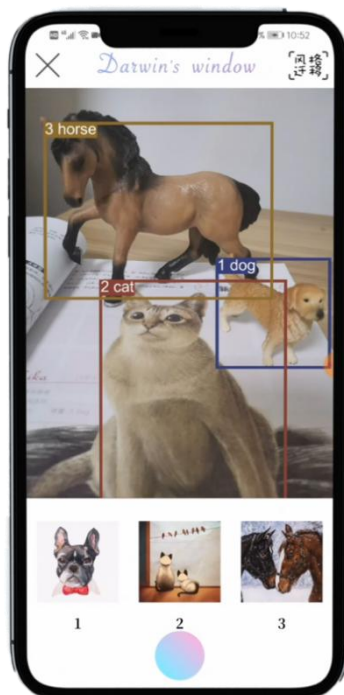


图 2-2 探索界面

风格迁移:

提供多种风格对所拍的照片进行风格迁移, 增加趣味性。



图 2-3 风格迁移

2.2.2 儿童端：知识体系建构

知识图谱：

根据识别得到物体，联想生成知识图谱，展示万物互联关系。

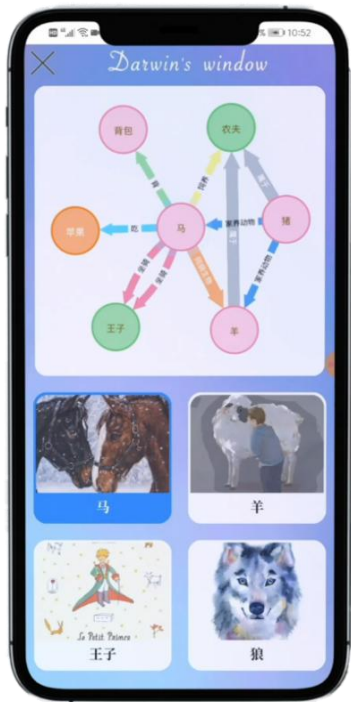


图 2-4 含有知识图谱的联想界面



图 2-5 知识卡片

知识卡片：

进入物体的知识卡片，展示该物体的中英文解释、相关童话、科普知识、古诗词。并提供丰富精美的图片。

为卡片提供收藏功能，方便以后回看。

智能点读：

人声分割，智能点读，利用 AI 为孩子绘声绘色的朗读知识卡片的内容及故事。

2.2.3 儿童端：个性推荐系统

根据儿童的学习数据，在首页个性化推荐相关的物体、童话或寓言故事。并根据孩子是否点击“喜欢”来分析孩子的兴趣分布，便于更加智能的推荐相应卡片。

2.2.4 家长端：儿童学习轨迹

家长可以看到儿童今日的学习轨迹，查看各类物品学习情况，历史统计图表以及对于儿童性格偏好的分析。



图 2-6 数据分析

2.2.5 家长端：亲子互动方式

根据学习轨迹为家长推荐睡前故事以及有关的趣味知识问答，方便亲子交流。

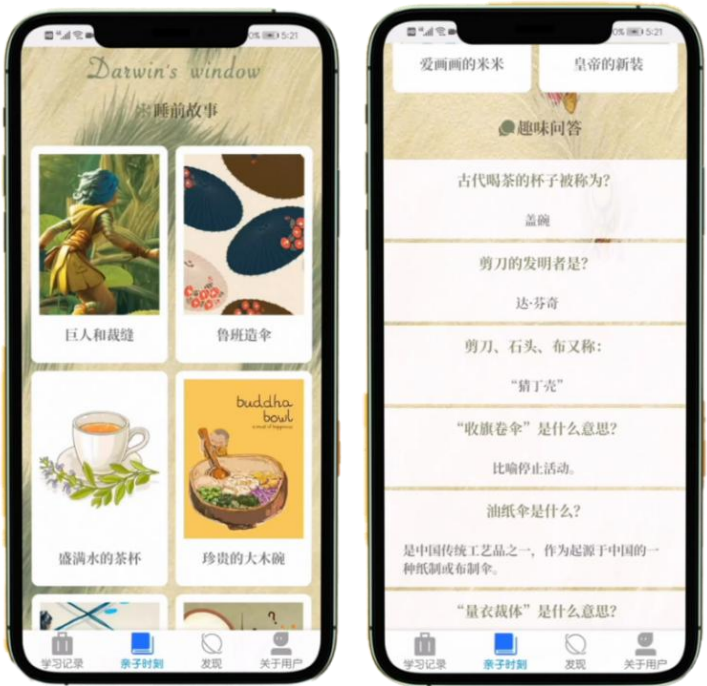


图 2-7 亲子互动方式

2.2.6 家长端：商业价值拓展

根据儿童的兴趣和行为为家长生成相关商品推荐，拓展商业价值。



图 2-8 发现商品

2.2.7 快应用入口：负一屏

为丰富本快应用的进入方式，达尔文之窗提供了负一屏。儿童端负一屏根据用户信息个性化推荐物品，点击物品进入快应用的知识卡片。点击拍照探索进入首页进行拍照。



图 2-9 儿童端负一屏

家长端负一屏为儿童的学习行为提供统计数据。可点击学习轨迹或商品推荐进入对应的快应用页面。



图 2-10 家长端负一屏

2.3 运行环境

项目最终运行环境为网络环境正常的快应用平台。

表 2-1 硬件环境

硬件环境	CPU (Intel)	GPU	Memory	Storage	基准频率
硬件设备 1 (DELL)	Xeon(R)E5-2630 v4	GTX 1080	32GB	256GB SSD 2TB HDD	2.20GHz
硬件设备 2 (RAZER)	Core(TM) i7-9750H	RTX 2070 MAX-Q	32GB	512GB SSD 512GB SSD	2.20GHz
系统环境	Ubuntu 18.04 LTS Windows Server 2019				
应用环境	后端框架: Django 数据库: MySQL + Redis 内存数据库 前端代理服务器: Nginx 文件共享服务器: ftp				

表 2-2 快应用环境

机型	Android	操作系统	快应用框架版本
OPPO Reno 2	9	ColorOS 6.1	1060
HUAWEI nova5 pro	10	EMUI 10	1060

2.4 验收标准

- (1) 代码验收: 完整的项目源码。代码结构清晰, 注释完整合理, 变量命名风格统一。
- (2) 文档验收: 包括项目开发文档, 技术研究文档, 测试文档, 项目创新性报告。
- (3) 服务验收: 项目预期功能应全部实现, 通过稳定性和安全性等测试, 能够保证项目的可靠性。

2.5 关键问题

- 问题 1: 物体检测模型是否能够达到预期的效果。
- ✓ 对策: 目标检测算法近几年取得了很大突破, 其中该项目使用的 YOLOv3 算法就是其中效果较好的一个。该算法训练完成后, 只需将目标导入模型中即可得出结果, 所以算法运行的非常快, 且随着版本的不断更新, 其标注的准确率也越来越高, 且基本涵盖了生活中常见的物体, 能够满足我们的开发需求。
- 问题 2: 构建个性化知识体系是否能够达到预期的效果。
- ✓ 对策: 利用 JRNN 模型进行精确的事件抽取, 构建了合理的知识图谱。并利用知识图谱提供科学的联想。基于协同过滤思想, 利用知识图谱构建物体关系矩阵, 解决用户数据少的问题。
- 问题 3: 风格迁移的速度是否满足要求
- ✓ 对策: 快速风格迁移算法每个模型文件只需要训练一次就可以多次使用, 所以在调用该算法时只需加载模型文件即可, 因此速度极快; 本项目选择这个算法是可以满足速度要求的。
- 问题 4: 快应用交互是否能够流畅友好地进行, 且利于儿童使用。
- ✓ 对策: 采用知识卡片、简单按钮、滑动操作等设置来设计界面, 使得儿童能够简单操作。
- 问题 5: 预期功能是否能够全部应用在快应用平台。
- ✓ 对策: 调查快应用的操作限制, 在限制范围内选择可以实现且符合项目需求的功能。

2.6 进度安排

表 2-3 项目进度日程表

项目重要里程碑	完成日期
组建团队	2019.09.20
前期调研与分析	2019.09.28
项目创意和主题确定	2019.10.20

软件架构及分工确定	2019.10.27
通过学院答辩	2019.11.08
完成初赛文档与视频	2019.11.20
快应用及后端初步搭建	2019.12.15
快应用和后端进一步完善	2020.01.10
图片识别和语言模型训练	2020.01.15
智能推荐模型训练	2020.02.10
文本摘要模型训练	2020.02.19
图片分类模型训练	2020.03.25
系统测试和完善	2020.04.21

决赛进度规划:

表 2-4 决赛进度日程表

进度	开始时间	预计完成时间
完善基础功能	2020.05.01	2020.05.20
集成测试	2020.05.20	2020.05.25
系统测试	2020.05.25	2020.05.30
性能/压力测试	2020.05.30	2020.06.05

2.7 开发预算

开发人员: 4 人

时间: 5 个月

服务器费用: 8000 元

训练用服务器费用: 4000 元

其他费用 (公网 IP, 域名, 内网穿透等): 1000 元

3 可行性分析

3.1 技术可行性分析

(1) 开发平台及软件包

项目基于快应用平台进行系统开发。使用了基于 Python 的 Django 后端, 以便于调用众多 Python 第三方库。同时, 为了训练基于深度学习的物体识别模型和图像风格迁移模型, 采用 Tensorflow 和 PyTorch 深度学习框架实现相关智能学习算法。

(2) 图像识别的可行性

近几年目标检测算法取得了很大的突破,其中最突出的技术之一是最新版本的 YOLO v3。YOLO 算法训练完成后,只需将目标导入模型中即可得出结果,因此算法运行的非常快;其识别的准确率与速度,完全能够满足我们的开发需求,并且训练物体速度较其他网络较快,因此是可行的。

(3) 构建知识体系的可行性

构建知识图谱的难点在于获取三元组,我们采用基于循环神经网络的联合事件抽取(JRNN)方法来提取知识三元组,该方法发表在自然语言处理领域的顶级会议 NAACL2016 上,是非常有影响力的模型。JRNN 解决了泛化能力差和误差传播大的问题,提出了离散特征 BRNN 模型。该模型用两个 RNN 分别正/反向学习一个句子的表示;另一方面,使用了一个记忆向量与两个记忆矩阵来分别存储相关参数以及两者之间的依赖关系。为了训练 JRNN,我们的数据集来源自爬虫获取的百度百科以及百科全书资料来预训练词嵌入表,并使用 ACE 2005 语料库训练和评估事件抽取模型,得到三元组,最后利用 neo4j 数据库进行可视化,并构建知识图谱。因此,构建知识体系的方法是可行的,并且相关方法所涉及的模型训练数据集是可获取的。

(4) 个性化内容推荐的可行性

基于上面所建立的知识图谱,通过协同过滤方法来实现个性化推荐。在协同过滤推荐方法和理论的研究方面,国内外学者们已取得了很多优秀的研究成果,并且这些成果中的一部分已在电子商务和新闻推送等得到了广泛的应用。但这些成果在用户数据较少以及偏斜性较大时效果欠佳,难以获得关系矩阵。我们的解决方案是利用知识图谱,对物体之间的关系进行分类,获取物体间的信息。在此基础上,协同过滤时增加权值构建物体之间的相似度矩阵,再根据用户是否浏览以及是否喜欢代回协同过滤模型训练。在冷启动问题时会为用户推荐最热门的内容。可见,实现个性化内容推荐是可行的。

(5) 语音及文本处理的可行性

将采用的方案是利用 spleeter 和 pydub 来构建智能语音处理模型。pydub 进行语音切割时,需要静默时间来进行句子划分,一般录制的音频会有背景音乐或噪音,难以进行精确的划分,也难以保证语音识别的质量。一个可行的解决方案是利用 2019 年 12 月份由 Deezer 公司开发的基于 tensorflow 的音轨分离项目 spleeter master 对语音进行降噪处理以及人声音轨分离,可以很好的监测到静默时间,进行句子的分割以及语音识别。

(6) 风格迁移的可行性

风格迁移技术是图像处理技术的一种,在近期也取得了很大的成果。我们的项目将使用其中效果较好的固定风格任意内容的快速风格迁移技术,只需将目标导入模型中便可以快速的风格迁移。我们将选择选择合适的图片作为模型训练的风格图片,供用户选择不同的转换风格。据测试,风格转换可以在 2 秒左右完成,速度极快,效果较好,可以达到项目的要求,因此是可行的。

3.2 资源可行性分析

- (1) 人力资源: 开发团队共四人, 有四个月充足的开发时间。项目组成员熟练掌握 c++、Python。一位同学有 web 前端开发经历, 一位同学有后端服务器开发经历, 两位同学有机器学习模型训练经历。
- (2) 计算资源: 项目得到了来自英特尔云平台与学院的支持, 二者为项目提供了足够的计算资源。
- (3) 硬件依赖: 学院拥有可用于本项目研究的多种资源, 包括实验室、软硬件环境与资料等。为项目特别提供了 Dell 图形工作站一台, 硬件资源充足。

3.3 市场可行性分析

国内外已出现如“形色”, “智能识图王”等根据图片描述物体的软件, 虽然对于物体的描述非常详尽, 但提供的信息对于儿童而言复杂冗长, 如同百度百科, 不利于儿童学习。儿童教育软件也大多数按照固定的模式, 无法为孩子个性化的构建知识体系。这些软件对于儿童来说有一些显著的缺点, 而达尔文之窗针对这些问题做了合理改善:

- (1) 市面上的儿童学习软件个性化程度低。如一些儿童故事软件、儿歌软件, 识字软件, 但是大都靠内置好的固定场景让儿童进行操作, 无法针对儿童进行个性化服务, 很难让儿童有持久的兴趣也很难激发儿童的创造力, 而达尔文之窗可以根据儿童的学习轨迹, 为儿童智能推荐与之相关的知识卡片、童话故事, 真正做到“善解人意”。
- (2) 市面上的产品缺乏家长与儿童的有机互动。市面上的教育软件很多都是儿童在操作, 与家长的互动性很少, 达尔文之窗可以让家长看到儿童的学习轨迹, 并且还能收到智能推荐的睡前故事与趣味问答, 增加亲子感情。
- (3) 市场上的软件与环境交互性低。不能与儿童身边的环境进行联系。儿童总是对身边对事情怀有好奇心, 家长不能总是在儿童身边为儿童进行解答, 如果好奇心没有得到及时的满足, 儿童对知识渴望就会有一定的下降。而达尔文之窗可以让儿童通过手机相机与身边的物体或者图画产生联系, 从其中学习到有趣的知识。从身边的事物学习, 充分满足儿童对身边事物的好奇心。
- (4) 市面上软件质量和成本不能兼得。儿童需要了解的知识应该是广泛而简单易懂的, 而且展现的形式还要生动有趣。而达尔文之窗借助快应用这个新平台, 操作简便; 且利用智能语音处理技术, 使得提供的知识不仅形式多样、生动形象, 并且可以让儿童广泛的学习, 激发儿童的联想能力。

达尔文作为一款基于万物互融的亲子个性化知识探索快应用, 目标人群为学龄儿童, 研发理念来源于皮亚杰的儿童教育心理学“鼓励儿童在探索中主动学习”, 并增加亲子互动, 将满足提升亲子幸福感、儿童探索学习的需求, 填补市场空白, 具有较高的市场竞争力, 一定会收到欢迎。

表 3-1 达尔文之窗与提供童话故事软件的对比

评判标准	凯叔讲故事	常青藤爸爸	宝宝巴士	达尔文之窗
亮点:	凯叔自己录制的故事, 声音好听。给家长的一些音频课程, 关于如	主打英文故事和阅读。	是一个综合性的游戏 app。	为儿童构建个性知识体系, 提供童话故事和声情并茂的朗读, 让

	何教育孩子			儿童自己探索世界。并且提供了多种亲子互动方式。
缺点:	只具备讲故事的功能	大部分故事仅对会员开放。	如果作为游戏App, 它没有其他游戏APP做的有趣, 如果作为讲故事的APP, 自己编写的故事内容不够专业, 教育性弱, 娱乐性更多。	个性化创作还有提升的空间
用户体验	功能单一, 不能充分满足用户需求, 用户体验感一般。	趣味性较低, 免费功能较少, 用户体验感一般。	教育性差, 内容不够专业, 用户体验感一般。	知识体系完善, 亲子交互良好, 鼓励个性化探索学习, 用户体验感较好。

4 需求分析

4.1 数据需求

4.1.1 静态数据

静态数据是指在运行过程中主要作为控制或参考用的数据,它们在很长的一段时间内不会变化,一般不随运行而变。

表 4-1 静态数据配置

定位	名称	默认值	说明	格式	类型
数据库配置信息	setting.DATABASES		配置访问数据库的信息	String	String
接口调用信息	URL.py	Path=’ ’	定义访问远程服务器接口地址	String	String
数据库查询语句	SQL	SQL = ‘SELECT * FROM ***’	用于查询数据库	String	String
静态文件保存地址	STATIC_URL	FILE_LOCATION = ‘/static/’	django 静态文件的地址	String	String

文件保存地址	MEDIA_URL	FILE_LOCATION = '/meida/'	设置保存文件的地址	String	String
服务器配置信息	Setting		配置服务器	String	String
日志配置	LOG		配置日志打印信息	json	json

4.1.2 动态数据

动态数据是指在系统应用中随时间变化而改变的数据，如库存数据等。动态数据的准备和系统切换的时间有直接关系。

表 4-2 动态数据

	名称	定义	格式	类型
服务器操作部分	图片信息	根据用户所上传图片，经过算法得到图片信息和童话故事	URL	URL
	图片获取	取得数据库中的图片	URL	URL
	用户数据	用户浏览卡片信息，学习童话故事、科普知识等记录	String	String
	用户收藏	用户收藏卡片记录	String	String

4.1.3 数据词典

User 用户表存储用户信息

Music 存出相应的音频

Images 图片表用户上传的所有图片

Favorite 收藏表存储用户喜欢的图片

表 4-3 用户信息表

字段	描述	类型	允许空值	默认值
id	用户 id	int(10) UNSIGNED	No	
name	用户名	varchar(255)	No	
password	用户密码	varchar(255)	No	
remember_token	找回密码令牌	varchar(100)	Yes	
created_at	创建时间	timestamp	Yes	
updated_at	更新时间	timestamp	Yes	
avatar_url	头像地址	varchar(255)	No	/avatar/default_avatar.png
birthday	生日	date	No	2003-05-21

gender	性别	tinyint(1)	No	0
--------	----	------------	----	---

4.1.4 数据采集

达尔文之窗中，用户的数据来自于用户的注册，登录，上传图片，上传文本等操作，通过快应用与用户交互获取数据。模型训练数据主要来自清华文本分类新闻语料以及我们爬取的各种数据。

4.2 功能需求

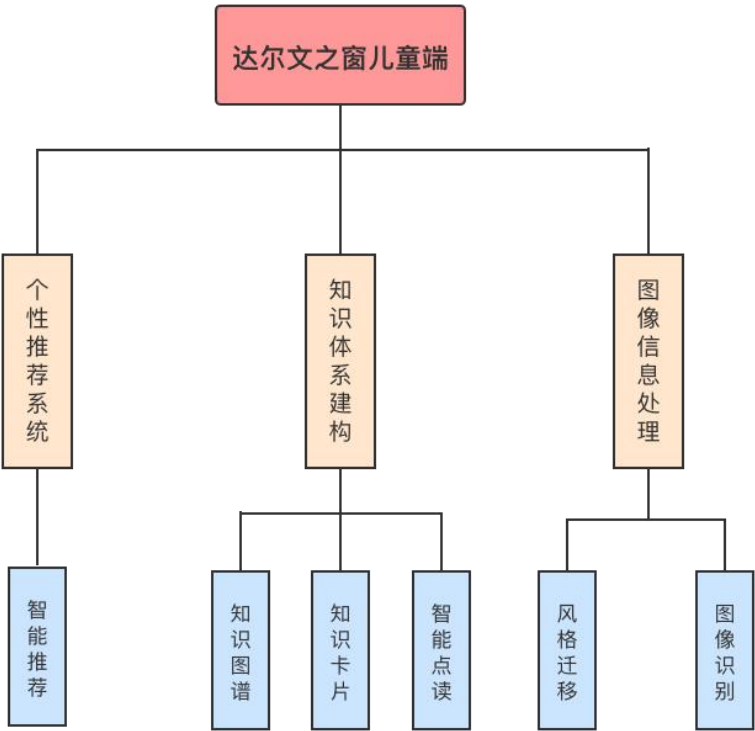


图 4-1 达尔文之窗儿童端

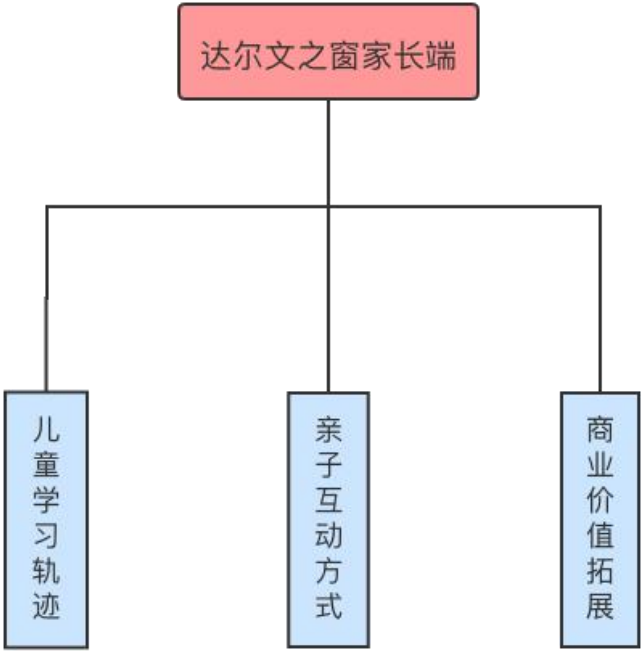


图 4-2 达尔文之窗家长端

4.2.1 图片信息功能模块

表 4-4 图像信息功能模块描述

功能模块	功能	功能描述	优先级
图像信息处理	图像识别	根据儿童拍摄的图片，识别图中的单一或多个物体。孩子可以选择感兴趣的物体并进行下一步探索。	0
	风格迁移	提供多种风格对所拍的照片进行风格迁移，增加趣味性。	1

表 4-5 图像识别用例规约

用例名称	图像识别
功能简述	根据儿童拍摄的图片，识别图中的单一或多个物体。孩子可以选择感兴趣的物体并进行下一步探索。
用例编号	IF001
执行者	Server
前置条件	用户上传图片到服务器
后置条件	服务器处理完图片后，讲数据反馈给快应用
涉众利益	用户：希望得到图片的信息

	图形识别系统：识别图片中的物体特征，做出相应的概率判断
基本路径	① 用户上传目标图片 ② 识别图片细类别
扩展路径	图像无法加载，提示用户
字段列表	图像是否合格(是否符合规则)
设计规则	单独模块
未解决的问题	无
备注	无

表 4-6 风格迁移用例规约

用例名称	风格迁移
功能简述	提供多种风格对所拍的照片进行风格迁移，增加趣味性。
用例编号	IF002
执行者	Server
前置条件	用户上传图片到服务器
后置条件	服务器处理完图片后，将数据反馈给快应用
涉众利益	用户：希望得到迁移风格后的图片 风格迁移系统：切换风格
基本路径	① 用户上传目标图片 ② 风格迁移系统处理，由后台返回成功图片
扩展路径	图像无法加载，提示用户
字段列表	图像是否合格(是否符合规则)
设计规则	单独模块
未解决的问题	无
备注	无
备注	无

4.2.2 知识体系建构功能模块

表 4-7 知识体系建构功能模块描述

功能模块	功能	功能描述	优先级
知识体系建构	知识图谱	根据识别得到物体，联想生成知识图谱，展示万物互联关系。	0

	知识卡片	进入物体的知识卡片，展示该物体的中英文解释、相关童话、科普知识、古诗词。并提供丰富精美的图片。 为卡片提供收藏功能，方便以后回看。	1
	智能点读	人声分割，智能点读，利用 AI 为孩子绘声绘色的朗读知识卡片的内容及故事。	2

表 4-8 知识图谱用例规约

用例名称	知识图谱
功能简述	根据识别得到物体，联想生成知识图谱，展示万物互联关系。
用例编号	E001
执行者	Server
前置条件	拍照识别出的物品
后置条件	服务器调用算法生成相关知识图谱，并返回。
涉众利益	用户：希望得到该物体的知识图谱
基本路径	用户点击感兴趣的物体
扩展路径	图谱无法加载，提示用户
字段列表	图片是否合格
设计规则	单独模块
未解决的问题	无

表 4-9 知识卡片用例规约

用例名称	知识卡片
功能简述	进入物体的知识卡片，展示该物体的中英文解释、相关童话、科普知识、古诗词。并提供丰富精美的图片。为卡片提供收藏功能，方便以后回看。
用例编号	E002
执行者	Server
前置条件	拍照识别的物体
后置条件	服务器依据知识体系，返回该物体相关信息
涉众利益	用户：希望得到物体的知识信息，希望能将卡片保存 知识体系构建系统：返回知识卡片内容
基本路径	用户点击相关物体
扩展路径	图片、文字无法加载，提示用户
字段列表	音频是否合格、图片是否合格，文字是否合格
设计规则	单独模块

未解决的问题	无
--------	---

表 4-10 智能点读用例规约

用例名称	智能点读
功能简述	人声分割, 智能点读, 利用 AI 为孩子绘声绘色的朗读知识卡片的内容及故事。
用例编号	E003
执行者	Server
前置条件	童话故事或者诗词内容
后置条件	服务器调用算法人声分割, 将单行语音返回
涉众利益	用户: 希望得到特定的语音朗读
基本路径	用户点击相应句子
扩展路径	音频无法加载, 提示用户
字段列表	音频是否合格
设计规则	单独模块
未解决的问题	无

4.2.3 个性推荐功能模块

个性推荐	智能推荐	根据儿童的学习轨迹, 在首页个性化推荐相关的物体、童话或寓言故事。并根据孩子是否点击“喜欢”来分析孩子的兴趣分布, 便于更加智能的推荐相应卡片。	0
------	------	--	---

表 4-11 智能推荐用例规约

用例名称	智能推荐
功能简述	根据儿童的学习轨迹, 在首页个性化推荐相关的物体、童话或寓言故事。并根据孩子是否点击“喜欢”来分析孩子的兴趣分布, 便于更加智能的推荐相应卡片。
用例编号	E004
执行者	Server
前置条件	用户学习行为产生的数据
后置条件	服务器根据用户分类得到信息, 对相关内容进行个性化智能推荐
涉众利益	根据用户学习记录, 推荐用户的感兴趣的内容
基本路径	用户历史操作记录、服务器进行相关推荐内容的计算
扩展路径	无
字段列表	无
设计规则	单独模块

未解决的问题	无
--------	---

4.2.4 儿童学习轨迹功能模块

表 4-12 儿童学习轨迹功能模块描述

功能模块	功能	功能描述	优先级
儿童学习 轨迹	学习轨迹	家长可以看到儿童今天的学习轨迹，查看统计数据而对于儿童性格偏好的分析。	0

表 4-13 学习轨迹用例规约

用例名称	学习轨迹
功能简述	家长可以看到儿童今天的学习轨迹，查看统计数据而对于儿童性格偏好的分析。
用例编号	E005
执行者	Server
前置条件	用户学习行为产生的数据
后置条件	服务器根据用户的数据，绘制统计图表，调用算法分析。
涉众利益	家长对儿童的学习轨迹和兴趣偏好进行了解
基本路径	用户历史操作记录
扩展路径	无
字段列表	无
设计规则	单独模块
未解决的问题	无

4.2.5 亲子互动功能模块

表 4-14 亲子互动功能模块描述

功能模块	功能	功能描述	优先级
亲子互动	趣味问答与睡前故事	会并根据学习轨迹为家长推荐睡前故事以及有关的趣味知识问答，方便亲子交流。	0

表 4-15 亲子互动用例规约

用例名称	亲子互动
功能简述	会并根据学习轨迹为家长推荐睡前故事以及有关的趣味知识问答，方便亲子交流。
用例编号	E006

执行者	Server
前置条件	用户学习行为产生的数据
后置条件	服务器根据用户的数据, 对相关内容进行个性化智能推荐
涉众利益	用户希望得到亲子互动相关推荐, 包括相关睡前故事和趣味问答等
基本路径	用户历史操作记录、服务器进行相关推荐内容的计算
扩展路径	无
字段列表	无
设计规则	单独模块
未解决的问题	无

4.2.6 商品推荐功能模块

表 4-16 商品推荐功能模块描述

功能模块	功能	功能描述	优先级
商品推荐	商品推荐	根据儿童偏好, 智能推荐相关商品, 拓展商业价值	0

表 4-17 商品推荐用例规约

用例名称	亲子互动
功能简述	根据儿童偏好, 智能推荐相关商品, 拓展商业价值
用例编号	E007
执行者	Server
前置条件	用户学习行为产生的数据
后置条件	服务器根据用户的数据, 对相关内容进行个性化智能推荐
涉众利益	用户希望得到相关商品推荐
基本路径	用户历史操作记录、服务器进行相关推荐内容的计算
扩展路径	无
字段列表	无
设计规则	单独模块
未解决的问题	无

4.3 性能需求

4.3.1 时间特性

1. 快应用有缓存时少于 50ms, 无缓存时少于 2s;

- 2. 基础的用知识体系构建业务，以下简称“基础业务”，必须在不超过 10s 的时间内响应，90% 响应时间为 5s；
- 3. 用户上传、下载（2200*1080）品质的图片的响应时间不超过 5s；
- 4. 在 100 个并发用户的高峰期，基本业务处理能力至少达到 10 TPS；创作时智能提示等功能响应时间最多延长 1 倍；
- 5. 系统能在高于实际系统运行压力 1 倍的情况下，稳定运行 12 小时；
- 6. 知识体系占有所有业务的 80% 以上，集中在每天的 20% 时间内；
- 7. 事务成功率：事务通过数/（事务通过数+事务失败数）达到 99.999%以上。

4.3.2 适应性

系统应该能动态感知负载变化，自适应调整服务器配置参数，确保应用的服务质量。达尔文之窗应该具有的适应能力包括以下几个方面：

操作方式：

达尔文之窗的用户群体主要为儿童，因此，达尔文之窗的操作方式需要尽可能地简单，且易于儿童使用。

运行环境：

达尔文之窗后端服务器可以部署在 Linux 或 Windows 上，前端快应用运行在快应用联盟的手机厂商上。

接口：

达尔文之窗的开发采用了前后端分离思想，后端服务器为快应用提供必要接口，并对应了相关算法。在开发过程中维护了前端接口文档和算法接口文档。

4.4 界面需求

见 5.4 用户界面设计

4.5 接口需求

4.5.1 硬件接口

无

4.5.2 软件接口

表 4-18 软件接口

接口	功能
百度语音识别	绘声绘色模块中音频识别训练
基于 YOLOv3 的图像识别	识别图片中的物体并实现编号，同时能够返回对应物体的名称列表
自定义风格迁移	用户可以在提供的风格选项中选择，实现指定图片的风格转换

4.6 其他需求

达尔文之窗关注用户使用的便捷性、安全性、可移植性。

便捷性: 达尔文之窗会在未来更新中将编辑器作为一个单独插件来开发，其他开发者可将其嵌入到自己的程序中，享受人工智能带来的便利。

安全性: 系统采用双机热备，基于高可用系统中的两台服务器的热备份，采用主备方式，保证数据的安全，同时提高容错性。有完整的日志系统，能够快速定位出错方位。在开发时要考虑跨站脚本攻击、sql 注入等安全隐患。

可移植性: 采用快应用这一平台，在不同系统环境下提供不同的兼容性方案，保证系统的通用性。当服务器环境发生变化时，不会影响到用户的访问体验。

5 概要设计

5.1 处理流程

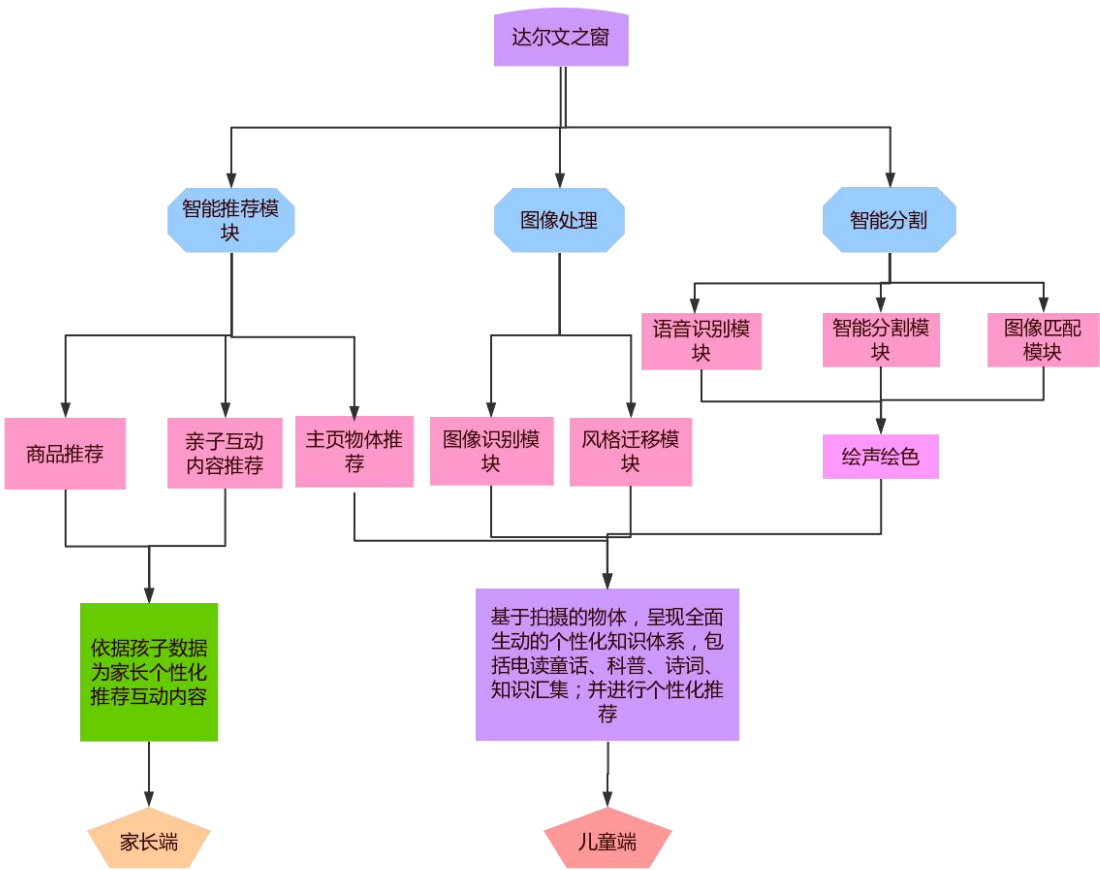


图 5-1 程序流程图

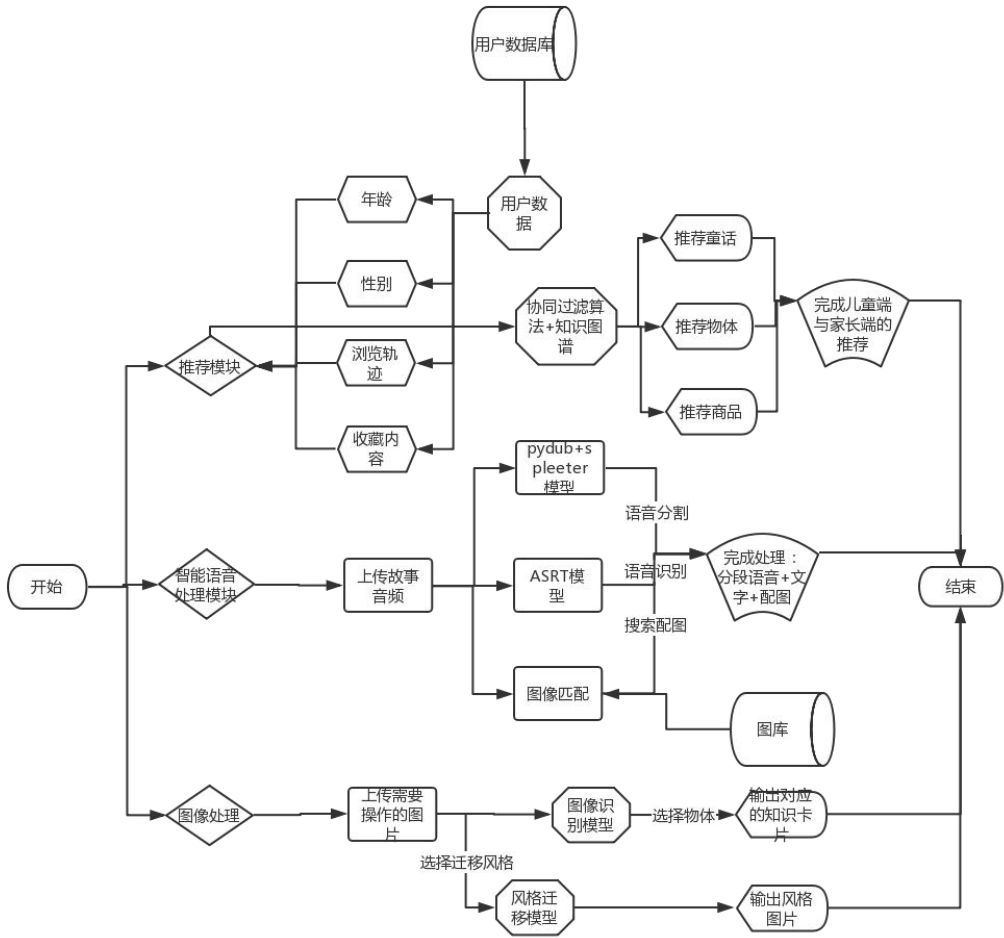


图 5-2 数据流程图

5.2 总体结构设计

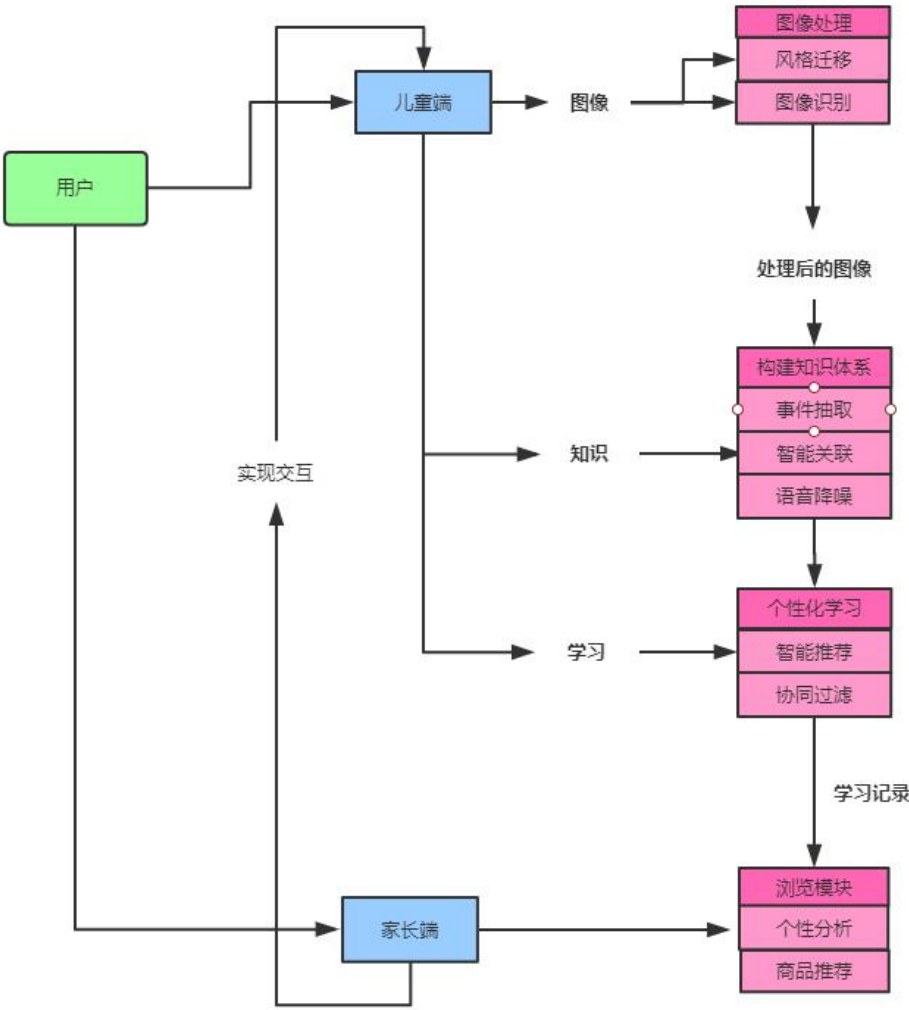


图 5-3 模块架构图

达尔文之窗（Darwin's window）整体架构主要分为儿童客户端和家长客户端，儿童客户端包括图像处理模块、知识体系构建模块、个性化学习模块，家长端包括浏览模块。图像处理模块主要包括物体检测、图像处理、风格迁移等技术；知识体系构建模块主要为 JRNN 事件抽取、智能关联、语音降噪、人声分离，主要涉及推荐技术和自然语言处理技术；个性化学习模块主要包括智能推荐，涉及的技术为协同过滤。浏览模块主要为个性化分析、商品推荐，主要涉及的技术为协同过滤。

5.3 功能设计

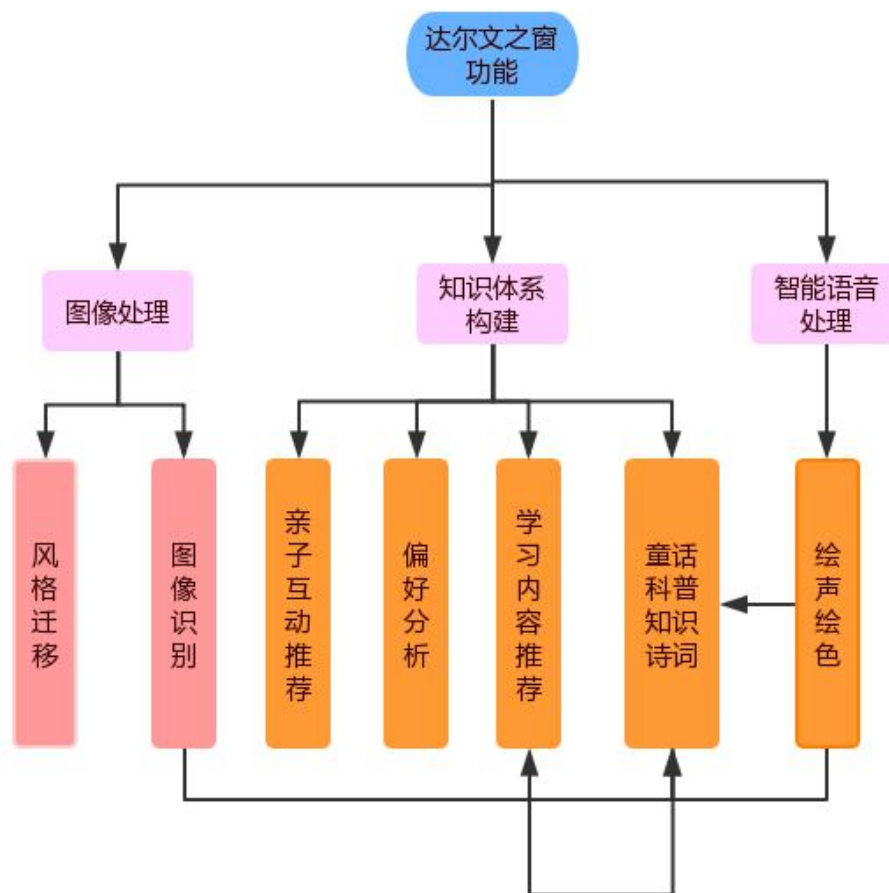


图 5-4 功能模块结构

功能结构整体如上图，主要包括知识体系构建模块、图像处理模块、智能语音处理模块：

- (1) 知识体系构建是为每一个儿童及其家长定制专属的知识体系，个性化推荐学习以及互动内容，学习的内容包括经过智能语音处理的绘声绘色的童话、科普知识以及知识汇集；个性化的学习和互动内容依据协同过滤进行推荐，并依据浏览轨迹给家长呈现孩子的偏好分析。
- (2) 智能语音处理呈现在知识体系框架中的童话模块，将简单的音频转化为绘声绘色的点读呈现，孩子想听哪里点哪里，还可以看到精美的配图，适合孩子学习。
- (3) 图像处理主要包含两个功能，物体检测和风格迁移；物体检测主要是识别儿童拍摄的照片中的物体并进行标注，并且构建照片中的物体列表；风格迁移主要是使用快速风格迁移技术对儿童选择的照片进行风格转换，儿童可在卡通类和油画类共 12 中风格中选择不同的风格，创造出自己喜欢的个性图片。

5.4 用户界面设计

5.4.1 UI 设计思路

- (1) 注重用户定位: 用户群体定位为学龄儿童, 因此整体界面采用卡通设计, 彩色插画居多, 为儿童提供趣味学习环境。设计时注重视觉层次感和美观性。
- (2) 注重交互性: 拍照和对多物品的选择等功能, 让页面有良好的交互性。
- (3) 注重功能性: 每个页面都有一个特定的主题, 功能明确。
- (4) 注重逻辑性: 页面动作主次划分清晰。完整的逻辑链实现特定功能。页面跳转逻辑设计贴合用户使用习惯。
- (5) 注重引导状态: 用户对软件的初体验很重要。设计时页面逻辑考虑到为用户提供方向和指导, 以此来帮助用户快速适应设计。
- (6) 注重效率性: 家长端采用底部 tab 使功能划分清晰, 更加简洁, 提高效率。

5.4.2 儿童端

登陆界面

如果没有存储的 Token, 则会进入这个界面。可以选择登录、注册和以游客方式进入首页

登录后进入此页面, 左上角可进入用户中心, 右上角查看收藏; 相机拍照开启拍照探索; 点击“推荐物品”和“推荐童话”模块分别查看推荐的知识卡片和童话故事



图 5-4 登入界面



图 5-5 首页推荐物品

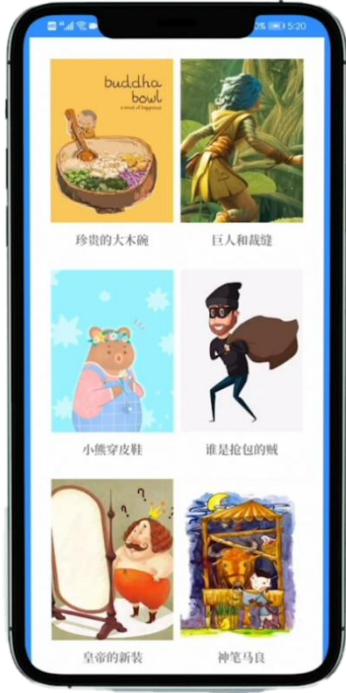


图 5-6 首页推荐童话

探索界面

拍照后进入此界面，点击右上角可进入风格迁移；图片下方可选择感兴趣的物体进入知识图谱；点击底部圆形按钮可以重新拍照

知识图谱界面

页面顶部呈现该物品的知识图谱，底部可以选择本物体或关联物品进入知识卡片

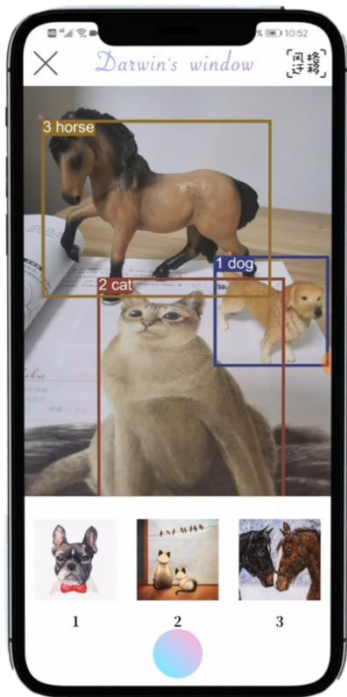


图 5-7 探索界面

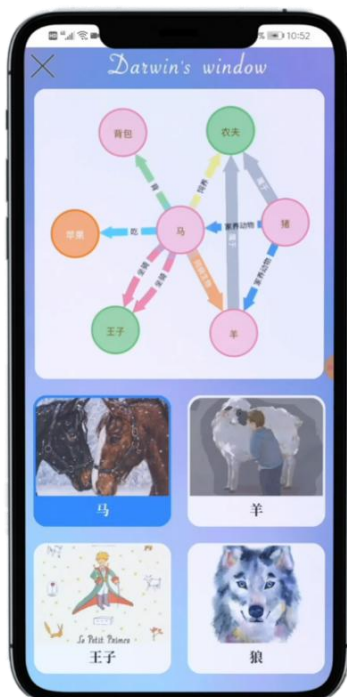


图 5-8 知识图谱界面

知识卡片界面

可通过左右滑动或者点击来切换童话、知识、科普、诗词模块，左上方相机按钮返回首页，右上方用户按钮进入用户中心

知识界面

呈现该物品的拼音、中英文解释；点击播放按钮朗读；点击心形按钮收藏，

童话界面

点击切换按钮切换不同童话，点击单句进行朗读或者播放全文朗读；底部左右滑动查看该物体的插画

科普界面

顶部是该物品的图片，底部是该物体的科普文字

诗词界面

呈现与物品相关的诗词，点击单句进行朗读或者播放全文朗读



图 5-9 知识界面



图 5-10 童话界面



图 5-11 童话界面底部插画



图 5-12 科普界面



图 5-13 诗词界面

风格迁移界面

选择对应的风格对图片进行风格切换；底部圆形按钮可以重新拍照

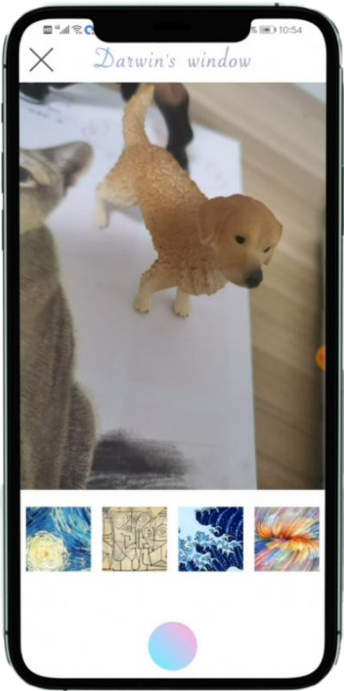


图 5-14 风格迁移界面

收藏界面

展现收藏内容，点击相应的物品可进入该物品的知识卡片



图 5-15 收藏界面

5.4.3 家长端

学习记录界面

记录孩子学习轨迹，依次为：今日孩子所学物品、孩子历史学习分类统计、扇形图和条形图展示统计结果、孩子的性格分析



图 5-16 学习轨迹



图 5-17 性格分析

亲子时刻

提供亲子互动方式。睡前故事，点击进入相应童话；趣味问答



图 5-18 睡前故事



图 5-19 趣味问答

发现界面

相关商品的推荐



图 5-20 商品推荐

5.4.4 负一屏

儿童端负一屏

根据用户数据推荐相关物品，点击物品进入儿童端对应的知识卡片界面。点击拍照探索进入儿童端首页拍照。



图 5-21 儿童端负一屏

家长端负一屏

呈现儿童学习的统计数据。点击学习轨迹进入家长端学习记录界面，点击商品推荐进入家长端发现界面。

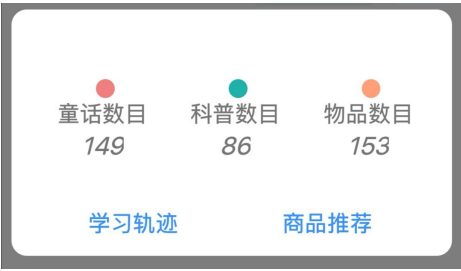


图 5-22 家长端负一屏

5.5 数据结构设计

5.5.1 逻辑结构设计

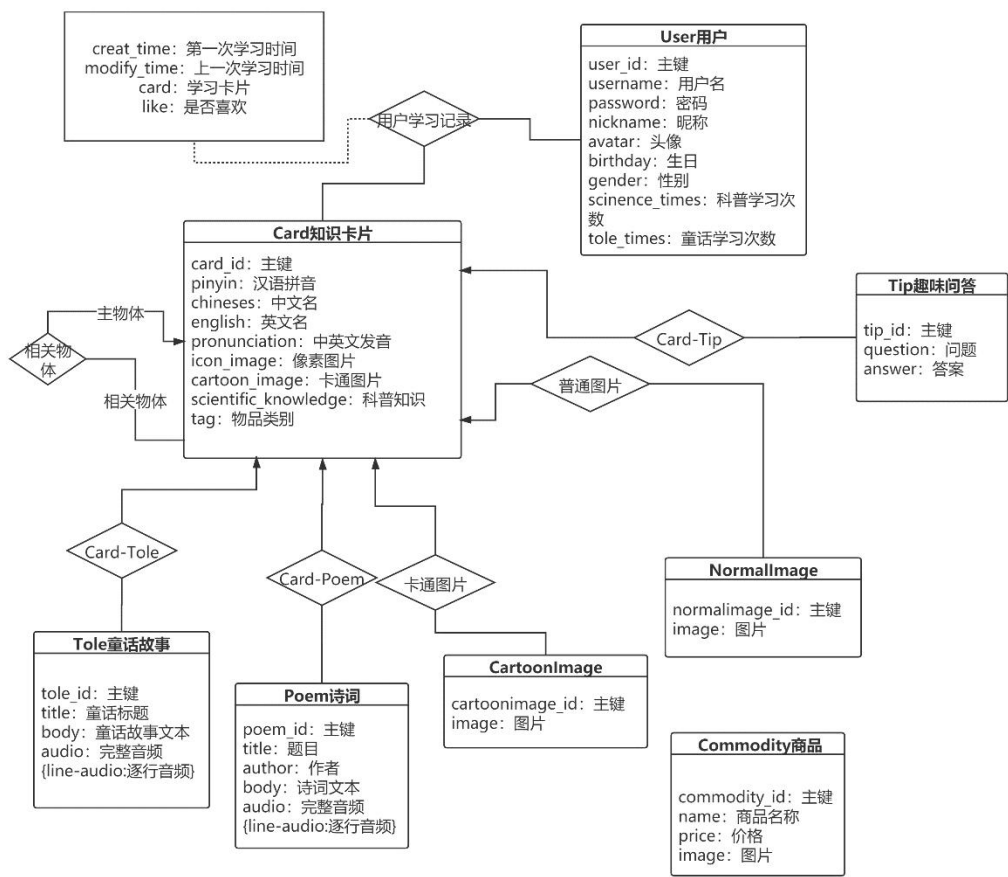


图 5-23 数据库 E-R 图

5.5.2 物理结构设计

数据库物理设计阶段的任务是根据具体计算机系统(DBMS 和硬件等)的特点，为给定的数据库模型确定合理的存储结构和存取方法。所谓的“合理”主要有两个含义：一个是要使设计出的物理数据库占用较少的存储空间，另一个对数据库的操作具有尽可能高的速度。

达尔文之窗在数据库的物理结构上主要做了以下的设计:

- (1) 对于字符和数值数据, 采用关系数据库 (MySQL), 降低读写硬盘频率, 加快程序运行速度。
- (2) 对于音频, 图片等数据采用文件服务器。并于关系数据库建立相对路径间的联系。
- (3) 数据库满足 3NF, 减少了冗余数据。并尽可能的使用外键, 减少表的连接。

5.6 接口设计

5.6.1 外部接口

- (1) 百度 AI 开放平台 -- 获取 Access Token

接口描述:

向授权服务地址发送请求

请求说明:

Method: POST

URL: <https://openapi.baidu.com/oauth/2.0/token>

表 5-1 请求说明

参数	值
grant_type	固定为: client_credentials
client_id	应用的 API Key
client_secret	应用的 Secret Key

返回说明:

表 5-2 返回说明

参数名称	类型	描述
access_token	string	要获取的 Access Token
expires_in	int	Access Token 的有效期

- (2) 百度 AI 开放平台 -- 短语音识别

接口描述:

将 60 秒以内的完整音频文件识别为文字, 支持上传完整的录音文件, 录音文件时长不超过 60 秒。实时返回识别结果

请求说明:

Method: POST

URL: https://vop.baidu.com/pro_api

表 5-3 请求说明

参数	值
access_token	账户获取的全局 access_token

表 5-4 Header 参数说明

字段名	数据类型	所需	描述
format	string	必填	语音格式: pcm、哇 v、m4a。不区分大小写, 推荐使用 pcm 文件

rate	int	必填	采样率 16000.固定值
------	-----	----	---------------

表 5-5 url 参数说明

字段名	所需	描述
cuid	必填	用户唯一标识, 用来区分用户, 计算 UV 值。建议填写能区分用户的机器 MAC 地址或 IMEI 码, 长度为 60 字符以内。
token	必填	开放平台获取到的开发者[access_token]获取 Access Token "access_token")
dev_pid	选填	不填写 lan 参数生效, 都不填写, 默认 80001, dev_pid 参数见本节开头的表格
lan	选填, 废弃参数	历史兼容参数, 使用 dev_pid。如果 dev_pid 填写, 该参数会被覆盖。语种选择, 输入法模型, 默认中文 (zh)。中文=zh、粤语=ct、英文=en, 不区分大小写。

返回说明:

表 5-6 返回说明

字段名	数据类型	所需	描述
err_no	int	必填	错误码
err_msg	string	必填	错误码描述
sn	string	必填	语音数据唯一标识, 系统内部产生
result	array ([string,string])	选填	识别结果数组, 提供一到五个候选结果, 优先使用第一个结果

5.6.2 内部接口

表 5-7 内部接口

模块	接口名称	接口功能
儿童端——探索模块	/explore/uploadphoto/	调用图像识别算法识别用户上传图片, 返回识别后的图片和识别出的物品列表。
	/explore/relatedobject/id/	根据物品 ID 获取物体知识图谱、返回原物体及其相关物体
	/explore/cardinfo/id/	根据物品 ID 返回物体知识汇集
	/explore/scientific/id/	根据物品 ID 返回物品科普知识
	/explore/likecard/id/	用户收藏或取消收藏物品
	/explore/likecardlist/	返回用户收藏的所有物品

	/explore/styletransfer/styleid/	调用风格迁移算法, 迁移用户上传照片
	/explore/poeminfo/id/	根据诗词 ID 返回诗词信息
儿童端——首页	/home/recommen/	根据用户 token 从数据库检索出用户学习记录, 调用推荐算法, 返回推荐物品列表和童话故事列表
	/home/toleinfo/id/	根据童话故事 ID 进入童话故事界面 (会记录学习行为)
登入、注册	/users/register/	用户注册
	/users/login/	用户登入
	/users/token/refresh/	更新用户 token
	/users/info/	根据用户 token 获取用户信息
	/users/changeinfo/	修改用户信息
	/users/changeavatar/	修改用户头像
家长端	/parent/trajectory/	获取用户学习轨迹
	/parent/storytip/	根据用户 token 获取推荐算法运算得到的睡前故事和趣味问答
	/parent/toleinfo/id/	根据童话 ID 获取童话信息 (不会记录学习行为)
	/parent/recocommodity/	根据用户 token 获取学习记录, 调用物品推荐算法获取推荐物品列表

5.7 错误/异常处理设计

5.7.1 错误/异常输出信息

错误和异常的输出信息在调试阶段和维护系统阶段都是必要的。达尔文之窗采用两种方法来处理错误和异常。

- (1) 后端基于 Django REST Framework (DRF) 框架, 因此采用了 DRF 中内置的 `exception_handler` 方法和自定义的异常处理来捕获视图层执行过程中的异常。
- (2) 同时每一个请求, 会先通过中间件判断是否合法 (是否包含需要的内容), 若不合法则直接返回错误信息。

表 5-8 后端服务器错误信息

出错模块	错误位置	错误信息
探索界面	图片上传	图片格式错误
风格迁移界面	图片上传	图片格式错误
用户信息模块	更改密码	原密码错误/验证密码和新密码不一致

	登入	token 过期
	更改头像	图片格式错误
	注册	用户名已被注册

5.7.2 错误/异常处理对策

- **问题:** 运行时错误。

对策:

- ✓ 运行时的异常和错误主要分为两类，一类是系统的错误，一类是用户操作错误。系统错误达尔文之窗尽量避免，用 `try`, `catch` 等方法将这类错误的发生率降低。第二类是用户操作错误，这类错误达尔文之窗选择中断执行过程，直接将错误返回给用户。

- **问题:** 预期功能是否能够全部实现。

对策:

- ✓ 数据备份: 当发生系统级的错误时，有可能会对数据库和用户的数据造成很大影响，所以需要设备做备份操作，定期备份数据库和用户的资料。
- ✓ 分布式架构: 多台设备同时提供一个功能，当一台机器出现故障时，另外的机器可以接替这台机器的工作，继续提供服务，出现故障的机器迅速排查错误，再重新启动服务，这样保证不会由于一台设备的问题影响整个网站。

5.8 系统配置策略

达尔文之窗的我们后端使用框架 Django+Celery+Redis+MySQL 的组合作为业务处理逻辑采用 Nginx+Gunicorn 进行部署。

选取 Django 的原因:

- (1) 利用 Django 在用户的 request 与 response 处理之间处理一些底层（如用户身份校验、缓存）等业务。
- (2) Django 是基于 python 语言的，可以很方便的预加载深度学习模型，缩短了模型预加载时间。

用户业务处理所需的静态数据固化在一台服务器的 MySQL 关系型数据库上，并采用 Redis 做缓存。

使用 Redis 的方法如下:

- (1) 对于纯静态数据，我们直接对 URL 请求进行缓存，在 Django 处理视图前查询是否存在 URL 匹配，相当于对当前调用的视图做缓存，一旦命中，直接返回数据。
- (2) 对于需要调用算法动态处理的视图，仍采取 Redis 做缓存。

采用 Redis 后，大大提高了服务器运行效率，与为采用 Redis 配置做对比，使我们的核心算法效率提升 30%，静态数据处理速度提高 170%。详细测试数据可参考测试文档。

由于我们的逻辑业务需求会产生用户临时文件，因此我们采用 Celery 配置了定时人物。到晚上服务器流量不大的时间段，执行 Celery 设置的动态任务，清理缓存文件。

5.9 系统部署方案

当前阶段达尔文之窗的系统部署方案是分布式在两台服务器配置两例 Django 实例。并采用 Nginx+Goicorn 作为 HTTP 服务器。

Gunicorn 依赖 Nginx 的代理行为，同 Nginx 进行功能上的分离。由于不需要直接处理用户来的请求（都被 Nginx 先处理），Gunicorn 不需要完成相关的功能，其内部逻辑很简单，接受从 Nginx 来的动态请求，处理完之后返回给 Nginx，由后者返回给用户，同时 Guicorn 具有高并发的特点，解决了系统的并发问题。综上，我们采取 Nginx+Guicorn 的组合来作为 HTTP 服务器，Nginx 用来处理用户请求静态文件两者相互配合，提高了项目运行效率。

我们采用 Nginx 搭建 HTTP 服务器。并做 L7 请求分法式负载均衡，在服务器 A 和 B 上分布式部署两例 Django 实例。由于服务器 B 上有较强的 GPU，因此所有图像识别请求（图片识别、风格迁移）均交给服务器 B 上的实例来处理。对于相似度分析和智能推荐等其它请求随机分发给服务器 A 和服务器 B 处理，数据库部署在服务器 A 上，并配置了 RAID 1。两台服务器并行执行有利于整体效率的提升。



图 5-24 后端架构图流程图

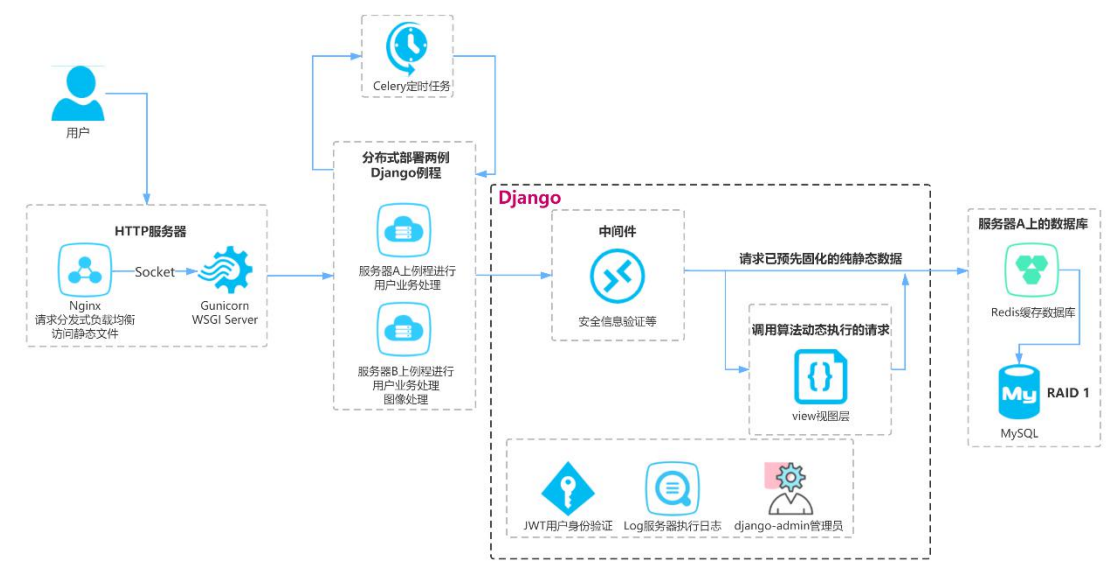


图 5-25 后端配置策略

表 5-9 两台服务器的配置表

硬件环境	CPU (Intel)	GPU	Memory	Storage	基准频率
硬件设备 A (RAZER)	Core(TM) i7-9750H	RTX 2070 MAX-Q	32GB	512GB SSD 512GB SSD	2.20GHz
硬件设备 B (DELL)	Xeon(R)E5-2630 v4	GTX 1080	32GB	256GB SSD 2TB HDD	2.20GHz

5.10 其他相关技术与方案

5.10.1 跨域资源共享

跨域资源共享 (CORS) 是一种机制, 它使用附加的 HTTP 标头来告诉浏览器以使 Web 应用程序在一个来源运行, 并从另一个来源访问选定的资源。Web 应用程序请求其来源 (域, 协议或端口) 不同的资源时, 将执行跨域 HTTP 请求。

对于开发者来说, CORS 通信与同源的 AJAX 通信没有差别, 代码完全一样。浏览器一旦发现 AJAX 请求跨源, 就会自动添加一些附加的头信息, 有时还会多出一个附加的请求, 但用户不会有感觉。

因此, 实现 CORS 通信的关键是服务器。只要服务器实现了 CORS 接口, 就可以跨源通信。

- (1) 我们利用 Django 的中间件 corsheaders 来解决跨域问题。它向响应添加跨域资源共享 (CORS) 标头。这允许浏览器从其他来源向 Django 应用程序发出请求。
- (2) 使用 Nginx 对不同的请求进行代理, 如果请求的时候打包好的静态资源, 直接将请求路由到静态资源, 如果请求的是后端的接口, 则将请求转发到后端服务器, 这样能有效的避免跨域问题而不需要再前后端现有请求上进行额外的修改。

5.10.2 错误日志

对 Nginx 和 Django 均配置错误日志, Nginx 用于收集错误的请求, Django 用于收集具体 (调用算法, 访问数据库) 的错误信息。

6 数据库设计

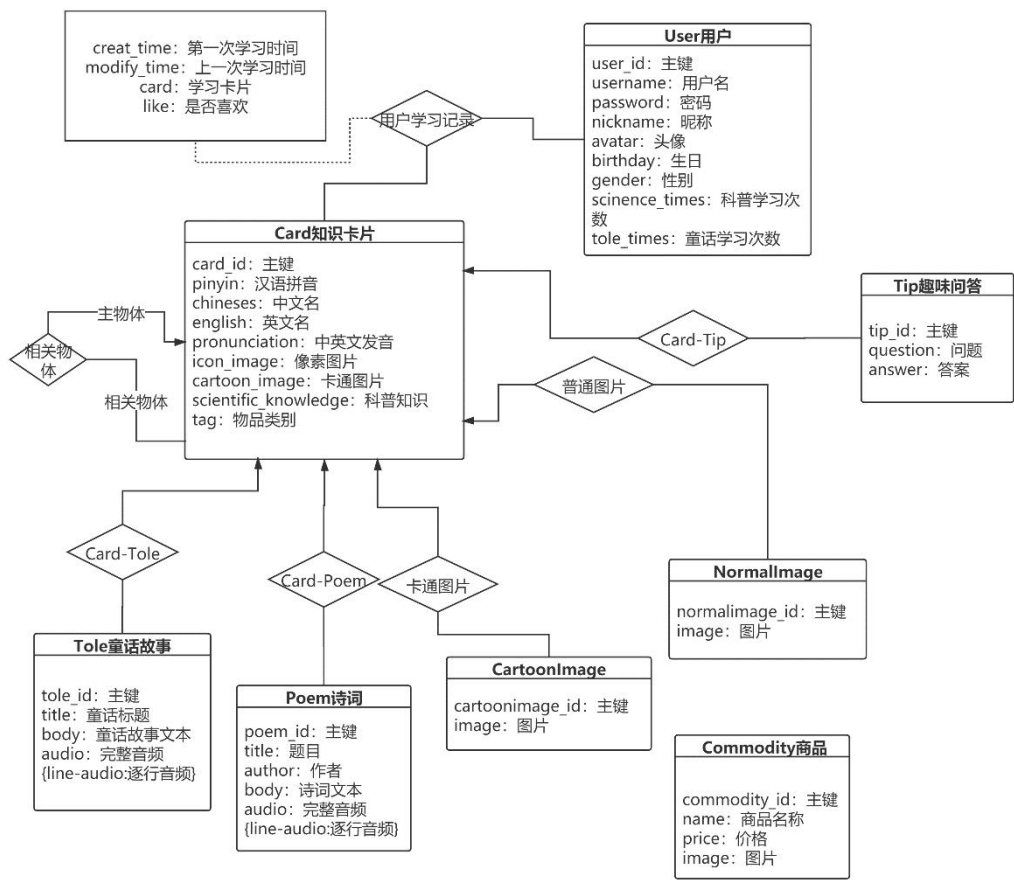


图 6-1 数据库设计

7 详细设计

7.1 智能语音处理功能模块

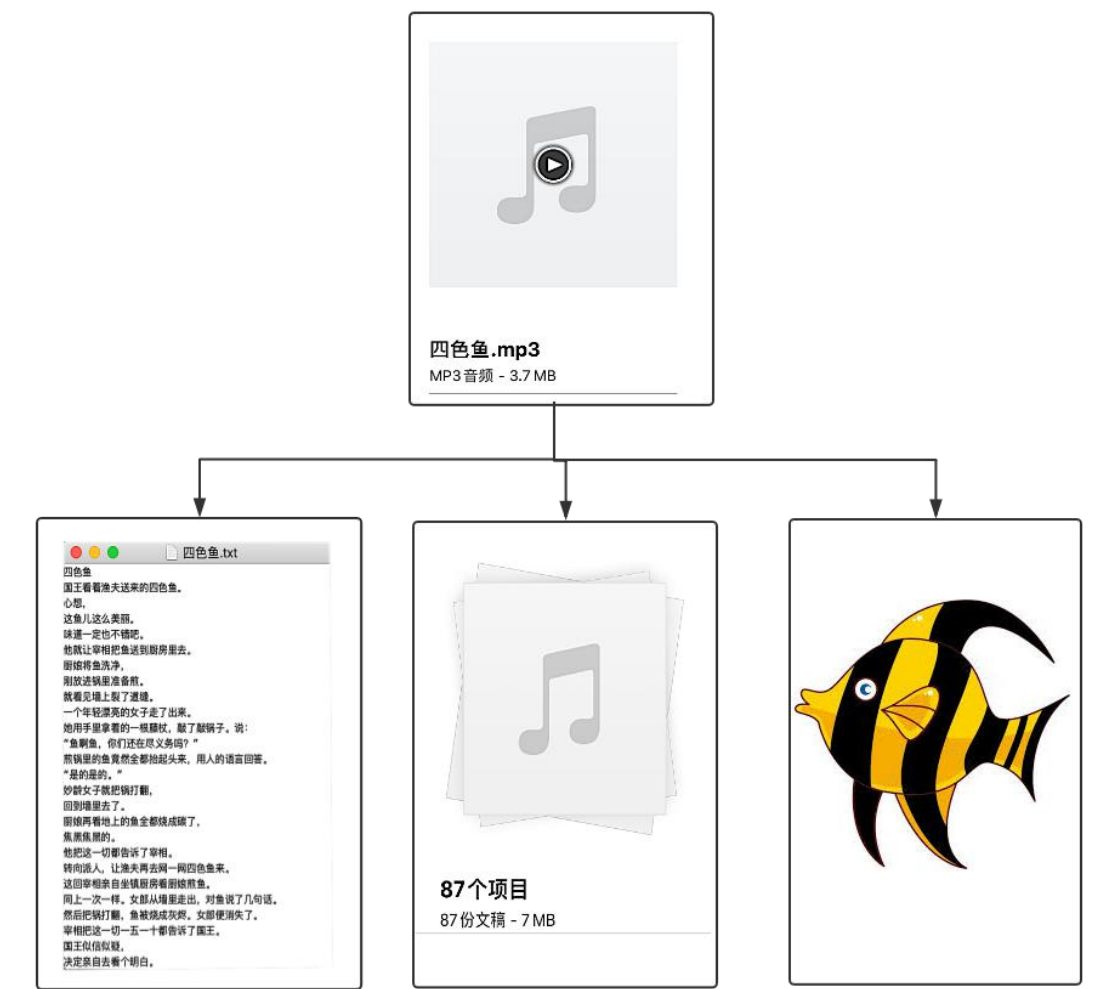


图 7-1 功能描述图

7.1.1 功能描述

首先，上传需要进行智能分割的音频，接着进行音轨分离操作，处理之后，再进行分割，最后对每一句话进行语音处理，将文字与图片进行匹配。在利用 spleeter master 进行音轨分离的同时，我们对语音进行了降噪处理，来保证语音清晰，中断明显。在语音识别时，我们对童话中的特殊用词如常见人名，著名人物进行了训练。

7.1.2 性能描述

整体处理时间性能：<5s 内存占用：<100MB

7.1.3 输入

以语音形式输入优质音频

7.1.4 输出

一段段的语音以及对应的文字及推荐图片

7.1.5 程序逻辑

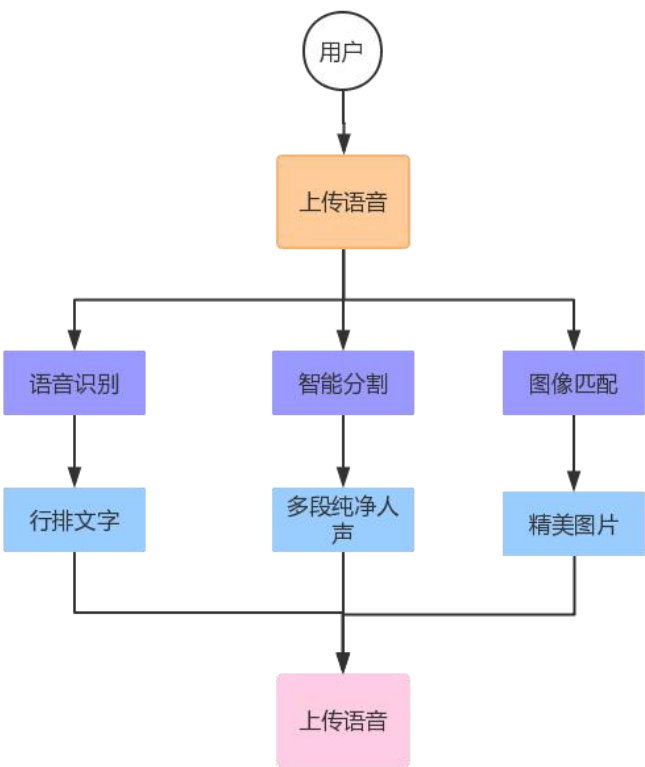


图 7-2 模块运行逻辑

7.1.6 限制条件

输入的语音的每一句话不要超过 1 分钟

7.2 个性化推荐功能模块

7.2.1 功能描述

为了让儿童个性化学习，项目采用了个性化推荐模块，根据孩子的浏览轨迹与兴趣偏向（喜欢的物体）来进行知识卡片的推荐；童话的推荐；基于知识图谱来量化物体之间的相关系数，并利用协同过滤来优化推荐系统，进行回测。

7.2.2 性能描述

表 7-1 性能描述表

功能单元	时间性能	单元鲁棒性	局限性	可扩展性	效果描述
推荐系统	<0.2 s	根据物体之间的相关系数以及群体倾向推测个人喜好,对于个别用户的极端数据分布具有一定的耐受性	对于异常极大数据耐受性存在局限,处理时将此种情况按冷启动看待	可移植到模型的服务器	及时没有用户数据,根据构建的物体之间的关系矩阵也可以进行合理推荐,推荐变化比较迅速。

7.2.3 输入

表 7-2 推荐系统输入

单元名称	输入
浏览轨迹	物体、童话故事的名称与个数; 童话、科普、拍照探索的个数
收藏夹	名称
个人信息	性别、年龄

7.2.4 输出

表 7-3 推荐系统输出

名称
首页推荐物体、童话

7.2.5 程序逻辑

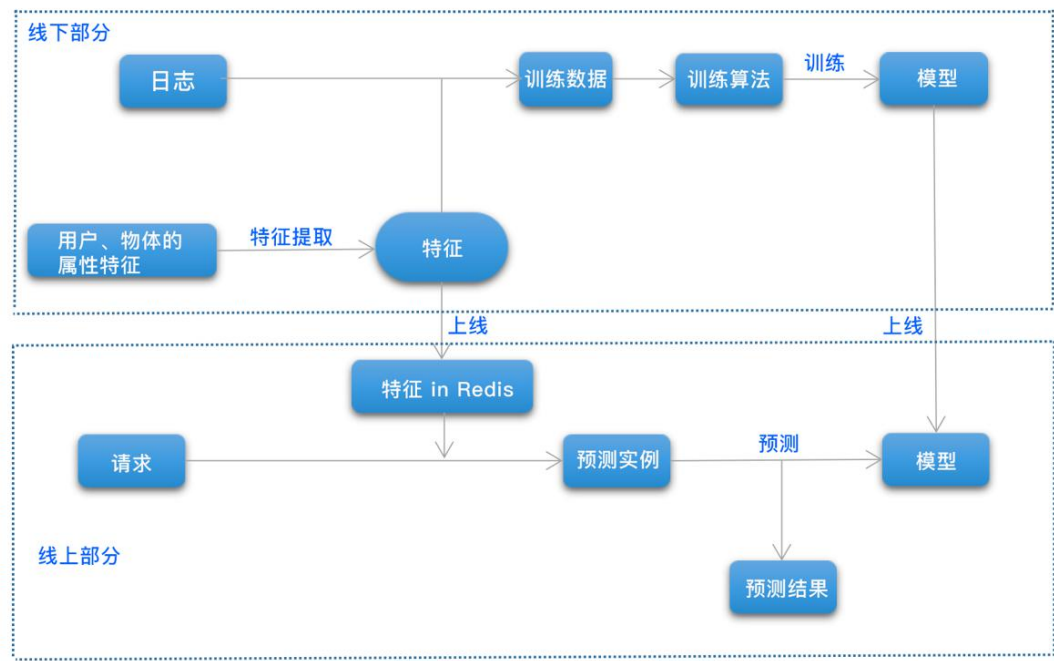


图 7-3 用户运行逻辑

7.2.6 限制条件

在用户较多时，回测模型运算量较大。

7.3 图像识别模块

7.3.1 功能描述

图像识别可以对儿童拍摄的照片进行识别，可以检测照片中的物体，能够对标注的物体进行标注，并将标注后的图片进行返回，同时能够返回与图片中物体相对应的物体名称列表。

7.3.2 性能描述

表 7-4 图像识别性能描述

功能单元	时间性能	单元鲁棒性	局限性	可扩展性	效果描述
图像识别	<3s	对数据库中的物体能够准确的检测和标注。对图片中贴近真实物体的物体识别准确。	对非数据库中的物体识别准确率下降；对类似简笔画的图像识别准确率下降	可移植到模型的服务器	能够检测照片中的物体，对标注的物体进行标注，并将标注后的图片进行返回，同时能

					够返回与图片中物体相对应的物体名称列表。
--	--	--	--	--	----------------------

7.3.3 输入

表 7-5 图像识别输入

名称
Name.jpg 图片或 Name.png 图片

7.3.4 输出

表 7-6 图像识别输出

名称
Name1.jpg
物体名称列表

7.3.5 程序逻辑

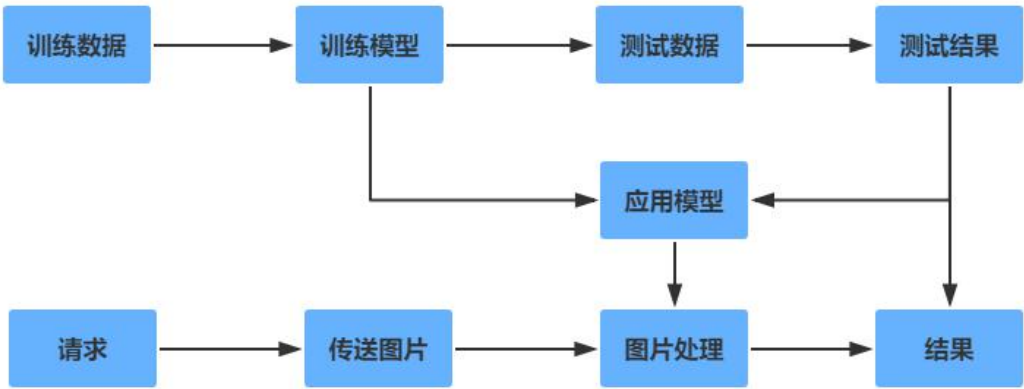


图 7-4 图像识别流程图

7.3.6 限制条件

图片格式要求为 JPG 格式，识别的物体种类数目受到数据库限制。

7.4 风格迁移模块

7.4.1 功能描述

儿童可以在提供的卡通类、油画类共 12 种风格中任意选择风格，再选择之前拍摄的照

片或者选择重新拍摄一张照片，该功能会将照片转换成儿童选择的风格。

7.4.2 性能描述

表 7-7 风格迁移性能描述

功能单元	时间性能	单元鲁棒性	局限性	可扩展性	效果描述
风格迁移	<3s	对传入的图片都能够再较短的事件内完成迁移,并且能够保证图片的质量。	对不同内容进行不同风格的迁移时,效果难统一。	可移植到模型的服务器	可以在提供的风格中任意选择风格,再选择之前拍摄的照片或者选择重新拍摄一张照片,该功能会将照片转换成儿童选择的风格。

7.4.3 输入

表 7-8 风格迁移输入

名称
Name.jpg 图片或 Name.png 图片
风格 id

7.4.4 输出

表 7-9 风格迁移输出

名称
Transfer.png 图片

7.4.5 程序逻辑

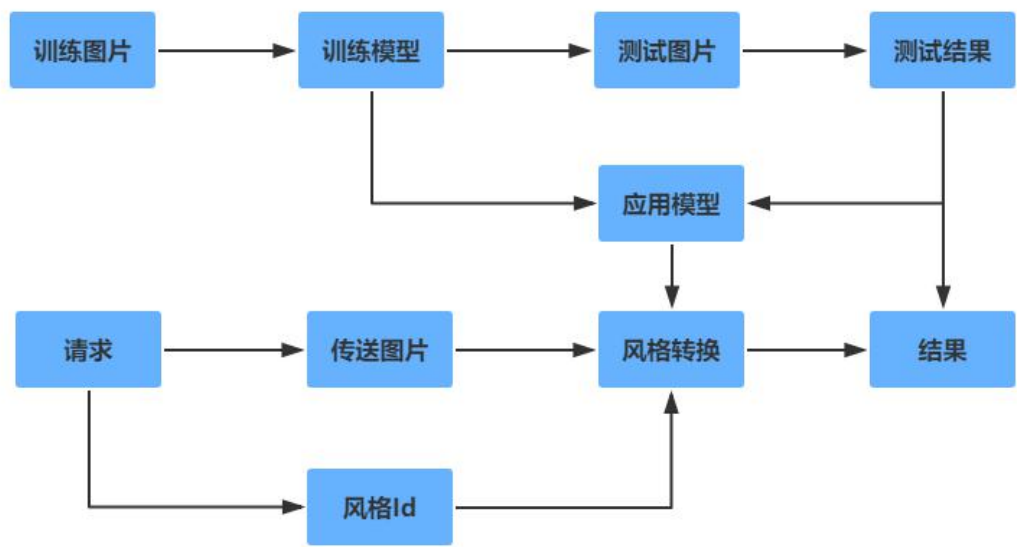


图 7-5 风格迁移流程图

7.4.6 限制条件

风格种类选择须在置定的风格种类范围内，某些个别图片（如图片内容过多）在某些风格上可能会导致效果欠佳。

7.5 亲子互动推荐功能模块

7.5.1 功能描述

亲子互动模块通过商品推荐实现亲子之间的有效互动，项目采用了个性化推荐模块，根据孩子的浏览轨迹与兴趣偏向（喜欢的物体）来进行家长睡前故事的推荐；趣味问答的推荐；基于知识图谱来量化物体之间的相关系数，并利用协同过滤来优化推荐系统，进行回测。

7.5.2 性能描述

表 7-10 性能描述表

功能单元	时间性能	单元鲁棒性	局限性	可扩展性	效果描述
推荐系统	<0.2 s	根据物体之间的相关系数以及群体倾向推测个人喜好，对于个别用户的极端数据分	对于异常极大数据耐受性存在局限，处理时将此种情况按冷启动看待	可移植到模型的服务器	及时没有用户数据，根据构建的物体之间的关系矩阵也可以进行合理推荐，推荐变化

		布具有一定的耐受性			比较迅速。
--	--	-----------	--	--	-------

7.5.3 输入

表 7-11 推荐系统输入

单元名称	输入
浏览轨迹	物体、童话故事 的名称与个数；童话、科普、拍照探索的个数
收藏夹	名称
个人信息	性别、年龄

7.5.4 输出

表 7-12 推荐系统输出

名称
家长端儿童偏好分析
家长端睡前故事
家长端趣味问答

7.5.5 程序逻辑

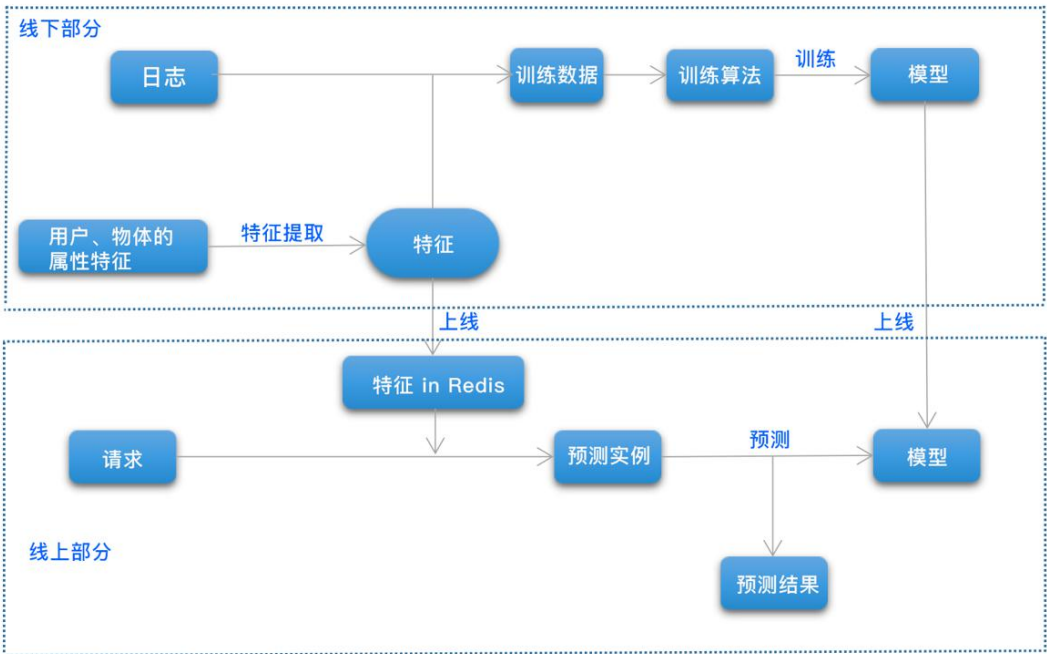


图 7-6 用户运行逻辑

7.5.6 限制条件

在用户较多时，回测模型运算量较大。

7.6 商品推荐功能模块

7.6.1 功能描述

为了让亲子互动更有效，同时提升产品的商业价值，项目采用了个性化推荐模块，根据孩子的浏览轨迹与兴趣偏向（喜欢的物体）来进行商品推荐；基于数据分析进行商品的推荐。

7.6.2 性能描述

表 7-13 性能描述表

功能单元	时间性能	单元鲁棒性	局限性	可扩展性	效果描述
推荐系统	<0.2 s	根据物体之间的相关系数以及群体倾向推测个人喜好，对于个别用户的极端数据分布具有一定的耐受性	对于异常极大数据耐受性存在局限，处理时将此种情况按冷启动看待	可移植到模型的服务器	及时没有用户数据，根据构建的物体之间的关系矩阵也可以进行合理推荐，推荐变化比较迅速。

7.6.3 输入

表 7-14 推荐系统输入

单元名称	输入
浏览轨迹	物体、童话故事的名称与个数；童话、科普、拍照探索的个数
收藏夹	名称
个人信息	性别、年龄

7.6.4 输出

表 7-15 推荐系统输出

名称
家长端商品推荐

7.6.5 程序逻辑

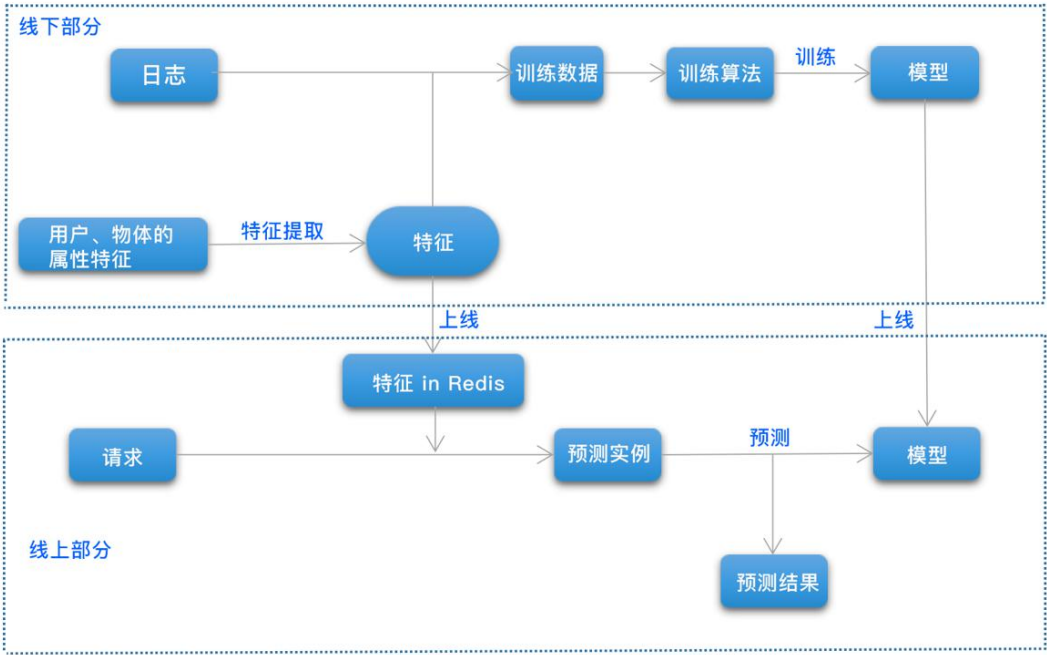


图 7-7 用户运行逻辑

7.6.6 限制条件

在用户较多时，回测模型运算量较大。