



项 目 视障人士友好的咨询辅助软件

组 长 曾平

组 员 刘一璇 汤耀轮 彭瑞杰 陈雨延

课 程 软件项目管理

指导教师 刘洪

二 Ο 二 二 年 四 月 十三 日

# 一、可行性分析

## 1.1、引言

**1.1.1、项目背景**

在人工智能技术不断发展的今天，各种智能APP应用让我们的生活变得越来越便捷。我们希望AI技术能够覆盖更多需要帮助的人群。比如视障人群[1]（半盲和低视力），他们存在视力障碍，无法通过佩戴眼镜等方式进行矫正，我们希望能够利用云上资源以及人工智能的帮助，让他们享受到科技进步的成果。

目前移动云上有着大量的人工智能API能力及算力资源，但是单纯的API能力无法直接供一般的用户使用，必须创新方法将这些功能结合起来，形成应用场景。

**1.1.2、项目目标**

本项目以“视障人士友好”为主要方向，以视障人士在使用掌上移动设备获取新闻资讯为基本应用背景。分析视障人士在使用软件时的使用困难，并在客户端提供友好的交互方式与使用理念，提高本软件面向人群的使用体验。分析用户的新闻浏览习惯，服务端向用户推送更适合用户的新闻内容；使用用户上传的图片，服务端提供图片文字识别以及拍照实时提示定位功能。服务端还将可以对新闻的文本与语音进行文字纠错，减少新闻内容的错误。

## 1.2、可行性研究前提

**1.2.1、功能要求**

（1）软件服务端需要具备用户行为信息、新闻数据信息的统计功能。

（2）软件服务端可支持新闻的实时输入，并推送至客户端。

（3）软件客户端需要具备新闻内容展示以及新闻推荐功能，推荐使用九天深度平台进行模型训练。

（4）软件客户端可基于手机自带的辅助功能实现拍照功能，以及对照片实现文字语音提示功能。比如：书籍的拍照阅读，可以通过手机支架+书籍固定器+手机辅助功能，结合OCR的API接口实现视障人士的拍照阅读。更便捷与有创意的交互模式可以额外加分。

（5）软件服务端具备基础的web系统界面，并可以运行在windows/linux平台。

1. 软件客户端（APP）功能完整，包括但不限于可使用小程序/APP软件实现。

**1.2.2、非功能要求**

（1）软件需要使用到移动云云上AI能力

（2）演示时新闻数据量不少于1000，新闻类型不得少于10类

（3）创新功能需具备自主知识产权/不产生产权纠纷

（4）作品需要标明使用的开源数据/软件，并标明出处

（5）开发环境：使用移动云相关环境、资源进行开发

**1.2.3、可行性研究方法**

利用现阶段项目成员所掌握的知识，先以最简洁、最容易的办法，开发的同时进行测试，实现初步的开发平台，后期进行优化。

在分析和开发阶段，编写所需文档，采用Github管理项目的形式呈现，文档包括：

（1）可行性分析报告

（2）项目计划书

（3）需求规格及使用说明书

（4）概要设计文档

（5）详细设计文档

（6）用户手册

以确保项目开发可行，保证项目顺利进行。

## 1.3、技术可行性

本项目的开发主要分为客户端与服务端，其中服务端的开发包括基础的应用协议支持以及利用AI模型提供智能服务的部分。具体描述为：

**1.3.1、客户端的开发**

**1.3.1.1、客户端技术要求分析**

分析视障人士与掌上设备的交互模式，提供一套更适合视障人士的软件交互方案。团队将尽可能采用大致屏幕手势以及语音控制来进行与软件的交互，在展示与播报新闻时提供切换、调节音量、改变语速、重读等友好操作，同时界面也将展示文字以充实应用，提供不同的用户使用不同的方式获取新闻的功能。

另外，软件对视障人士的线下阅读也会提供支持，客户端将拥有拍照功能，并且在与服务端交互中进行拍照位置与姿势的校准提示，拍照之后经过服务端处理，客户端将输出图片中可能的内容。

**1.3.1.2、客户端技术实现分析**

服务端的Web界面将支持提示在线人数、服务器运行状况，更改与浏览新闻源，更换服务模型，系统消息推送，查阅新闻功能。通过vue，js，css，html等技术构造前端管理页面。

手机app服务端通过uniapp部署开发，uni-app 是一个使用 Vue.js (opens new window) 开发所有前端应用的框架，开发者编写一套代码，可发布到iOS、Android、Web（响应式）、以及各种小程序、快应用等多个平台。基于通用的前端技术栈，采用vue语法+微信小程序api，无额外学习成本。

**1.3.2、服务端的开发**

服务端将从指定的网站实时获取最近的新闻并自动更新新闻数据，依据用户向其方案化推送相关新闻；同时需要能够承载大量用户访问，快速完成用户的请求。团队将提高服务端的并行处理效率，更好更快地使用相关服务。

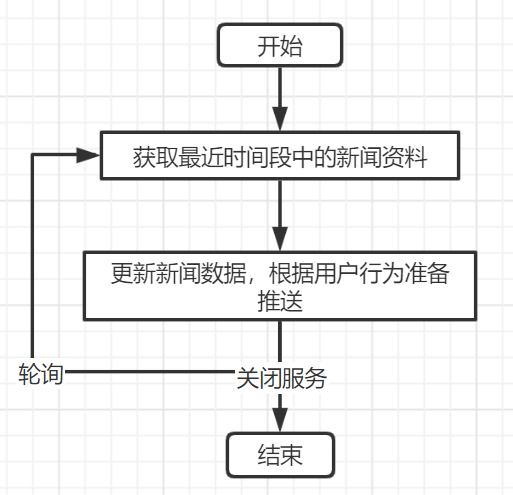
**1.3.2.1、AI模型的研究**

书籍阅读：充分利用云上AI资源，使用移动云九天深度学习平台的OCR接口，来实现书籍的拍照阅读。并结合各种有趣的语音包为用户打造个人性化阅读方式。其中还可提供包括关键词屏蔽以及语速调整等功能。仍有余力的情况下，可以加入Transformer、BERT以及MacBert等模型提供中文文本纠错功能，防止现有接口无法应对个别异常情况。

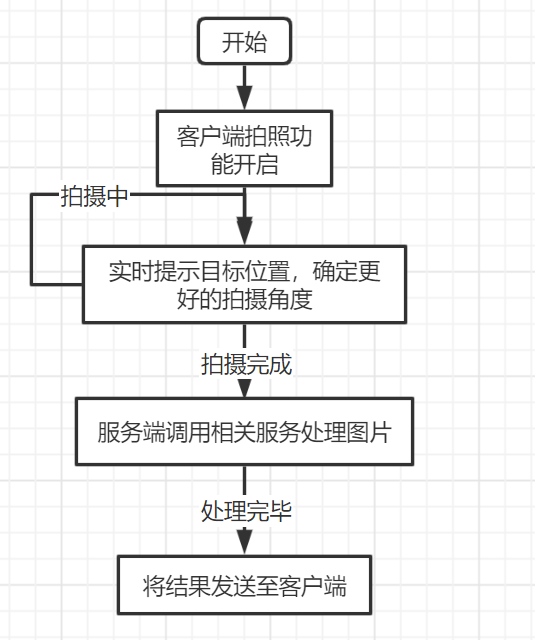
目标检测与识别：由于实物识别涉及范围太广，需要大量数据集，所以也使用云上AI资源。而目标检测的应用场景主要是，视障人士在拍照阅读时，会出现手机离目标太近或者太远以及书籍位置错位等情况，系统通过检测书籍和手机的相对位置，为用户提供提示指导。目标检测常用模型有Faster R-CNN、SSD、YOLO v3。

处理流程

服务端获取新闻：

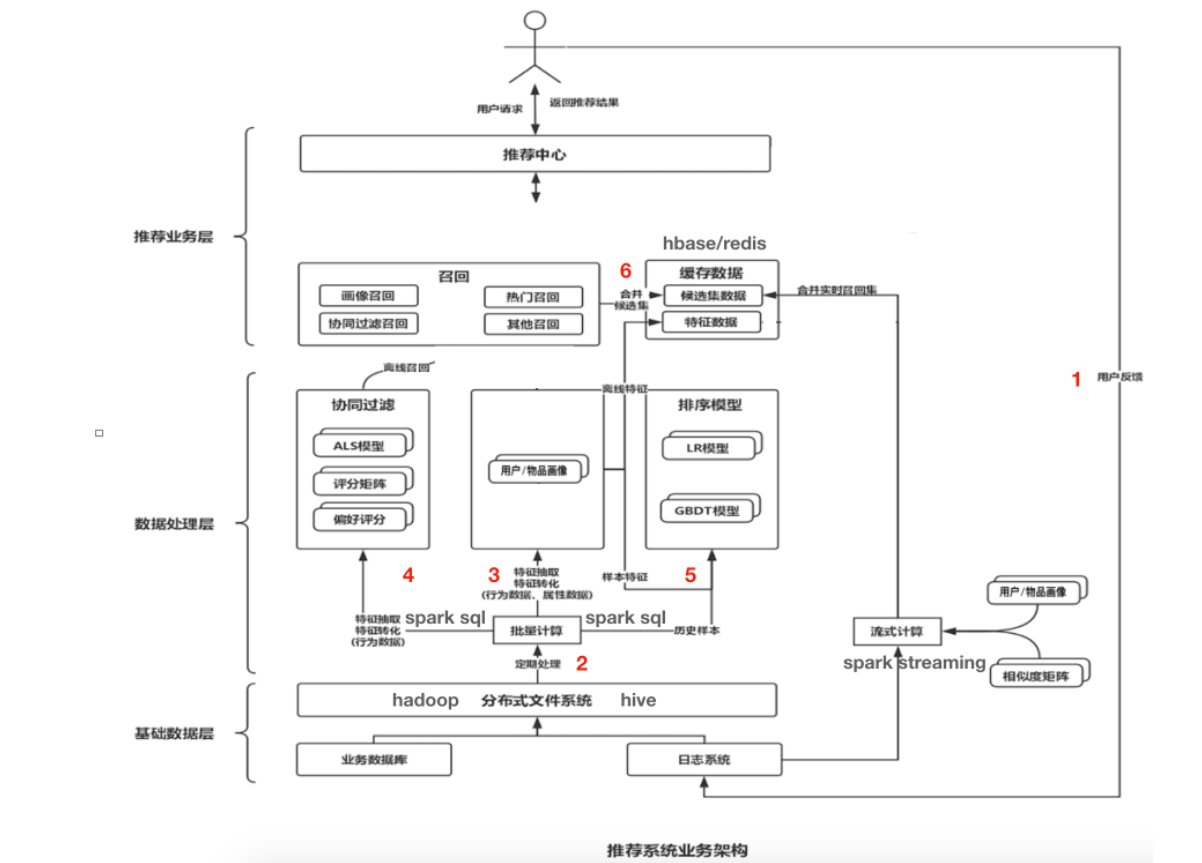


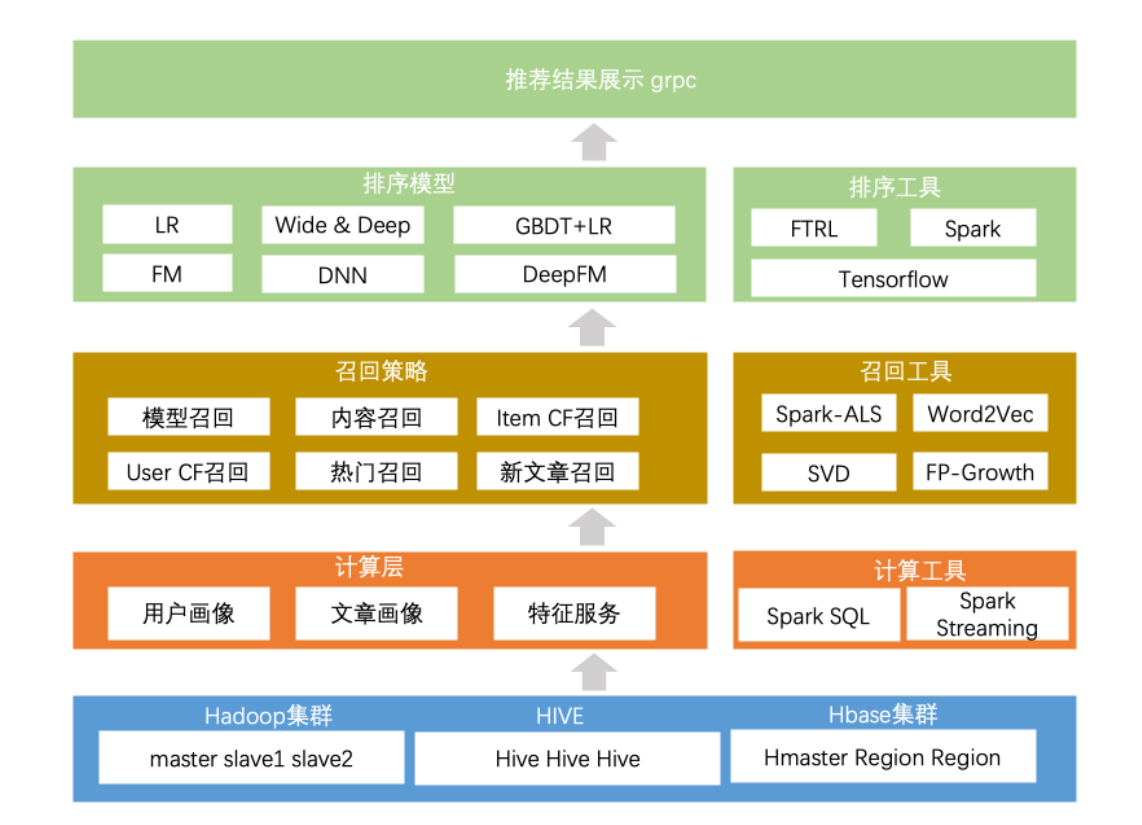
线下处理图片：



**1.3.2.2、算法模型构建的流程**

推荐系统框架分为基础数据层、数据处理层以及业务层。具体结构如下：





**1.3.2.2.1、基础数据层**

以Hbase作为数据库，使用Hadoop分布式文件存储系统，Hadoop在大数据处理上具有高可靠性、高扩展性、高效性和高容错性的优势。在数据层会储存数万条新闻数据、用户数据以及辅助关联表的数据。

**1.3.2.2.2、数据处理层**

利用Spark SQL的DataFrame和DataSet进行批量计算，将数据转化为需要的特征数据，传入召回阶段和排序阶段。

**1.3.2.2.3业务层**

业务层分为召回阶段和排序阶段。

召回阶段：根据用户的历史行为和短期行为，分析用户的兴趣偏好，从数万级的文章库中挑选出一个小的候选集（通常几十篇文章），这些候选集都是用户感兴趣的内容集合。

而召回阶段又分为三个通道：

(1)基于内容的召回：根据文章内容本身的语义特征或者字面特征做召回，这类特征通常有关键词，topic，分类，媒体来源等。在对新闻文本做数据标签时，如添加关键词，可以利用jieba分词以及TF-IDF提取关键词等技术，为每条新闻贴好标签。

(2) 基于协同过滤的召回：协同过滤算法的原理是汇总所有用户（user）和文章（item）的行为，利用集体智慧进行推荐。主要分为两大类，User-CF和Item-CF。

User-CF，就是找到和用户兴趣相似的群体，然后把这个群体喜欢的推荐给这个用户，这里重点是找出用户的相似用户。

Item-CF，分析每篇文章的点击行为，可以得出所有文章之间的相似度，找出与当前点击的文章最相似的N篇文章推荐给用户。

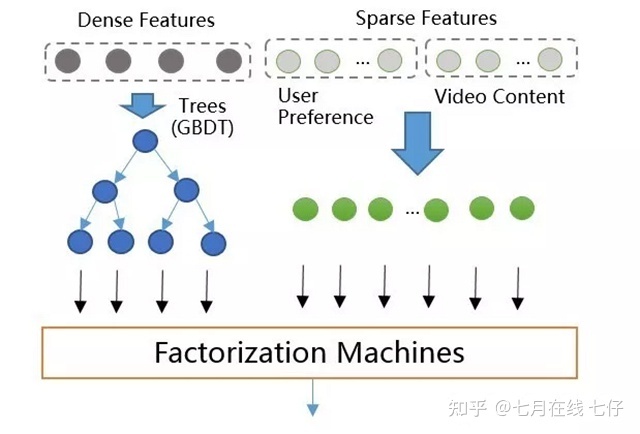
(3)基于热点的召回：基于热点的召回在冷启动阶段有着重要意义，即一个新用户的历史行为数据太少，协同过滤效果较差，那么热点内容的占比将会更高。

不管是User-CF，还是tem-CF，都可以通过传统的计算相似度矩阵或者聚类来实现。

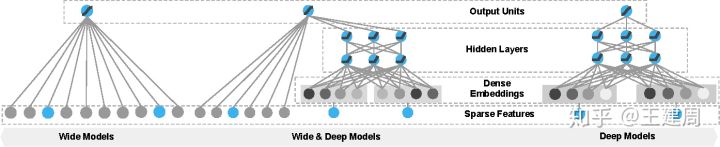
排序算法：在召回集的基础上进行更加精准的个性化计算，给每篇文章进行精确打分，这个分值就是文章与用户的个性化匹配分值，利用该分值进行排序，进而从几十篇文章中选出用户最感兴趣的少量高质量内容。

现在的推荐系统使用较多的模型是GBDT+FM，不过随着DNN的引入，推荐模型的准确率有所提高。但是模型的表现效果还是会因不同的实际问题而有所差异。

GBDT是不支持高维稀疏特征的，而加入Factorization Machines使得模型能够处理高危稀疏矩阵，而且模型的返回能力强。模型结构如下：



DNN模型中较为突出的有Google推出的wide&deep模型，模型左半部分，就是一般线性模型；模型右半部分，一半来说就是多层的DNN网络，因为输入的特征可能是离散型，会加入embeddings层随模型一起训练；联合部分将wide输出和deep的输出联合输入到全连接网络进行目标函数拟合。Google将网络设计成这个样子，说是Wide部分有”记忆“能力，Deep部分有”泛化“能力。这种方法的优点在于：通过神经网络可以拟合高阶的非线性关系，同时减少了人工特征的工作量。模型结构如下：



## 1.4、经济可行性

**1.4.1、支出**

经济费用：系统开发费用，系统安装、运行和维护费用；云上资源使用费用

**1.4.2、效益**

软件商店出售app效益；用户内部功能付费效益

**1.4.3、收益/投资比**

投资收益与各项收入的比例，即投资占比=投资收益/各项收入x100%

**1.4.4、社会因素可行性**

**1.4.4.1、法律可行性**

该产品开发过程中使用正版软件，无滥用或无授权使用情况；

项目代码除去框架外均为自行开发，无抄袭或违法情况；

项目中所使用技术均未被申请专利，无侵权情况；

所有的用户数据均作了脱敏处理，只保留用户邮箱手机等信息识别用户，所有的用户邮箱手机信息均由提出方

保管，合同制定确定违约责任；

经以上分析，该项目不存在法律风险。

**1.4.4.2、用户使用可行性**

大部分手机系统都有自带的无障碍阅读工具，可以帮助视障人士正常使用手机功能，但是单独的辅助软件无法直接帮助他们实现资讯的阅读，大部分已有阅读软件主要针对的还是视力正常的人群，我们要做的是通过软件让特殊人群更好地获取现实资讯。

**1.4.5、竞品分析对比**

**1.4.5.1、苹果辅助阅读功能**

竞品内容：iPhone在浏览新闻资讯时，不需要始终注视和划动手机屏幕，直接利用自带的朗读功能即可听完整的文章。

本产品相对优点：该功能需要用户在设置中通过一系列操作打开，对于半盲人士非常不友好；且该功能并不是单独针对半盲人士使用，用户体验不足；该功能并不能帮助半盲人士进行新闻推送阅读

**1.4.5.2、苹果辅助阅读功能**

竞品内容：iPhone在浏览新闻资讯时，不需要始终注视和划动手机屏幕，直接利用自带的朗读功能即可听完整的文章。

本产品相对优点：该功能需要用户在设置中通过一系列操作打开，对于半盲人士非常不友好；且该功能并不是单独针对半盲人士使用，用户体验不足；该功能并不能帮助半盲人士进行新闻推送阅读

暂无其他主流辅助阅读软件

**1.4.6、结论意见**

市场竞品较为空白，市场环境好，开发成本仅限软件部分，并且使用云上资源训练，成本较低，技术实现有可行性，可以开发。