

一、小程序作品说明

见微实现近 100 万条医护信息的集成，其中数据词条由**国家中医药管理局名词术语成果转化与规范推广项目**审核认证，减少了冗余信息的干扰，方便用户更快更准确的对日常医疗护理、保健知识和各个地区药店药物价格内容进行查阅，尤其是老年人群体，更加需要关于医护的相关信息的获取。因此该小程序的出现，为更加准确的医护信息获取提供了一种可能性。

二、应用场景

在用户购买药品时提供药品参考价格，在用户需要了解医疗医护信息时给予详细的信息提供，在用户对于不了解的中药材时可拍照上传查看对应的详细信息。

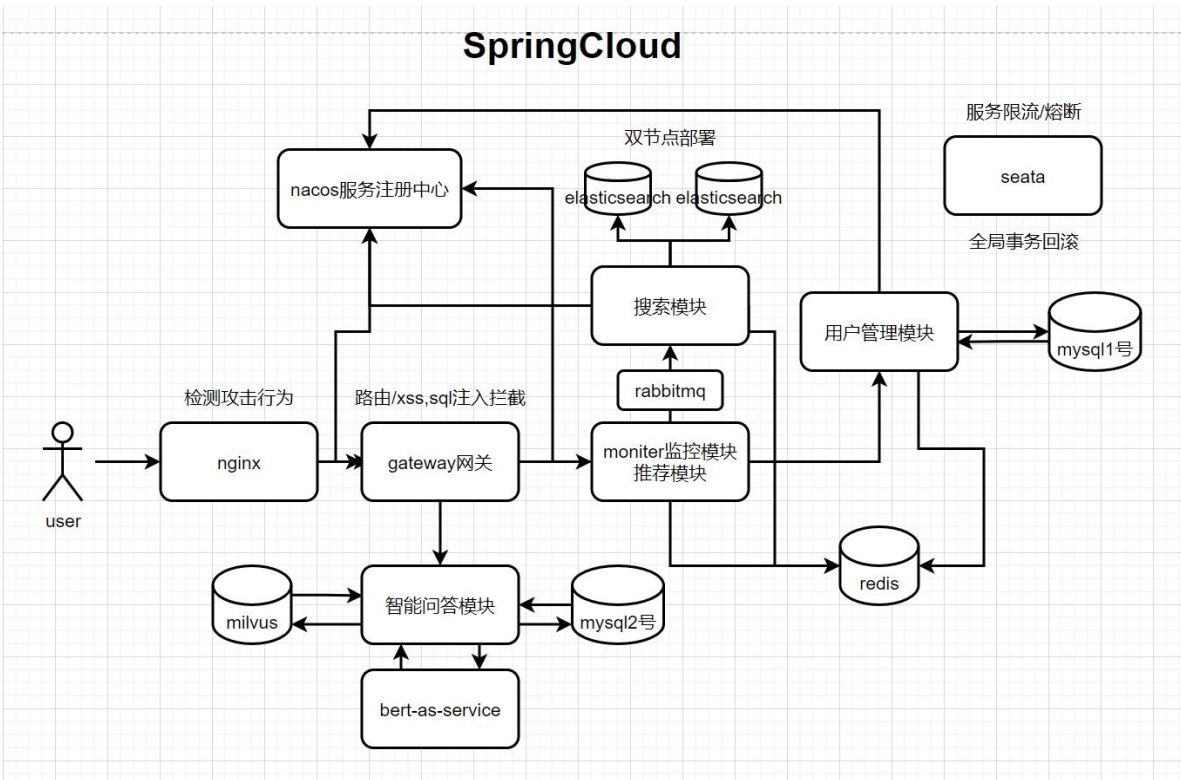
三、解决的实际问题

当代人群对医疗医护信息掌握较少，并且在药店购药时不了解药品的真实价格，可能存在被无良商家坑骗钱财的情况。该小程序目前已集成 10 余万个医疗医护以及药品药价信息和 80 余万个相关医疗问题问答对，致力于宣传医疗医护信息和药品药价信息。并且由于部分医药物品信息对于用户来说并不熟悉，可以通过扫码或者是搜索的方式，来了解该药品的具体使用方式，禁忌等其余有效信息。

四、产品设计和技术实现方案

(1) 基于 SpringCloudAlibaba 的分布式微服务架构

考虑到项目未来可扩展性，采取将业务模块独立出来。本项目使用 5 台服务器尽可能分离各个业务模块，提高每个模块的可利用资源，相比于单机架构，减少了服务器负载压力。



架构概述:

使用 springcloudAlibaba 中 Nacos 组件作为服务注册中心, 对用户模块, 搜索模块, 监控与推荐模块进行注册, 智能问答模块是 python 代码编写作为单独模块运行。
全局使用 5 台云服务器提供分布式服务。

完善的全局事务处理:

使用 seata 组件作为各个微服务分布式事务服务, 对恶意访问服务做限流熔断处理, 对服务中事务操作出错做全局事务回滚处理。

周全的安全性设计:

1. 使用 nginx 对全局用户访问量进行控制, 防止 QPS 过大直接使内部业务服务器宕机。
2. 使用 gateway 网关基于 Nacos 做服务路由, 并在 GlobalFilter 中对请求体中内容进行 Xss 和 Sql 注入过滤, 提高整体服务安全性, 还有在部分 url 请求过滤 token 进行身份鉴权。

Elasticsearch 搜索引擎:

	zhongyaocai	zhongchengyao	xuewei	xiyao	terminology	prescriptions	other	hospital	drug	diseasequestion	disease	diet
★ node-1	0	0	0	0	0	0	0	0	0	0	0	0
● node-2	0	0	0	0	0	0	0	0	0	0	0	0

1. 双节点部署:

使用双节点防止单一节点崩溃导致整体业务不可用, 提高高可用性。

2. 更高效的搜索速度:

底层采用倒排索引提高了搜索效率, 在约 10 万条医疗数据条件下进行模糊搜索用时在 0.001 秒内, 相比于传统模糊搜索提高近百倍速度。

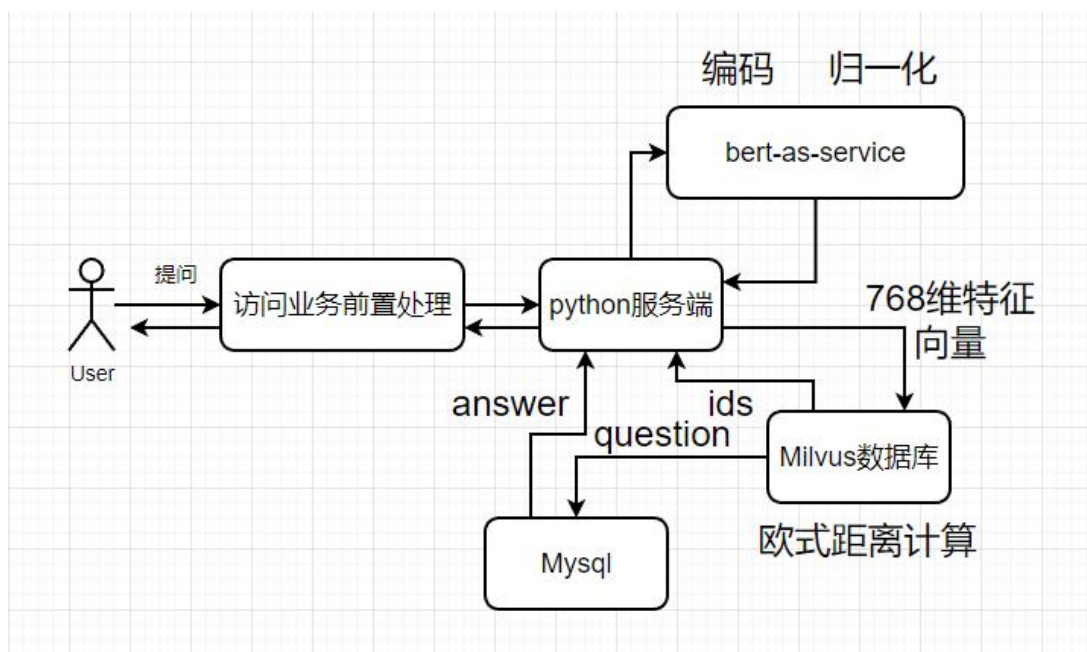
Redis 缓存:

- 1.对用户登录产生的 token 进行存储,并在 gateway 中配合过滤器实现身份鉴权。
- 2.对高请求的数据进行缓存,提高查询效率。
- 3.在扫二维码调用 api 时对药品信息进行缓存,解决了多用户在请求同一个十三位药品一维码多次调用 api 浪费调用次数的问题。

(2) 基于检索式的智能问答系统

基本流程:

使用 Bert 预训练模型作为基本模型,搭建 Bert-as-service,并将 80 万个问答对进行问句特征提取存入向量数据库 milvus 生成对应 id,再将 id,问题,回答存入 Mysql。每次用户提问时先通过使用 Bert 对用户问题进行特征提取并配合 sklearn 的 L1 范式归一化获取 768 维向量与向量数据库中存储的特征向量使用欧式距离算法相似匹配,获取距离最近的几个特征向量以及对应问题 id,再从 Mysql 中获取对应的回答。



丰富的数据集:

1. 本软件共收集 792099 个问答对,其中 94596 个男科,220606 个内科,183751 个妇产科,75553 个肿瘤科,101602 个儿科,115991 个外科。
2. 本软件共收集 117347 个医疗专业词汇预料,其中 11,170 个口腔科病历词汇,54,304 个国际疾病分类 ICD 全库,12109 个疾病诊断编码库 ICD-10,471 个医院固定资产词汇,37,308 个药品名称词汇,1985 个电子病历常见词汇。

Bert 服务:

使用较成熟的 Bert 模型,拥有强大的语言表征能力和特征提取能力以及很好的预训练模型,能更好的对用户语义进行理解,降低了开发问答系统的训练成本。

Milvus 向量数据库:

Milvus 有完善和快速的向量搜索服务,经测试 220607 个问答对在构建基于量化的索引(IVF_SQ8)

情况下仅用时 0.01 秒。

(3) 基于卷积神经网络的中药材识别介绍功能

基本流程:

用户将现实中中药材照片通过微信的接口 wx.uplordFile 上传，通过 python 端卷积神经网络进行分类获取对应中药材名称，后发送至 java 业务端在中药材大类数据中查找对应名称的具体信息后返回给用户。

数据申明:

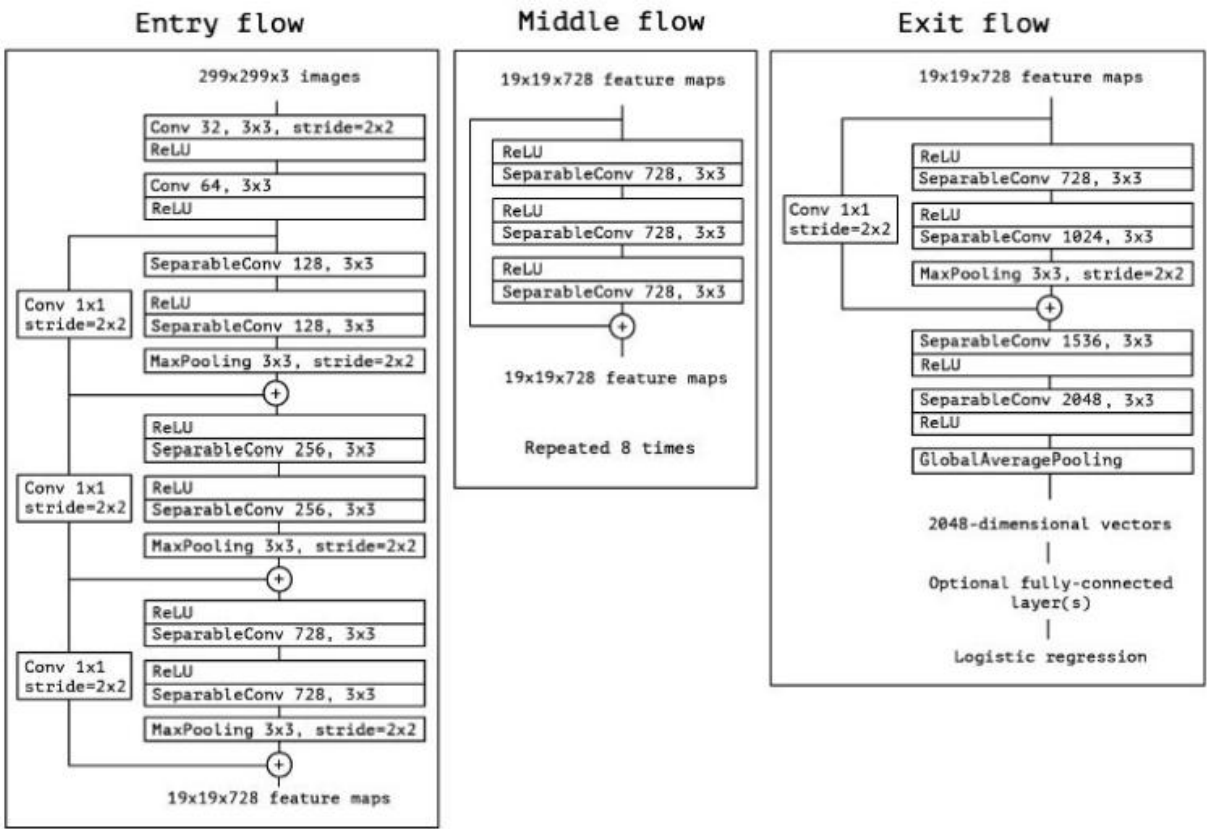
数据集来源于百度人工智能社区，有 100 类不同中药材共 8200 张相关图片。

目前可识别种类有:

安息香,白扁豆,白矾,白藜,白茅根,白前,白芍,白芷,柏子仁,北沙参,草拨,草澄茄,鳖甲,槟榔,苍术,草豆蔻,沉香,川楝子,川木香,川牛膝,大腹皮,淡豆豉,稻芽,地龙,冬虫夏草,防风,番泻叶,蜂房,甘草,干姜,甘松,藁本,赤石脂,枸杞子,桂枝,谷精草,谷芽,海龙,海螵蛸,合欢皮,黄柏,黄芪,黄芩,贝母,僵蚕,芥子,鸡冠花,锦灯笼,鸡内,荆芥穗,金果榄,金钱白花蛇,九香虫,橘核,地丁,莱菔子,莲房,莲须,莲子,莲子心,灵芝,荔枝核,龙岩肉,芦根,路路通,麦东,母丁香,羌活,千年健,秦皮,蝎子,忍冬藤,人参,肉豆蔻,桑寄生,桑螵蛸,桑葚,山慈菇,山奈,山茱萸,沙,石榴皮,丝瓜楼,酸枣仁,苏木,太子参,天花粉,天麻,土荆皮,瓦楞子,五加皮,细辛,银柴,胡,薏苡仁,郁金,浙贝母,枳壳,竹茹,猪牙皂,自然铜

技术介绍:

采用 Xception 深度可分离的卷积神经网络.相比与 VGG16 在该项目上测试集准确率提高了 15%，且更为轻量。



训练过程:

将 100 种中药材类别向量转换为独热编码的形式. 将所有图像处理成 (299, 299, 3) 形状放入图像增强函数 ImageDataGenerator 中进行图像变换以提高模型泛化能力再批量经过 keras 搭建的 Xception 卷积神经网络中训练约 20 个 Epoch. 最终在训练集的准确率为 86%, 在测试集准确率约 78%. (数据集作者并不是很官方, 训练集和测试集图片上广告和干扰图案噪声等很多, 对测试准确率产生很大影响)

(4) 权重智能推荐系统

基本流程:

当用户访问时具体类别时, 对应搜索业务模块向消息队列插入用户访问的具体类别和次数, 监控模块监听对应的消息队列频道并与缓存 (redis) 交互增加用户对应类别浏览次数, 当浏览次数达到预设值时 (50 次) 或用户收藏某一具体类别则向数据库中更新用户权重。当用户刷新推荐时会去数据库中获取用户在各大类中的权重并根据权重给用户推荐一定比例内容。

基于消息队列监控用户浏览数据:

使用 Rabbitmq 在解决了大量用户浏览时搜索服务调用监控服务的时间占用。用户浏览具体信息时不需要等待监控服务统计完权重才能获取结果, 搜索服务只需向消息队列中插入对应用户的信息和访问的大类信息即可完成操作。

(5) 扫条形码查看药品信息功能

功能背景和说明:

考虑到许多人购买药品时习惯性的丢弃说明书或药盒, 再次想查看药品详细信息时较为麻烦。所以用户可以将药品上的条形码进行扫描, 通过对微信 api 的调用, 即可查看药品具体信息

技术实现:

主要是通过接收前端给出的 13 位数字码调用 api 接口获取药品具体信息。

五、运营计划

将会在小程序所有功能测试稳定后上线试营, 并且推广让用户更方便的获取医疗医护信息。

六、系统开发工具与技术

开发工具: idea, pycharm, 微信小程序开发者工具

开发技术:

后端:

Rabbitmq: 用于监控用户浏览与收藏的中间件

Keras: 用于图像识别的深度学习库

Opencv: 用于图像识别的图像处理

Mysql:用于存储一些业务相关内容

Flask:部署 python 服务端

MybatisPlus:用于 java 操作 mysql 事务

Milvus:开源向量数据库，用于智能问答模块提供向量相似算法计算

Redis:对用户一些业务内容进行缓存。

Nginx:用于监控是否存在攻击服务器行为，设置黑名单，将请求路由至 gateway

Fastapi:用于 python 服务端的智能问答服务部署

Bert:提供预训练模型，对语句或词语进行编码提取特征向量

Elasticsearch:搜索引擎，存储大部分医护资料。

SpringBoot:用于快速部署单一 java 服务。

SpringCloud-Gateway:对接 nginx,将请求路由至服务注册中心上的服务。过滤 xss，sql 注入语句。

springCloud-Seata:微服务的全局事务管理，对事务操作错误进行全局回滚。

SpringCloudAlibaba-Nacos:服务注册中心，所有除 python 编写的服务外均注册到 nacos 上。

SpringCloud-OpenFeign:用于服务间的服务调用，负载均衡。

前端：

微信开发者工具：通过使用微信开发者工具中原生的 API，进行部分接口的调试与对接，通过 wx.uploadFile, wx.request 等原生的 API 方式，进行数据的访问和接收。并且通过对接口的请求和调用，来从后台中获取所需要展示的数据，将数据进行整理之后，在小程序中进行展示和分类。

七、作品服务对象

本应用服务于任何有医护信息查询需求的全年龄段用户，更偏向于年龄较大或是缺乏医护知识人群。

八、系统运行环境

服务端：Ubuntu 22.04，CentOS7.9。

客户端：微信小程序