



# 本 科 生 毕 业 论 文

题    目：基于云平台开发个人健康助手

院    系：数据科学与计算机学院

专    业：软件工程

学生姓名：冯张弛

学    号：13354068

指导教师：郑贵锋 讲师

二〇一七 年 三 月

## 学术诚信声明

本人所呈交的毕业论文，是在导师的指导下，独立进行研究工作所取得的成果，所有数据、图片资料均真实可靠。除文中已经注明引用的内容外，本论文不包含任何其他人或集体已经发表或撰写过的作品或成果。对本论文的研究作出重要贡献的个人和集体，均已在文中以明确的方式标明。本毕业论文的知识产权归属于培养单位。本人完全意识到本声明的法律结果由本人承担。

本人签名：

日期：

## 摘 要

随着移动医疗的高速发展，移动医疗服务在我们生活中扮演着越来越重要的角色。在这篇论文中，本文设计并实现了一个基于云平台支持的更高效且更加人性化的移动医疗类应用---WeBaymax。在云平台被广泛应用的今天，利用云平台与本地系统的通讯构筑的应用可以有效地解决移动端上数据储存量有限的问题。这套系统立足于Android平台，主要功能包括疾病查询、用药提醒与基于J48决策树算法的疾病自检。用户无论身处何处都可以通过本文介绍的应用连接云数据库获取即时的最新的服务。本文介绍了这套系统的实现细节并对最终结果进行展示。

**关键词：**移动医疗、基于云、移动健康服务、通信技术、决策树。

## **ABSTRACT**

With the rapid growth of mobile medical applications, mobile healthcare services play a vital role in our daily life. In this paper, we design a mobile medical application called WeBaymax based on cloud platform and Android that we tested to enhance higher quality and better performance of healthcare services. The author motivated to develop and apply communication technologies between cloud platform and mobile applications so as to solve the problem of limited data storage in mobile devices. The main features include disease query, medication reminder and auxiliary medical diagnosis based on J48 decision tree algorithm. Users can use this cloud based medical application and access to the personal data on the cloud database through any mobile devices anytime and anywhere. This article summaries the implementation details and presents the results of our application.

**Keywords:** mobile medical application, cloud based, mobile health- care services, communication technology, decision tree

# 目 录

第一章	概述/引言 .....	1
1.1	移动医疗应用的背景与意义.....	1
1.2	目前移动医疗遇到的问题.....	1
1.3	本文的工作.....	2
1.4	本文结构介绍.....	2
第二章	综述.....	3
2.1	系统相关介绍.....	3
2.2	相关工作.....	5
第三章	WeBaymax 系统概览 .....	6
3.1	系统总体功能分析与设计理念.....	6
3.2	系统主要功能需求.....	7
3.3	系统主界面.....	8
第四章	WeBaymax 系统设计 .....	9
4.1	系统总体功能结构设计.....	9
4.2	系统主要功能设计及实现.....	9
第五章	项目分析与结果 .....	13
第六章	结论 .....	18
参考文献	.....	19
致 谢	.....	20
附 录	.....	21

# 第一章 概述/引言

## 1.1 移动医疗应用的背景和意义

在互联网时代，移动医疗应用被越来越广泛地应用在临床治疗、疾病研究以及其他领域中<sup>[1][2]</sup>。移动医疗应用致力于减少病患在医院中所需消耗的时间以及非必要的检查<sup>[3]</sup>。这些应用为用户提供了许多有效且及时的功能，包括临床向导、医院挂号、相关知识查询以及基于当前位置的相关紧急服务<sup>[2][4][5]</sup>。他们有效提升了病人就医的治疗效果，并借助移动设备以及互联网缓解了目前社会医疗资源短缺的尴尬现状。

基于智能手机、平板电脑的移动医疗 App 已逐渐并必将成为健康信息化数据流的起点，同时也是终点。智能手机、平板电脑与移动式健康检测设备配合，获取用户体重、血压、心率、血糖等监测数据，并传输至数据库或云平台；经数据分析得到的健康评估结果及相应的健康干预内容也由短信、微信等反馈给用户。根据全球移动医疗市场调查及预测分析，到 2017 年，移动医疗服务的市场规模将达 260 亿美元。用户利用智能手机，通过移动医疗 App，不仅能满足其日常保健、在线问诊、慢性病调养等需求，还能检测多种疾病，如精神病、癌症、心脏病等。针对消费者，苹果公司在 IOS8 中集成 health kit 套件，具备计步、运动轨迹记录等功能；Well-Doc 公司主打糖尿病管理平台，提供个体化实时辅导；国内的“春雨掌上医生”、“掌上药店”等提供掌上问诊和咨询服务；“口袋体检”提供体检服务。

此外，借助移动医疗 App，可以提高医护专业人员工作效率。数据显示，专为医护人员设计的移动医疗 App 占移动医疗 App 的 15%，主要包括医学继续教育、远程监控及医疗保健管理等功用。<sup>[6]</sup>

## 1.2 目前移动医疗应用遇到的问题

医疗行业是一个“性命攸关”的行业，国家对此的管理和审核都比较严格，市场门槛也较高。当前，政策明确规定，医生在移动医疗 App 如“春雨掌上医生”、“快速问医生”等，只能提供健康咨询服务，不允许进行治疗或开具处方。这就限制了移动医疗的发展。对移动医疗设备制造商和 App 开发者而言，要兼顾国家法律法规以及相关政策，如果针对移动医疗的规定和标准不确定，则很难推进设备和应用创新。目前，医院已建立起较为先进的信息系统，如医学影像存档与通信系统、医院信息管理系统，给医院工作带来了极大便利的同时，也形成了比较固定的利益格局，涉及网络运营商、设备制造商、用户、服务提供商及医院等诸多利益相关方。医院作为移动医疗的主导者，打破这种固有而封闭的格局需要费一番工夫。另外，使用新的移动技术需要改变一些工作习惯，医生也许不愿接受这个改变。在信息化时代，移动医疗 App 必将涉及大量用户隐私信息，而医疗行业的信息保护手段远落后于其他行业，健康数据安全问题、远程控制风险较其他行业大。比如，通过“后门”漏洞窃取用户身份、位置、健康及医疗等信息；医疗健康设备的数据存储、传输未加密处理，数据在传输过程中易被截取；医疗健康数据上传到云端后被云服务提供商滥用牟利，等等。一旦发生信息泄露，

可能对移动医疗产业发展带来毁灭性打击，不仅是经济损失，而且可能会威胁患者的生命。<sup>[6]</sup>

冠状动脉性心脏病（CAD，简称冠心病），是冠状动脉血管发生动脉粥样硬化病变而引起血管腔狭窄或阻塞，造成心肌缺血、缺氧或坏死而导致的心脏病，主要表现为胸痛。冠心病通常伴随着例如高胆固醇、肥胖、高血压和糖尿病等等不可忽视的高风险、治疗难度较大的疾病。因此，心脏病的诊断及治疗要求医师具备丰富的临床经验及精准的操作技术。对于这个现状，一个可行的方案是将机器学习算法应用到临床诊断上。

目前，人工智能领域的一些发展促使了辅助医疗诊疗系统的出现。机器学习<sup>[1-4]</sup>作为人工智能领域的一个重要分支，也是发展最快的一个分支，在近年来已经逐渐被应用到医疗数据集的分析中。通过获取医院过去诊断的病例，进行整理后将这些诊断结果作为训练集导入计算机中对机器学习模型进行训练，机器将能够自动得到一个经过训练的分类模型以用于该疾病的辅助诊断。这个分类器能够帮助医生对新病人进行初步诊断，极大地提高医生的诊断效率以及准确性，甚至能够帮助非专业人士在不需要到达医院的情况下对自身状况有初步的了解。

移动医疗应用作为一个较为新兴的领域尚处于探索期，移动终端最大的弱点在于硬件的限制之下相比计算机而言计算能力要差许多。移动终端的不稳定性也很大程度上制约了需要大量数据交互的人工智能算法的应用。

### 1.3 本文的工作

出于改善医疗服务的需要，本文以冠心病为主要研究对象，结合云平台与手机终端设计并建立了一套辅助诊疗系统---WeBaymax，云平台作为承载数据计算与储存的载体，手机终端作为与用户交互的手段。整体开发主要分为四个阶段：收集合适的医疗临床数据作为训练集、利用 Clementine 对收集到的数据进行分类、比较几个不同的可监督机器学习算法的效果（例如贝叶斯网络、决策树 C5.0、神经网络等）、对最终的系统进行测试以及评估。

### 1.4 论文结构介绍

第一章是概述与引言，对于目前医疗应用的背景与意义、目前遇到的挑战与问题做了初步的介绍，提出了本文的研究内容并对下文的内容进行了介绍。

第二章的内容是综述，介绍本文提出的方法相关的工作并对他们进行分析与评价；对本文提出的系统 WeBaymax 所使用的相关工具与技术进行一系列介绍。

第三章是 WeBaymax 的概览，分析该系统的整体结构。

第四章将具体解析该系统的相关技术细节与实现方法。

第五章是项目分析与结果，将展示该项目的测试过程与其评估结果。

第六章是结论，对论文中完成的主要工作进行总结，并对以后所需的工作进行了展望。

## 第二章 综述

本章将对项目中使用的相关技术与现阶段该领域其他相关工作进行介绍。

### 2.1 系统相关介绍

#### 2.1.1 Android 系统

Android 是一个真正意义上的开放性移动设备综合平台，它包括操作系统、用户界面、中间件和应用程序，拥有移动电话工作所需的全部软件，同时其开放性保证该平台不存在任何阻碍移动产业创新的专有权障碍。<sup>[9]</sup> Android 应用程序开发基于框架和组件。Android 本身已在其框架中提供了许多组件供应用程序调用，开发者也可在开发应用程序时顺带开发新的组件，并将该组件放入应用程序框架中，以供自己和其它应用程序调用。Android 应用程序框架之下是一套 C/C++ 函数库，它们服务于 Android 应用程序组件，其功能通过组件间接提供给开发者。这些函数库包括：标准 C 函数库、媒体功能库、浏览器引擎、D 和 D 图形库及 Android 的 Java 程序运行环境包含一组 Java 核心函数库及 Dalvik 虚拟机，它们有效地优化了 Java 程序的运行过程。Android 系统平台基于优化了的 Linux 内核，它提供诸如内存管理、进程管理、设备驱动等服务，同时也是手机软硬件的连接层。<sup>[10]</sup>

Android 平台良好的开放性既能够促进技术与平台本身的创新与发展，也有利于降低开发的成本与难度，更可以方便地高度定制符合用户需要的特色产品，比起 iOS 更加符合作者的开发需求。

#### 2.1.2 Django (Python)

Python 是一种简单易学、功能强大的编程语言，它有高效率的高层数据结构，简单而有效地实现面向对象编程。Python 简洁的语法和对动态输入的支持，再加上解释性语言的本质，使得它在大多数平台上的许多领域都是一个理想的脚本语言，特别适用于快速的应用程序开发。

Django 是一个高级的、支持快速开发的、简洁实用的 python 网络框架，从 2005 年发布就被许多站点广泛使用。在自由开源的 Django 的帮助下，开发者可以把更多的精力投入到 app 的编写中。本文使用 Django 创建了一个数据库驱动的 Web 应用并使用 python 的 API 以便对储存的数据进行操作。Web 应用将会为客户端提供一系列有效方便的接口。

Hypertext Transfer Protocol (HTTP) 是一种被设计用于客户端与服务器之间通讯的请求-反馈协议。当客户端向服务器提交一个 HTTP 请求时，服务器会对客户端的请求进行反馈。建立连接的主要步骤大致如下：首先在管理页面创建并登记本文设计好的模型；之后设计一些能够让用户使用以接入服务器的 URL。当用户请求一个页面时，Django 将会按顺序遍历所有的页面以此寻找与该 URL 对应的页面并在匹配到之后停止遍历。如果遍历之后未发现匹配项，Django 将



会返回 404 页面。一旦请求被确认，Django 会加载并调用对应页面，这一切都是靠 python 函数来实现的。在确认流程之后，在客户端上会接收到一个完整的 HTTP 反馈，其中包含了用户请求的页面。如此，部署在远程服务器上的数据库与网页之间的连接就建设完成了。页面通过包括在 URL 请求中的参数对数据库中的内容进行检索，系统需要分析这些参数再进行数据的查询。在安卓平台上，当连接建立之后，用户可以通过点击按钮或输入信息进行信息查询，后台进程将会对请求进行重组并向云服务器提交数据。当反馈收到之后，后台进程将会对收到的数据格式（例如 json、xml、键值对）进行解析以提取信息，并对其进行分析以更新至界面上。通过这种方式，用户可以获取到自身需要的即时信息。

以一个简易 Blog 开发为例，来讲述基于 Python 和 Django 的 Web 开发流程。该 Blog 的开发流程如图 2.1.2 所示。<sup>[11]</sup>

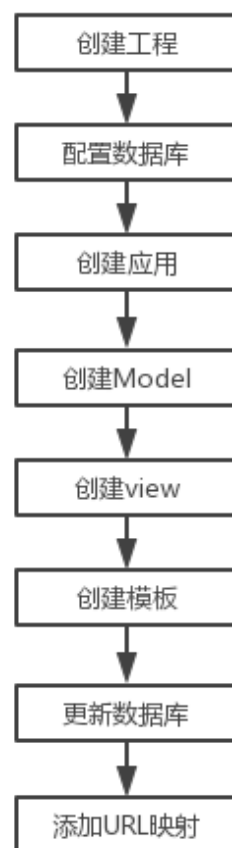


图 2.1.2 Django 框架下简单 blog 的开发流程<sup>[11]</sup>

### 2.1.3 开发工具

本文中使用的开发工具为 Android Studio。作为近年来被用于替代 Eclipse 进行 Android 开发的新工具，Android Studio 比起 Eclipse 要专一的多。Android Studio 是一种基于 IntelliJ IDEA 的免费 Android 开发环境，IntelliJ IDEA 是一种类似于 Eclipse 的 Java 语言的集成开发环境，在智能代码完成、代码自动提示、重构、J2EE 支持、Ant、JUnit、CVS 整合、代码审查、创新的 GUI 设

计等方面有优异表现。IntelliJ IDEA 与 Eclipse 相比，更佳的智能化思想已经深入到代码重构、自动完成和调试等方方面面，极大减轻了开发者负担提高了开发效率。<sup>[12]</sup>

## 2.2 相关工作

近年来，在机器学习领域的学者们探索之下，有数种机器学习算法已经被应用到高效率医学诊断系统的开发中。对于之前乳腺癌诊疗中机器学习算法表现不佳的状况，一种全新的决策树辅助诊疗算法（CADx）应运而生，其利用超声波病理图像作为训练集，采用他们的协方差作为他们的特征参数来构建决策树模型。结果，CADx 系统在灵敏度与准确度方面表现相当出色甚至略优于一个有经验的医生<sup>[7]</sup>。在神经网络领域中，专家给出了三个例子分别在心肌梗死,严重受伤后背部疼痛和生存概率方面来证明他的结论。这些例子说明，神经网络的表现几乎能够与一个经验丰富的专家相媲美。用机器进行医疗分析诊断不仅能够减轻目前医生的负担更能够提高治疗的效率以及准确性。<sup>[8]</sup>由此可见，将机器学习算法应用到医学诊断中是较为可行的。在大量数据的积累与总结方面，机器学习算法的效率要远远高于人类，培养成本也要更低。除此之外，机器比起人类对事物的看法更为客观，表现也更为稳定。

### 第三章 WeBaymax 系统概览

本章主要对本文提到的系统 WeBaymax 的设计理念与主要功能结构做相关介绍。

#### 3.1 系统总体功能分析与设计理念

本文的设计理念是融合 Android 与云平台的优点，建立一套基于云的辅助诊疗应用---WeBaymax。在 Android 平台上应用将为用户展示一个人性化的操作界面，让用户能够随时随地在手机端上方便全面地使用服务；云平台将发挥它大数据量储存、高效率计算与集成的数据管理等优点为移动端的服务提供保障。通过云计算与移动应用的结合，线上的数据管理与线下的医疗保健将会有更多更加有效的方式为用户提供帮助。

随着可获取的健康数据的大量增加与从历史病例数据中规范提取有效数据的迫切需求，在临床研究中逐渐形成了一套规范化的数据标准。本文从 UCI 的机器学习数据集中获取到医疗数据。之后将一些已知经过正确诊断的医疗数据应用到一些机器学习算法（例如贝叶斯网络、决策树 C5.0、神经网络）并生成预测分类器。不同算法的表现将被比较并评估。最终，分类器将被储存在云平台上并在移动终端上进行测试。

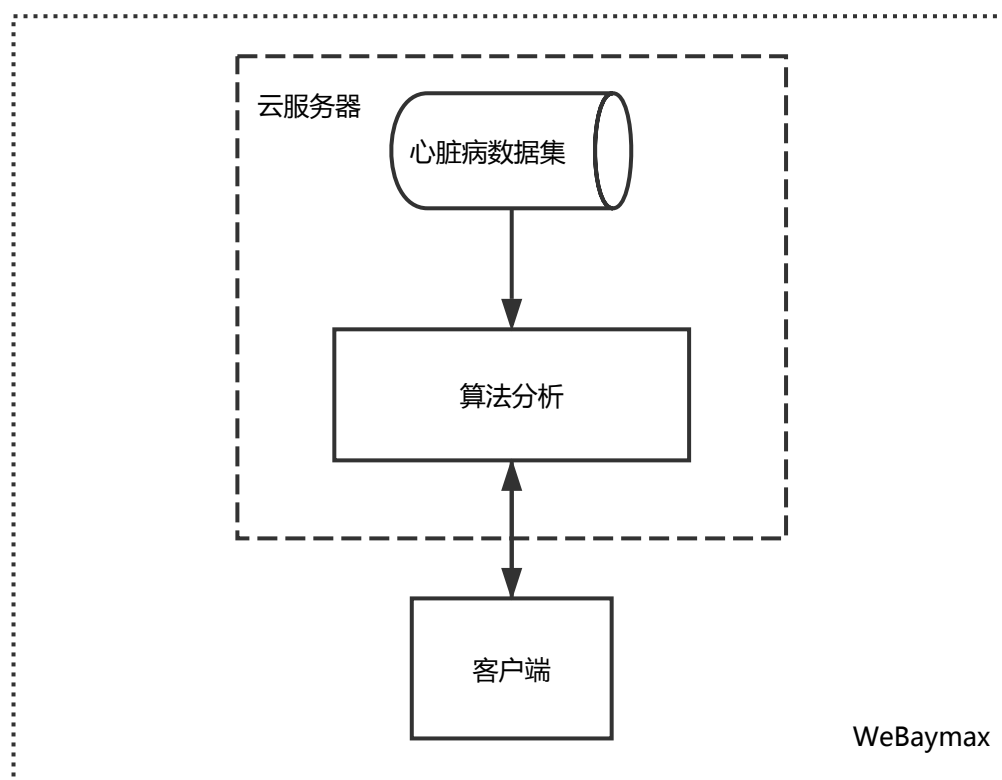


图 3.1 WeBaymax 结构

WeBaymax 的设计动机是让用户能够自由充分地管理自己的个人健康数据并进行有意义的数据分析。此应用包含了多个功能，图 3.1 展示了 WeBaymax 的设计框架。

## 3.2 系统主要功能需求

### 辅助医疗诊断

与以往通过文字输入来描述患病位置的方式不同，WeBaymax 提供了一个 3D 人体模型来帮助用户更加直观准确地定位患病位置，并给出了一些常见症状供用户进行选择。用户可以在人体模型上直接指出不适区域并与系统进行进一步交互。在定位可能患病的区域之后，系统会通过询问问题了解用户的病情，并根据患者的情况给出一些临床建议方便用户进行进一步治疗。以心脏病为例，在患者选择胸口位置之后进入心脏分支，系统将会从云端获取相关问题推送给用户，在问题的答案上传到云端之后，系统将会根据用户给出的回答缩小用户所患疾病的范围并给出下一个问题，直到系统得到一个较为可靠的结果。

### 疾病查询

针对目前互联网上医疗资讯缺乏规范化管理、医学谣言随处可见的问题，本文在系统中结合了规范化的疾病数据库供用户查询。借助云数据库的大容量，开发人员与用户将不需要担心本地应用程序的数据储存问题。用户可以通过输入疾病名称搜索得到相关疾病的详细信息，其中包括疾病名称、症状、治疗方法以及注意事项。此外，考虑到疾病名称的种类繁多，提供了关键字模糊查询，以使用户能够更快查询到所需的信息。

### 服药提醒

服药提醒功能将会按照设定的时间弹出提醒用户服药。众所周知，对于一些苛刻严格的服药要求，用户不仅需要花时间记住服药的时间，更需要记住服药的剂量与药品名，这是非常繁琐的。对此，本文提供了一些比较实用的功能以防患者忘记用药。为了更加方便快捷地记录药品信息，本文采用了拍照的形式来取代以往传统的文字形式对药物进行记录。照片带来的图像信息将更加直观，防止用户错误用药。在设置时间之后，应用将会在对应的时刻提醒用户使用正确的药物。不仅如此，在登录之后用户可以上传备份自己的用药提醒信息到云服务器，备份之后用户在其他设备上也能够随时随地获取到自己的提醒信息，在云服务的帮助下进行在线共享与管理。这种以图片为载体的提醒模式比起以往的文字提醒要直观许多。

## 3.3 系统主界面

在打开 app 之后，系统在经过欢迎页面之后将会进入主页面。主页面如图 3.3。主页面的设计模仿了当前较为流行的样式，分为三个部分：顶部导航栏、中间内容展示区与底部的标签栏，可以根据不同情况的需要调整显示的内容。

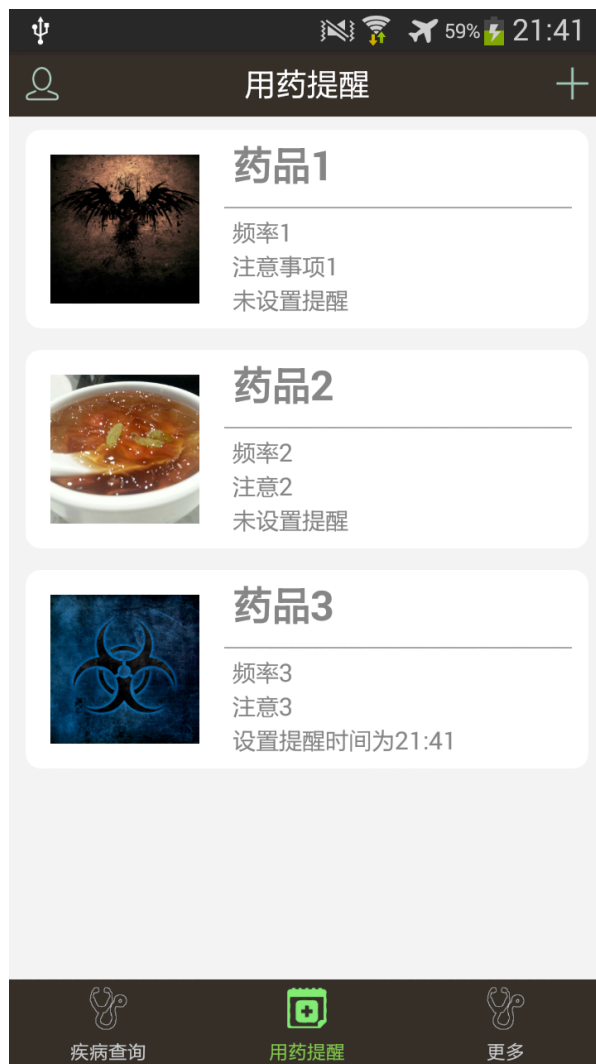


图 3.3 系统主界面

## 第四章 WeBaymax 系统设计

继前文对该系统的初步介绍,本章将更进一步对本系统的主要功能的具体使用方法及实现细节进行说明。

### 4.1 系统总体功能结构设计

系统总体功能设计如图 4.1。主要功能一共分为 4 个：疾病查询、用药提醒、天气预报与疾病自检。

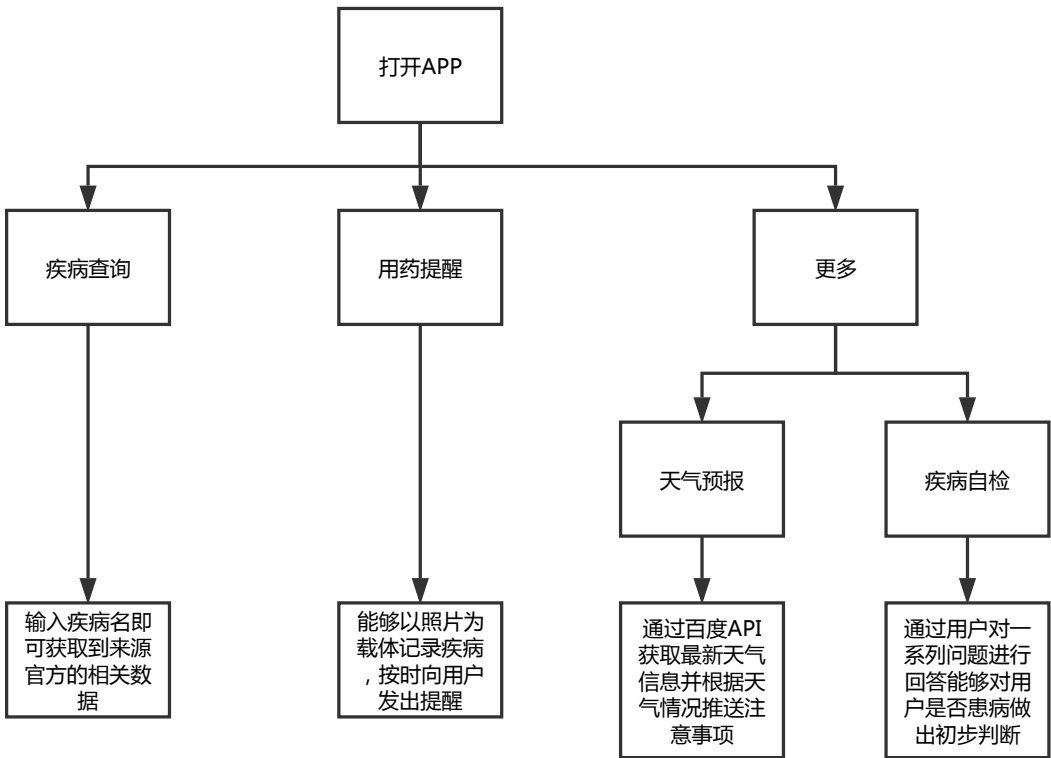


图 4.1 系统流程图

### 4.2 系统主要功能设计

#### 疾病查询设计

交互过程如图 4.2.1 所示。用户在输入疾病名称并点击搜索按钮之后，应用将会向服务器发送一个 GET 请求以获取相应的数据。服务器在接收到请求之后将会调用对应 python 函数对 mysql 数据库中的 eyedisease 表（因为资源的限制，当前只提供眼科疾病的相关数据）进行查询操作。在获取到资料之后服务器将对资料进行打包并返回 json 格式的信息。InspectorFragment.java 将对 json 格式进行解析并呈现到 inspector\_fragment.xml 上。

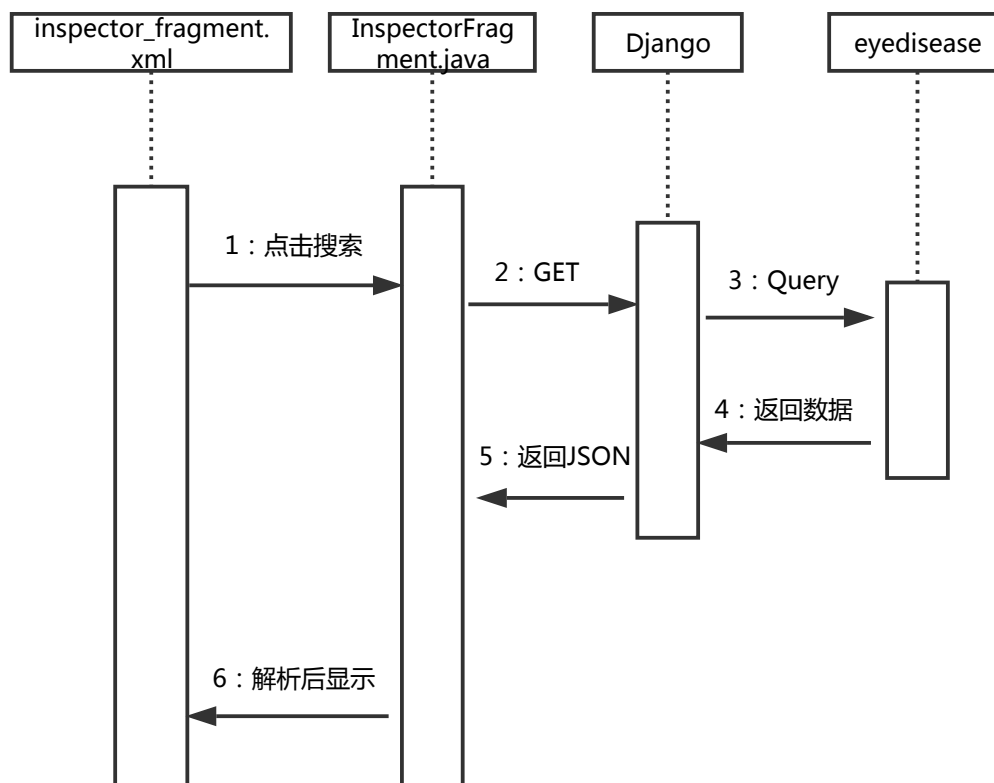


图 4.2.1 疾病查询用例顺序图

## 用药提醒设计

交互过程如 4.2.2 所示。用户可从导航栏右侧的“+”号处新建新的服药提醒事项或直接点击现有事项对参数进行编辑。在编辑完成之后点击确认按钮之后，ReminderFragment.java 会将用户填写的内容通过 MyDatabaseHelper 实例添加或修改到本机的 SQLite 数据库中。系统创建了一个名为 MedicationReminder 的表用于储存药品信息例如药品名、图片名（图片 id）、使用频率以及提醒时间等，并通过用户提供的手机号码作为外键连接该表与 User 表。因为提示图像不需要过高的显示精度，通过减少图像的采样点，本文将手机拍摄的图像进行压缩使其占用的空间更少，反应的速度也更加快。除了正常的列表操作之外，本文还为用户提供了一些进阶的交互操作，例如长按项出现上拉菜单等。

在每次修改之后，MyDatabaseHelper 同时会向系统的 AlarmManager 提交对应的定时事件。AlarmManager 是 Android 系统中的一项提示服务，会在特定的时刻广播一个指定的 intent，本文使用的提醒功能主要由它实现。

ReminderFragment 在每次调用 onResume 方法时会向数据库获取最新的数据并将最新的信息更新到控件中。同理，对当前事项进行删除操作时调用顺序与增加一致。

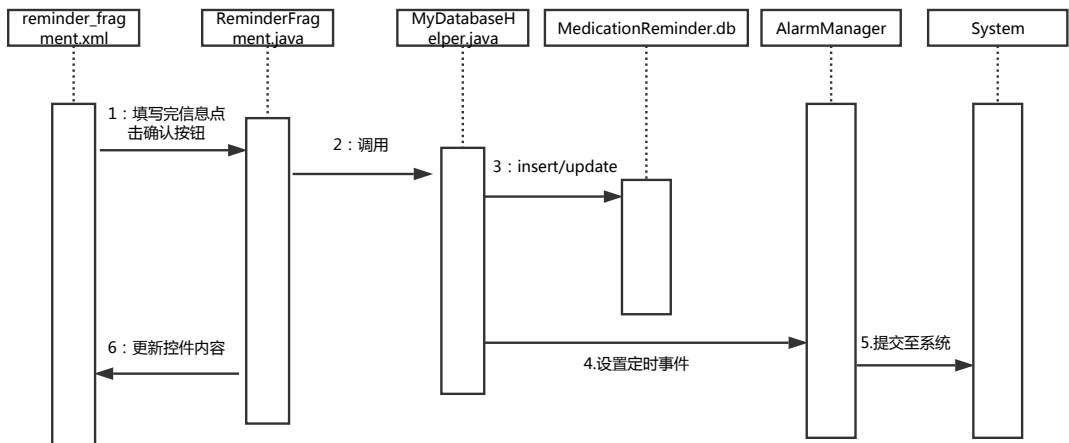


图 4.2.2 用药提醒用例顺序图

## 天气预报设计

交互过程如 4.2.3 所示。对比不同的天气服务，本文最终选择了使用百度提供的天气 API 作为信息来源。百度天气能够为系统提供即时的天气信息以及后续多达 15 天的天气预报，还有多种可选的相关信息，为开发者提供了高效、稳定、丰富多样的天气数据服务。该功能直接使用了百度天气的 API，在点击刷新按钮或到达预设时间（默认频率为一天更新一次）时，WeatherActivity 将会通过手机的定位功能获取当前所在城市再通过 API 获取近日该地区的天气情况，并对 API 返回的 json 数据进行解析并呈现给用户。

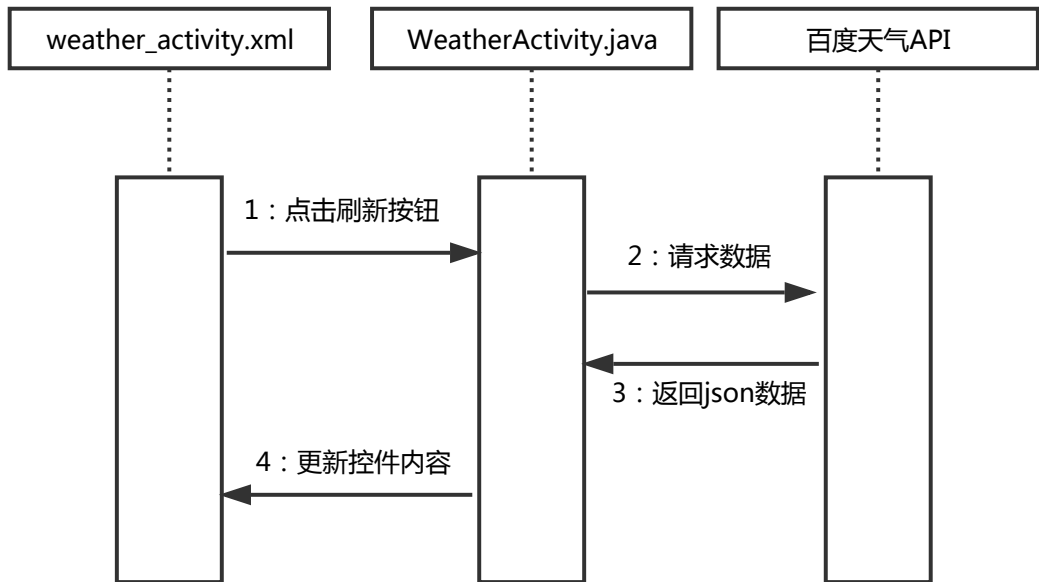


图 4.2.3 天气预报用例顺序图

## 疾病自检设计

交互过程如 4.2.4 所示。在进入疾病自检入口之后，首先呈现的是一个使用 Unity3D 创建的人体模型界面。在 Unity3D 完成模型之后将模型生成的 jar 包导入到 Android Studio 中，可以得到 UnityPlayerNativeActivity.java 用于显示模型，



ModelActivity 即是其子类。在点击模型不同位置之后，ModelActivity 会将点击的位置信息封装在 intent 中传递给 QueryActivity 以便对点击事件进行处理。QueryActivity 将根据点击部位的不同向服务器发送不同的请求以获取不同部位的相关问题。服务器收到请求之后结合事先建立的模型与数据库的资料将合适的问题返回给客户端。如此便完成了一次迭代，在多次迭代之后服务器中的模型会得出结果并向客户端返回结果，终止迭代。

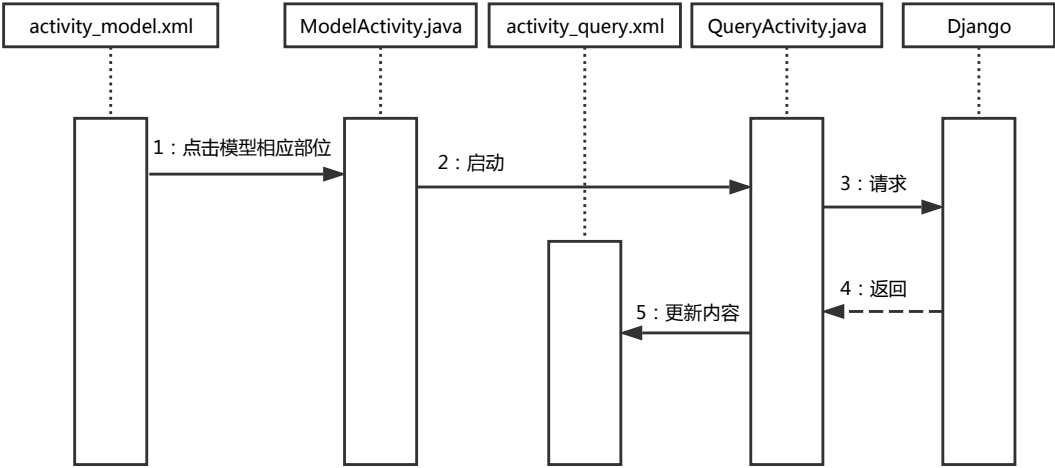


图 4.2.4 天气预报用例顺序图

为了提高辅助心脏疾病诊断的准确性，本文在临床数据上应用机器学习算法。本文比较了几个机器学习算法的效率，包括贝叶斯网络、J48，KNN 和 C4.5，分析了这些算法的结果。这些算法可以根据训练集和类标签对对象进行分类，然后使用规则对测试集中的新对象进行分类。出于可解释性和性能的综合考虑，本文选择决策树 J48 作为最终算法来构建模型。决策树算法构筑的节点模型更为形象更容易被用户理解并接受，提升用户对产品的信任，让结果更具有可信度。本文在 Waikato Environment for Knowledge Analysis(WEKA)<sup>[13]</sup>工具的帮助下构建了本文使用的决策树模型，并通过创建表来储存问题从而将树节点结构储存到云服务器上，每个节点存储的一个问题对应了疾病的一个症状。但是由于过拟合的问题，树结构很可能在训练集上表现得非常好但在测试集上表现糟糕。这个问题应该被重视因为模型只能依赖训练集进行构筑修正，过拟合的模型不能被应用到其他情况中。针对这个问题本文提出的解决方案是对树进行剪枝。本文引入惩罚项来惩罚树从而减少叶与子树的数量。通过这种方式，决策树将可以在合理范围内对自身进行修正以获得在测试集上的更好表现。通过统计每个叶片上的错误数来决定是否对其进行剪枝。

系统储存了一棵经过剪枝的心脏疾病数据相关的决策树。交互用的问题将在云平台与 Android 平台之间进行传递。在 Android 平台上作出回答之后，应用将对服务器发送 URL 请求，其中问题的回答将被一同打包上传。服务器根据 URL 请求提交的参数对数据进行处理，根据当前节点位置的不同（从根节点到叶节点），向用户反馈不同的信息。

WeBaymax 为用户提供了一个人体模型以使用户进行交互，方便操作也能够更直观地反馈数据。这个界面设计充分迎合了用户的操作习惯。原框架被设计用于通过患者过去的表现预测患者患心脏病的可能性及风险。为了提高疾病诊断的

准确性，用户需要与应用进行交互，通过回答来自于服务器的问题向应用提交自己对应的特征信息。当用户想要测试自己是否有患心脏病的风险时，应用会弹出问题询问患者相关问题并给出一些可选的特征。在患者作出选择之后，服务器会根据患者的回答检索得到下一个问题。在问题的最后，服务器将对所有回答获取到的特征进行分析并将最终的结果推送给用户。

## 用户身份验证

WeBaymax 将允许用户上传他们的个人健康数据并进行医疗数据的管理，为此，对用户的身份进行验证将成为必要的用户隐私信息的保障手段。在用户第一次使用 WeBaymax 时，他需要注册并提交一些必要的信息，例如手机号码。在云服务器上，因为手机号的唯一性，本文使用手机号作为标识符，沟通不同表上的信息。同时，手机号能够方便即时地让用户与系统进行沟通，将其直接作为账号不仅能够减少用户的记忆负担，在服务器角度看此举亦可减少无效用户的存在，更大程度地保障了用户的信息安全。除此之外，服务器上的私人数据不能被其他方式读取，在某种程度上确保了信息安全。在用户登录之后，用户能够在服务器上对自己的个人数据进行备份（例如用药提醒等）。

## 第五章 项目分析与结果

相比于传统的服务器，云服务器有着更高的可靠性，更加有效且易于使用<sup>[15]</sup>。因此，本文在阿里云服务器上测试了之前设计好的模型并对 WeBaymax 的表现进行评估。

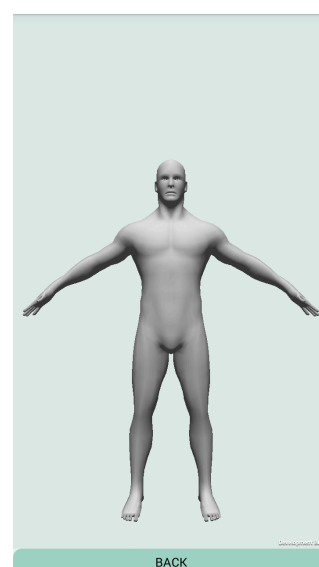
本文在 Android 手机上运行测试得到结果如图 5。其中，注册界面如图 5(a)。用户可以在该页面创建他们的个人账户或输入账户密码登录他们的现有账号，服务器会根据你输入的信息进行识别与判断。如果验证通过，用户便可以自由的享受医疗服务。用户可以通过搜索框搜索经过认证的储存在远程数据库中的疾病相关信息如图 5(b)。如图所示，只需要用户按下按钮，用户所需的信息将会被系统性地组织好并呈现给用户。图 5(c)(d)(e)(f) 展示了心脏病相关的诊断过程。当用户点击这个模型的指定位置时，一个对应的问题将会弹出。经过若干个问题的交互，用户将可以得到大致的诊断结果。另一个相当实用的功能是服药提醒，如图 5(g) 和图 5(h)。用户可以填写药物的详细信息，例如药物的照片、名称、频率等。闹铃将会在用户设置的时间按时提醒用户用药。不仅如此，用户可以上传或下载自己的个人数据到云服务器上。另外，在点击已有事项时可以浏览详细的事项信息并对其进行修改，长按事项会弹出高级菜单（如图 5(k)）。图 5(i)(j) 展示了登录前与登录后侧滑栏的不同表现，作为一种比较流行的 UI 设计，本文也将其加入系统。在每次启动时系统会展示欢迎页面，如图 5(l)。



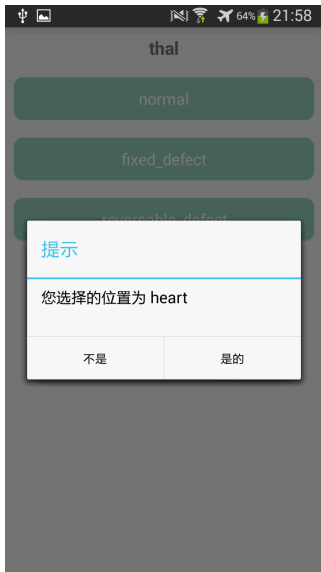
(a)



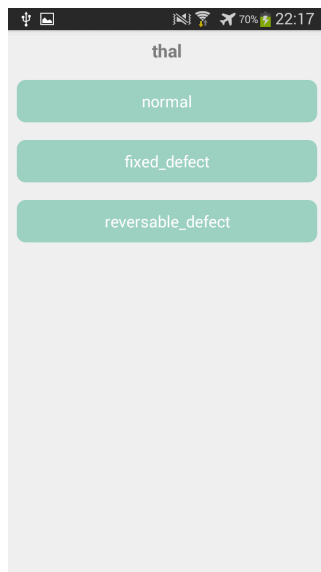
(b)



(c)



(d)



(e)



(f)



(g)



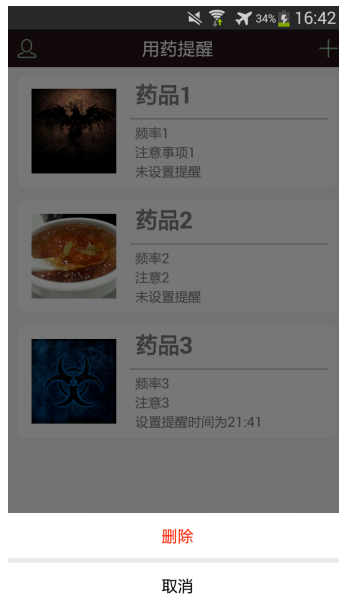
(h)



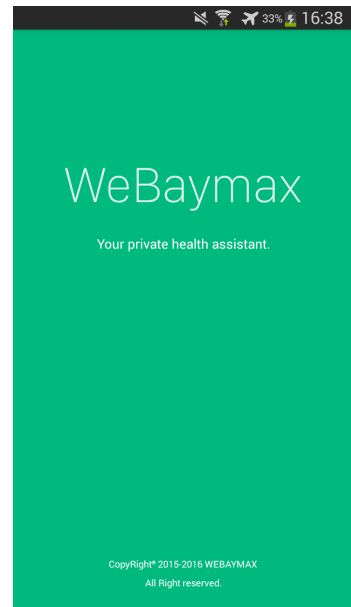
(i)



(j)



(k)



(l)

图 5 手机运行效果图

如前文所介绍，本文的项目使用 J48 决策树算法构建决策树，并在心脏病数据集上使用十倍交叉验证对其表现进行测试<sup>[16]</sup>。实验数据集由心脏病的 13 个主要属性（如表 1）组成，这 13 个主要属性是由研究人员从原始数据集中的 75 个属性中筛选提取得到，描述了心脏疾病中一些共性的特征。本文认真地分析了这些属性内容并根据他们的数据类型进行了预处理。在模型测试时，有效数据集被分成两个部分：测试集与训练集。分类器通过训练集进行学习并再测试集上进行结果评估。最终，分类器将对所有的数据分为两个大类，分别为有心脏病隐患与无心脏病隐患。

序号	类型	属性
1	实数	年龄
2	布尔	性别
3	选项	胸部疼痛类型（4 种）
4	实数	平静状态血压
5	实数	血清胆固醇（mg / dl）
6	布尔	空腹血糖
7	选项	静息心电图结果（0， 1， 2）
8	实数	最大心率
9	布尔	运动引发的心绞痛
10	实数	Oldpeak=ST depression induced by exercise relative to rest
11	范围	The slope of the peak exercise ST segment
12	实数	通过荧光透视着色的主要血管数量（0-3）
13	选项	Thal: 3 = 正常； 6 = 固有缺陷； 7 = 可逆缺陷

表 1. 数据集的 13 项主要属性

结果	实例数量	占比
正确分类的实例	216	80%
错误分类的实例	54	20%

表 2. 树在数据集上的表现

TP Rate	FP Rate	Precision	Recall	F-Measure	Class
0.853	0.267	0.8	0.853	0.826	1
0.733	0.147	0.8	0.733	0.765	2

表 3. 评估模型表现的 5 个指标

减少过拟合与提高智能疾病诊断的准确性是使用剪枝决策树的关键因素。最终在 WEKA 的帮助下系统成功生成了需要的树并将其结构储存到云上，树的具体表现详见表 2。本文通过分析 5 个标签来评估模型的表现，其中包括 TP rate、FP rate、Precision、Recall 和 F-Measure（表 3）。从最终的测试结果中，可以看到伴随着部分节点的减少，准确性有了一定程度的提升；互动问题的减少在合理范围内缩小了决策树的规模，树结构不至于过于臃肿。

## 第六章 结论

这篇论文介绍并展示了基于云的健康助手 WeBaymax 的各项功能，包括在线数据存取、数据管理与健康服务。该应用让用户能够方便地管理自己的医疗记录并获取即时的医疗资讯，通过云储存极大的扩展了手机上有限的数据储存量。在 WEKA 的帮助下，本文已经初步实现了心脏病相关的疾病诊断，也成功地在服务器与客户端之间实现了可靠的连接。本文向用户提供可靠可用可信的即时服务，致力于提升用户体验，造福广大患者、医生甚至更多的其他用户。未来的工作中，作者计划从以下方面入手：因为缺乏医疗或其他机构的支持，目前本文用于疾病查询的数据库只有部分眼科疾病，数量非常有限；疾病自检方面的数据集由于算法的计算需求要求更为严苛，各项指标类型对于普通人而言过于晦涩而难以回答，目前作者只有关于心脏病的数据集。另外，在疾病自检的最后，尽管可以得到大致预测结果，但作者目前并没能给出具体的建议与进一步处理的方案，在专业的医学数据方面还有很多欠缺。未来，作者希望能够积极寻找合适的数据库或专业的第三方机构支持。由于缺乏 unity3D 的编写经验，目前的人体模型略显简陋，操作不够流畅，未来会进一步优化模型。

## 参考文献

1. Thomas Lorchan Lewis and Jeremy C Wyatt. mhealth and mobile medical apps: A framework to assess risk and promote safer use. *Journal of medical Internet research*, 16(9), 2014.
2. Upkar Varshney. Pervasive healthcare and wireless health monitoring. *Mobile Networks and Applications*, 12(2-3):113–127, 2007.
3. Michael Armbrust, Armando Fox, Rean Griffith, Anthony D Joseph, Randy Katz, Andy Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica, et al. A view of cloud computing. *Communications of the ACM*, 53(4):50–58, 2010.
4. Upkar Varshney. *Pervasive healthcare computing: EMR/EHR, wireless and health monitoring*. Springer Science & Business Media, 2009.
5. Ilias Maglogiannis, Charalampos Doukas, George Kormentzas, and Thomas Pliakas. Wavelet-based compression with roi coding support for mobile access to dicom images over heterogeneous radio networks. *Information Technology in Biomedicine, IEEE Transactions on*, 13(4):458–466, 2009.
6. 钟姝雅, 任建萍, 刘肖肖,等. 移动医疗的应用终端(App)规范问题[J]. 卫生经济研究, 2016(10):49-51.
7. Kuo W J, Chang R F, Chen D R, et al. Data mining with decision trees for diagnosis of breast tumor in medical ultrasonic images.[J]. *Breast Cancer Research and Treatment*, 2001, 66(1):51-7.
8. Brause R W. Medical Analysis and Diagnosis by Neural Networks[M]// *Medical Data Analysis*. Springer Berlin Heidelberg, 2001:1-13.
9. 姚昱旻, 刘卫国. Android 的架构与应用开发研究[J]. *计算机系统应用*, 2008, 17(11):110-112.
10. 公磊, ZHOU Cong. 基于 Android 的移动终端应用程序开发与研究[J]. *计算机与现代化*, 2008, 2008(8):85-89.
11. 王冉阳. 基于 Django 和 Python 的 Web 开发[J]. *电脑编程技巧与维护*, 2009(2):56-58.
12. 陈甫. Android Studio 应用 [J]. *电脑知识与技术: 学术交流*, 2014(24):5659-5661.
13. Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H Witten. *The weka data mining software: an update*. *ACM SIGKDD explorations newsletter*, 11(1):10–18, 2009.
14. Panagiotis Germanakos, Constantinos Mourlas, and George Samaras. A mobile agent approach for ubiquitous and personalized ehealth information systems. In *Proceedings of the Workshop on Personalization for e-Health of the 10th International Conference on User Modeling (UM 2005)*, pages 67–70, 2005.
15. Chunye Gong, Jie Liu, Qiang Zhang, Haitao Chen, and Zhenghu Gong. The characteristics of cloud computing. In *Parallel Processing Workshops (ICPPW), 2010 39th International Conference on*, pages 275–279. *IEEE*, 2010.
16. Uci machine learning repository: Heart disease data set. <http://archive.ics>.



[uci.edu/ml/datasets/Heart+Disease](https://uci.edu/ml/datasets/Heart+Disease).

## 致 谢

感谢中大四年来对我的辛苦培育，让我在大学这四年来学到很多东西，特别感谢移动信息工程学院（现数据科学与计算机学院）为我提供了良好的学习环境、感谢领导、老师们四年来对我无微不至的关怀和指导，让我得以在这四年中学到很多有用的知识。在此，我还要感谢在班里同学和朋友，感谢你们在我遇到困难的时候帮助我，给我支持和鼓励，感谢你们。

特别感谢我的指导老师郑贵锋，在本项目开发中给予我悉心指导，从系统开发到结束中过程遇到很多困难都是他给我鼓励与指引，使我能够克服重重困难，将系统完成，在此谨向郑老师致以诚挚的谢意和崇高的敬意。谢谢！

## 毕业论文成绩评定记录

指导教师评语：

成绩评定：

指导教师签名：

年 月 日

答辩小组或专业负责人意见：

成绩评定：

签名（章）：

年 月 日

院系负责人意见：

成绩评定：

签名（章）：

年 月 日

