# 结题报告

1. **项目背景与需求**
   1. 项目背景描述

在步入21世纪以来，医疗行业借着互联网的东风，获得了蓬勃的发展。但是传统的医院在患者的眼中，看病难，看病贵的问题长期无法解决。一个感冒却要化验血液，拍片，传统医院给患者的用户体验度太差。阿里巴巴于2014年开始向医药健康产业进军，并于之后成立了阿里健康。2017年，腾讯结合人工智能成立了AI医疗影像平台“腾讯觅影”。如今，各个互联网巨头公司，包括百度，微软，谷歌也已经介入互联网医疗领域。由此可见，互联网+医疗给企业带来了巨大的商业前景。

MORE health 采取的是一种签约国内外一流医生，可以通过网络远程共同对病情进行会诊，主要针对的是比较疑难的杂症。而近期Babylon宣布其聊天机器人的医疗问诊软件可以在人类医生考试中取得与人相当的成绩，采用人工智能对症状进行分析并生成个人健康评估并就是否去医院给出合理化建议。

丁香医生最开始主要是基于检索模式的医学知识科普和疾病问答的平台，目前专家在线坐诊以及疾病小组论坛也已上线；平安健康于2014年推出平安好医生平台，app中的问诊大厅科室齐全，存在数百位全职医生在线坐诊并支持图文语音与实时对话；

尽管国内的市面上存在诸如丁香医生app,平安好医生app之类的软件，但是它们采用的是一种医师坐诊的形式。诸如，聘请医院的医师，在网上作为客服为患者进行诊断。尽管他们得到的诊断结果可能较之于智能机器人更加确切，但是仍然有几个缺点。首先，是所耗费的成本太大，其中专业医师的支付费用占据大部。其次，并没有实现特色性诊断，没有对用户进行画像，而导致每次诊断，都当作是一个全新的用户，诊断或者给出建议的时候并没有考虑到患者的体质状况或者其他身体特征。

* 1. 国内外调研

通过对国内外互联网医疗行业的调研，针对医疗机器人，现阶段主要都是以一些简单实时的聊天系统为基础的，主要有以下三个方向：

* + 1. 从聊天机器人的聊天来入手，通过对医疗领域自动问答领域的研究，基于传统检索技术的问答系统，利用关键字匹配即对问题答案进行检索。或者基于语义技术的问答系统，从而语义层面理解用户提出的问题，同时将数据以资源矿建三元组（资源-属性-属性值）形式进行存储，从而实现医学知识的共享和利用。
    2. 从用户的角度入手，将机器人作为一个助手来监控病人情况。通过病人的态度，基于条件熵来进行构建一种检索模型，分析并提供在患者和护理人员之间的假设对话场景中的实验结果。
    3. 从用户的潜在需求入手，主要争对药物相关的询问与检索，主要就是构建一个简单实时的聊天系统，文本、图片或者其他统一的规定实现交流，以回复内容为王，不要用自己的幻想或者冗余消息去打扰用户。尽可能能每时每刻都能为客户服务。
  1. 项目需求分析

功能性需求

* + - 1. 实现机器人与人进行对话，能根据症状对疾病进行诊断
      2. 构建用户画像模型，并通过与患者进行对话，实现对患者的健康画像，包括年龄，性别，慢性病，体质状况等。
      3. 根据患者的画像以及症状给出更加个性化的合理建议。

1. **设计思路与技术路线**
   1. 项目设计思路

本项目系统的设计主要涉及到两个方面的问题：

* + 1. 系统数据库的创建

由于在项目初期所拿到的数据库仅仅只有一个包含7000多条数据的疾病库，无法满足于该项目设计的要求，所以首要工作也就是将该疾病库进行人为的处理，划分为供项目需要的各个子数据库。

* + 1. 检索逻辑的设计

根据项目的目标，设计检索逻辑，有效的展示出系统特性。

智能医疗信息检索系统

系统数据库创建

检索逻辑的设计

1. **详细设计与功能实现**
   1. 数据库设计与实现

对于所拥有的数据库，包含有7000多条数据，每条数据包含有22个表项。从中挑选出能为该项目所用的部分表项，其中包括id、所属科室（rej\_departement）、中文名（chinese\_name）等，首先从这22个表项中选取中所能用到的表项，譬如chinese\_name、classical\_symptom、intro等，这是对于数据库内部信息的选取。然后还需要对整个数据库进行调整，为了提高系统的质量，首先从数据库中人为地筛选出二十多条常规疾病，构建一个常规疾病库。然后从中间筛选出十多条慢性病，构建一个慢性疾病库，最后还有一个用户健康画像的数据库，根据用户的诊断结果进行动态的增加。其中三个疾病库具有可扩充性，可通过添加更多的疾病信息来完善该系统。

下图为各个数据库：

* 1. 检索逻辑的实现

首先进行就身体部位进行选择，待用户选择完身体部位后，显示出该部位的相关症状，然后用户选择症状之后，优先使用症状查找常规疾病库，其次查找总库，然后返回症状所涉及的疾病，对于常规疾病库，会返回出所有并发症，而相对于疾病总库，按照频率返回最有可能的并发症，供用户选择以后，同样的方法再进行一次，然后结合两个症状查找两个库，对疾病进行排序后显示，排序的优先顺序就是优先显示常规疾病库所查询出来的可能疾病，再显示由总库所查询出来的可能疾病。然后，用户可以通过输入疾病名称，查询其详细信息。最后提醒用户就医，过一段时间后，询问用户是否就医，如果就医，输入医院所诊断的疾病，给出用户治疗的相关事项。没就医，询问是否出现好转，若并没有出现好转，则催促病人就医。对于用户问诊，若为新人，则为其立病历，存入到用户信息疾病库，若为老用户，则在其病历中添加起诊断结果即可。

流程图如下：

开始

选择身体部位

选择症状

检索疾病，返回并发症

从并发症中选择症状

返回可能疾病

建议就医

是否就医

询问疾病，给出建议

是否好转

存入病历

结束

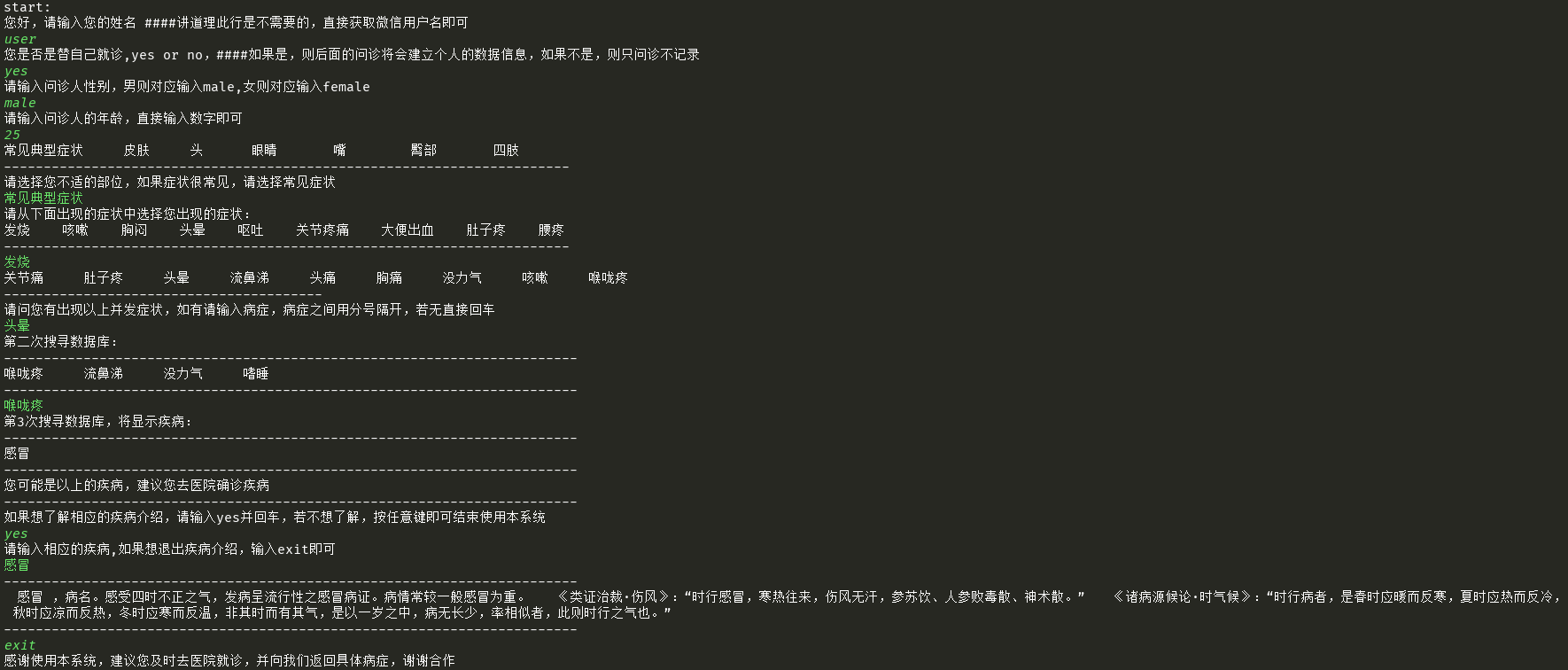
否

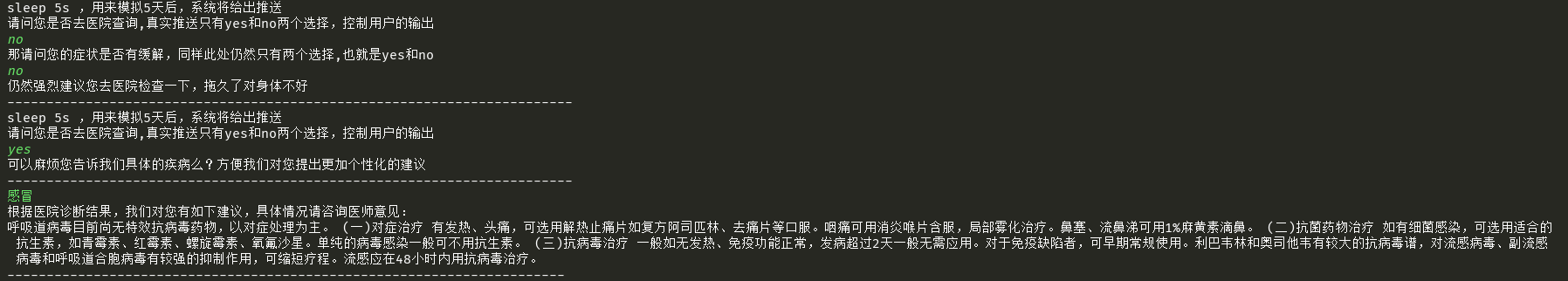
否

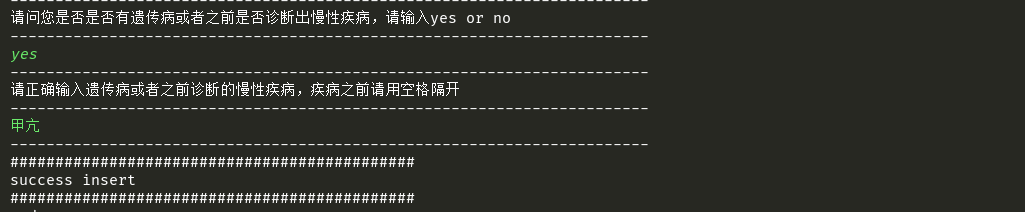
是

是

1. **项目实现结果展示**







1. **未来展望**

通过从更多的数据来源获得更多的数据，完善数据库，可以更加完善该系统。其次就是希望结合自然语言处理等技术来使得该聊天机器人成为真正的智能机器人。然后能智能地处理用户的输入，并给出正确的答案。当然由于数据有限的问题，目前只能达到通过一整套逻辑的设计来达到某些需求。未来还希望可以运用人工智能技术，结合真实的病历数据来更加准确地推测出用户可能患有疾病，使其成为真正的智能医生。

1. **参考文献**

[1]王智囊. 基于用户画像的医疗信息精准推荐的研究[D].电子科技大学,2016.

[2]谢刚,吴高巍,任俊宏,张似衡,牛景昊,张文生.面向患者的智能医生框架研究[J].计算机科学与探索,2018,12(09):1475-1486.

[3] Fischer M. (2017) Symptom-Diagnosis-Care: A Framework for a Collaborative Medical Chat Bot. In: Hassanien A., Shaalan K., Gaber T., Azar A., Tolba M. (eds) Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2016. AISI 2016. Advances in Intelligent Systems and Computing, vol 533. Springer, Cham

[4] P. Meropi, A. S. Billis, N. D. Hasanagas, C. Bratsas, I. Antoniou and P. D. Bamidis, "Conditional Entropy Based Retrieval Model in Patient-Carer Conversational Cases," *2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS)*, Thessaloniki, 2017, pp. 576-581.  
doi: 10.1109/CBMS.2017.145

[5] W. Lu, J. Cheng and Q. Yang, "Question Answering System Based on Web," *2012 Fifth International Conference on Intelligent Computation Technology and Automation*, Zhangjiajie, Hunan, 2012, pp. 573-576.  
doi: 10.1109/ICICTA.2012.150

[6] D. Madhu, C. J. N. Jain, E. Sebastain, S. Shaji and A. Ajayakumar, "A novel approach for medical assistance using trained chatbot," *2017 International Conference on Inventive Communication and Computational Technologies (ICICCT)*, Coimbatore, 2017, pp. 243-246.