

# 研究目的

在当今时代背景下，随着生活、工作的节奏都较于以往有了巨大的改变，带来的影响会时时刻刻映射到人体健康上，因此让每个人及时、正确、全面地了解到自身身体健康情况具有重大的意义。但是社会中医疗资源存在上限，无法保证每个人都能及时到医院就医，且许多情况下不必要地去医院就医会造成医疗资源的浪费，因此如何让用户自己清晰知晓到自身健康的变化就是该研究课题的主要目的，即希望通过设计一款疾病诊疗对话系统，方便用户随时随地自查身体健康。

# 研究内容

本次研究的主要内容在于如何打造一款疾病诊疗对话系统，这个系统的对话机制需要足够的智能化、人性化，而不能是干硬的机器人对话。例如在用户使用遇到系统在某刻无法解决的问题时，采取的方式是委婉更换询问方式，通过后续问答得到更多的有利信息；而并非是干硬的“I don't know”。而更为具体的内容则在于：当用户的疾病诊疗的结果是自己可处理的轻微疾病时，系统将会提供相应的治疗方法和供参考的药物使用及注意事项；而当诊疗结果较为严重时或是系统无法判别的疾病时，将会为用户提供其所在地医院的预约方法等。

鉴于本次项目周期缘故，具体针对的疾病如下：流行性感冒和肠胃炎。这两种疾病在 2020 中国人最易得的十大疾病中是排到了第一和第二的位置，这意味着这类疾病的解决是最为被需求的。并且这两

类疾病的细分下又有多类情况，因此需要系统对用户病情的描述来进行准确高效地判定，从而给出相应的解决方案。而在本项目后期，条件允许的情况下，会尝试将这两类疾病再一次扩大至消化系统和呼吸系统当中。

## 国内外研究现状和发展动态

国内目前在基于深度学习的对话系统的发展并不算得上迅速。对话系统询问的症状应与潜在疾病相关，并与医学知识相一致。然而当前面向任务的对话系统高度依赖于复杂的 **belief tracker** 和纯数据驱动的学习，由于这样的诊断缺乏医学知识的支撑所以它无法直接适用于自动诊断。现有的对话系统主要依赖于数据学习，无法对额外的专用知识图谱进行编码。**Task-oriented Dialogue System for Automatic Diagnosis(2018 ACL)** 率先构建了一个用于自动诊断的对话系统，将对话系统转换为马尔可夫决策过程，并通过强化学习训练对话策略。但这项工作只是利用 **DQN** 网络通过数据驱动学习来管理主题转换（确定应该询问哪些症状），而且结果复杂且重复。此外，此工作仅针对对话管理进行对话状态跟踪和策略学习，使用基于模板的自然语言处理模型，无法很好地匹配实际的自动诊断场景。

市面上存在的对话系统更多的是属于自然语言对话系统，即针对当前用户的语言输入实时提供建议，但是在上下文当中并不能有很好的衔接，在这方面比较具有代表性的就是各手机厂商的智能语言系统。

而在国内市面上的疾病诊疗对话系统仍处于初步发展阶段，并没有具有代表性的相关系统的出现。但是基于深度学习的多轮对话系统必定是时代发展的趋势所在，为用户提供更为深层次的、更具有现实价值意义的解决方法是对话系统存在的意义所在。

而针对广泛的任务型对话系统，目前主流的任务型对话系统有如下两类：（1）管道方法和（2）端对端方法。

所谓的管道方法，又称流水线方法。大概可以分为语言理解（NLU）、对话管理（DM）、语言生成（NLG）三个部分。而在其中的NLU模块里，早期的统计学习模型中，用到了诸如支持向量机（SVM）、朴素贝叶斯、K近邻等，而在实践中逐渐证明了以上的方法多或少存在一定的缺陷，并不适宜推广使用。而发展至今，Zhou以及Lee等人提出了利用循环神经网络（RNN）来配合之前提出的卷积神经网络模型（CNN），相互利用好彼此的优势，不但可以抓取任意长度的序列，分析长句之间的关系，而且可以进行时间和空间扩展，同时具有记忆功能。

DM是整个对话系统的“大脑”所在，而在深度学习的背景下，发展至今有如下几个主流的方法类型：传统强化学习、深度强化学习、DDQ等。但是以上模型都需要大量的标注工作，且都缺乏一定的情感。现今，情感计算在情感识别、人性对话生成等方向取得了丰硕的研究成果，但基于此的智能机器人大部分是集中在用户的自然语言、动作或表情等简单的层面，没有基于语音层面的情感交互，导致了目前大多数的智能机器人的情感识别和表达能力有限。由于缺乏有效的情感

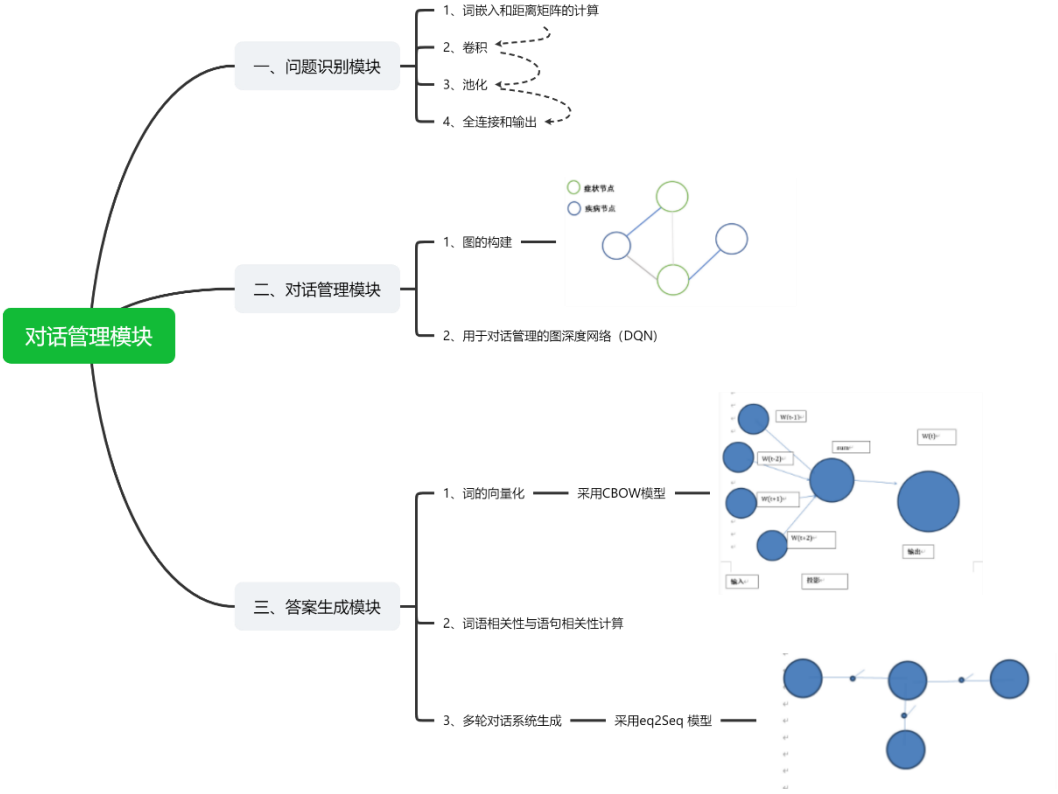
交互策略，现有的智能机器人距离真实人类之间的和谐、自然的交互尚有一定的差距。

NLG 模块是对话系统中重要的一个部分，创建一个结构良好的对话系统可以给用户带来良好的使用体验，而其发展趋势大致可归纳为如下：基于检索模型和生成模型相结合；有监督的端到端生成以及多样化的安全回复。

而另一个方向是端对端方法，由于本次项目拟采用的方法是（1）管道方法，故端对端方法在此不加以赘述。

# 技术路线

总体：



## 问题识别模块

我们通过卷积神经网络来实现对话系统中意图的识别。采用类卷积的方法来计算待识别的句子和意图类别之间的距离。

### 1、词嵌入和距离矩阵的计算

首先我们对文本进行向量化，将待识别的句子中的词组和意图类别中的关键词进行词向量嵌入表示，为了防止出现词向量的维度分布不均匀的情况，我们将普通的欧式距离进行标准化，再使用标准化后的欧氏距离计算公式来计算待识别的文本中各个词组和意图类别中文本特征值比较高的词组之间的距离，得到距离矩阵，并用计算得到的距离矩阵来代替原本的 TextCNN 模型中的词向量拼接矩阵，准备卷积；

### 2、卷积

通过设置 3 个不同大小的 Filter 来对距离矩阵进行卷积操作，并采用 Relu 函数来作为非线性激活函数，以使神经网络等效于非线性网络，来进行非线性卷积；

### 3、池化

将卷积过程之后得到的 feature map 进行最大池化的操纵，并提取出卷积层所输出的特征映射中的最重要的特征来减少参数量，以此提高模型的训练速度；

### 4、全连接和输出

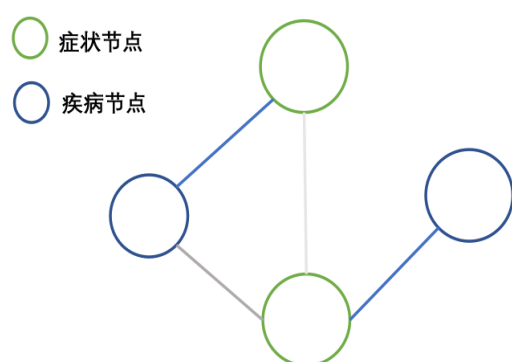
对最大池化所得到的特征值进行串联，以最终形成一个融合的向量来表示待识别文本的特征，最后，输出层使用 Tanh 函数来作为非线性激活函数，计算得到对于每一个意图类别的得分，从而获得最终意图分类的结果。

## 对话管理模块

对话管理模块是一个对话系统的大脑，它决定系统询问哪一个症状做出判断。例如系统识别出“咳嗽”和“发烧”两个症状，对话管理模块将决定下一步询问的症状。我们预计采取一个基于图的深度网络用于对话管理。接下来的内容将会介绍图的构建和基于图的深度网络。

### 1、图的构建

该图为双向有权图，不同节点的边权重不同表示节点之间的联系有强弱之分。给定一个疾病诊断对话资料，出现在对话中的症状和疾病将被看作图中的节点。图中有两种类型的边，一种是症状节点与疾病节点之间的边，另一种是症状与症状节点之间的边。



### 2、用于对话管理的图深度网络（DQN）

对话管理模块用于决定下一步对话系统的动作，是一个连续的决策过程，因此可以被建模为一个马尔可夫决策过程，并使用强化学习方法来训练对话管理模型。具体来说，DQN 使用神经网络模型对智能体的状态进行编码和预测智能体动作，相较于传统算法具有更强的状态建模能力和动作预测能力。其大致流程为将图提供的信息进行向量化表示，进行向量运算后决定下一步动作，并获得奖励

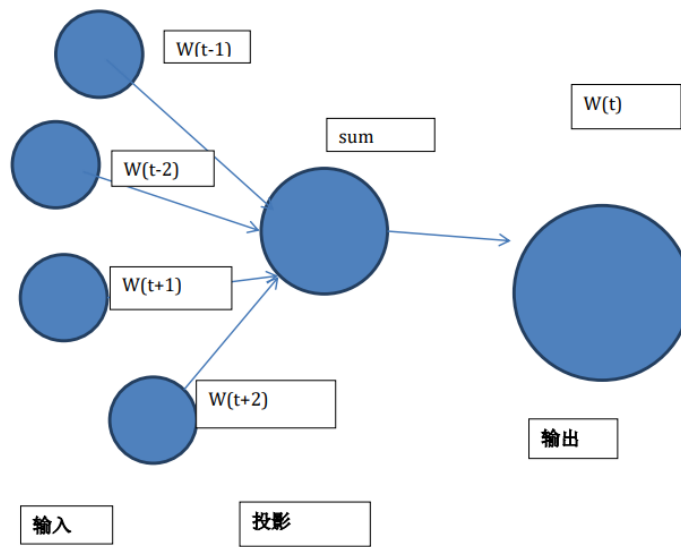
反馈，根据用户的进一步对话优化计算过程。

在这个模型中，我们引入图卷积网络来学习之前构建的图中疾病与症状的向量表示。利用图的邻接矩阵和矩阵的特征向量进行计算，最终选定 Q-value 最高的候选动作作为下一步的动作。同时我们也会采用目标网络和经验回放等技巧来优化训练过程。

## 答案生成模块

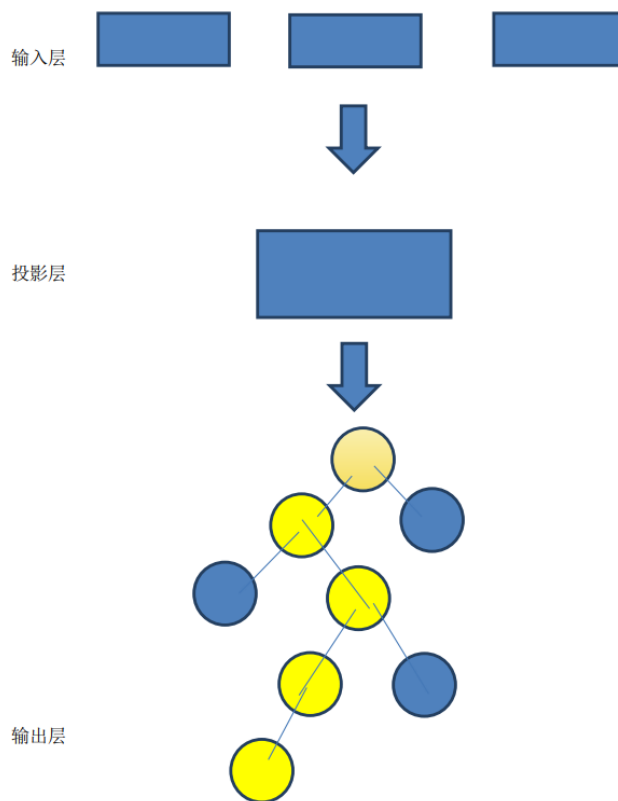
### 1、词的向量化

本文词的向量化部分使用的是 word2vec 的方法，word2vec 中有两个模型：CBOW 模型和 Skip-gram 模型。其中 CBOW 模型是利用句子中的周围词预测中心词，它需要训练的时间更少，适用于更广泛的领域；而 Skip-gram 模型是利用句子中的中心词预测周围词，它需要训练的时间更多，针对特定的领域效果更好。由于本部分是针对闲聊任务，所以使用的是 CBOW 模型。CBOW 模型如图所示：



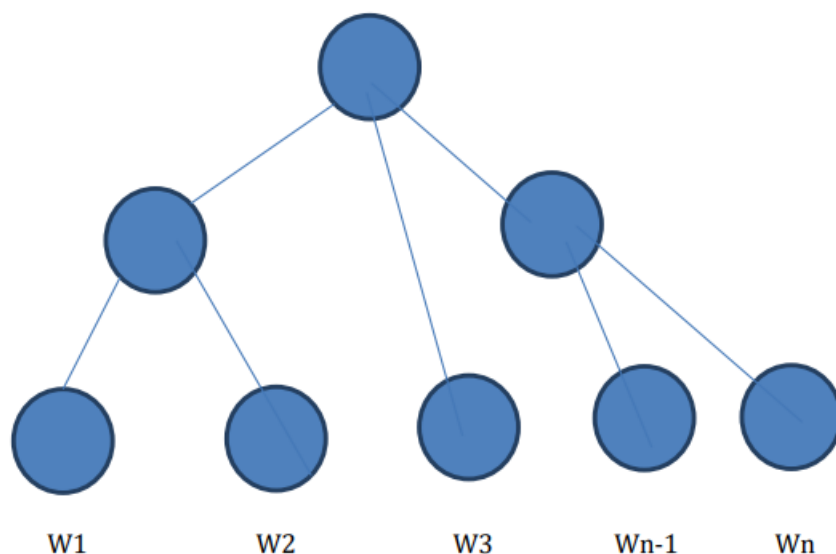
由图可知，CBOW 模型包含输入层、投影层、输出层三层。并且该模型是通过  $w(t-1), w(t-2), w(t+1), w(t+2)$  来预测  $w(t)$ ，其中，从投影到输出通过树来完成，树的每个节点都可以看作一个二分类问题：





## 2、 词语相关性与语义相关性的计算：

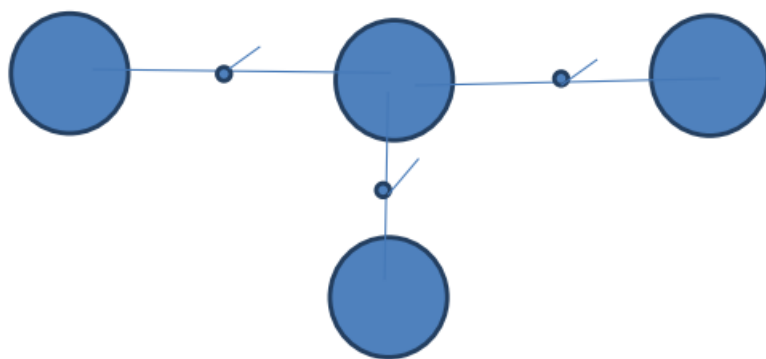
我们首先对词语的相关性和词语的相似性进行了对比，然后提出了在词向量空间中计算词相关性的方法，最后通过分析对话语句上下文的相关性来检测对话中是否出现了话题转移。相似性计算的是词语之间的欧式距离或余弦距离，值越小说明两个词语意思越相近。而相关性表示的是两个词语相关联的程度，是指在出现 A 的情况下，B 出现的概率，是一种后验概率。词语相似性反应了词语之间的距离关系。将具有相关性的词语表示为 Huffman 编码树，非叶子结点存储的是中间向量，可以通过该向量找到所有子节点，所以可以通过计算节点之间的概率找到与目前节点 相关的节点。



最后，通过相应公式及实验所得参数，计算每个节点的条件概率，然后再将每个概率相乘就得到了两个词语的似然概率，这个概率就衡量了两个词语的相关性。

### 3、 多轮对话系统生成

为了使多轮对话系统能根据用户输入的句子和对话的上文信息来生成更符合逻辑的句子，本部分搭建了一个基于序列到序列结构的神经网络模型，该模型结合了编码器（Encoder）、解码器（Decoder）模型思想，并在解码阶段加入了注意力机制（Attention），从而增加了前面的编码信息对后续生成结果的影响。eq2Seq 模型是一种 Encoder-Decoder 模型，在此模型的基础上在其解码部分加入了注意力机制(Attention)。而在 Seq2Seq 模型中的 LSTM 的核心就是三个开关，这三个开关都是为了维持长期的状态  $C$ ，三个开关和 LSTM 的三个状态构成的简图如下图所示，通过权重及相应计算可拟合出符合人类对话结构的多轮对话：



## 创新点和项目特色

该项目的特色主要在于将对话系统与医疗领域的进行一个统一有机的结合，并且更进一步地采用了深度学习相关领域的技术，而本次项目将会着重解决流行性感冒和肠胃炎患者的病情咨询问题，将在这二者领域内为用户提供较为权威的医疗指导。

## 拟解决的问题

包括但不限于：

- 1、三大模块的设计
- 2、类卷积模型的解决
- 3、DQN 模型的解决
- 4、多轮对话模型的设计等

# 预期成果

- 1、设计建立一个基于深度学习的具有一定情感关怀的疾病诊疗对话系统，并利用 web 开发技术进行呈现；
- 2、撰写一份关于该项目的一份技术报告，对本项目进行系列的总结。
- 3、在项目成果允许的情况下，拟为本项目申请相关专利。

# 项目研究进度安排

第一阶段：12 月上旬——2 月上旬系文献阅读阶段，在此阶段我们会着力于对不同模型进行了解查阅，并具体完善我们本次项目技术路线的细节；同时也要积累相应的技术层面的知识点，准备好相关的资源等。

第二阶段：2 月中旬——7 月上旬系算法分析以及算法编写阶段，在这一阶段内，我们会对系列模块进行搭建，并且最终实现基本的整体模型。

第三阶段：7 月中旬——8 月中旬系算法优化以及算法测试阶段，在这一阶段内我们会对已有的系统进行更细节的优化和对其可能存在的漏洞进行相关的处理。

第四阶段：8 月中旬——结束系对本次项目进行相关的技术报告撰写阶段，会对本次项目收尾阶段遇到的问题加以解决。