

参赛队员姓名：徐芷涵 何宝琦 徐学强

中学：杭州外国语学校

省份：浙江

国家/地区：中国南部

指导教师姓名：陈铁明

论文题目：面向育儿领域的智能问答系统

# 面向育儿领域的智能问答系统

**摘要:** 本文提出了一个面向育儿领域的智能问答系统。首先, 基于爬虫技术从专业的育儿平台上爬取数据, 作为问答系统的知识库。然后, 结合关键词检索和语义相似度匹配的方式, 从知识库里找出最符合用户问题的答案。结合关键词检索和语义相似度匹配的方式可以在一定程度上平衡系统效率和准确度。

**关键词:** 智能问答; 育儿; 相似度计算

## 目录

一、背景介绍.....	2
二、研究现状.....	3
2.1 基于传统检索技术的问答系统.....	3
2.2 基于语义技术的问答系统.....	4
2.3 本项目采取的思路.....	4
三、问答系统的技术实现.....	4
3.1 数据获取及预处理.....	4
3.2 词向量训练.....	5
3.3 关键词提取.....	6
3.4 句子的相似度计算.....	7
四、实验及结果.....	7
五、结论及将来工作.....	9
参考文献: .....	11
附录: .....	12
致谢.....	13

## 一、背景介绍

在大数据时代, 各种各样的信息飞速涌入人们的生活。大众获取信息已经变得越来越方便, 已经有很多人习惯借助互联网来获取各种各样的信息。例如, 只要在搜索框中输入一些关键字, 就能找到大部分包含所需信息的网页。

然而, 传统的搜索引擎仍有许多缺点, 例如返回太多相关页面同时没有过滤掉重复的信息。特别是在专业领域中, 面对大量的信息资源, 即使是专业人士会花很多时间来精准地获取所需的内容。而且, 搜索的成功在很大程度上还取决于用户的搜索技能与关键词选

取<sup>[1]</sup>。另外, 在面向专业领域时, 信息驳杂而且质量参差不齐, 用户在寻找, 理解和获取信息方面仍然留存许多困难和障碍, 很难满足用户的需求。所以, 与给予搜索者信息相比, 自动问答系统可以更加有效地利用信息并反馈给搜索者。

当下流行的智能问答系统, 主要是面向开放领域的问答系统。故而我们提出了一个面向专业领域(育儿)的智能问答系统。为了给用户一个较好使用体验, 我们提出了一个结合关键词检索和语义相似度匹配的方法来对智能问答系统进行设计与实现。该方法可以实现系统效率和准确度之间的平衡, 不仅仅解决了语义相似度匹配的效率过低的问题, 同时也解决了匹配关键词可能无法返回最佳答案的问题。

## 二、研究现状

随着自然语言技术的飞速发展, 智能问答成为了自然语言和信息检索领域的一个重要的研究热点<sup>[2]</sup>。同时技术的发展也衍生出各种问答系统模型。基于模板匹配的方法或者是根据专家的先验经验制定的规则来获取问题的答案, 如 LUNAR<sup>[3]</sup>, 但是模板和规则的构建代价过大, 且可扩展性差, 对于模板规则之外的问题具有较差的适用度; 基于常用问题集的问答系统 FAQ<sup>[4]</sup>, 它是一种依托用户建立的知识库, 通过与候选问题集的相似度计算, 得到最相似问题, 并将其答案返回给用户, 优点在于可以将匹配到的相似问题答案, 直接反馈给用户, 而且具有问题答案库大、问题质量高和组织好等优点<sup>[5]</sup>; 基于互联网搜索的智能问答系统和社区问答系统, 这是一种无结构化文本的问答技术, 依托互联网来检索与用户问题的相似问题, 然后将答案反馈给用户, 例如微软的 Encarta3<sup>[6]</sup>; 最后, 当下比较流行的基于用户数据语料库的问答系统, 比如微软小冰<sup>[7]</sup>就是根据用户聊天记录作为语料的。

### 2.1 基于传统检索技术的问答系统

基于传统检索技术的问答系统, 即在问题分析中将问题中的关键词与数据资源中的关键词进行匹配, 进而获取可能相关的答案片段[8]。这种直接匹配的方法思路简单, 而且在理想的条件下, 可以得出及其精确的结果。但是正因如此, 缺点也同样明显。对以下常见情况, 传统方法都无法很好解决:

- 1、搜索的问题极其常见, 导致可匹配答案极其繁多。即搜索引擎针对用户查询返回的结果仍然很零散[9]。对于使用者来说只是缩小了人工检索的范围, 仍需大量时间来检索, 并未解决实际问题。

- 2、搜索的问题很小众, 没有可以完全匹配的答案。

## 2.2 基于语义技术的问答系统

基于语义的问答系统，能够在语义层面上对用户提出的问句进行处理，从而达到了真正的理解用户所提出的问题。相比于传统检索技术，语义技术能够从统计层面上“理解”用户提出的问题，再从候选的库中找到语义上最相近的问题，甚至“主动”回答问题。智能问答系统只有知识远远不够，还需要理解人提出的问题，将自然语言表述的问题转化为计算机可以理解的形式化语言<sup>[10]</sup>。

## 2.3 本项目采取的思路

事实上，此项目是结合传统方法与语义技术而形成的产物。总的来说分为两步，首先使用了传统方法的关键词匹配在数据库（作者在网络上爬取的育儿问答数据库）中进行第一遍粗略的筛选，缩小语义分析范围。接着使用语义分析的方法，通过对用户输入问题与粗略筛选出的几十到几百条问题进行相似度匹配，找到语义最为接近的一条，反馈给用户。在两个步骤中具体用到的技术，会在下文具体阐释。

# 三、问答系统的技术实现

该问答系统的实现主要分为以下几个方面：数据获取、数据预处理、词向量训练、关键词提取、句子向量表示及相似度计算。系统结构图如图 3.1 所示。

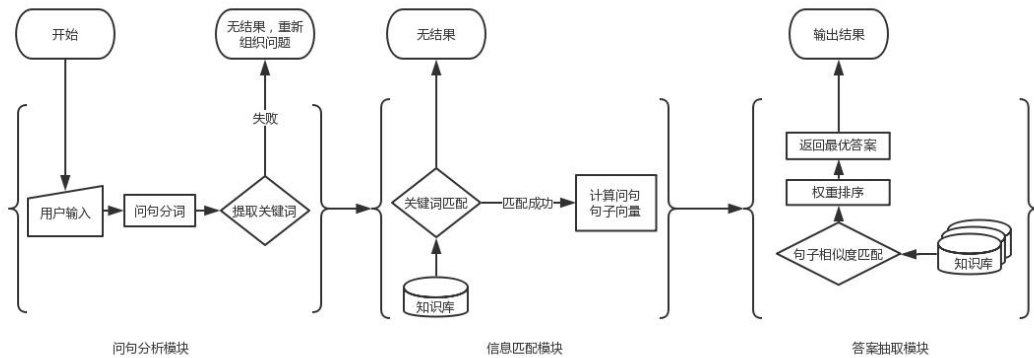


图 3.1 系统结构图

## 3.1 数据获取及预处理

随着信息时代的不断发展，基于互联网的线上育儿平台不断涌现，这些平台招募了大量的专业的相关领域的从业人员，为用户提供专业性的咨询服务。由于互联网的普及以及客服人员的专业素质较高，所以从平台上获取的问题答案的质量较高。本文采用了基于

python 的 Scrapy 框架<sup>[11]</sup>的爬虫, 通过这些爬虫获取了 40 多万条数据, 每条数据都包括一个问题答案。然后清洗数据去掉无关信息, 只保留问题和答案这两部分文本, 然后整理存入数据库, 建立一个知识库。

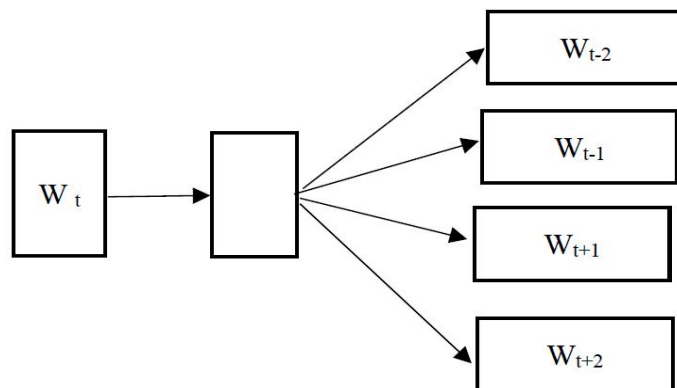
数据预处理是自然语言处理任务中必须的步骤, 而根据不同的需求处理的工作也是不同的。在本文中主要包含以下几个部分: 分词、去停用词、词性标注。首先, 对爬取的数据进行分词。目前有较多的开源中文分词工具, 如 jieba[12], Stanford Tokenizer[13] 等, 在本文, 我们选用 jieba 作为分词工具, 对数据进行分词。接着是去停用词, 之所以要去除停用词是因为原始语聊在分词之后存在很多高频的无意义的词。比如“得”、“的”、“地”, 或是“宝宝”、“医院”等。这些高频词的存在会在接下来的提取关键词环节对实验结果产生影响。正因如此, 我们针对育儿领域数据专门建立了一个停用表, 里面除了基础的停用词以外, 还加入了如“怎么办”, “医生”, “孩子”等在寻医问药的过程中没什么意义的词。使用停用表可以有效的防止关键词的遗漏或提取错误。

### 3.2 词向量训练

在自然语言处理任务中, 只有将文本符号转化成数字符号, 才能作为下一步骤的输入, 进行后续的计算。将文本转化成数字符号往往采用 one-hot 编码的方法。然而通过 one-hot 编码得到的向量是很稀疏的。如果直接使用这样的向量作为输入来进行相似度计算开销过大。因此我们使用了 word2vec<sup>[14]</sup>模型。该模型提供了一个减少特征维度的策略, 以 one-hot 编码的向量作为输入, 得到维度较低的新词向量, 以此来进行相似度计算。

word2vec 模型实际上是一个含有隐含层的神经网络。它是将我们语料库中所有的词通过 one-hot 编码后得到的词汇表向量当作输入来训练这个神经网络。收敛之后, 将从输入层到隐含层的那些权重, 当作每一个词语的词向量。比如, 第一个词的向量是  $(w_{1,1} w_{1,2} w_{1,3} \dots w_{1,m})$ ,  $m$  是表示向量的维度,  $w_1$  则表示相应维度上的权重, 将这些权重取出就是该词的向量表示。我们根据需要就可以得到有限维度的向量表示, 然后就可以用到其它的应用当中。

Word2vec 包括两个模型 CBOW 和 Skip-gram。以 Skip-gram 为例介绍 word2vec 方法。



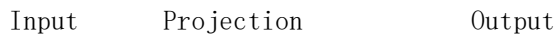


图 3.2 Skip-gram 模型示意图

Skip-gram 的核心思想就是意思相近的词也可能具有相似的上下文。所以该模型在训练的过程中，需要输入当前词来训练预测得到上下文窗口中的相邻词。模型如图 3.2 所示。

由图 3.2 可以看出模型的输入以及输出都对应着相应词的词向量。而词向量则是通过单词的 one-hot 编码得到的，模型通过输入的词向量进行训练，利用 softmax 函数使得上下文的临近词的词向量尽量相近。

### 3.3 关键词提取

为了提高系统的回答问题的准确率，我们采用了一个结合关键词检索和语义相似度匹配的方法，通过这个想法我们可以实现系统效率和准确度的平衡。所以对知识库中问题的关键词的进行提取，并建立一个问题关键词库用以匹配。然后通过用户所提问题中提取的关键词来跟关键词库中每个问题的关键词进行逐一匹配，将两个及以上关键词相同的问句提取出来，得到一个候选问句集，方便将用户所提问题与候选问题集中的问句一一进行相似度计算，得到最为相似的问句。

传统的关键词的提取算法都是基于词频统计，没有充分的考虑词与词之间语义关系，故而本文使用了 TextRank 算法<sup>[15]</sup>。TextRank 算法源于经典的 PageRank 算法<sup>[16]</sup>。PageRank 是来自 Google 的著名算法，在评判特网页价值时经常使用。其主要思想是：如果有大量网页链接到某个页面或有某些重要页面链接到该页面，就表示此页面比较重要值。相互连接的页面能够相互打分。然后通过多次迭代，最终能够获得一个表示页面重要性的分数。

TextRank 的主要原理就是将每个词语都作为图上的一个节点，窗口中的任意两个词语对应的节点存在着一条无向边。然后利用投票的原理，将每一条无向边都看成是单词之间的互相投票，经过不断迭代，最终每个单词的得票数都会趋于稳定。然后根据词语的得票数进行排序，得票数越多则认为这个单词越重要。

利用 TextRank 来提取关键词，第一步就是要将处理过的文本形成一个中文词的集合。设某句子分别由以下的一系列单词组成，窗口为 k: w1, w2, w3, w4, w5, ..., wn，故可以得到 [w1, w2, ..., wk]、[w2, w3, ..., wk+1]、[w3, w4, ..., wk+2] 等窗口。在同一窗口中，任意两个词语所对应的节点之间都存在着一条无向无权的边。基于以上窗口构成图，通过公式(1)算出节点处每一个词的权值，然后按照大小顺序选出权值较大的几个作为关键词。

$$WS(V_i) = (1 - d) + d * \sum_{V_j \in In(V_i)} \frac{w_{ji}}{\sum_{V_k \in Out(V_j)} w_{jk}} WS(V_j) \quad (1)$$

其中 Vi 表示某个词，Vj 表示链接到 Vi 的词，WS(Vi) 表示词 Vi 的 TextRank 值，

$In(V_i)$  表示词  $V_i$  的所有入链的集合,  $Out(V_j)$  即集合中词的个数,  $d$  表示阻尼系数, 这里设为 0.85,  $W_{ij}$  表示两个词语节点之间的权重。

### 3.4 句子的相似度计算

为了进行句子的相似度计算, 首先要得到每个句子的向量表示。假设句子由  $n$  个词组成, 词向量为  $m$  维。然后我们利用公式(2), 对句子中的每个词的第  $k$  维的数据进行计算, 作为句子的第  $k$  维的数据, 然后重新组合得到句子的向量表示。

$$X_k = \frac{\sum_{i=1}^n x_i}{n} \quad (2)$$

本文采用了计算比较余弦相似度的方法来计算句子之间的相似程度。所谓余弦相似度, 就是在向量空间中将两个向量夹角的余弦值作为衡量两个向量差异的标准。余弦值与 1 相近, 就意味着夹角越趋向于 0 度。反之, 两个向量之间夹角越大或者方向相反, 说明两个句子之并不相似, 甚至毫无关系。

假设问题 A 和问题 B 分别用向量  $(x_1, x_2, x_3, \dots, x_n)$ ,  $(y_1, y_2, y_3, \dots, y_n)$  表示。现在通过余弦相似度公式(3)进行计算, 然后将得到的余弦值进行归一化处理, 最终得到问题 A 与问题 B 的相似值。然后将相似值进行排序得到匹配度最高的问题, 通过该问题在库中找到对应的答案。

$$\cos \theta = \frac{\sum_{i=1}^n x_i \times y_i}{\sqrt{\sum_{i=1}^n (x_i)^2} \times \sqrt{\sum_{i=1}^n (y_i)^2}} \quad (3)$$

## 四、实验及结果

由于没有一个公开的中文问答系统测试集, 我们从互联网上爬取了约 462761 条一一对应的医疗问答数据, 使用其中的问句集训练词向量模型。测试集采用复述问句的方式 (即用不同的表达方式改变问句的语言) 来人工生成, 如:

数据库问题 1: 婴儿出生脚上有一块很小的红色, 现在一岁半了, 都没有退, 今天发现就在红色的地方长了几颗硬的东西, 其中有一颗大一点, 请问这是什么来的?

复述问句的问题 1-1: 婴儿从小脚上就长着一块很小的红色, 一岁半了都没有退, 最近红色的地方长了几颗硬硬的东西, 有一颗略大一些, 请问这是怎么回事?

复述问句的问题 1-2: 婴儿出生的时候脚上长了块红色的东西, 很小, 我们都没在意, 现在一岁半了还在, 今天发现红色的那里长了几颗硬东西, 其中有一颗稍微大一点, 请问这个要不要紧?

数据库问题 2: 宝宝没有感冒现象, 喉咙哑了有小半个月了, 该怎么处理? 需要就医吗?

复述问句的问题 2-1: 孩子没得感冒, 但喉咙一直不好, 哑了小半个月了, 要怎么办?

复述问句的问题 2-2: 孩子没感冒症状, 只是小半个月来喉咙一直哑, 该去医院吗?

我们随机抽取了 100 个问题, 基于复述的方式一共生成了 255 个问题。使用复述的方式主要是为了便于之后的正确率统计。由于本系统的问题比较开放, 因此在某些问题上难以界定“正确”与“错误”。为使评价结果更加客观, 我们认为复述问句的答案只有完全与原问题答案一致才认定为正确, 无论是相似的结果或是在人工判断下正确但并非完全与原问题匹配的结果, 均认为是不正确的。因此, 该评价标准比系统现实运行中要严格很多。

我们使用了两项评价指标:

(1) 问题匹配率

给定一个测试样本集, 无论对错, 能够匹配到答案的比例: 77.6%

大约有 1/4 的句子无法匹配到答案, 或是无法提取关键词, 或是有了关键词而无法获得词向量。在用来做测试的 255 个问题中, 有 57 个问题无法匹配到相应的答案。出现这样的问题有两个原因。第一、用以训练的数据集不够大, 虽然有近 50 万条问答数据用以训练。但是, 医学领域的表述涵盖很广泛, 数字组合, 英语单词, 拉丁字母, 或是特殊的中文术语组合都很常见, 因此有一些词从未在训练集中出现过, 无法计算词向量。第二、提问句子过短, 表意不清, 夹杂市井俚语, 导致在关键词提取环节就出现问题, 自然语言处理算法无法得到这一类句子的关键词, 因而匹配答案失败。

(2) top-k 准确度

在能够匹配到答案的测试样本中选取: 给定一个测试样本, 只要系统输出最相关的前 k 个答案中有正确答案, 则该测试样本为正确, 否则为错误。k 取 5、7、10、15、20 计算所有测试样本中正确的比例。

表 1 测试集 Top-k 结果

	k=5	k=7	k=10	k=15	k=20
Top-k	0.45	0.588	0.87	0.912	0.936

为了检验上述方法对问题相似度计算的影响, 我们对测试集的 Top-k 进行了测验, 将精度分为五级, 结果如上表 1-1。附表给出了部分问句与答案的示例。

分析结果, 可以得到以下几个方面的特征:

1、随着窗口增大, 正确率上升。

这就是说, 某些时候那些靠前的问题候选不见得就是匹配度最高的答, 主要原因是提问的句子太过冗余。长问题由于存在的词很多, 在取平均值之后句子的向量就比较居中, 也就是说, 很长的问题之间容易匹配到, 哪怕他们并不是十分相似。这时候, 相对排名靠



前的而且文本较长的候选来说, 也许排名靠后的句子中包含的关键词更多。因此假设将 K 放大一些, 且仅仅对那些长度适中的问句做考量, 那么输出结果还是能够令人满意的。

2、无论将 K 方法到多少, 有一部分句子仍然得不到最恰当的答案。

在测试中, 有一部分问题无法匹配到正确的答案。这实际上牵扯到另一个有难度的课题——专业名词识别<sup>[17]</sup>。分词算法无论是对俚语还是对医疗专业名词, 辨认率都不高, 导致分词结果很凌乱。如果这些育儿领域的专有名词都能被正确识别, 那么系统的识别成效将显著提升。

根据得到的测试集结果, 画出折线图 4.1 如下。

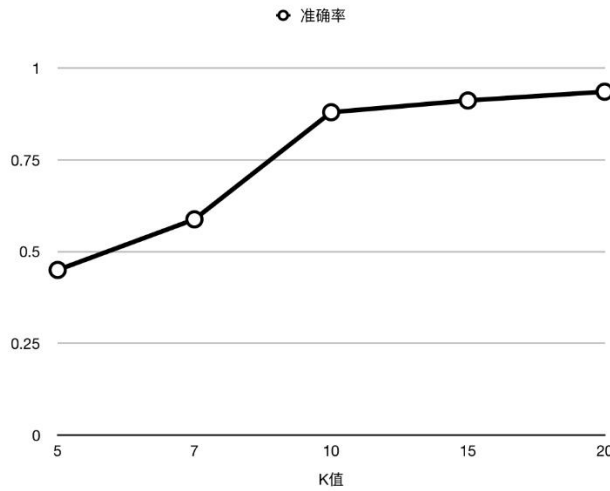


图 4.1 top-k 准确率

通过折线图可以得到另一个结论。由此图斜率不断变化, 即随着 k 值的增加, 准确率的上升趋势减慢, 且在窗口值由 7-10 时上升最快, k 超过 10 之后再增大对结果提升不再显著。可知大量问题的正确答案均分布在前十个候选问题中。根据已得到的结论可知, 有一部分问题无法匹配到正确答案。对于这些问题, 本系统寻找到的答案分布是比较随机的, 在折线图上表现为斜率小, 趋势平缓。

## 五、结论及将来工作

本文介绍了一个基于问题集的育儿问答系统, 通过对具体实验结果内容的分析, 我们发现尚有若干个有待进一步研究解决的议题。1) 具体句子关键词的提取对检索的结果有着极其重大的影响, 并且这种影响对最答案抽取起到了决定性的干扰。例如附录中的问句 6, 两个相关度很低的句子, 但是抽取出的关键词很相似, 因此产生了错误的匹配。2) 育儿领域专有名词难以准确地识别。3) 句子向量的计算方法仍需改进, 如何在不受关键词影响的前提下, 使长句的句子向量有差异。

因此, 以后的工作重心将放在改进关键词提取算法上, 以算法为落脚点来步提高句子

匹配的正确率, 与此同时, 且在现有的系统中增加一个可信度评价的机制, 将有效地帮助用户得到更为精确的答案。最后, 还需要对那些存在着语法问题的问句以及育儿领域专有名词的识别响作进一步的研究。

## 参考文献

- [1] Wren JD. Question Answering Systems in Biology and Medicine—the Time is Now [J]. *Bioinformatics*, 2011, 27 (14):2025–2026.
- [2] 郑实福, 刘挺, 秦兵, 等. 自动问答综述[J]. *中文信息学报*, 2002, 16(6):47–53.
- [3] Wood J A, Dickey J S, Marvin U B, et al. Lunar anorthosites and a geophysical model of the moon[C]//*Proceedings of the Apollo 11 Lunar Science Conference*, Houston, TX, Jan 5–8, 1970. New York: Pergamon Press, 1970: 965–988.
- [4] 秦兵, 刘挺, 王洋, 等. 基于常问问题集的中文问答系统研究 [J]. *哈尔滨工业大学学报*, 2003, 35(10):1179–1182
- [5] 毛先领, 李晓明. 问答系统研究综述[J]. *计算机科学与探索*, 2012, 6(3):193–207.
- [6] Dong X, Gabrilovich E, Heitz G, et al. Knowledge vault: a web-scale approach to probabilistic knowledge fusion [C] // *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2014: 601–610.
- [7] 微软小冰 <http://www.msxiaoice.com/>.
- [8] 张芳芳, 马敬东, 王小贤, 等. 国外医学领域自动问答系统研究现状及启示[J]. *医学信息学杂志*, 2017, 38(3):2–6.
- [9] 白宇, 周俏丽, 蔡东风. 问答式信息检索技术[J]. *沈阳航空航天大学学报*, 2008, 25(5):63–67.
- [10] 屠要峰. 基于电信业务场景的智能问答机器人关键技术[J]. *计算机科学与应用*, 2017, 7(4): 291–300.
- [11] Scrapy 框架 <https://scrapy.org/>
- [12] jieba <https://github.com/ficsjy/jieba/>
- [13] Manning C D, Surdeanu M, Bauer J S et al. The Stanford CoreNLP Natural Language Processing Toolkit[C]//*ACL (System Demonstrations)*. 2014: 55–60.
- [14] Mikolov T, Sutskever I, Chen K, et al. Distributed representations of words and phrases and their compositionality[J]. *arXiv preprint arXiv:1310.4546*, 2013.
- [15] Rada, Mihalcea, and Paul Tarau. "TextRank: Bringing Order into Texts." *empirical methods in natural language processing* (2004): 404–411.
- [16] A N Langville, C D Meyer, *Google's PageRank and beyond: the science of search engine rankings*, Princeton: Princeton University Press, 2006.
- [17] 崔桓, 蔡东风, 苗雪雷. 基于网络的中文问答系统及信息抽取算法研究[J]. *中文信息学报*, 2004, 18(3):25–32.

## 附录

编号	问句	提取的关键词	匹配到的问题	匹配到的问题的关键词	相似度
1	孩子前几天感冒去医院, 刚吃完药咳嗽好一些, 但这几天咳嗽又开始反复, 而且有加重的趋势, 现在该怎么办啊?	咳嗽加重反复	幼儿感冒引起的反复咳嗽是支气管炎还是肺炎	引起咳嗽反复	0.81 7
2	婴儿从小脚上就长着一块很小的红色, 一岁半了都没有退, 最近红色的地方长了几颗硬硬的东西, 有一颗略大一些, 请问这是怎么回事?	小脚红色婴儿	婴儿出生脚上有一块很小的红色, 现在一岁半了, 都没有退, 今天发现就在红色的地方长了几颗硬的东西, 其中有一颗大一点, 请问这是什么来的	红色婴儿地方	0.62 4
3	我家宝宝5个月13天大了, 从小到大都混合喂养, 喝奶粉比母乳多。大便一天四次, 吃完就拉, 该怎么办?	混合母乳喂养	47天宝宝, 母乳奶粉混合喂养, 奶粉吃的	奶粉混合喂养	0.93 2
4	之前不知道得了什么病, 以为感冒就连吃了九天的药, 后来发烧了才去医院, 查出是支原体感染, 打点滴输了十五天红霉素, 前天又发烧了, 嗓子还发炎, 结果今天又复发, 还伴随呕吐症状, 还要去医院吗?	嗓子查出呕吐	丁医生, 12岁小孩, 发烧, 一吃药就呕吐, 嗓子有点疼, 头重, 没精神, 吃了药也不好, 应该吃什么药, 还是打点滴, ??	嗓子发烧呕吐	0.71 6
6	打扰一下, 我家女儿三岁, 这几天走路右腿突然一瘸一拐, 问她她说不疼, 大约20天前得了场感冒, 发烧血象高, 吃消炎抗病毒药物。	发烧抗病毒走路	三个月的宝宝发烧医生给开阿莫西林克拉维酸钾干混悬剂和抗病毒口服液, 吃一次就退烧了, 还要吃多久才可以停药?	抗病毒发烧悬剂	0.75 8

## 致谢

团队成员在研究报告撰写中承担的工作以及贡献如下：

论文议题选择；资料搜集与整合：徐芷涵

算法研究及编写：徐学强

论文书写及校对：何宝琦

实验模拟和结果分析：全队

在本论文的写作过程中，本团队的导师陈铁明教授倾注了大量心血，从选题到拟定论文大纲，相关背景资料的引入到依据实验结果的修正，都提供了极大的帮助。指正我们的不足，协助我们突破桎梏，引导我们在独立研究中找到方向。在此，谨向导师陈铁明教授致以崇高的敬意和衷心的感谢！

本参赛团队声明所提交的论文是在指导老师指导下进行的研究工作和取得的研究成果。尽本团队所知, 除了文中特别加以标注和致谢中所罗列的内容以外, 论文中不包含其他人已经发表或撰写过的研究成果。若有不实之处, 本人愿意承担一切相关责任。

参赛队员: 徐芷涵 何宇琦 徐学新 指导老师: 陈铁刚

2018年9月14日