

面向医疗设备的深度问答系统的设计

饶林尚 吴怡 冯前进*

(南方医科大学生物医学工程学院 广东 广州 510515)

(广东省医学图像处理重点实验室 广东 广州 510515)

摘要 提出辅助医疗设备维修保养的深度问答系统的设计方案。为医院设备工程师提供智能化的设备信息咨询平台,提供日趋复杂而广泛的设备知识服务,增加医院设备的效益。系统包括算法模块和应用模块,算法模块通过深度学习卷积神经网络实现。通过设计实验进行答案搜索任务测试,在问题相似度前三的反馈信息里面,包含搜索目标的准确率达 65%,证明算法可搜索到有效信息。将算法模型嵌入到 Web 应用中,进一步实现问答的功能。

关键词 深度学习 CNN 深度问答系统 Spring MVC 医疗设备

中图分类号 TP3 **文献标识码** A **DOI**:10.3969/j.issn.1000-386x.2019.06.033

DESIGN OF DEEP QUESTION ANSWERING SYSTEM FOR MEDICAL EQUIPMENT

Rao Linshang Wu Yi Feng Qianjin*

(School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, Guangdong, China)

(Guangdong Provincial Key Laboratory of Medical Image Processing, Guangzhou 510515, Guangdong, China)

Abstract This paper presented a design of deep question answering system to assist medical equipment maintenance. It provided an intelligent equipment information consulting platform for hospital equipment engineers, provided increasingly complex and extensive equipment knowledge services, and increased the benefits of hospital equipment. The system included algorithm module and application module. The algorithm module was implemented by deep learning convolutional neural network. We designed the experiment and tested the answer search task. The accuracy of the search target contained in the feedback information with the first three questions' similarity was up to 65%. This proves that the algorithm can search for effective information, and embeds the algorithm model into the Web application to further realize the function of question answering.

Keywords Deep learning CNN Deep QA system Spring MVC Medical equipment

0 引言

近年来随着中国经济高速发展及社会老龄化加剧,人们对医疗资源的需求不断增长。医院巨资引进大量先进医疗设备,但是同时,医院用于对设备售后管理的费用也以更高的速度增长^[1]。主要原因有:(1)大型医疗设备例如 CT、MRI 设备,属于集成度较高的精密电子设备,随着设备的复杂化、模块化,及厂家提供的资料有限,设备工程师越难以找出故障和短时间

内给出合适的解决方案。因此,相对于医疗仪器技术的发展,医院对设备的自主保养维修能力反而下降。设备不能及时维修影响到医院的运营和甚至引发不良的医患纠纷。(2)医疗设备销售市场竞争激烈,设备的售后市场的利润也逐渐引起了设备厂商的注意,导致医院对设备的保养和维修成本随之增加^[2]。

近年来,深度学习被应用于各行各业,在问答系统方面也有很广泛的研究和应用。自 2011 年,IBM 的深度问答系统(Deep QA System)在美国电视竞答节目 Jeopardy 中,以绝对优势击败该节目的两位前冠军以

来,越来越多的企业开始重视研究 Deep QA System,纷纷把自身的业务与自然语言处理、深度学习结合起来^[3]。2014年,微软推出“微软小冰”,实现开放领域人工智能对话。它衍生的产品形态涉及对话式人工智能机器人、智能语音助手和一系列垂直领域解决方案。2016年3月,京东推出京东 JIMI 客服机器人,能够代替人工客服完成一般的咨询服务,涵盖售前咨询、售后服务等电子商务环节。2016年,百度推出百度医疗大脑,结合深度学习和大数据实现在线的健康问题咨询。随着互联网技术的发展和网络信息量的大爆炸,人们快速获取资料并准确定位信息的愿望促进了自动问答系统的发展^[4]。但是目前还没有发现有关针对医疗设备信息检索方面的数据分析和应用系统的研究。

本文提出一个面向医疗设备和基于深度学习的问答系统的设计方案,第一,实现了医疗设备信息的资源共享,第二,可以给工程师提供专业的辅助信息。对医院来说,设备投入增加,其设备售后产生的花销倍受重视。同时,医疗设备能够及时、合理地维修与保养,可以延长设备的使用寿命,增加医院的经济效益,降低医院的运行成本。目前,大部分医院通过以下方式加强设备管理和节省设备的开销:(1) 协商设备购买合同,在购买设备前尽可能保证自身利益;(2) 加强设备工程师队伍建设;(3) 构建医疗设备信息资源共享平台等^[5]。但现实是医疗设备行业内信息的交流和工程师可以咨询的途径非常少。而本文提出的问答系统借助当前信息化和深度学习技术的优势,既实现设备信息共享,也给工程师开辟了一条专业化咨询的途径。文献^[6]介绍了六种基于深度学习的问答系统算法框架。本文在此基础上设计模型,通过实验证明深度学习可以有效应用在医疗设备文本数据上。

面向医疗设备的深度问答系统的出现可以帮助医院的工程师从医疗设备大数据分析中得到相似的参考方法或者解决方案。通过线上获取设备维修保养经验交流的方式,及时对设备进行合理的保养或维修,避免医院因为设备故障导致的经济和声誉上的损失,增加了设备的可使用价值和效益。另一方面,设备工程师也有一个学习和咨询的平台,从中学习技术和经验,提高医院工程师队伍的实力,一定程度避免通过设备厂商提供的高额售后服务,节省医院开支。

1 系统可行性分析

对医疗设备实行现代信息化管理,保证了医疗设备的使用年限和提高医院工作效率,增加了医院的经

济效益^[7]。医院引入数字化、信息化的设备管理系统是大趋势,并在该背景下产生了大量的有关设备的数据,如设备检测时产生的大量的记录。可利用大数据手段充分挖掘这些数据背后的潜在价值,可以给医院带来更大的效益^[8]。随着人工智能的发展,借助自然语言处理和深度学习的手段更深层次分析和处理这些数据,可实现更准确的智能信息检索。

1.1 基于深度学习的问答系统

深度问答系统本质上是一个信息检索(IR)系统,只是它从文本数据中获取更多深层的语义信息,返回更加精准化的答案。这样的问答系统功能上与现在的搜索引擎,如百度、搜狐基本类似,但是在性能上优于搜索引擎。现在甚至有人认为深度问答系统会成为以后搜索引擎的新型态^[9]。

以往的搜索引擎实现方式一般利用网页中所标注的关键词匹配出得分高的数据资源,它并没有揭示出这些数据之间深层的关系,所以难以把检索过程与用户实际想要的答案结合起来。因此,需要反馈大量相关的信息给到用户,由用户自己筛选^[10]。随着深度学习的发展,加深了计算机对语言文字的理解和分析的能力。问答系统中采用深度学习对目标文本进行更细致的语义分析,从中抽取出知识,并根据用户的问题准确匹配、推理相对应的答案,实现为用户提供更高精度的搜索反馈^[11]。深度学习理论和技术的发展日益成熟,并且被应用到各行各业中。因此,利用深度学习算法实现医疗设备的面向限定域的信息搜索具备可行性。

1.2 Spring MVC 框架

Spring MVC 是一个分层的 Java Web 应用开发框架,为 Web 应用开发人员提供了一个快速、方便开发的分层体系结构。分层体系结构既能保证代码的整洁性,又能保证代码的可维护性。框架也具有高解耦性,所以系统下不同功能模块之间耦合性低,降低了算法模型嵌入到其他已开发医疗设备管理系统的难度。该框架突出了 HTTP 协议中的请求/响应特性,在该框架中,用户的每一个请求都声明了一个需要执行的动作。主要是通过将每个请求 URI 映射到一个可执行的方法来实现。同时,也将请求参数映射到对应方法的参数列表中。在该框架下,用户以对话形似在前端输入问题,生成对应问答服务的 URL(统一资源定位符),后台根据该 URL 执行算法模型,再将答案反馈到前端推荐给用户,即完成一次问答行为。基于 Spring MVC 框架实现问答的需求,现已经不成问题。

2 系统设计

2.1 整体设计

本文提出的深度问答系统包含算法模块和应用模块两部分。算法模块是系统的核心部分,实现基于深度学习的信息检索策略。算法模块的实现需要数据的支持,通过爬虫技术,从网站“影像之家”爬取有关医疗设备的数据,例如 CT、MRI 和 DR 等设备的维修保养等记录。应用模块将算法模块嵌入到基于 Spring MVC 框架下的 Java Web 应用中,前端实现与用户的交互操作,后台实现算法的运行和信息检索及反馈。应用模块 Spring MVC 框架由 6 个核心组件组成:前端控制器扮演枢纽角色完成请求的接收和分发,视图数据的包装和传输;处理器映射器匹配服务 URL 地址;处理器适配器通过服务 URL 地址调用算法模块的 Handler 处理器,Handler 执行程序并返回结果数据;视图解析器及视图负责视图的构建和向用户显示计算结果^[12]。各个组件在一次问答行为的业务流程如图 1 所示。

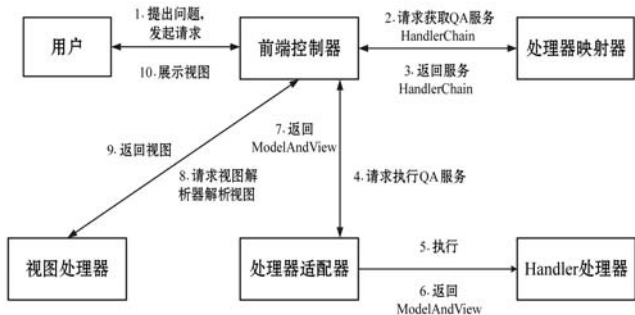


图 1 Web 应用框架示意图

2.2 算法流程设计

整个算法模型的流程包括:医疗设备相关数据获取,文本预处理,文本分词和量化,模型实现。具体的步骤说明如图 2 所示。

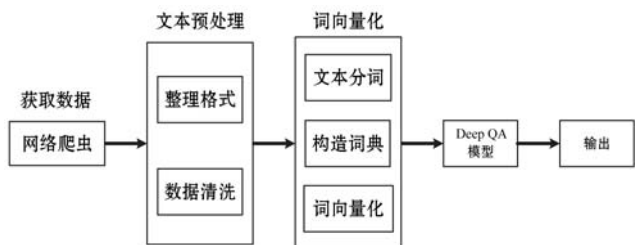


图 2 文本数据处理流程

通过网络爬虫框架 Scrapy 从网站“影像之家”挖取 228 条有效医疗设备的记录数据,并将爬取到的数据整理成如图 3 所示的格式。

["answer": "BC 检查线路连接并重新拔插所有接头,感觉连接并无异常。重新启动系统,故障依旧。根据 BC-Status-1 的测试内容(在 BC Pin Diodes 两端加上 15V 电压,该线路会产生 5A 的电流,如果不能产生则测试不能通过)。此测试仅涉及 RFIS, Cable, BC, 测量 RFIS 电源都在正常范围内,所以初步判定为 DysCon_5A 故障。次日更换 DysCon_5A 后故障依旧。此时,怀疑 BC 线圈故障的可能性更大,但不敢贸然断定。再次进行线路及 BC 线圈的检测。检测过程中经过数次线路重连及重启,设备状态皆未改变。但在最后一次重新连接 DysCon_5A 及 RFIS Mother Board 连接 X3 与 X15 后系统启动正常,判断问题出在这里。拔掉该线路检查,单看此线针脚位置并无异常,拉动黑线也无退针现象,只是针脚在塑料槽里有轻微的正常晃动,再次连上该线路,通过推拉黑色线路部分反复试验,证明问题确实是该线路的接头连接不够紧密。至此找到问题出现的原因,下次再出现 cable 问题,一定要小心卡槽接头的线。DysCon_5A 及 RFIS Mother Board 连接街头。", "question": "故障现象:BC 开机测试不能通过,报错显示 BC-Status-Test-1,查看 EventLog,提示 RFIS 至 BC 之间线路故障或 RFIS 错误故障分析与排除:测量 BC 到 RFIS 之间的连线。参考 Function 的线路图,分别测量 DysCon_5A 的 X3 与 X15, X15 与 X9, X9 与 BC 的 X3 之间线阻正常,从 X15 向 BC 测去,正向电压 0.38V,反向截止,初步判断 BC 正常。"]

图 3 医疗设备样本数据示例

实验假设爬取到的每一个答案只与其配对的答案相关,认为是正确答案,其他答案认为是负答案,并且问题之间不存在相关性。每一个训练样本由一个问题描述匹配一个正确答案和一个从答案集合中随机采样得到的负答案组成。按照这样的方式,训练集中每一个问题描述随机采样不同的负答案匹配生成训练样本。

数据清洗:文本数据中任何与数据上下文和最终输出无关的文本都可被判作噪声。例如停止符号(/n),标点符号和 URL 地址,HTML 符号等。因为不同人记录的习惯不一样,导致可利用的文本数据的绝大多数具有不规范性。同时也因为人记录的随意性,导致文本数据夹杂噪声,例如错别字,无意义字符等。为了训练得到更好的算法模型,对数据进行预处理非常必要。

词向量化:文本分词基于开源的使用广泛的中文分词工具 Jieba。对面向限定域的数据构建专有的词典,可以提高分词的准确率和模型的泛化能力。针对挖取到的文本数据构建专属词典,以提高分词的准确性。但是构建完备的基于医疗设备的词典需要花费大量的人力,而且并不一定保证能收录所有词。因此如何识别未登录词是分词中的一个重要问题。Jieba 采用的是 Unigram + HMM(隐马尔科夫模型)的方法实现分词。Unigram Model 原理通过求解下式,实现最优的分词效果:

$$y(x_1, x_2, \dots, x_n) = \operatorname{argmax} \left(\prod_{i=1}^n p(x_i) \right) \quad (1)$$

在 Unigram 分词后,HMM 以序列标注的方式识别未登录词,在一定程度上修正分词结果。词向量化基于 Word2Vec 实现,Word2Vec 是 Google 在 2013 年推出的一个 NLP 工具,应用深度神经网络算法,转化为对应的词向量。中文文字数量庞大,使用一般的 bag of words 模型表达文本,很容易造成维度灾难,并且语义信息无法准确表示,无法揭示词与词之间的潜在联系^[13]。而采用 Word2Vec 基于神经网络的方式得到的

空间词语的语义表达,不仅解决了维度灾难的问题,而且挖掘了词语之间的关联属性,从而提高了词向量语义上的表示准确度。一般 Word2Vec 包含两种训练模式,一种是基于 CBOW(Continuous Bag of Words Model) 框架,另外一种 Skip-gram 框架^[14]。本文采用 CBOW 模式,其原理如图 4 所示。

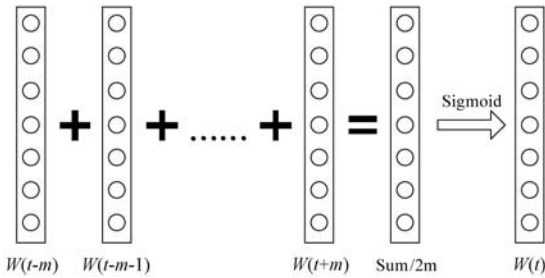


图 4 CBOW 原理示意图

通过上下文预测中心词 $w(t)$ 的向量表达。 w 表示每个词的 one-hot representations。one-hot representations 是对每个词在词典位置的向量表示,例如“设备”在词典中的位置为 1,则其对应的向量为 $\{1, 0, 0, 0, \dots\}$,除了第一位为 1,其余位置全部取 0。这种表示的缺点是只考虑了词的位置关系,忽略了词与词之间的语义关系。CBOW 经过累加过程,将上下文每一个词的 one-hot representations 相加求平均后,经过 Sigmoid 将数据映射到 0-1 区间,最后得到预测词 $w(t)$ 的向量表达。此时的向量表达加强了词与词之间的语义关系。

2.3 算法框架设计

Deep QA 框架如图 5 所示。

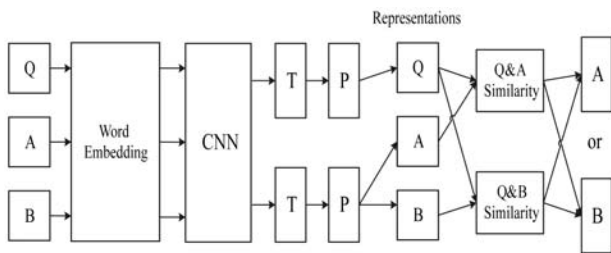


图 5 Deep QA 卷积神经网络框架

Q 表示序列化的问题描述;A 表示序列化的答案 A 选项,B 表示序列化的答案 B 选项。Word Embedding 层表示词向量层,经过 Word Embedding 层以后,Q、A、B 分别转换成词向量矩阵。CNN 层表示卷积神经网络层,P 表示池化层,T 表示激活层。算法均采用 Relu 函数作为激活函数。Q、A、B 的词向量化矩阵经过 CNN 层、P 层和 T 层以后,分别得到维度一样的问题 Q,选项 A 和选项 B 的向量化表示。求取问题 Q 分别和选项 A 和选项 B 的欧式距离,表示它们之间的相似性。最后,经过全连接层得到 Z 输出,再由如下 Sigmoid 函数得到概率值最大的选项,作为模型最后的

预测:

$$P_i = \frac{1}{1 + e^z} \quad (2)$$

分别以 0,1 代表结果 A 和 B。实验设计的原理在于,通过不断迭代,实现模型对关联答案的识别能力的同时提高模型对非关联答案的辨别能力。

对问题和答案文本序列化,是指将已分词文本的每个词用其在字典中的索引表示。Word Embedding 层通过词索引,得到词向量。具体步骤如图 6 所示。

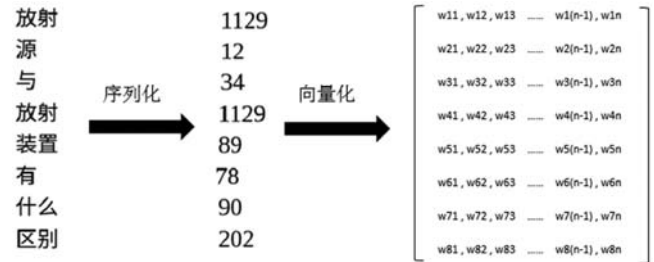


图 6 文本量化过程

文本中每个词在字典中对应一个位置索引,将文本中每个词用索引代替实现序列化。文本经过 Word2Vec 向量化以后,每个词也会对应一个词向量,词向量按字典顺序存储。在向量化过程中,每个词按照字典索引获取到对应的词向量。最后,文本量化成矩阵。按照该方式向量化文本大大节省了存储空间。

CNN 对序列化向量的空间卷积实现原理如图 7 所示。在卷积层,每个卷积核分别对文本向量化矩阵卷积,得到 feature map 向量。在池化层,取 feature map 全局最大值,将多个 feature map 的全局最大值连接,通过全连接层,得到文本的向量表达。

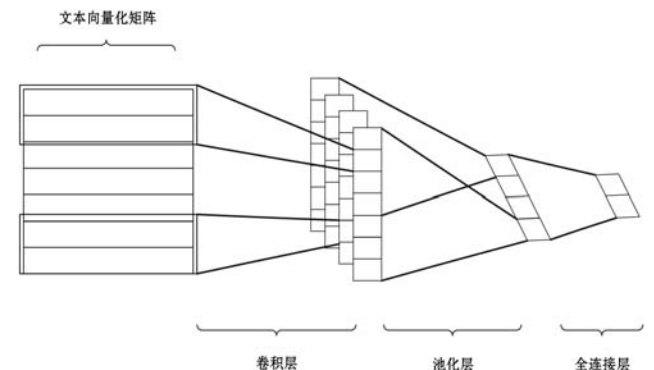


图 7 文本序列卷积原理示意图

3 实验和分析

3.1 模型训练结果和分析

实验一共生成 13 300 个训练样本,其中验证样本 2 872 个。训练过程采用交叉熵作为调整模型的代价函数:

$$loss = \sum_i y_i \log(y_{i_predicted_A}) + (1 - y_i) \log(y_{i_predicted_B}) \quad (3)$$

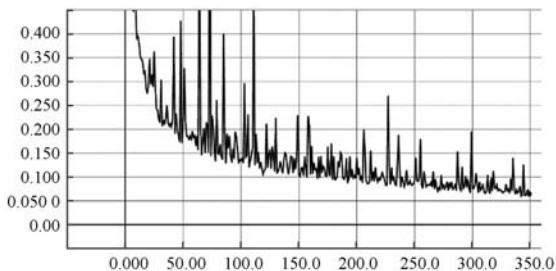
使用 Cosine Similarity 来度量问题和答案之间的相关性,其中 A_i 、 B_i 分别表示问题和选项的向量表达。

$$\cos\theta = \frac{\sum_{i=1}^n A_i * B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \quad (4)$$

Cosine Similarity 通过向量夹角的余弦值衡量两个文本间的相似度,相比欧式距离度量,更加注重向量在方向上的差异。

通过算法框架,预测出答案 A 或者 B。若预测结果与实际的答案选项相符,认为预测正确;若预测结果与实际答案不一致,则认为预测错误,反向传播错误率,调整模型参数。再通过迭代训练,正负选项的方式有利于模型预测结果更加偏向正答案,远离负答案,以此提高模型对文本数据的认知和辨别能力。

由图 8 可知,训练集和验证集的 acc 和 loss 曲线趋势一致。训练迭代 350 次以后,训练集和验证集的 acc 和 loss 曲线也达到收敛,并且 acc 可以达到 90% 以上,同时 loss 降低到 0.05。在测试集的准确率也达到 90%。实验结果证明了模型从医疗器械文本数据中学习到了知识,具有一定的分析和辨别能力。



(d) 验证集 loss 曲线

图 8 模型训练 acc 和 loss 情况

在信息搜索测试中,需要对训练的网络结构进行修改,具体的框架如图 9 所示。实际应用中,我们只需要计算每个问题和答案的相似性,然后根据得到的相似性进行排序,将相似性程度高的信息反馈回来。通过将 228 条问答分别输入系统,测试反馈的回复的准确性。模型预测相似性程度前三的回答里面,包含目标答案的准确度是 65%,而预测相似性程度前五的回答包含目标答案的准确度可以达到 82%。进一步证明了系统在实际的问答任务中可以反馈给用户有效的信息。

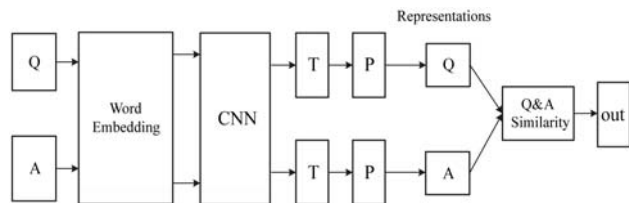
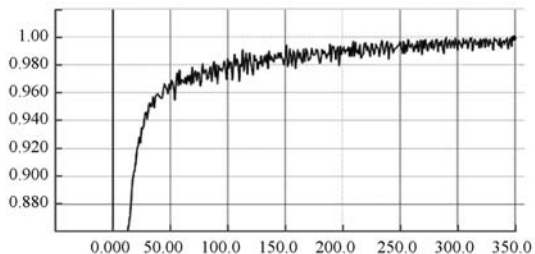


图 9 系统算法框架

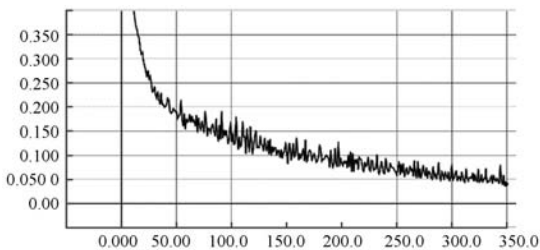
3.2 应用模块设计

前端页面基于 Javascript 开发,并采用 Ajax 技术连接后台。用户在前端页面提交请求后,前端控制器接收,并访问处理器映射器获取到 QA 服务的对应的 URL 地址。前端控制器再通过该 URL 服务地址,命令处理器适配器调用 Handler 处理器执行 QA 服务。Handler 处理器与数据库交互,执行算法模块服务,执行完将检索结果返回到前端控制器。前端控制器将交由视图处理层,包装数据成视图,并展示给用户,完成一次问答行为。

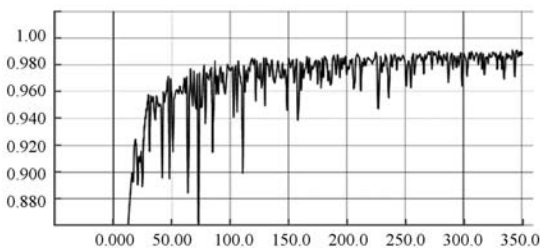
根据算法模块设计,数据存储基于 MySQL 数据库,数据库表格式以键值匹配原始答案文本记录方式存储,键值设计为以 0、1、2、... 的递增形式。算法模块检索答案时,SQL 语句通过键值逐条访问文本记录。词典及 Word Embedding 矩阵以文本格式存储在本地电脑磁盘。计算机访问数据库获取文本记录后,文本记录通过匹配词典和 Word Embedding 矩阵实现向量化。



(a) 训练集 acc 曲线



(b) 训练集 loss 曲线



(c) 验证集 acc 曲线

数据访问层设计依赖 Hibernate 框架。Hibernate 是一个轻量级 JavaEE 应用的持久化层解决方案,它提供了强大、高性能的对象到关系数据库的持久化服务。它能将实体对象与数据库表建立映射关系,无需写 SQL 语句,就完成答案实体对象到对应的数据库表的映射^[15]。

最终将设备问答的功能作为一个模块嵌入到设备管理系统中,效果示例界面如图 10 所示。



图 10 问答界面简图

4 结 语

医院设备管理的效率和成本问题将越来越得到医院的重视。如何借助目前的深度学习技术提高医院设备管理的效率和成本,辅助医院设备工程师完成设备管理的任务,在未来将会是更多医院去研究和挖掘的方向。设计基于深度学习算法框架的问答实验,证明了深度学习算法应用于医疗设备文本数据的可行性。并且将算法模型与互联网技术整合起来,提出一个面向限定域的辅助问答系统设计方案。目前算法模型尚存在的问题是因为医院设备管理相关的数据不足,上述算法训练得到的模型泛化能力有限,尚不能广泛应用于实际工程应用当中。在未来,随着医院信息云平台的发展,越来越多的医院参与其中,利用数据分析等技术挖掘和处理平台产生的设备管理数据,将会大大节省医院的设备成本,产生不可估量的价值。

参 考 文 献

[1] 童其荣,刘伟,亢德洪. 医疗设备维修成本分析与管理的[J]. 中国医疗设备, 2011, 26(8):84 - 85.
 [2] 黄鑫,刘尚军. 现代医院大型医疗设备维修存在问题与对策探讨[J]. 医疗卫生装备, 2012, 33(12):103 - 105.
 [3] 黄昌宁. 从 IBM 深度问答系统战胜顶尖人类选手所想到的[J]. 中文信息学报, 2011, 25(6):21 - 26.

[4] 陈谷川,陈豫. 自动问答系统 QA 的研究与实现[J]. 数字图书馆论坛, 2006(5):37 - 41.
 [5] 丰国庆. 降低医疗设备维修成本有效途径的探讨[J]. 中国设备工程, 2018(17):34 - 35.
 [6] Feng M, Xiang B, Glass M R, et al. Applying Deep Learning to Answer Selection: A Study and An Open Task[C]// 2015 IEEE Workshop on Automatic Speech Recognition and Understanding(ASRU). IEEE, 2015:813 - 820.
 [7] 严军. 医院医疗设备维修信息化管理现状和发展[J]. 医疗装备, 2016(16):76 - 77.
 [8] 徐海琴,王雅军,南玉萍. 大数据在医疗设备管理中的应用[J]. 医疗卫生装备, 2017(6):136 - 137.
 [9] 刘康,张元哲,纪国良,等. 基于表示学习的知识库问答研究进展与展望[J]. 自动化学报, 2016, 42(6):807 - 818.
 [10] 赵洁. 基于搜索引擎的中文自动问答系统的设计与实现[D]. 北京:北京工业大学, 2016.
 [11] 刘宝瑞,郭宏娇. 基于 Deep QA 的图书馆数字参考咨询问答系统研究[J]. 情报科学, 2017(4):103 - 108.
 [12] Deck P. Spring MVC 学习指南[M]. 林仪明,崔毅,译. 2 版. 北京:人民邮电出版社.
 [13] 郑文超,徐鹏. 利用 word2vec 对中文词进行聚类研究[J]. 软件, 2013(12):160 - 162.
 [14] Mikolov T, Chen K, Corrado G, et al. Efficient Estimation of Word Representations in Vector Space[EB]. arXiv:1301.3781, 2013.
 [15] 张少应,程传旭. 基于 Hibernate 持久化层的设计与实现[J]. 计算机技术与发展, 2014(12):101 - 104.

~~~~~  
 (上接第 100 页)

[ 4 ] 李洪奇,徐青松,朱丽萍,等. 基于数据集相似性的分类算法推荐[J]. 计算机应用与软件, 2016, 33(8):62 - 66.  
 [ 5 ] 张培根,黄树成. 一种用于中文数据清洗的近邻排序算法[J]. 计算机应用与软件, 2018, 35(8):286 - 288, 333.  
 [ 6 ] 奉国和,郑伟. 国内中文自动分词技术研究综述[J]. 图书情报工作, 2011, 55(2):41 - 45.  
 [ 7 ] 赵秦怡,王丽珍. 一种基于互信息的串扫描中文文本分词方法[J]. 情报杂志, 2010, 29(7):161 - 162, 172.  
 [ 8 ] 丁麒,庄志画,刘东丹. 基于文本数据挖掘技术的 95598 业务工单主题分析应用[J]. 电力需求侧管理, 2016(S1):55 - 57.  
 [ 9 ] 盛华,张桂珠. 一种融合 K-means 和快速密度峰值搜索算法的聚类方法[J]. 计算机应用与软件, 2016, 33(10):260 - 264, 269.  
 [ 10 ] 苏毅娟,邓振云,程德波,等. 大数据下的快速 KNN 分类算法[J]. 计算机应用研究, 2016, 33(4):1003 - 1006, 1023.