

智能对外汉语学习系统的设计与研究

李 斌¹, 王浩畅²

(1. 湖南科技大学, 湖南 湘潭 411201;

2. 东北石油大学 计算机与信息技术学院, 黑龙江 大庆 163318)

摘 要:近年来,人工智能飞速发展并在社会各个方面都得到了广泛应用,随之也影响着教育向智能化方向发展。该文首先分析对外汉语教学对于智能聊天机器人对话系统的实际需求;然后在总结国内外研究成果和相关会议的基础上,对基于聊天机器人的智能对外汉语学习系统的相关理论和系统结构进行了详细阐述,以情境学习理论、建构主义学习理论、人工智能中的自然语言处理技术为指导,使用句子词语切分和基于向量空间模型的计算实现句子语义相似匹配,利用 TF-IDF 算法进行关键词匹配,从而构建基于人工模板和规则的聊天机器人,来达到自动问答的目的;接着介绍智能对外汉语学习系统在对外汉语教学上的独特功能以及应用实践效果;最后指出了今后需要进一步完善和改进的地方。

关键词:聊天机器人;汉语学习;智能学习系统;人机对话;人工智能;自然语言处理

中图分类号:TP18;G434

文献标识码:A

文章编号:1673-629X(2022)03-0015-06

doi:10.3969/j.issn.1673-629X.2022.03.003

Design and Research of Intelligent Foreign Chinese Learning System

LI Bin¹, WANG Hao-chang²

(1. Hunan University of Science and Technology, Xiangtan 411201, China;

2. School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China)

Abstract: In recent years, artificial intelligence has been developed rapidly and widely applied in various aspects of human society, which also affects the development of education towards intellectualization. Firstly, we analyze the actual demand of intelligent chatbot dialogue system for Teaching Chinese as a foreign language. Then, based on summarizing the research results and relevant conferences at home and abroad, we elaborate the related theories and system structure of the intelligent Teaching Chinese as a foreign language system based on chatbot in detail. Guided by situational learning theory, constructivist learning theory and natural language processing technology in artificial intelligence, sentence semantic similarity matching is achieved by using sentence segmentation and vector space model based computation, and the TF-IDF algorithm is used for keyword matching to build a chatbots based on artificial templates and rules to achieve the purpose of automatic question-and-answer. Next, we introduce the unique function and practical effects of the intelligent foreign language learning system in Teaching Chinese as a foreign language. Finally, it is pointed out that the next work is to improve this system.

Key words: chatbot; Chinese learning; intelligent instruction system; human computer interaction; artificial intelligence; natural language processing

0 引言

近年来,随着人们对人工智能技术的开发与研究,它在全球都受到了广泛关注,人工智能技术不仅在科技方面有着不可替代的作用,而且在教育教学领域也有着非常广阔的前景。中国对人工智能的重视已上升为国家战略层面,不仅于2017年发布的《新媒体联盟地平线报告:2017 高等教育版》中表示:人工智能对今后的教育发展将产生重要影响^[1],而且于2018年4月

召开的全国教育信息化工作会议中也提到“人工智能化领跑教育信息化2.0”^[2]。不仅国内对人工智能教育领域应用十分重视,国外还专门针对这一领域设有权威学术期刊——《人工智能教育应用国际杂志》(International Journal of Artificial Intelligence in Education)。在这些信息的推波助澜下,越来越多的高校和科研机构都对“人工智能+教育”进行了系统开发和深入研究。未来,人工智能技术与教育的融合将

收稿日期:2021-03-24

修回日期:2021-07-26

基金项目:国家自然科学基金项目(61402099)

作者简介:李 斌(1993-),男,硕士研究生,研究方向为智能教育、自然语言处理;王浩畅,博士,教授,硕导,研究方向为自然语言处理、信息抽取。

成为必然趋势。

人工智能技术、信息技术和教育技术是实现“人工智能+教育”的三大技术支持。其中,机器学习、自然语言处理、图像识别等技术都是人工智能技术和信息技术的整合;相关的教育技术有:教学设计、课程建设、教学方法等内容^[3]。这些技术为人工智能在教育领域中的应用奠定了技术基础。随着中国国际地位的不断提高,以及汉语国际教育在全球范围的大力推广,全世界掀起一股汉语热的风潮,促使越来越多的外国学生选择来华进行汉语学习。虽然学习汉语已经成为一种趋势,但是对外汉语教师极为匮乏,不仅如此,单一的汉语教学模式、课上课下缺少强化练习过程、教师教学任务繁重以及缺少真实的语言练习环境等限制条件,从根本上制约了汉语学习者的学习。

由于计算机在语言教学中的应用,即计算机辅助语言学习(Computer Assisted Language Learning, CALL)可以使学习者突破时间和空间的限制进行有效的语言学习和训练,便于学习者充分利用碎片时间进行个性化学习。一时间内,国内外学者和科学家开发了许多汉语学习软件,并进行深入研究,研究内容主要集中在应用开发与设计、具体功能设计与实现、应用调查分析、具体内容分析与设计、应用效果分析等方面。

目前,汉语学习软件种类繁多,大多数软件都涵盖了课件音频、视频的播放、汉字书写和发音教学、词汇学习等方面的内容,但由于在设计、技术、语言情境等方面上仍存不足,造成对外汉语学习者的学习支持力度不够。对于语言学习来说,语境是至关重要的,学习语言的最佳路径就是在真实的语境中进行一对一对话,然而利用智能聊天机器人系统就可以实现创设真实的或模拟的汉语学习环境,主要运用人工智能中的自然语言处理技术,创设出真实的语境,让学习者与机器人用自然语言进行对话。智能聊天机器人是人工智能领域中的一个重要研究方向,它是通过模拟自然语言来与人类进行对话。人们学习语言的目的是为了交流,而学习语言的最佳途径也是通过对话来学习交流,将聊天机器人与汉语学习相结合,开发具有理解功能的交互式人机对话汉语学习环境,对于提高对话和汉语表达能力,以及提高对外汉语教学质量有着举足轻重的地位和实践价值。基于以上原因,笔者开发了基于聊天机器人的智能对外汉语学习系统。

1 研究现状

随着人工智能技术的热潮席卷全世界各个领域,人工智能技术在教育领域的应用逐渐具有了不可替代的作用,因此,越来越多的国内外研究者对在语言教学领域中使用人工智能技术产生了浓厚兴趣。探讨人工

智能技术如何与语言教学领域进行有机结合,不仅能极大地推动人工智能技术的进步,还能够促进整个语言教学科研领域的发展。

黄晓洁^[4](2006)提出辅助阅读系统和辅助教材编写系统的设计构想,从理论层面和模型构想上探索利用自然语言处理技术实现对外汉语辅助教学系统。卢丹丹^[5](2009)提出基于人工智能技术的大学英语辅助教学系统,主要包括两个部分:一是基于专家系统理论的教师英语辅助教学专家系统模块;二是基于BP神经网络模型的学生自主学习系统模块。安静^[6](2011)利用数据挖掘技术及人工智能中的专家系统理论,实现英语辅助教学系统。微软首席研究官瑞克·拉希德(Rick Rashid)在2012年11月演示了计算机“同声传译”系统,该系统不仅能将英文翻译成中文,还可以模拟说话者的语音语调,该技术对实时口译和语言学习等方面有很大帮助。陈颖^[7](2013)利用自然语言理解、机器学习和智能搜索等人工智能技术,尝试设计针对中学英语教学的人工智能系统,并探索该系统的具体实施条件和实现。郝丹^[8](2015)通过将多媒体技术与自然语言处理技术相结合,开发汉语学习系统,实现汉语基础知识学习模块、专题浏览模块以及语句学习、新闻摘要和情感分析三个结合自然语言处理技术的工具助手模块。微软亚洲研究院于2016年开发的微软小英,是一款英语口语学习和训练软件,融合了自然语言处理、语音识别和语音合成等人工智能技术,主要通过情景模拟、情境对话和跟读训练等功能,帮助英语学习者完善发音、熟练口语,以及快速建立日常英语沟通能力。孙雅琳^[9](2018)利用人工智能中的机器学习、深度学习等技术,实现对大学英语技能训练系统中的口语主观表达题的智能批改,将对减轻教师批改压力、指导学生学习有很大帮助。贾积有博士^[10]自主设计研发的“希赛可”,是一个模拟英语语言交流情境的智能英语学习系统,将自然语言处理、智能代理等人工智能技术应用于英语教学领域,该系统主要包括“希赛可”聊天机器人、“希赛可”教学平台和“希赛可”背单词软件三个部分^[11]。杨晓琼等^[12]介绍的批改网,是一款基于自然语言处理、人工智能等技术实现的在线自动英语作文批改系统,通过分析学生英语书面作文,检测出其中的错误,并即时给出评分、改善建议以及内容分析结果,学生可根据提示自主修改多次提交,直至达到满意为止,这种方式将有助于提升学生的英语写作能力。不仅研究者在语言教学领域上运用人工智能技术非常重视,国内外研究机构和学者也组织了许多相关会议和语言评测,进一步促进人工智能技术在语言教学领域的应用。

计算机辅助语言学习的自然语言处理研讨会

(Natural Language Processing for Computer Assisted Language Learning, NLP4CALL) 由北欧语言技术协会 (Northern European Association of Language Technology, NEALT) 组织, 从 2012 年开始, 每年举行一次, 致力于将自然语言处理和语音技术集成到 CALL 系统中, 更好地理解和应用语言, 从而使 CALL 智能化, 将该研究领域取名为智能计算机辅助语言学习 (ICALL), 主要探讨研究内容如下: 使用 NLP 工具来测试第二语言习得 (SLA) 理论和教学理论; 利用 SLA 理论和教学实践在 ICALL 工具中建模。从 2016 年起, 将语言习得研究的自然语言处理研讨会 (Natural Language Processing for Research in Language Acquisition, NLP4LA) 联合在一起举办, 目的是为研究语言习得过程中出现的理论和方法问题的研究人员提供一个会议场所, 包括经验和实验研究以及基于 NLP 的应用。语言与智能高峰论坛 (Language & Intelligence Summit) 由中国中文信息学会和中国计算机学会联合举办, 2016 年在北京首次召开, 每年举行一次。该高峰论坛主要探讨语言与智能领域的新发展和新技术, 目的在于向全世界介绍国际语言与人工智能领域的发展趋势和创新成果, 进一步推动国内语言与人工智能技术领域的发展。

北京洪堡论坛 (BHF) 于 2013 年 11 月创办, 由德国洪堡基金会和对外经济贸易大学共同主办, 每年 9 月在中国举行。第六届北京洪堡论坛于 2018 年 9 月在对外经济贸易大学举行, 以“绿色经济·文化传承·教育 2035”为主题, 将从多角度、多方位、多领域探讨, 其中包含“人工智能与语言研究”主题论坛, 主要有以下议题: 人工智能技术在语言研究和翻译领域的应用、语言研究对人工智能研究的可能贡献、人工智能与语言学发展的历史关系等。

语言测试与评价国际研讨会 (International Conference on Language Testing and Assessment, ICLTA) 是由教育部考试中心举办的年度学术会议, 从 2015 年开始, 该会议旨在及时传递国内外语言测试领域的最新研究成果和研究动态及未来的发展趋势, 将成为具有中国特色的现代外语能力评测体系。该会议主要围绕语言能力标准的制定、语言教师的教学能力及测评素养、外语测试改革与发展以及人工智能技术在语言测评中的应用等议题进行深入的交流和研讨。

通过以上研究成果发现, 人工智能技术应用于语言教学有利于学习者自主学习和训练; 便于教师教学和研究, 以及减轻教师日常工作中面临的各项任务, 例如语法错误标记、论文评分、创建测试和练习等; 及时反馈有效增强学生学习。目前, 大部分研究者利用人工智能技术对英语辅助学习进行探索和研究, 然而对

汉语辅助学习的研究少之又少, 更没有以对话形式系统地对学习进行汉语辅助学习。因此, 该文主要研究如何将智能聊天机器人和汉语学习相结合, 利用聊天机器人系统辅助汉语学习, 提供即时反馈, 创设真实、交互的语言学习环境, 让学习者在模拟真实的语境中通过与聊天机器人对话来提高语言交流能力, 实现学习效率最大化。

2 系统设计的理论基础

系统主要以情境学习理论、建构主义等教育理论, 以及人工智能中的自然语言处理技术为指导, 设计开发了基于聊天机器人的智能对外汉语学习系统。

情境学习理论认为, 学习不仅仅是一个个体性的意义建构的心理过程, 而更是一个社会性的、实践性的、以差异资源为中介的参与过程^[13-14]。它关注学习对认知条件的依赖性, 强调创建合适的情境来激发学习动机, 鼓励学习者参与到相关的社会活动中来进行练习, 有效地掌握语言^[15-16]。情境学习理论强调学习者需要在知识实际应用的真实情境中学习, 对于汉语语言学习, 最好的方法是让学习者融入真实的生活情境中, 借助恰当的语言情境交流促成汉语学习。该智能对外汉语学习系统为学习者创设了一个类似真实的汉语学习情境, 在模拟真实的情境下, 通过与聊天机器人进行场景对话练习, 从而能有效地促进汉语学习者的汉语学习水平, 并灵活应用于实际生活中。

建构主义学习理论认为, 知识不是通过教师传授得到的, 而是学习者在一定的情境, 即社会文化背景下, 借助其他人 (包括教师和学习伙伴) 的帮助, 利用必要的学习资料, 通过意义建构的方式而获得^[17]。建构主义学习理论是以学习者为中心, 强调学习者的主观能动性, 然而基于聊天机器人的智能对外汉语学习系统正体现学习者是学习主体的思想, 强调学习过程是学习者积极主动构建的过程。建构主义还强调情境对汉语习得的意义构建的重要性, 学习总是与一定的情境相联系的, 即应创建与学习者的实践生活相近的学习情境和相应的交流活动, 可以使学习者在这样的语言环境和交流中完成意义建构。该智能对外汉语学习系统将学习者置于特定的情境之中, 通过和聊天机器人的相互交流, 使学习者在交互的情境中利用自己已有的经验基础去主动获取当前学习到的新知识, 从而完成意义建构, 进而提高学习者对汉语学习的积极性和主动性, 以及汉语语言运用能力。

聊天机器人系统的核心技术是自然语言处理技术, 自然语言处理是将人类之间交流沟通所用的语言经过处理转化为机器所能理解的机器语言, 进而实现人间相互交流^[18]。它是一门综合性的交叉学科, 涉

及到计算机科学、语言学 and 数学等领域的知识。该聊天机器人涉及到中文分词、词性标注、命名实体识别、关键词提取等自然语言处理技术。利用自然语言处理技术的对外汉语学习系统的特点是不仅具有自然语言生成与理解能力,并且在学习过程中模拟真实的语言环境,利用简单流畅的人机交互界面方式与机器人进行汉语对话交流,激发学习者的自主学习能力,通过创设自然的语言交流环境,达到对学习者的所学知识的巩固与反馈,提高学习效率。

3 系统结构

智能对外汉语学习系统主要侧重于对话练习,为汉语对话提供场景模拟,为学习者提供对话练习的模拟语境,通过聊天机器人技术来实现对话质量良好的

汉语对话练习,帮助用户提升汉语水平。该文主要使用句子词语切分和基于向量空间模型的计算实现句子语义相似匹配的对话检索方法,利用 TF-IDF 算法将句子切分转为利用关键词表征的向量,构造语句的空间向量模型后,通过计算两个向量的余弦相似度来获取用户输入文本和语料库中问题 Questions 的相似度,根据相似度得分进行排序,选择得分最高的 Questions 对应的回复 Answer 作为回答反馈给用户,达到聊天的目的。基于聊天机器人的智能对外汉语学习系统是一个能够检索语料库产生回复和用户进行汉语聊天的系统,系统结构和流程如图 1 所示。系统构成主要包括:输入预处理、自然语言理解、中文-拼音匹配、对话管理、答案生成、输出处理。

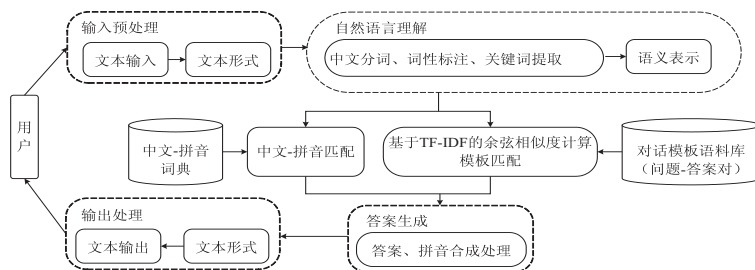


图 1 基于聊天机器人的智能对外汉语学习系统结构

3.1 输入预处理

输入预处理模块主要负责对学习者的输入文本进行前期的处理,然后交由自然语言理解模块进行处理。

3.2 自然语言理解

自然语言理解模块主要通过中文分词、词性标注、关键字提取等技术,理解用户输入问句的语义,从完整语句中提取出需要的信息,产生一个特定的语义表示。此语义表示是一个或多个关键字,并将其转换为向量的形式。该系统采用了中国科学院计算机技术研究所研发的 NLP-ICTCLAS 汉语分词系统^[19],实现将语句进行分词、词性标注和关键词提取等功能。NLP-ICTCLAS 汉语分词系统采用基于层叠隐马尔可夫模型 CHMM(Cascaded Hidden Markov Model),其分词正确率达 97.58%^[8],所以最终采用该汉语分词系统。

3.3 中文-拼音匹配

中文-拼音匹配模块首先检测学习者输入的文本是否符合格式要求,利用正则表达式匹配所有中文字符,然后根据规则将中文字符与词典中的词条进行逐一匹配。本模块主要是按照顺序方向来对语句进行扫描,按最大词优先的匹配优先策略,完成中文字符切分,最后从微软提供的语言包 ChnCharInfo.dll 读取拼音,实现汉字转拼音及显示声调的功能。该语言包库功能强大,对于多音字都能转换出来。

3.4 对话管理

对话管理模块是整个系统的核心,主要通过检索

关键词,利用 TF-IDF 算法计算词语权重,并将用户输入的句子进行向量表达,计算句子的相似度。接着进行知识匹配,在对话模板语料库中寻找到与之相匹配的问句模板,然后按照对应的应答模板将此答案模型交由答案生成模板进行进一步的处理。该系统聊天机器人对话模板语料库(问题-答案对)的构建主要以基于人工模板和规则的方法实现,模板匹配的任务就是从对话模板语料库中检索出与学习者输入的文本最相关的问题,检索通过基于关键词的匹配方法,从中选择回复效果最好的语句返回给用户。

TF-IDF(即词频-逆文档频率)经常应用在信息检索和文本处理等领域,由 Jones^[20]首次提出。该算法的实现原理是将句子分成若干个词语,每个词语对句子的重要程度是由语料库中出现该词语的文档数和该词语在句子中出现的次数共同决定的,如果一个词语在句子中出现的次数越多,在语料库中含有这个词语的文档数越少,则表明该词语对句子具有很强的代表性。其中,TF 指的是词在句子中出现的次数,IDF 是一个词语普遍重要性的度量,利用 $TF * IDF$ 来表示词语在句子中的权重,其计算公式如下:

$$\text{词频}(TF) = \frac{\text{某个词在某文档中出现的次数}(n)}{\text{该文档的总词数}(N)} \quad (1)$$

逆文档频率(IDF) =

$$\log\left(\frac{\text{语料库中文档总数}(M)}{\text{包含该词的文档数}(m) + 1}\right) \quad (2)$$

词语在句子中的权重(TF-IDF) = TF × IDF (3)

在完成用户输入文本和语料库中问题 Questions 文本的空间向量构造后,就可以采用余弦夹角来计算向量之间的相似度,若两个句子包含的相同特征词较多时,说明匹配度较高,可以认为这两个句子相似,则语料库中问题 Questions 对应的应答也符合对用户输入文本的应答。 $T = \langle T_1, T_2, \dots, T_n \rangle$ 是语料库中的问题 Questions 向量, $T' = \langle T'_1, T'_2, \dots, T'_n \rangle$ 是用户提问的问句向量, n 表示不重复的词语的个数,为向量的维数。得到 T 和 T' 后,这两个问句之间的相似度就可用这两个向量之间夹角的余弦值来表示,计算公式如下:

$$\text{Sim}(T, T') = \frac{\sum_{i=1}^n T_i \times T'_i}{\sqrt{\sum_{i=1}^n T_i^2 \sum_{i=1}^n T'^2_i}} \quad (4)$$

由式(4)可知, $\text{Sim}(T, T')$ 的值越大,说明两者的相似度越高,反之则越低。相似性的得分范围在 $[0, 1]$, 将得分最高的问题 Questions 文本相对应的回复 Answer 输出给用户。对话模板语料库中包含特定主题语料和开放主题语料,特定主题语料适合于初级学习者,该库增加了回答引导功能,在特定对话场景中通过机器人提问引导学习者进一步学习汉语,机器人回复的内容是多样化的,有效增加了聊天的趣味性。开放主题语料的聊天话题没有限制,聊天方式也没有限制,可由机器人提问,也可学习者提问,模拟了在现实生活中人与人之间的聊天形式的语境,更有利于学习者对话训练,并灵活应用于实际生活中。聊天机器人的语料库规模越大,质量越高,返回的答案将越精准,效果越好。语料库中数据的获取主要通过以下两种方法实现:一是通过对软件或网站的解析,利用八爪鱼采集器进行数据爬取,整合所有数据之后进行清洗和清洗,然后存储在数据库中;二是从网络中查找一些相关开源的中文聊天语料,在此基础上进行修改和优化,从而进一步扩充语料库的规模以及提高语料数据的质量。

3.5 答案生成

答案生成模块的主要任务是从当前对话答案模型中提取出最终答案,然后将其与中文-拼音匹配模块中输出的拼音进行合成,将回复文本交由输出处理模块。

3.6 输出处理

输出处理模块主要负责将文本答案输出给用户,这将完成了学习者与聊天机器人一次完整的对话流程。该智能对外汉语学习系统的设计充分考虑到了情境学习在汉语学习中的重要作用,融合情境学习和建

构主义学习理论,突出以学习者为中心,学习者可以随时与聊天机器人进行交流,在情境对话中,系统能模拟现实生活中的真实对话情境,使学习者进行对话练习时能更好地掌握对话的实际使用方法,能有效地激发学习者的学习动机,提高学习兴趣,帮助学习者在构建的语言学习情境中对拼音、汉字、词语及情境对话形成直观的认识,并完成对汉语对话的意义建构。

4 系统教学特点及应用效果

基于聊天机器人的智能对外汉语学习系统应用在对外汉语教学上,具有如下特点:

(1)实时聊天。学习者可以在任何时间通过电脑打开系统,并和该系统进行聊天,前提是此电脑必须安装此系统软件。学习者在使用系统的过程中不受时间、地点及聊天内容的约束,目前该系统仅支持文本输入、输出方式,聊天方式侧重于对话训练,根据学习者输入的语句,理解此语句的语义,然后在模板库找到相对应的问题-答案对,生成答案返回给学习者,保障对话的流畅进行。

(2)灵活即时反馈。学习者向系统输入文本时,聊天机器人都会即时给出合适的输出响应,当学习者连续提出相同问题时,聊天机器人输出的答案是灵活多样的,不是单一死板的,这样,学习者可从众多反馈结果中选择自己认为合适的答案,并用于日常对话中。

(3)拼音声调显示。此功能根据学习者输入的文本产生对应的拼音和声调。对于那些汉语初学者,可以通过拼音和声调来进行对话和学习,从而达到较好的学习效果。

(4)对话场景选择。学习者可以根据自己的汉语能力水平选择相应的对话场景来进行学习,包括特定主题的聊天和开放主题的聊天,通过选择不同的对话场景,能使学习者沉浸在不同的情境中。特定主题的聊天适合于初级学习者,由于汉语水平不高、不愿多谈、缺乏自信,跟聊天机器人的对话则不多,所以需要由聊天机器人引导,在某一特定主题进行限制性对话。对于高级学习者则适合于开放主题的聊天,由于这些学习者对汉语熟练、比较健谈,所以选择开放主题无限制的聊天,进一步提高汉语会话水平。

(5)真实、轻松的语言学习环境。该系统为学习者创建了一个轻松愉快的语言学习环境,能够在一对一的交流过程中,有效地减少因面部表情、性别、种族、社会地位等因素的影响,缓解学习者紧张情绪,增强自信心^[21-22]。因此,该系统不仅能调动汉语学习者的学习主动性,更能激发学习热情和对话练习动机,使他们能够流利地使用汉语。通过与聊天机器人之间的对话,模拟真实情景的学习,帮助学习者更好地掌握对话

的实际使用方法,从而达到学以致用目的。为了检验智能对外汉语学习系统的设计及应用实践效果,主要采用问卷调查的形式对系统进行评价与反馈,随机抽取本校国际教育学院多名留学生作为对外汉语学习系统的试应用对象,并要求他们对系统使用效果与收获做出评价。

学习者使用效果与收获主要从情感态度、知识与技能两个层面展开,从调查结果可以看出,学习者的汉语学习态度和学习能力都得到了一定的提高。学习者能够在汉语对话中增强自信,提高学习汉语的兴趣,并且能够学到实用的会话句型和知识,对汉语交流的流利程度,以及汉语表达都有所提高。从中可以发现,学习者使用系统进行汉语学习时,让学习者处于一个类似真实的交互式人机对话汉语学习环境中,能有效地提高学习者的汉语水平,但在场景主题的个数、场景对话的长度、对话回答的质量等方面还需要加以强化。

5 结束语

在基于聊天机器人的智能对外汉语学习系统的设计和开发中,要注重其系统功能方面的设计和实现,同时,也要注重运用教育理论和教学方法,在这些基础上,使整个系统结构更加科学、合理,为学习者提供一个个性化、真实的语言学习环境。该系统的创新之处表现在理论和实践两个方面。理论上将情境学习理论和建构主义学习理论应用到了对外汉语学习系统中,实现了一个直观形象可互动的语言交流学习环境。实践上将聊天机器人与汉语学习系统相结合,也是将人工智能技术应用到教育领域,通过自然语言的方式进行人机对话,辅助学习者汉语会话的学习。

就目前而言,该智能对外汉语学习系统仍需要进一步进行完善和提高。例如:提高对话的准确率,扩充对话语料库的规模,以及使用关键词检索与基于语义相似度匹配结合的方法,使聊天机器人回复的效果更好、质量更高;增加对话场景的个数,并控制如何展开聊天话题;根据用户各异的学习风格创建适合的、独特的谈话类型,从而使谈话具有个性化,以此发挥学习者的自主性,支持个性化学习,增加学习者的信心,并激发其学习兴趣;对于不同水平的汉语学习者,如何让他们更快和更好地适应;研究语音识别技术,并将其结合到该系统中,使用户能够用语音功能进一步进行汉语学习,提高汉语对话和口语水平等。

参考文献:

[1] S·亚当斯贝克尔, M·卡明斯, A·戴维斯, 等. 新媒体联盟地平线报告:2017 高等教育版[J]. 北京广播电视大学学报, 2017(2):1-20.

- [2] 王珠珠. 教育信息化 2.0:核心要义与实施建议[J]. 中国远程教育, 2018(7):5-8.
- [3] 吴永和, 刘博文, 马晓玲. 构筑“人工智能+教育”的生态系统[J]. 远程教育杂志, 2017(5):27-39.
- [4] 黄晓洁. 自然语言处理技术在对外汉语教学中的应用研究[D]. 洛阳:中国人民解放军外国语学院, 2007.
- [5] 卢丹丹. 基于人工智能的大学英语辅助教学专家系统[D]. 北京:北京化工大学, 2009.
- [6] 安 静. 基于专家系统的英语辅助教学系统的设计与实现[D]. 北京:北京化工大学, 2011.
- [7] 陈 颖. 人工智能在中学英语教学中的应用探索[D]. 南昌:南昌大学, 2013.
- [8] 郝 丹. 结合 NLP 技术的汉语学习系统设计与实现[D]. 武汉:华中师范大学, 2015.
- [9] 孙雅琳. 人工智能在语言训练指导和批改中的应用[D]. 北京:北京邮电大学, 2018.
- [10] 贾积有. 人工智能技术的远程教育应用探索——“希赛可”智能型网上英语学习系统[J]. 现代教育技术, 2006(2):26-29.
- [11] 丁竹卉, 贾积有, 陈宇溟, 等. 中学英语教学中应用智能教学系统的效果评估——以“希赛可”为例[J]. 现代教育技术, 2012(4):68-72.
- [12] 杨晓琼, 戴运财. 基于批改网的大学英语自主写作教学模式实践研究[J]. 外语电化教学, 2015(3):17-23.
- [13] LAVE J, WENGER E. Situated learning: legitimate peripheral participation [M]. Cambridge: University of Cambridge Press, 1991:49-52.
- [14] 王世鹏. 高职日语专业商贸课程情境化教学实验研究——以江苏经贸职业技术学院为例[D]. 长春:吉林农业大学, 2018.
- [15] BARAB S, DUFFY T. Theory foundation of learning environment[M]. USA: Laurence Erlbaum, 2000:25-56.
- [16] BROWN J S, COLLINS A, DUGUID P. Situated cognition and the culture of learning [J]. Educational Researcher, 1989, 18(1):32-42.
- [17] 何克抗. 建构主义——革新传统教学的理论基础(上)[J]. 电化教育研究, 1997(3):3-9.
- [18] 高 源. 自然语言处理发展与应用概述[J]. 中国新通信, 2019(2):117-118.
- [19] 刘 群, 张华平, 俞鸿魁, 等. 基于层叠隐马模型的汉语词法分析[J]. 计算机研究与发展, 2004, 41(8):1421-1429.
- [20] JONES K S. A statistical interpretation of term specificity and its application in retrieval [J]. Journal of Documentation, 1972, 28(1):11-21.
- [21] 贾积有, 陈 霏, 陈宇灏, 等. 从聊天机器人到单词测试和课程管理——“希赛可”智能英语教学系统的进一步研发[J]. 现代教育技术, 2011(6):86-90.
- [22] SUN Y C. Voice blog: an exploratory study of language learning [J]. Language Learning & Technology, 2009, 13(2):88-103.