



(21) 申请号 202011384225.3

G06Q 10/0639 (2023.01)

(22) 申请日 2020.11.30

G06Q 50/20 (2012.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 112395403 A

(56) 对比文件

CN 105868313 A, 2016.08.17

CN 110704746 A, 2020.01.17

(43) 申请公布日 2021.02.23

CN 111506721 A, 2020.08.07

(73) 专利权人 广东国粒教育技术有限公司

CN 111753098 A, 2020.10.09

地址 519000 广东省珠海市高新区唐家湾

镇金唐路1号港湾1号科创园24栋C区1

层367室(集中办公区)

审查员 李荣荣

(72) 发明人 孙永毫 徐强

(74) 专利代理机构 广州三环专利商标代理有限

公司 44202

专利代理师 卢泽明

(51) Int. Cl.

G06F 16/332 (2019.01)

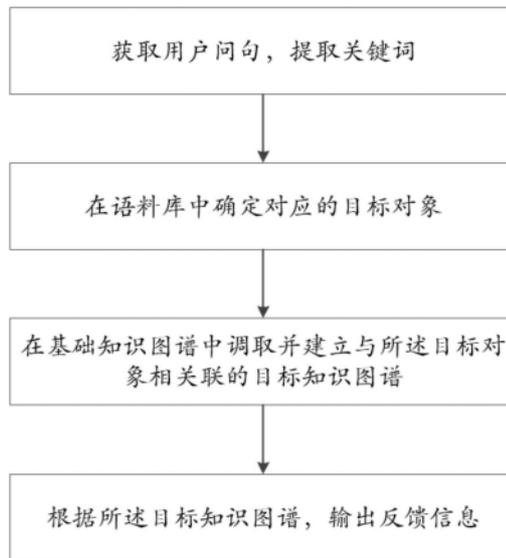
权利要求书3页 说明书10页 附图3页

(54) 发明名称

一种基于知识图谱的问答方法、系统、电子设备及介质

(57) 摘要

本发明提供一种基于知识图谱的问答方法、系统、电子设备及介质,属于网络智能教育技术领域。其中问答方法为:获取用户问句,提取关键词;在语料库中确定与所述用户问句对应的目标对象;在基础知识图谱中调取并建立与所述目标对象相关联的目标知识图谱;根据所述目标知识图谱,输出反馈信息。其中,基础知识图谱的建立需要采集学生每次考试的个性化数据;根据所述个性化数据计算每个知识点所对应的个性化掌握度;根据所述个性化掌握度,对知识点进行层级划分,并输出层级划分结果;根据所述层级划分结果,建立所述基础知识图谱。本发明使学生可根据薄弱知识点进行自我补强,包括课程的学习或者试题的练习,实现精准提升,有效消除薄弱知识点。



1. 一种基于知识图谱的问答方法,供学生、老师和家长共同使用,其特征在于,包括以下步骤:

步骤S1:获取用户问句,提取关键词;

步骤S2:在语料库中确定与所述用户问句对应的目标对象;

步骤S3:在基础知识图谱中调取并建立与所述目标对象相关联的目标知识图谱;

步骤S4:根据所述目标知识图谱,输出反馈信息;

在步骤S3之前,先建立基础知识图谱,包括以下步骤:

步骤S31:采集学生每次考试的个性化数据;

步骤S32:根据所述个性化数据计算每个知识点所对应的个性化掌握度;

步骤S33:根据所述个性化掌握度,对知识点进行层级划分,并输出层级划分结果;

步骤S34:根据所述层级划分结果,建立所述基础知识图谱;

所述反馈信息包括薄弱知识图谱、优势知识图谱、薄弱知识点课程、薄弱知识点试题中的至少一种;

在进行薄弱知识点试题的练习后,重复步骤S31至步骤S34,更新所述基础知识图谱;

对每一套薄弱知识点试题的错题进行收集,根据所述错题优化所述薄弱知识点试题的组建算法,以推荐与所述错题相似度高的试题。

2. 如权利要求1所述的一种基于知识图谱的问答方法,其特征在于,在步骤S2中:

根据所述知识图谱中的三元组,基于设定模板生成问答对;

每一个所述问答对包括一设定问句和一设定答案,对所述设定问句进行分词处理和向量化处理,得到与所述设定问句对应的设定词向量,根据所述设定词向量构建语料库;

对所述用户问句进行分词处理和向量化处理,得到用户词向量;

对所述用户词向量和设定词向量进行对比计算,根据相似度确定目标对象;所述目标对象包括N个问答对,其中N为大于等于1的自然数。

3. 如权利要求2所述的一种基于知识图谱的问答方法,其特征在于,对所述设定问句和所述用户问句进行向量化处理时,将所述设定问句和所述用户问句的文本进行分布式向量化表示,使用word2vec模型计算每一个所述设定问句和所述用户问句所对应的词向量。

4. 如权利要求3所述的一种基于知识图谱的问答方法,其特征在于,在所述语料库中,所述问答对存储有与其对应的问答信息,所述问答信息包括所述设定问句的五十维词向量和与所述设定问句对应的设定答案信息,所述问答信息组成所述语料库。

5. 如权利要求4所述的一种基于知识图谱的问答方法,其特征在于,根据知识图谱中的三元组关系种类,输入人工问答对至所述语料库中。

6. 如权利要求5所述的一种基于知识图谱的问答方法,其特征在于,在对所述用户词向量和设定词向量进行对比计算时,计算所述用户词向量和设定词向量之间所组成夹角的余弦值,并根据所述余弦值判断所述用户问句和设定问句的相似度,其中,所述余弦值越接近1,所述用户问句和设定问句的相似度越高;所述余弦值越接近0,所述用户问句和设定问句的相似度越低。

7. 如权利要求6所述的一种基于知识图谱的问答方法,其特征在于,所述个性化数据包括考试时间信息、考试试题信息、考试试卷信息或考试成绩信息中的至少一项,所述考试试卷信息包括所属科目、试卷类型、试卷名称、试题满分分值和原始题目信息中的至少一项,

所述原始题目信息包括所属题型、关联章节、关联知识点、题目本身所属能力层次、题目适用类型、题目难度系数、试题题干及答案、题目分值、题目年份、区域、来源、考点、分析和点评中的至少一项。

8. 如权利要求7所述的一种基于知识图谱的问答方法,其特征在于,在步骤S32中,所述个性化掌握度通过以下方法计算得到:

$f = [(a_1 + a_2 + a_3 + \dots + a_n) \div n] \times g$, 其中, f 是所述知识点的个性化掌握度, g 是所述知识点重要度, a_1 、 a_2 、 a_3 ... a_n 是第1个、第2个、第3个...第 n 个和所述知识点相应的题目评分值, n 是所述知识点的错题个数;

所述题目评分值 a 通过以下方法得到: $a = \frac{x\%b \times y\%(c+1) \times z\%(d+1)}{p}$, 其中, a 是所述题目评分值, b 是试题满分分值, $x\%$ 是试题满分分值的权重, c 是试题区分度, $y\%$ 是试题区分度的权重, d 是知识点个人得分率和班级得分率的差值, $z\%$ 是得分率差值的权重, p 是难度系数, 难度系数 p 的范围在 $0 \sim 1$ 之间, 难度系数 p 的值越趋近于 0 难度越大;

所述试题区分度 c 通过以下方法得到: $c = (v_1 - v_2) / b$, 其中, v_1 是成绩排名前 20% 的所述考试成绩平均分, v_2 是成绩排名后 20% 的所述考试成绩平均分;

所述知识点重要度 g 通过以下方法得到: $g = 70\%i + 30\%j$, 其中, i 是大纲要求学生对所知识点的掌握度, j 是本试卷中有关所述知识点的题目分值占总分值的百分比。

9. 如权利要求8所述的一种基于知识图谱的问答方法,其特征在于,在步骤S33中,所述层级划分结果包括薄弱知识点层级和优势知识点层级,所述知识点与所述个性化掌握度一一对应,将所述个性化掌握度与设定的薄弱知识点阈值和优势知识点阈值相比较,将与小于所述薄弱知识点阈值的个性化掌握度相对应的知识点定义为薄弱知识点,将所述薄弱知识点划分至薄弱知识点层级,将与大于所述优势知识点阈值的个性化掌握度相对应的知识点定义为优势知识点,将所述优势知识点划分至优势知识点层级。

10. 如权利要求9所述的一种基于知识图谱的问答方法,其特征在于,在步骤S34中,所述基础知识图谱包括薄弱知识图谱和/或优势知识图谱;

其中,根据薄弱知识点层级信息,在设定的原始知识图谱中标记每个薄弱知识点的个性化掌握度,建立薄弱知识图谱;

根据优势知识点层级信息,在设定的原始知识图谱中标记每个优势知识点的个性化掌握度,建立优势知识图谱。

11. 如权利要求10所述的一种基于知识图谱的问答方法,其特征在于,所述原始知识图谱基于学科权威的教材教辅资料来确定,并根据知识点概念和百科信息框来确定关系;所述薄弱知识图谱还包括多个薄弱知识点之间的关系;所述优势知识图谱还包括多个优势知识点之间的关系。

12. 一种应用于如权利要求1至11任一项所述的问答方法的系统,其特征在于,包括:

获取单元,所述获取单元被配置为获取用户问句,提取关键词;

判断单元,所述判断单元被配置为在语料库中确定与所述用户问句对应的目标对象;

生成单元,所述生成单元被配置为在基础知识图谱中调取并建立与所述目标对象相关联的目标知识图谱;

反馈单元,所述反馈单元被配置为根据所述目标知识图谱输出反馈信息。

13. 一种电子设备,其特征在于,包括处理器和存储器,所述存储器中存储有至少一条指令、至少一段程序、代码集或指令集,所述至少一条指令、至少一段程序、代码集或指令集由所述处理器加载并执行以实现如权利要求1至11任一所述的基于知识图谱的问答方法。

14. 一种计算机可读存储介质,其上存储有计算机指令,其特征在于,所述计算机指令被处理器执行时实现如权利要求1至11中任一项所述方法的步骤。

一种基于知识图谱的问答方法、系统、电子设备及介质

技术领域

[0001] 本发明属于网络智能教育技术领域,尤其涉及一种基于知识图谱的问答方法、系统、电子设备及介质。

背景技术

[0002] 在“互联网+教育”的背景下,随着国家教育信息化的持续推进,教育数据的持续累积与深度挖掘,大数据在构建新型教学生态、助力教学结构变革、再造教学流程方面的驱动作用日益凸显。

[0003] 传统教学环境中,老师以班为单位统一下发作业,无视学生对知识消化和接受的快慢程度,直接同质化施教,出现有的学生觉得作业简单而有的学生觉得作业难的问题,作业与学生知识水平的匹配度不高,因此老师在掌握班级的教学质量的时候难免会出现偏差,无法精准定位到每位学生的个性化学习情况。因此在这种同质化施教的方式下,很难调动学生的学习积极性,无法针对每个学生自身知识水平进行区别化分层,学生学习兴趣不高,学习效率低下。

发明内容

[0004] 本发明的目的在于克服上述现有技术存在的不足,提供一种基于知识图谱的问答方法、系统、电子设备及介质,解决了现有技术中对学生同质化施教所带来的不良影响。

[0005] 为了实现上述目的,第一方面,本发明提供一种基于知识图谱的问答方法,包括以下步骤:

[0006] 步骤S1:获取用户问句,提取关键词;

[0007] 步骤S2:在语料库中确定与所述用户问句对应的目标对象;

[0008] 步骤S3:在基础知识图谱中调取并建立与所述目标对象相关联的目标知识图谱;

[0009] 步骤S4:根据所述目标知识图谱,输出反馈信息。

[0010] 进一步地,在步骤S2中:

[0011] 根据所述知识图谱中的三元组,基于设定模板生成问答对;

[0012] 每一个所述问答对包括一设定问句和一设定答案,对所述设定问句进行分词处理和向量化处理,得到与所述设定问句对应的设定词向量,根据所述设定词向量构建语料库;

[0013] 对所述用户问句进行分词处理和向量化处理,得到用户词向量;

[0014] 对所述用户词向量和设定词向量进行对比计算,根据相似度确定目标对象;所述目标对象包括N个问答对,其中N为大于等于1的自然数。

[0015] 进一步地,对所述设定问句和所述用户问句进行向量化处理时,将所述设定问句和所述用户问句的文本进行分布式向量化表示,使用word2vec模型计算每一个所述设定问句和所述用户问句所对应的词向量。

[0016] 进一步地,在所述语料库中,所述问答对存储有与其对应的问答信息,所述问答信息包括所述设定问句的五十维词向量和与所述设定问句对应的设定答案信息,所述问答信

息组成所述语料库。

[0017] 进一步地,根据知识图谱中的三元组关系种类,输入人工问答对至所述语料库中。

[0018] 进一步地,在对所述用户词向量和设定词向量进行对比计算时,计算所述用户词向量和设定词向量之间所组成夹角的余弦值,并根据所述余弦值判断所述用户问句和设定问句的相似度,其中,所述余弦值越接近1,所述用户问句和设定问句的相似度越高;所述余弦值越接近0,所述用户问句和设定问句的相似度越低。

[0019] 进一步地,在步骤S3之前,先建立基础知识图谱,包括以下步骤:

[0020] 步骤S31:采集学生每次考试的个性化数据;

[0021] 步骤S32:根据所述个性化数据计算每个知识点所对应的个性化掌握度;

[0022] 步骤S33:根据所述个性化掌握度,对知识点进行层级划分,并输出层级划分结果;

[0023] 步骤S34:根据所述层级划分结果,建立所述基础知识图谱。

[0024] 进一步地,所述个性化数据包括考试时间信息、考试试题信息、考试试卷信息或考试成绩信息中的至少一项,所述考试试卷信息包括所属科目、试卷类型、试卷名称、试题满分分值和原始题目信息中的至少一项,所述原始题目信息包括所属题型、关联章节、关联知识点、题目本身所属能力层次、题目适用类型、题目难度系数、试题题干及答案、题目分值、题目年份、区域、来源、考点、分析和点评中的至少一项。

[0025] 进一步地,在步骤S32中,所述个性化掌握度通过以下方法计算得到:

[0026] $f = [(a_1 + a_2 + a_3 + \dots + a_n) \div n] \times g$,其中, f 是所述知识点的个性化掌握度, g 是所述知识点重要度, a_1 、 a_2 、 a_3 ... a_n 是第1个、第2个、第3个...第 n 个和所述知识点相应的题目评分值, n 是所述知识点的错题个数;

[0027] 所述题目评分值 a 通过以下方法得到: $a = \frac{x\%b \times y\%(c+1) \times z\%(d+1)}{p}$,其中, a 是所

述题目评分值, b 是试题满分分值, $x\%$ 是试题满分分值的权重, c 是试题区分度, $y\%$ 是试题区分度的权重, d 是知识点个人得分率和班级得分率的差值, $z\%$ 是得分率差值的权重, p 是难度系数,难度系数 p 的范围在0~1之间,难度系数 p 的值越趋近于0难度越大;

[0028] 所述试题区分度 c 通过以下方法得到: $c = (v_1 - v_2) / b$,其中, v_1 是成绩排名前20%的所述考试成绩平均分, v_2 是成绩排名后20%的所述考试成绩平均分;

[0029] 所述知识点重要度 g 通过以下方法得到: $g = 70\%i + 30\%j$,其中, i 是大纲要求学生所述知识点的掌握度, j 是本试卷中有关所述知识点的题目分值占总分值的百分比。

[0030] 进一步地,在步骤S33中,所述层级划分结果包括薄弱知识点层级和优势知识点层级,所述知识点与所述个性化掌握度一一对应,将所述个性化掌握度与设定的薄弱知识点阈值和优势知识点阈值相比较,将与所述小于薄弱知识点阈值的个性化掌握度相对应的知识点定义为薄弱知识点,将所述薄弱知识点划分至薄弱知识点层级,将与所述大于优势知识点阈值的个性化掌握度相对应的知识点定义为优势知识点,将所述优势知识点划分至优势知识点层级。

[0031] 进一步地,在步骤S34中,所述基础知识图谱包括薄弱知识图谱和/或优势知识图谱;

[0032] 其中,根据薄弱知识点层级信息,在设定的原始知识图谱中标记每个薄弱知识点的个性化掌握度,建立薄弱知识图谱;

[0033] 根据优势知识点层级信息,在设定的原始知识图谱中标记每个优势知识点的个性化掌握度,建立优势知识图谱。

[0034] 进一步地,所述原始知识图谱基于学科权威的教材教辅资料来确定,并根据知识点概念和百科信息框来确定关系;所述薄弱知识图谱还包括多个薄弱知识点之间的关系;所述优势知识图谱还包括多个优势知识点之间的关系。

[0035] 进一步地,所述反馈信息包括薄弱知识图谱、优势知识图谱、薄弱知识点课程、薄弱知识点试题中的至少一种。

[0036] 进一步地,在进行薄弱知识点试题的练习后,重复步骤S31至步骤S34,更新所述基础知识图谱。

[0037] 进一步地,对每一套薄弱知识点试题的错题进行收集,根据所述错题优化所述薄弱知识点试题的组建算法,以推荐与所述错题相似度高的试题。

[0038] 第二方面,本发明提供一种应用于上述问答方法的系统,包括:

[0039] 获取单元,所述获取单元被配置为获取用户问句,提取关键词;

[0040] 判断单元,所述判断单元被配置为在语料库中确定与所述用户问句对应的目标对象;

[0041] 生成单元,所述生成单元被配置为在基础知识图谱中调取并建立与所述目标对象相关联的目标知识图谱;

[0042] 反馈单元,所述反馈单元被配置为根据所述目标知识图谱输出反馈信息。

[0043] 第三方面,本发明提供一种电子设备,包括处理器和存储器,所述存储器中存储有至少一条指令、至少一段程序、代码集或指令集,所述至少一条指令、至少一段程序、代码集或指令集由所述处理器加载并执行以实现上述的基于知识图谱的问答方法。

[0044] 第四方面,本发明提供一种计算机可读存储介质,其上存储有计算机指令,所述计算机指令被处理器执行时实现上述方法的步骤。

[0045] 本发明的有益效果:

[0046] 1、本发明对学生每次考试的个性化数据进行分析处理,计算出每个知识点的个性化掌握程度,然后对所有知识点进行层级划分,识别出薄弱知识点和优势知识点,在原始知识图谱中结合此个性化掌握程度进行标记,并形成与之对应的薄弱知识图谱和/或优势知识图谱,从直观画面中让学生清楚了解到其自身对各个知识点的个性化掌握程度,并可以根据薄弱知识图谱找出自己的知识弱点及与其有关联的其他知识点,形成一个知识网,不再是在众多知识点中靠自己一个一个去摸索,提高了学生归纳整理的效率;

[0047] 2、本发明基于知识图谱,实现跟学生之间的互动,学生只需输入包含有其疑问知识点的用户问句,即自动将用户问句与语料库中设定问句进行相似度比对,并确定N个相似度最高的问答对,通过此目标对象反查基础知识图谱中与之对应的那部分,以形成具有学生疑问知识点的目标知识图谱,学生或者家长或者老师可以直观看到学生对该知识点的掌握度,方便不同的用户清楚了解学生的个性化学习掌握水平,并以此确定下一步教育学习方针;

[0048] 3、本发明还可同时输出薄弱知识点课程和薄弱知识点试题,让学生自主地针对其薄弱知识点进行自我补强,包括课程的学习或者试题的练习,实现精准提升,有效消除薄弱知识点,其中知识图谱、薄弱知识点试题都会根据学生每次的做题情况进行自我优化和更

新,实现知识图谱与实际学生知识掌握水平的同步更新,优化薄弱知识点试题的组建算法,帮助学生攻克难关。

附图说明

[0049] 利用附图对本发明作进一步说明,但附图中的实施例不构成对本发明的任何限制,对于本领域的普通技术人员,在不付出创造性劳动的前提下,还可以根据以下附图获得其它的附图。

[0050] 图1是本发明提供的一种基于知识图谱的问答方法在实施例1中的流程框架示意图。

[0051] 图2是本发明提供的一种基于知识图谱的问答方法在实施例1中建立基础知识图谱的流程框架示意图。

[0052] 图3是本发明提供的一种基于知识图谱的问答方法在实施例1中薄弱知识图谱的示意图

具体实施方式

[0053] 下面将结合附图对本发明的技术方案进行清楚、完整地描述,显然,所描述的实施例是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0054] 在本发明的描述中,需要说明的是,术语“中心”、“上”、“下”、“左”、“右”、“竖直”、“水平”、“内”、“外”等指示的方位或位置关系为基于附图所示的方位或位置关系,仅是为了便于描述本发明和简化描述,而不是指示或暗示所指的装置或元件必须具有特定的方位、以特定的方位构造和操作,因此不能理解为对本发明的限制。此外,术语“第一”、“第二”、“第三”仅用于描述目的,而不能理解为指示或暗示相对重要性。

[0055] 实施例1:

[0056] 参照图1,本实施例提供一种基于知识图谱的问答方法,包括以下步骤:

[0057] 步骤S1:获取用户问句,提取关键词;

[0058] 步骤S2:在语料库中确定与用户问句对应的目标对象;

[0059] 步骤S3:在基础知识图谱中调取并建立与目标对象相关联的目标知识图谱;

[0060] 步骤S4:根据目标知识图谱,输出反馈信息。

[0061] 需要说明的是,本方法可给学生、老师和家长共同使用,用户只需输入一个问句,这个问句包含学生想要查询的知识点,即根据此知识点与语料库中的信息进行分析比对,以确定与之对应的目标对象,此目标对象已经包含了用户知识点的相关信息,并且在基础知识图谱中选择跟此知识点相关联的目标知识图谱,最后根据此目标知识图谱输出反馈信息,由于目标知识图已经包含有用户知识点的相关信息,因此所推送的反馈信息更具有针对性、精准性。

[0062] 在以上方法中,最重要的是两方面,第一方面是在步骤S2中,如何有效建立这个问答机制,使得所推送的反馈信息即符合用户的需求;第二方面是在步骤S3中,基础知识图谱是如何建立的,怎样将学生的个性化数据体现在基础知识图谱中。

[0063] 下面先说明第一方面,在本实施例中,在步骤S2中:

[0064] 根据知识图谱中的三元组,基于设定模板生成问答对;

[0065] 每一个问答对包括一设定问句和一设定答案,设定问句和设定答案是预先设定好,是一一对应的关系,对设定问句进行分词处理和向量化处理,得到与设定问句对应的设定词向量,根据设定词向量构建语料库;

[0066] 对用户问句进行分词处理和向量化处理,得到用户词向量;

[0067] 对用户词向量和设定词向量进行对比计算,根据相似度确定目标对象;目标对象包括N个问答对,其中N为大于等于1的自然数。

[0068] 作为一种优选方式,对设定问句和用户问句进行向量化处理时,将设定问句和用户问句的文本进行分布式向量化表示,使用word2vec模型计算每一个设定问句和用户问句所对应的词向量。

[0069] 在本实施例中,在语料库中,问答对存储有与其对应的问答信息,问答信息包括设定问句的五十维词向量和与设定问句对应的设定答案信息,问答信息组成语料库,因此通过语料库,实现用户词向量到设定词向量再到设定答案信息之间的数据联系。

[0070] 另外地,根据知识图谱中的三元组关系种类,还可输入人工问答对至语料库中,即可实现人为地有针对性地添加或修改语料库中的信息。

[0071] 作为一种优选方式,在对用户词向量和设定词向量进行对比计算时,计算用户词向量和设定词向量之间所组成夹角的余弦值,并根据余弦值判断用户问句和设定问句的相似度,其中,余弦值越接近1,用户问句和设定问句的相似度越高;余弦值越接近0,用户问句和设定问句的相似度越低;在最接近1的结果中,选取排在前的N个余弦值,确定与之对应的N个设定问句,还有与此设定问句相对应的设定答案信息,完成目标对象的选取。

[0072] 参照图3,至此,在此目标对象中,已经包含有用户希望知道的知识点的相关信息,而选取N个设定问句,一是提高准确度,二是确定跟最相似知识点有密切关联的其他知识点,由于知识点不是单独存在的,知识点都是在一个网中的其中一个点,知识点之间存在密切的关联,例如搜“二次函数”,但是与其有关联的如平面直角坐标系、一元二次方程等知识点也有可能是影响此“二次函数”掌握程度的因素,所以这N个设定问句及其设定答案信息,可以组成知识点路径或者知识点区域关系网,各个知识点深浅串联,让用户了解学生的薄弱之处。

[0073] 为了实现以上问答机制,背后需要建立具有学生知识掌握差别化信息的基础知识图谱,这个基础知识图谱能针对每个学生,依据其个性化数据,确定每个知识点的个性化掌握度。

[0074] 下面再说明第二方面,如何建立基础知识图谱,怎样将学生的个性化数据体现在基础知识图谱中。

[0075] 参照图2,在本实施例中,在步骤S3之前,先建立基础知识图谱,包括以下步骤:

[0076] 步骤S31:采集学生每次考试的个性化数据;

[0077] 步骤S32:根据个性化数据计算每个知识点所对应的个性化掌握度;

[0078] 步骤S33:根据个性化掌握度,对知识点进行层级划分,并输出层级划分结果;

[0079] 步骤S34:根据层级划分结果,建立基础知识图谱。

[0080] 需要说明的是,学生的每次考试都会产生相应的个性化数据,这些数据包括学生在此次考试中的数据,反应学生在考试当中的相关表现;更具体地,个性化数据包括考试时

间信息、考试试题信息、考试试卷信息或考试成绩信息中的至少一项,考试时间信息包括考试的日期、考试前后的时间点、总花费时间、每道题花费时间、检查时间等,考试试卷信息包括所属科目、试卷类型、试卷名称、试题满分分值和原始题目信息中的至少一项,原始题目信息包括所属题型、关联章节(科目、年级、单元、课文)、关联知识点、题目本身所属能力层次、题目适用类型、题目难度系数、试题题干及答案、题目分值、题目年份、区域、来源、考点、分析和点评中的至少一项;当然个性化数据还可以包括其他能代表与本次考试相关的信息。

[0081] 在本实施例中,在收集了个性化数据后,在步骤S32中,个性化掌握度通过以下方法计算得到:

[0082] $f = [(a_1 + a_2 + a_3 + \dots + a_n) \div n] \times g$,其中,f是知识点的个性化掌握度,g是知识点重要度, a_1 、 a_2 、 $a_3 \dots a_n$ 是第1个、第2个、第3个...第n个和知识点相应的题目评分值,n是知识点的错题个数;

[0083] 题目评分值a通过以下方法得到: $a = \frac{x\%b \times y\%(c+1) \times z\%(d+1)}{p}$,其中,a是题目评

分值,b是试题满分分值,x%是试题满分分值的权重,c是试题区分度,y%是试题区分度的权重,d是知识点个人得分率和班级得分率的差值,z%是得分率差值的权重,p是难度系数,难度系数p的范围在0~1之间,难度系数p的值越趋近于0难度越大;

[0084] 试题区分度c通过以下方法得到: $c = (v_1 - v_2) / b$,其中, v_1 是成绩排名前20%的考试成绩平均分, v_2 是成绩排名后20%的考试成绩平均分;

[0085] 知识点重要度g通过以下方法得到: $g = 70\%i + 30\%j$,其中,i是大纲要求学生对该知识点的掌握度,j是本试卷中有关知识点的所有题目分值占总分值的百分比。

[0086] 作为一种实施方式,试题满分分值的权重x%、试题区分度的权重y%、得分率差值的权重z%加起来为1,即 $x\% + y\% + z\% = 1$ 。根据本次考试情况,定义:试题满分分值的权重 $x\% = 25\%$,试题区分度的权重 $y\% = 35\%$,得分率差值的权重 $z\% = 40\%$ 。

[0087] 在本次考试中,一共26题,满分120分,含有某个知识点的题目一共有6道题,占总分值的33%,大纲要求学生对该知识点掌握度为80%。其中:

[0088] 第1道题满分3分,难度系数0.9,班级得分率为92%,该学生该道题的得分率为100%。

[0089] 第2道题满分3分,难度系数0.88,班级得分率为90%,该学生该道题的得分率为0。

[0090] 第3道题满分3分,难度系数0.85,班级得分率为89%,该学生该道题的得分率为100%。

[0091] 第4道题满分6分,难度系数0.75,班级得分率为64%,该学生该道题的得分率为83%。

[0092] 第5道题满分10分,难度系数0.6,班级得分率为48%,该学生该道题的得分率为60%。

[0093] 第6道题满分15分,难度系数0.2,班级得分率为19%,该学生该道题的得分率为10%。

[0094] 计算这六道题的区分度得:

[0095] 第1道题满分3分,高分组平均得分3分,低分组平均得分0分,则该题区分度为

$$c1 = (3 - 0) / 3 = 1;$$

[0096] 第2道题满分3分,高分组平均得分3分,低分组平均得分0分,则该题区分度为

$$c1 = (3 - 0) / 3 = 1;$$

[0097] 第3道题满分3分,高分组平均得分3分,低分组平均得分0分,则该题区分度为

$$c1 = (3 - 0) / 3 = 1;$$

[0098] 第4题满分6分的试题,高分组平均得分4.5分,低分组平均得分2.5分,则该题区分度为

$$c4 = (4.5 - 2.5) / 6 = 0.33;$$

[0099] 第5道题满分10分,高分组平均得分8.8分,低分组平均得分2.4分,则该题区分度为

$$c5 = (8.8 - 2.4) / 10 = 0.6;$$

[0100] 第6道题满分15分,高分组平均得分12.8分,低分组平均得分2.8分,则该题区分度为

$$c6 = (12.8 - 2.8) / 15 = 0.7。$$

[0101] 另外地,大纲要求学生对该知识点掌握度为 $i=80%$,本试卷中有关知识点的所有题目分值占总分值的百分比 $j=33%$,因此 $g=70% \times 80% + 30% \times 33%=0.659$ 。

[0102] 由上述内容得到上述那道题的评分值 $a1$ 为:

$$[0103] \quad a1 = \frac{x\%b \times y\%(c+1) \times z\%(d+1)}{p} = \frac{25\% \times 3 \times 35\%(1+1) \times 40\%(100\% - 92\% + 1)}{0.9} = 0.252$$

剩下的计算出5道题评分值分别为:

$$[0104] \quad a2 = 0.024;$$

$$[0105] \quad a3 = 0.274;$$

$$[0106] \quad a4 = 0.443;$$

$$[0107] \quad a5 = 1.045;$$

$$[0108] \quad a6 = 4.061。$$

[0109] 则该学生在该次数学期中模拟考试对该知识点掌握度为:

$$[0110] \quad f = [(a1 + a2 + a3 + \dots + a6) \div 6] \times g = [(0.252 + 0.024 + 0.274 + 0.443 + 1.045 + 4.061) \div 6] \times 0.659 \approx 0.67。$$

[0111] 最后将本知识点的掌握度数据结合其他知识点的掌握度数据,建立并完善薄弱知识图谱。

[0112] 作为一种优选方式,在步骤S33中,层级划分结果包括薄弱知识点层级和优势知识点层级,知识点与个性化掌握度一一对应,将个性化掌握度与设定的薄弱知识点阈值和优势知识点阈值相比较,将个性化掌握度按照数值大小从小至大依次排列,在这个数列中分别以薄弱知识点阈值和优势知识点阈值为分界点,其中将小于薄弱知识点阈值的个性化掌握度分离出来,将与此个性化掌握度相对应的知识点定义为薄弱知识点,将薄弱知识点划分至薄弱知识点层级;将大于优势知识点阈值的个性化掌握度分离出来,将与此个性化掌握度相对应的知识点定义为优势知识点,将优势知识点划分至优势知识点层级。由此,通过量化的数值比较,将在考试试卷中出现的所有知识点分离提取出薄弱知识点和优势知识点。

[0113] 更进一步地,在步骤S34中,基础知识图谱包括薄弱知识图谱和/或优势知识图谱;

[0114] 其中,根据薄弱知识点层级信息,在设定的原始知识图谱中标记每个薄弱知识点

的个性化掌握度,建立薄弱知识图谱;

[0115] 根据优势知识点层级信息,在设定的原始知识图谱中标记每个优势知识点的个性化掌握度,建立优势知识图谱。

[0116] 不论是薄弱知识图谱还是优势知识图谱,均是建立在原始知识图谱中,并对各个知识点进行标记,区分开是薄弱知识点还是优势知识点,另外对应的标记上其个性化掌握度,并且根据知识点之间的联系,组成相关的路径或者网络的图谱。

[0117] 在本实施例中,原始知识图谱基于学科权威的教材教辅资料来确定,并根据知识点概念和百科信息框来确定关系,每个知识点概念下都有很多实例,实例在百科中都有对应的信息框,通过整合同一知识点下多个实例的信息框关系,得到该知识点下较为重要的关系;其中,薄弱知识图谱包括多个薄弱知识点之间的关系;优势知识图谱包括多个优势知识点之间的关系。

[0118] 作为一种优选方式,在确定好目标知识图谱后,需要输出反馈信息,反馈信息包括薄弱知识图谱、优势知识图谱、薄弱知识点课程、薄弱知识点试题中的至少一种。

[0119] 在本实施例中,在进行薄弱知识点试题的练习后,重复步骤S31至步骤S34,更新基础知识图谱,使得基础知识图谱实时更新,保证与学生的知识水平同步。

[0120] 在本实施例中,对每一套薄弱知识点试题的错题进行收集,形成错题本,根据错题优化薄弱知识点试题的组建算法,以推荐与错题相似度高的试题,使得在下次再进行此薄弱知识点的相关联系时,出题更有针对性,出题的质量更高,更人性化。

[0121] 参照图1至图3,综上,在建立好基础知识图谱后,当获取到用户问句时,通过将用户词向量和设定词向量进行对比计算,确定目标对象,此目标对象即包含有知识点的信息,通过此知识点在基础知识图谱进行搜索,以找到跟其相对应的目标知识图谱,根据目标对象中包含的知识点的层级,相应地输出跟此目标知识图谱相对应的反馈信息,包括薄弱知识图谱、薄弱知识点课程或薄弱知识点试题。所以用户可以得到与学生想要查询的有疑问的知识点相关的信息,特别是当遇到学习上的难点,例如是二次函数,即会出现与本二次函数相关的薄弱知识图谱,此薄弱知识图谱可能沿着此二次函数依次连接着多个知识点,且根据连接关系,可以分成一级相关知识点、二级相关知识点等,除了看到在二次函数这一对象中的个性化掌握度,还能看到与之有关联的其余相关知识点的个性化掌握度,并且可以在此路径上找到真正的问题,如与之相连的依次有平面直角坐标系、一元二次方程、因式分解和一元一次方程,平面直角坐标系、因式分解和一元一次方程的个性化掌握度较高,只有一元二次方程的个性化掌握度较低,因此是一元二次方程这一环有缺失,才导致二次函数这个知识点称为薄弱知识点,正因为利用这种知识点之间的联系,可以让用户清楚看到学生的弱点,能做到点对点攻克,提高学习效率,学生、家长和老师三方受益。

[0122] 实施例2:

[0123] 本实施例2提供一种应用于上述实施例1中问答方法的系统,包括:

[0124] 获取单元,获取单元被配置为获取用户问句,提取关键词;

[0125] 判断单元,判断单元被配置为在语料库中确定与用户问句对应的目标对象;

[0126] 生成单元,生成单元被配置为在基础知识图谱中调取并建立与目标对象相关联的目标知识图谱;

[0127] 反馈单元,反馈单元被配置为根据目标知识图谱输出反馈信息。

[0128] 需要说明的是,本系统可给学生、老师和家长共同使用,当学生、老师和家长输入其想要查询的知识点时,获取单元即会从此用户问句中提取关键词,判断单元在在语料库中确定与用户问句对应的目标对象,此目标对象已经包含了用户知识点的相关信息,然后生成单元再从基础知识图谱中调取并建立与目标对象相关联的目标知识图谱,最后反馈单元根据目标知识图谱输出反馈信息,由于目标知识图已经包含有用户知识点的相关信息,因此所推送的反馈信息更具有针对性、精准性。

[0129] 实施例3:

[0130] 本实施例3提供一种电子设备,包括处理器和存储器,存储器中存储有至少一条指令、至少一段程序、代码集或指令集,至少一条指令、至少一段程序、代码集或指令集由处理器加载并执行以实现上述的实施例1中基于知识图谱的问答方法。

[0131] 实施例4:

[0132] 本实施例4提供一种计算机可读存储介质,其上存储有计算机指令,计算机指令被处理器执行时实现上述实施例1中方法的步骤。

[0133] 上述实施例阐明的系统、装置、模块或单元,具体可以由计算机芯片或实体实现,或者由具有某种功能的产品来实现。一种典型的实现设备为计算机,计算机的具体形式可以是个人计算机、膝上型计算机、蜂窝电话、相机电话、智能电话、个人数字助理、媒体播放器、导航设备、电子邮件收发设备、游戏控制台、平板计算机、可穿戴设备或者这些设备中的任意几种设备的组合。

[0134] 在一个典型的配置中,计算机包括一个或多个处理器(CPU)、输入/输出接口、网络接口和内存。

[0135] 内存可能包括计算机可读介质中的非永久性存储器,随机存取存储器(RAM)和/或非易失性内存等形式,如只读存储器(ROM)或闪存(flash RAM)。内存是计算机可读介质的示例。

[0136] 计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括,但不限于相变内存(PRAM)、静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、其他类型的随机存取存储器(RAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器(CD-ROM)、数字多功能光盘(DVD)或其他光学存储、磁盒式磁带、磁盘存储、量子存储器、基于石墨烯的存储介质或其他磁性存储设备或任何其他非传输介质,可用于存储可以被计算设备访问的信息。按照本文中的界定,计算机可读介质不包括暂存电脑可读媒体(transitory media),如调制的数据信号和载波。

[0137] 相对于现有技术,本发明对学生每次考试的个性化数据进行分析处理,计算出每个知识点的个性化掌握程度,然后对所有知识点进行层级划分,识别出薄弱知识点和优势知识点,在原始知识图谱中结合此个性化掌握程度进行标记,并形成与之对应的薄弱知识图谱和/或优势知识图谱,从直观画面中让学生清楚了解到其自身对各个知识点的个性化掌握程度,并可以根据薄弱知识图谱找出自己的知识弱点及与其有关联的其他知识点,形成一个知识网,不再是在众多知识点中靠自己一个一个去摸索,提高了学生归纳整理的效率;

[0138] 本发明基于知识图谱,实现跟学生之间的互动,学生只需输入包含有其疑问知识点的用户问句,即自动将用户问句与语料库中设定问句进行相似度比对,并确定N个相似度最高的问答对,通过此目标对象反查基础知识图谱中与之对应的那部分,以形成具有学生疑问知识点的目标知识图谱,学生或者家长或者老师可以直观看到学生对该知识点的掌握度,方便不同的用户清楚了解学生的个性化学习掌握水平,并以此确定下一步教育学习方针;

[0139] 本发明还可同时输出薄弱知识点课程和薄弱知识点试题,让学生自主地针对其薄弱知识点进行自我补强,包括课程的学习或者试题的练习,实现精准提升,有效消除薄弱知识点,其中知识图谱、薄弱知识点试题都会根据学生每次的做题情况进行自我优化和更新,实现知识图谱与实际学生知识掌握水平的同步更新,优化薄弱知识点试题的组建算法,帮助学生攻克难关。

[0140] 最后需要强调的是,本发明不限于上述实施方式,以上仅为本发明的较佳实施例而已,并不用以限制本发明,凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

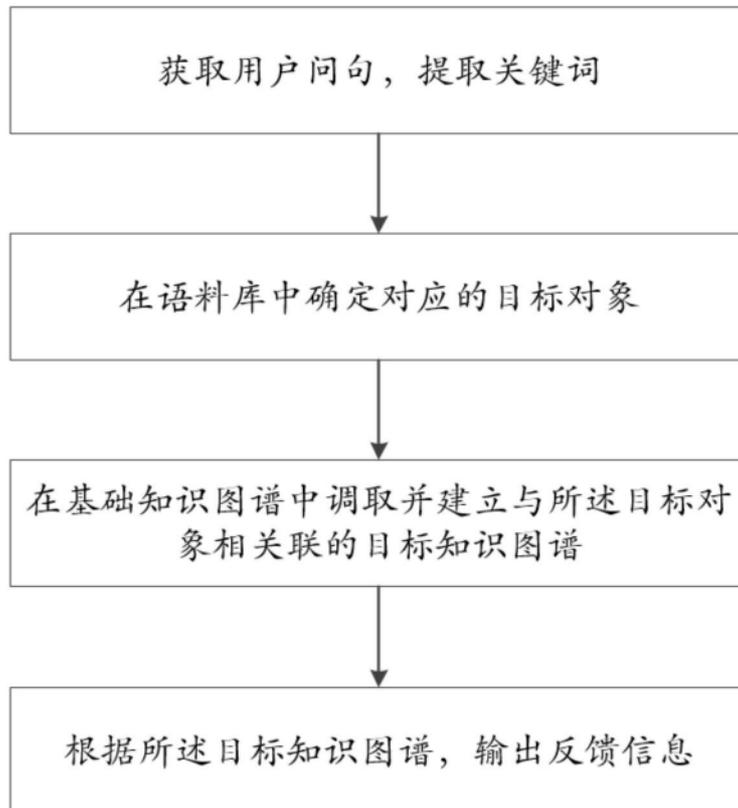


图1

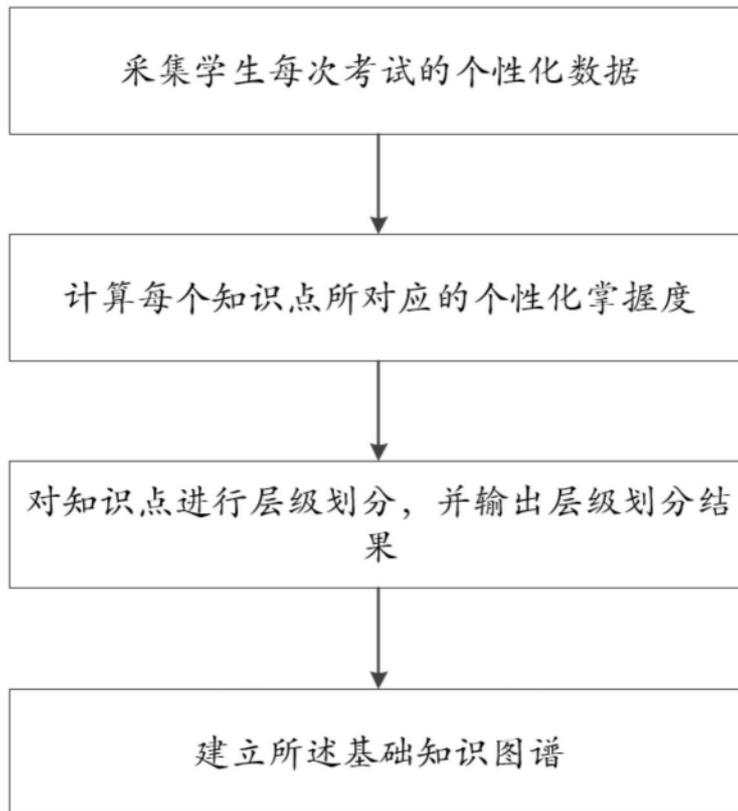


图2

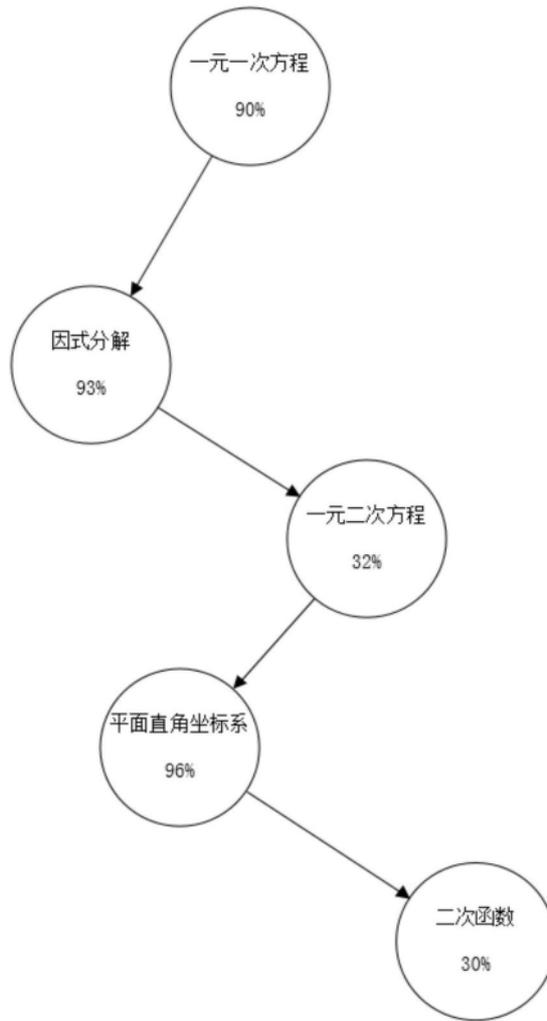


图3