(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 118245600 A
(43)申请公布日 2024.06.25

(21)申请号 202410316033.0
(22)申请日 2024.03.20
(71)申请人 佛山职业技术学院
地址 528000 广东省佛山市禅城区清水二街17号
(72)发明人 陈秀荣 陈月球 江雅琴 邓子念 梁茵
(74)专利代理机构 广东广盈专利商标事务所(普通合伙) 44339
专利代理师 李俊
(51)Int.Cl.
G06F 16/35(2019.01)
G06F 16/36(2019.01)
G06F 40/30(2020.01)
G06Q 50/26(2024.01)
G10L 15/26(2006.01)

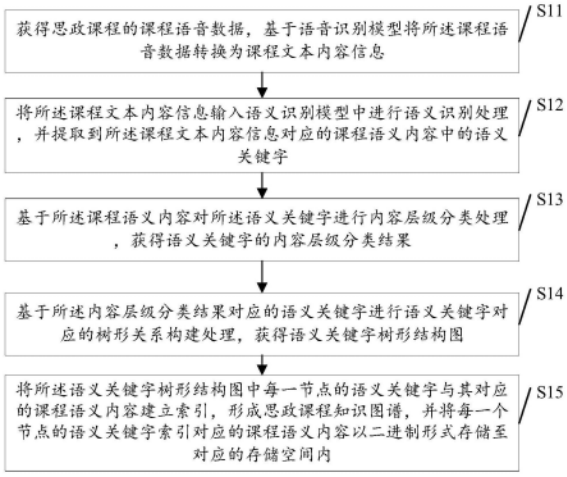
权利要求书3页 说明书10页 附图2页

(54)发明名称

一种基于数字化的思政课程知识图谱构建方法及相关装置

(57)摘要

本发明公开了一种基于数字化的思政课程知识图谱构建方法及相关装置,其中,所述方法包括:基于语音识别模型将所述课程语音数据转换为课程文本内容信息;将课程文本内容信息输入语义识别模型中进行语义识别处理,并提取到所述课程文本内容信息对应的课程语义内容中的语义关键字;对语义关键字进行内容层级分类处理,获得语义关键字的内容层级分类结果;进行语义关键字对应的树形关系构建处理;将语义关键字树形结构图中每一节点的语义关键字与其对应的课程语义内容建立索引,形成思政课程知识图谱,并存储至对应的存储空间内。在本发明实施例中,将思政课程内容构建成知识图谱,实现将课程内容的串联,提升思政课程的学习效率,提高学习效果。



CN 118245600 A

1.一种基于数字化的思政课程知识图谱构建方法,其特征在于,所述方法包括:

获得思政课程的课程语音数据,基于语音识别模型将所述课程语音数据转换为课程文本内容信息;

将所述课程文本内容信息输入语义识别模型中进行语义识别处理,并提取到所述课程文本内容信息对应的课程语义内容中的语义关键字;

基于所述课程语义内容对所述语义关键字进行内容层级分类处理,获得语义关键字的内容层级分类结果;

基于所述内容层级分类结果对应的语义关键字进行语义关键字对应的树形关系构建处理,获得语义关键字树形结构图;

将所述语义关键字树形结构图中每一节点的语义关键字与其对应的课程语义内容建立索引,形成思政课程知识图谱,并将每一个节点的语义关键字索引对应的课程语义内容以二进制形式存储至对应的存储空间内。

2.根据权利要求1所述的思政课程知识图谱构建方法,其特征在于,所述基于语音识别模型将所述课程语音数据转换为课程文本内容信息,包括:

基于所述课程语音数据进行课程中老师讲课时的环境音频信息提取提取处理,获得环境音频信息;

基于所述环境音频信息对所述课程语音数据进行降噪处理,获得降噪课程语音数据;

将所述降噪课程语音数据输入语音识别模型中,并在所述语义识别模型中将接收到的所述降噪课程语音数据转换为课程文本内容信息。

3.根据权利要求2所述的思政课程知识图谱构建方法,其特征在于,所述基于所述环境音频信息对所述课程语音数据进行降噪处理,获得降噪课程语音数据,包括:

提取所述环境音频信息中的环境音频信号特征,并基于提取到的环境音频信号特征生成降噪信号,所述环境音频信号特征包括环境音频信息中的相位数据和振幅数据,所述降噪信号的相位数据与所述环境音频信息中的相位数据相反,所述降噪信号的振幅数据与所述环境音频信息中振幅数据一致;

基于所述降噪信号对所述课程语音数据进行降噪处理,获得降噪课程语音数据。

4.根据权利要求1所述的思政课程知识图谱构建方法,其特征在于,所述将所述课程文本内容信息输入语义识别模型中进行语义识别处理,并提取到所述课程文本内容信息对应的课程语义内容中的语义关键字,包括:

将所述课程文本内容信息输入语义识别模型中进行语义识别处理,获得所述课程文本内容信息对应的初始课程语义,所述语义识别模型为利用思政课程中的课程文本内容以及对应的课程语义内容对语义表达提取网络训练收敛获得的模型;

将所述初始课程语义进行调整校对处理,获得所述课程文本内容对应的课程语义内容;

基于关键字提取模型对所述课程文本内容信息对应的课程语义内容进行关键字提取处理,获得所述课程文本内容信息对应的课程语义内容中的语义关键字。

5.根据权利要求1所述的思政课程知识图谱构建方法,其特征在于,所述基于所述课程语义内容对所述语义关键字进行内容层级分类处理,获得语义关键字的内容层级分类结果,包括:

获得所述语义关键字所对应的课程语义内容,并确认所述语义关键字所对应的课程语义内容在所述课程文本内容信息中的层级;

根据所述层级对所述语义关键字所对应的课程语义内容进行内容层级分类处理,获得语义关键字的内容层级分类结果。

6.根据权利要求1所述的思政课程知识图谱构建方法,其特征在于,所述基于所述内容层级分类结果对应的语义关键字进行语义关键字对应的树形关系构建处理,获得语义关键字树形结构图,包括:

根据所述内容层级分类结果对应的语义关键字确定树形结构中的根节点和其余各个节点的对应位置;

将所述内容层级分类结果对应的语义关键字填充至所述树形结构中的根节点和其余各个节点的对应位置中心,形成语义关键字树形结构图。

7.根据权利要求1所述的思政课程知识图谱构建方法,其特征在于,所述将所述关键字树形结构图中每一节点的关键字与其对应的课程语义内容建立索引,形成思政课程知识图谱,并将每一个节点的关键字索引对应的课程语义内容以二进制形式存储至对应的存储空间内,包括:

获得所述语义关键字树形结构图中每一个节点的语义关键字所对应的课程语义内容,并为每一个节点的语义关键字分配分配对应的第一元素和语义关键字对应的课程语义内容分配对应的第二元素;

基于对应的第一元素和对应的第二元素将每一节点的关键字与其对应的课程语义内容建立索引关联,形成思政课程知识图谱;

将一个节点的语义关键字索引对应的课程语义内容转化为二进制数据,并将所述二进制数据索引存储至对应的存储空间内。

8.一种基于数字化的思政课程知识图谱构建方法,其特征在于,所述方法包括:

转换模块:用于获得思政课程的课程语音数据,基于语音识别模型将所述课程语音数据转换为课程文本内容信息;

提取模块:用于将所述课程文本内容信息输入语义识别模型中进行语义识别处理,并提取到所述课程文本内容信息对应的课程语义内容中的语义关键字;

分类模块:用于基于所述课程语义内容对所述语义关键字进行内容层级分类处理,获得语义关键字的内容层级分类结果;

构建模块:用于基于所述内容层级分类结果对应的语义关键字进行语义关键字对应的树形关系构建处理,获得语义关键字树形结构图;

索引模块:用于将所述语义关键字树形结构图中每一节点的语义关键字与其对应的课程语义内容建立索引,形成思政课程知识图谱,并将每一个节点的语义关键字索引对应的课程语义内容以二进制形式存储至对应的存储空间内。

9.一种电子设备,包括处理器和存储器,其特征在于,所述处理器运行存储于所述存储器中的计算机程序或代码,实现如权利要求1至7中任一项所述的思政课程知识图谱构建方法。

10.一种计算机可读存储介质,用于存储计算机程序或代码,其特征在于,当所述计算机程序或代码被处理器执行时,实现如权利要求1至7中任一项所述的思政课程知识图谱构

建方法。

一种基于数字化的思政课程知识图谱构建方法及相关装置

技术领域

[0001] 本发明涉及教学辅助技术领域,尤其涉及一种基于数字化的思政课程知识图谱构建方法及相关装置。

背景技术

[0002] 在思政课程的讲课过程中,一般都是由老师根据相应的课程讲义内容进行讲课,但是由于课堂上,学生的接受能力是有限的,往往需要再课后进行复习巩固,但是学生在上课过程中无法准确的抓住课程的核心重点内容,无法精确的做好课程笔记,因此课后无法进行有效的学习巩固,导致思政课程的学习效率较低,无法有效的提升学习效果。

发明内容

[0003] 本发明的目的在于克服现有技术的不足,本发明提供了一种基于数字化的思政课程知识图谱构建方法及相关装置,可以将思政课程内容构建成知识图谱,实现将课程内容的串联,提升思政课程的学习效率,提高学习效果。

[0004] 为了解决上述技术问题,本发明实施例提供了一种基于数字化的思政课程知识图谱构建方法,所述方法包括:

[0005] 获得思政课程的课程语音数据,基于语音识别模型将所述课程语音数据转换为课程文本内容信息;

[0006] 将所述课程文本内容信息输入语义识别模型中进行语义识别处理,并提取到所述课程文本内容信息对应的课程语义内容中的语义关键字;基于所述课程语义内容对所述语义关键字进行内容层级分类处理,获得语义关键字的内容层级分类结果;

[0007] 基于所述内容层级分类结果对应的语义关键字进行语义关键字对应的树形关系构建处理,获得语义关键字树形结构图;

[0008] 将所述语义关键字树形结构图中每一节点的语义关键字与其对应的课程语义内容建立索引,形成思政课程知识图谱,并将每一个节点的语义关键字索引对应的课程语义内容以二进制形式存储至对应的存储空间内。

[0009] 可选的,所述基于语音识别模型将所述课程语音数据转换为课程文本内容信息,包括:

[0010] 基于所述课程语音数据进行课程中老师讲课时的环境音频信息提取提取处理,获得环境音频信息;

[0011] 基于所述环境音频信息对所述课程语音数据进行降噪处理,获得降噪课程语音数据;

[0012] 将所述降噪课程语音数据输入语音识别模型中,并在所述语义识别模型中将接收到的所述降噪课程语音数据转换为课程文本内容信息。

[0013] 可选的,所述基于所述环境音频信息对所述课程语音数据进行降噪处理,获得降噪课程语音数据,包括:

[0014] 提取所述环境音频信息中的环境音频信号特征,并基于提取到的环境音频信号特征生成降噪信号,所述环境音频信号特征包括环境音频信息中的相位数据和振幅数据,所述降噪信号的相位数据与所述环境音频信息中的相位数据相反,所述降噪信号的振幅数据与所述环境音频信息中振幅数据一致;

[0015] 基于所述降噪信号对所述课程语音数据进行降噪处理,获得降噪课程语音数据。

[0016] 可选的,所述将所述课程文本内容信息输入语义识别模型中进行语义识别处理,并提取到所述课程文本内容信息对应的课程语义内容中的语义关键字,包括:

[0017] 将所述课程文本内容信息输入语义识别模型中进行语义识别处理,获得所述课程文本内容信息对应的初始课程语义,所述语义识别模型为利用思政课程中的课程文本内容以及对应的课程语义内容对语义表达提取网络训练收敛获得的模型;

[0018] 将所述初始课程语义进行调整校对处理,获得所述课程文本内容对应的课程语义内容;

[0019] 基于关键字提取模型对所述课程文本内容对应的课程语义内容进行关键字提取处理,获得所述课程文本内容信息对应的课程语义内容中的语义关键字。

[0020] 可选的,所述基于所述课程语义内容对所述语义关键字进行内容层级分类处理,获得语义关键字的内容层级分类结果,包括:

[0021] 获得所述语义关键字所对应的课程语义内容,并确认所述语义关键字所对应的课程语义内容在所述课程文本内容信息中的层级;

[0022] 根据所述层级对所述语义关键字所对应的课程语义内容进行内容层级分类处理,获得语义关键字的内容层级分类结果。

[0023] 可选的,所述基于所述内容层级分类结果对应的语义关键字进行语义关键字对应的树形关系构建处理,获得语义关键字树形结构图,包括:

[0024] 根据所述内容层级分类结果对应的语义关键字确定树形结构中的根节点和其余各个节点的对应位置;

[0025] 将所述内容层级分类结果对应的语义关键字填充至所述树形结构中的根节点和其余各个节点的对应位置中心,形成语义关键字树形结构图。

[0026] 可选的,所述将所述关键字树形结构图中每一节点的关键字与其对应的课程语义内容建立索引,形成思政课程知识图谱,并将每一个节点的关键字索引对应的课程语义内容以二进制形式存储至对应的存储空间内,包括:

[0027] 获得所述语义关键字树形结构图中每一个节点的语义关键字所对应的课程语义内容,并为每一个节点的语义关键字分配分配对应的第一元素和语义关键字对应的课程语义内容分配对应的第二元素;

[0028] 基于对应的第一元素和对应的第二元素将每一节点的关键字与其对应的课程语义内容建立索引关联,,形成思政课程知识图谱;

[0029] 将一个节点的语义关键字索引对应的课程语义内容转化为二进制数据,并将所述二进制数据索引存储至对应的存储空间内。

[0030] 另外,本发明实施例还提供了一种基于数字化的思政课程知识图谱构建方法,所述方法包括:

[0031] 转换模块:用于获得思政课程的课程语音数据,基于语音识别模型将所述课程语

音数据转换为课程文本内容信息；

[0032] 提取模块：用于将所述课程文本内容信息输入语义识别模型中进行语义识别处理，并提取到所述课程文本内容信息对应的课程语义内容中的语义关键字；

[0033] 分类模块：用于基于所述课程语义内容对所述语义关键字进行内容层级分类处理，获得语义关键字的内容层级分类结果；

[0034] 构建模块：用于基于所述内容层级分类结果对应的语义关键字进行语义关键字对应的树形关系构建处理，获得语义关键字树形结构图；

[0035] 索引模块：用于将所述语义关键字树形结构图中每一节点的语义关键字与其对应的课程语义内容建立索引，形成思政课程知识图谱，并将每一个节点的语义关键字索引对应的课程语义内容以二进制形式存储至对应的存储空间内。

[0036] 另外，本发明实施例还提供了一种电子设备，包括处理器和存储器，所述处理器运行存储于所述存储器中的计算机程序或代码，实现如上述中任一项所述的思政课程知识图谱构建方法。

[0037] 另外，本发明实施例还提供了一种计算机可读存储介质，用于存储计算机程序或代码，当所述计算机程序或代码被处理器执行时，实现如上述中任一项所述的思政课程知识图谱构建方法。

[0038] 在本发明实施例中，通过将思政课程内容构建成知识图谱，实现将课程内容的串联，提升思政课程的学习效率，提高学习效果。

附图说明

[0039] 为了更清楚地说明本发明实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见的，下面描述中的附图仅仅是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其它的附图。

[0040] 图1是本发明实施例中的基于数字化的思政课程知识图谱构建方法的流程示意图；

[0041] 图2是本发明实施例中的基于数字化的思政课程知识图谱构建装置的结构组成示意图；

[0042] 图3是本发明实施例中的电子设备的结构组成示意图。

具体实施方式

[0043] 下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅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作出创造性劳动前提下所获得的所有其它实施例，都属于本发明保护的范围。

[0044] 实施例一，请参阅图1，图1是本发明实施例中的基于数字化的思政课程知识图谱构建方法的流程示意图。

[0045] 如图1所示，一种基于数字化的思政课程知识图谱构建方法，所述方法包括：

[0046] S11：获得思政课程的课程语音数据，基于语音识别模型将所述课程语音数据转换

为课程文本内容信息;

[0047] 在本发明具体实施过程中,所述基于语音识别模型将所述课程语音数据转换为课程文本内容信息,包括:基于所述课程语音数据进行课程中老师讲课时的环境音频信息提取提取处理,获得环境音频信息;基于所述环境音频信息对所述课程语音数据进行降噪处理,获得降噪课程语音数据;将所述降噪课程语音数据输入语音识别模型中,并在所述语义识别模型中将接收到的所述降噪课程语音数据转换为课程文本内容信息。

[0048] 进一步的,所述基于所述环境音频信息对所述课程语音数据进行降噪处理,获得降噪课程语音数据,包括:提取所述环境音频信息中的环境音频信号特征,并基于提取到的环境音频信号特征生成降噪信号,所述环境音频信号特征包括环境音频信息中的相位数据和振幅数据,所述降噪信号的相位数据与所述环境音频信息中的相位数据相反,所述降噪信号的振幅数据与所述环境音频信息中振幅数据一致;基于所述降噪信号对所述课程语音数据进行降噪处理,获得降噪课程语音数据。

[0049] 具体的,首先是需要获得思政课程的课程语音数据,一般情况下该课程语音数据由教师在思政课上进行录音获得或者从思政课程的课程视频中进行语音提取处理的方式获得课程语音数据;在获得课程语音数据之后,需要将课程语音数据转换为课程文本内容信息,这时候,需要利用语音识别模型实现,因此为了提高语音转换文字的识别准确率,需要对课程语音数据进行相应的降噪处理,即首先是根据课程语音数据进行课程中老师讲课时的环境音频信息提取提取处理,从而获得环境音频信息;具体是通过人工定义环境音频的音频特征,然后根据该环境音频的音频特征来对课程语音数据中的环境音频信息进行提取处理,即可得到环境音频信息;然后就需要根据环境音频信息来对课程语音数据进行降噪处理,获得降噪课程语音数据;最后则是将降噪课程语音数据输入语音识别模型中,并在语义识别模型中将接收到的降噪课程语音数据转换为课程文本内容信息;该语音识别模型可以为科大讯飞的语音识别模型,并且在使用语音识别模型转换为课程文本内容信息之后,还需要讲课老师对该课程文本内容信息内容进行相应的调整处理,使得最终的课程文本内容信息具有较高的准确性,使得后续构建的思政课程知识图谱的准确度得到相应的保证,有助于学生的学习巩固。

[0050] 在进行降噪处理时,则需要提取环境音频信息中的环境音频信号特征,其中环境音频信号特征为环境音频信息中的相位数据和振幅数据,然后根据环境音频信号特征来生成降噪信号,其中降噪信号的相位数据与环境音频信息中的相位数据相反,降噪信号的振幅数据与环境音频信息中振幅数据一致;然后将降噪信号与课程语音数据混合,实现对课程语音数据进行降噪处理,最终获得降噪课程语音数据。

[0051] S12:将所述课程文本内容信息输入语义识别模型中进行语义识别处理,并提取到所述课程文本内容信息对应的课程语义内容中的语义关键字;在本发明具体实施过程中,所述将所述课程文本内容信息输入语义识别模型中进行语义识别处理,并提取到所述课程文本内容信息对应的课程语义内容中的语义关键字,包括:将所述课程文本内容信息输入语义识别模型中进行语义识别处理,获得所述课程文本内容信息对应的初始课程语义,所述语义识别模型为利用思政课程中的课程文本内容以及对应的课程语义内容对语义表达提取网络训练收敛获得的模型;将所述初始课程语义进行调整校对处理,获得所述课程文本内容对应的课程语义内容;基于关键字提取模型对所述课程文本内容对应的课程语义内

容进行关键字提取处理,获得所述课程文本内容信息对应的课程语义内容中的语义关键字。

[0052] 具体的,首先是提取课程文本内容信息中的语义信息,这是需要将课程文本内容信息输入语义识别模型中进行语义识别处理,从而获得课程文本内容信息对应的初始课程语义;因为思政课程的语义识别和一般的文本语义识别具有差别,因此需要利用具体的语义识别数据来进行语义识别模型的训练,在此,该语义识别模型为语义表达提取网络模型,在训练时,所使用的训练数据为思政课程中的课程文本内容以及该课程文本内容对应的语义内容,在此需要准备不少于100份思政课程中的课程文本内容以及该课程文本内容对应的语义内容作为训练数据,并将这些训练数据输入语义表达提取网络训练,直至模型收敛时,则得到语义识别模型;在提取到课程文本内容信息对应的初始课程语义之后,则需要进行校对,校对的时候需要思政老师进行校对调整,即通过思政老师对初始课程语义进行调整校对处理,获得课程文本内容对应的课程语义内容;然后就是需要提取课程语义内容中的语义关键字,具体是通过关键字提取模型来对课程文本内容信息对应的课程语义内容进行关键字提取处理,获得课程文本内容信息对应的课程语义内容中的语义关键字。在本申请中,采用基于图排序算法的关键字提取模型,该模型对于单篇或者文档分割较少的文本进行关键字提取时,将更准确的提取到对应的关键字;在利用基于图排序算法的关键字提取模型进行课程语义内容的关键字提取时,首先是需要确定初始值,该初始值可以设置为非零值;在模型中的公式如下:

[0053] $$WS(V_i) = (1-d) + d\sum_{V_j \in in(V_i)} \frac{w_{ij}}{\sum_{V_K \in out(V_j)} w_{jk}} WS(W_j);$$

[0054] 其中,d表示阻尼系数,一般为0.85;V_i表示任意节点;$in(V_i)$表示指向顶点V_i的所有顶点集合;$out(V_j)$表示由顶点V_j连接出去的所有顶点集合;W_{ij}表示顶点V_i和V_j的连接权重;$WS(W_j)$表示顶点V_i的最终排序权重。

[0055] 具体步骤为:对课程语义内容进行切分为字或词的形式;将切分好的字或词构建图模型,即构建的一个字或词与字或词的连接矩阵,选择用滑动窗口的方式对每一个单词取邻居,并假设取一个长度为k的滑动窗口,则$w_1,w_2,\cdots,w_k$;$w_2,w_3,\cdots,w_{k+1}$;$w_3,w_4,\cdots,w_{k+2}$等等,在一个窗口中任意两个单词对应的节点之间村子一个无向无权的边,在邻居上构成图,并计算每一个单词节点的重要性;然后就是计算权重,即设定最大迭代次数,并一次进行迭代;按照连接矩阵,对每一个单词节点更新器排序权重;对连接到自身或连接为空的单词节点不进行计算;对于连接的其他单词节点,则按照模型公式进行逐步更新其权重;同时根据前后两次迭代之间单词的权重变化值来判断是否提前结束循环过程;同时上述算法即可提取到课程语义内容中的语义关键字。

[0056] S13:基于所述课程语义内容对所述语义关键字进行内容层级分类处理,获得语义关键字的内容层级分类结果;

[0057] 在本发明具体实施过程中,所述基于所述课程语义内容对所述语义关键字进行内容层级分类处理,获得语义关键字的内容层级分类结果,包括:获得所述语义关键字所对应的课程语义内容,并确认所述语义关键字所对应的课程语义内容在所述课程文本内容信息中的层级;根据所述层级对所述语义关键字所对应的课程语义内容进行内容层级分类处理,获得语义关键字的内容层级分类结果。

[0058] 具体的,需要获得语义关键字所对应的课程语义内容,并且通过思政课程的教师来确定语义关键字所对应的课程语义内容在课程文本内容信息中的层级;然后根据该语义关键字所对应的课程语义内容在课程文本内容信息中的层级来对语义关键字所对应的课程语义内容进行内容层级分类处理,获得语义关键字的内容层级分类结果。

[0059] S14:基于所述内容层级分类结果对应的语义关键字进行语义关键字对应的树形关系构建处理,获得语义关键字树形结构图;

[0060] 在本发明具体实施过程中,所述基于所述内容层级分类结果对应的语义关键字进行语义关键字对应的树形关系构建处理,获得语义关键字树形结构图,包括:根据所述内容层级分类结果对应的语义关键字确定树形结构中的根节点和其余各个节点的对应位置;将所述内容层级分类结果对应的语义关键字填充至所述树形结构中的根节点和其余各个节点的对应位置中心,形成语义关键字树形结构图。

[0061] 具体的,首先需要根据内容层级分类结果来确定每一个语义关键字在树形结构中的节点位置,即确定树形结构中根节点和其余节点所对应的语义关键字;然后将内容层级分类结果对应的语义关键字填充至树形结构中的根节点和其余各个节点的对应位置中心,形成语义关键字树形结构图。

[0062] S15:将所述语义关键字树形结构图中每一节点的语义关键字与其对应的课程语义内容建立索引,形成思政课程知识图谱,并将每一个节点的语义关键字索引对应的课程语义内容以二进制形式存储至对应的存储空间内。

[0063] 在本发明具体实施过程中,所述将所述关键字树形结构图中每一节点的关键字与其对应的课程语义内容建立索引,形成思政课程知识图谱,并将每一个节点的关键字索引对应的课程语义内容以二进制形式存储至对应的存储空间内,包括:获得所述语义关键字树形结构图中每一个节点的语义关键字所对应的课程语义内容,并为每一个节点的语义关键字分配分配对应的第一元素和语义关键字对应的课程语义内容分配对应的第二元素;基于对应的第一元素和对应的第二元素将每一节点的关键字与其对应的课程语义内容建立索引关联,形成思政课程知识图谱;将一个节点的语义关键字索引对应的课程语义内容转化为二进制数据,并将所述二进制数据索引存储至对应的存储空间内。

[0064] 具体的,需要获得语义关键字树形结构图中每一个节点的语义关键字所对应的课程语义内容,并且为每一个节点的语义关键字分配分配对应的第一元素和语义关键字对应的课程语义内容分配对应的第二元素;并且根据于对应的第一元素和对应的第二元素将每一节点的关键字与其对应的课程语义内容建立索引关联,形成思政课程知识图谱;在建立索引关联时,将第一元素与第二元素进行对应绑定的方式建立索引关联;最后将一个节点的语义关键字索引对应的课程语义内容转化为二进制数据,并将所述二进制数据索引存储至对应的存储空间内。

[0065] 在本发明实施例中,通过将思政课程内容构建成知识图谱,实现将课程内容的串联,提升思政课程的学习效率,提高学习效果。

[0066] 实施例二,请参阅图2,图2是本发明实施例中的基于数字化的思政课程知识图谱构建装置的结构组成示意图。

[0067] 如图2所示,一种基于数字化的思政课程知识图谱构建方法,所述方法包括:

[0068] 转换模块21:用于获得思政课程的课程语音数据,基于语音识别模型将所述课程

语音数据转换为课程文本内容信息;

[0069] 在本发明具体实施过程中,所述基于语音识别模型将所述课程语音数据转换为课程文本内容信息,包括:基于所述课程语音数据进行课程中老师讲课时的环境音频信息提取提取处理,获得环境音频信息;基于所述环境音频信息对所述课程语音数据进行降噪处理,获得降噪课程语音数据;将所述降噪课程语音数据输入语音识别模型中,并在所述语义识别模型中将接收到的所述降噪课程语音数据转换为课程文本内容信息。

[0070] 进一步的,所述基于所述环境音频信息对所述课程语音数据进行降噪处理,获得降噪课程语音数据,包括:提取所述环境音频信息中的环境音频信号特征,并基于提取到的环境音频信号特征生成降噪信号,所述环境音频信号特征包括环境音频信息中的相位数据和振幅数据,所述降噪信号的相位数据与所述环境音频信息中的相位数据相反,所述降噪信号的振幅数据与所述环境音频信息中振幅数据一致;基于所述降噪信号对所述课程语音数据进行降噪处理,获得降噪课程语音数据。

[0071] 具体的,首先是需要获得思政课程的课程语音数据,一般情况下该课程语音数据由教师在思政课上进行录音获得或者从思政课程的课程视频中进行语音提取处理的方式获得课程语音数据;在获得课程语音数据之后,需要将课程语音数据转换为课程文本内容信息,这时候,需要利用语音识别模型实现,因此为了提高语音转换文字的识别准确率,需要对课程语音数据进行相应的降噪处理,即首先是根据课程语音数据进行课程中老师讲课时的环境音频信息提取提取处理,从而获得环境音频信息;具体是通过人工定义环境音频的音频特征,然后根据该环境音频的音频特征来对课程语音数据中的环境音频信息进行提取处理,即可得到环境音频信息;然后就需要根据环境音频信息来对课程语音数据进行降噪处理,获得降噪课程语音数据;最后则是将降噪课程语音数据输入语音识别模型中,并在语义识别模型中将接收到的降噪课程语音数据转换为课程文本内容信息;该语音识别模型可以为科大讯飞的语音识别模型,并且在使用语音识别模型转换为课程文本内容信息之后,还需要讲课老师对该课程文本内容信息内容进行相应的调整处理,使得最终的课程文本内容信息具有较高的准确性,使得后续构建的思政课程知识图谱的准确度得到相应的保证,有助于学生的学习巩固。

[0072] 在进行降噪处理时,则需要提取环境音频信息中的环境音频信号特征,其中环境音频信号特征为环境音频信息中的相位数据和振幅数据,然后根据环境音频信号特征来生成降噪信号,其中降噪信号的相位数据与环境音频信息中的相位数据相反,降噪信号的振幅数据与环境音频信息中振幅数据一致;然后将降噪信号与课程语音数据混合,实现对课程语音数据进行降噪处理,最终获得降噪课程语音数据。

[0073] 提取模块22:用于将所述课程文本内容信息输入语义识别模型中进行语义识别处理,并提取到所述课程文本内容信息对应的课程语义内容中的语义关键字;

[0074] 在本发明具体实施过程中,所述将所述课程文本内容信息输入语义识别模型中进行语义识别处理,并提取到所述课程文本内容信息对应的课程语义内容中的语义关键字,包括:将所述课程文本内容信息输入语义识别模型中进行语义识别处理,获得所述课程文本内容信息对应的初始课程语义,所述语义识别模型为利用思政课程中的课程文本内容以及对应的课程语义内容对语义表达提取网络训练收敛获得的模型;将所述初始课程语义进行调整校对处理,获得所述课程文本内容对应的课程语义内容;基于关键字提取模型对所

述课程文本内容对应的课程语义内容进行关键字提取处理,获得所述课程文本内容信息对应的课程语义内容中的语义关键字。

[0075] 具体的,首先是提取课程文本内容信息中的语义信息,这是需要将课程文本内容信息输入语义识别模型中进行语义识别处理,从而获得课程文本内容信息对应的初始课程语义;因为思政课程的语义识别和一般的文本语义识别具有差别,因此需要利用具体的语义识别数据来进行语义识别模型的训练,在此,该语义识别模型为语义表达提取网络模型,在训练时,所使用的训练数据为思政课程中的课程文本内容以及该课程文本内容对应的语义内容,在此需要准备不少于100份思政课程中的课程文本内容以及该课程文本内容对应的语义内容作为训练数据,并将这些训练数据输入语义表达提取网络训练,直至模型收敛时,则得到语义识别模型;在提取到课程文本内容信息对应的初始课程语义之后,则需要进行校对,校对的时候需要思政老师进行校对调整,即通过思政老师对初始课程语义进行调整校对处理,获得课程文本内容对应的课程语义内容;然后就是需要提取课程语义内容中的语义关键字,具体是通过关键字提取模型来对课程文本内容信息对应的课程语义内容进行关键字提取处理,获得课程文本内容信息对应的课程语义内容中的语义关键字。在本申请中,采用基于图排序算法的关键字提取模型,该模型对于单篇或者文档分割较少的文本进行关键字提取时,将更准确的提取到对应的关键字;在利用基于图排序算法的关键字提取模型进行课程语义内容的关键字提取时,首先是需要确定初始值,该初始值可以设置为非零值;在模型中的公式如下:

[0076] $$WS(V_i) = (1-d) + d\sum_{V_j \in in(V_i)} \frac{w_{ij}}{\sum_{V_K \in out(V_j)} w_{jk}} WS(W_j);$$

[0077] 其中,d表示阻尼系数,一般为0.85;V_i表示任意节点;$in(V_i)$表示指向顶点V_i的所有顶点集合;$out(V_j)$表示由顶点V_j连接出去的所有顶点集合;W_{ij}表示顶点V_i和V_j的连接权重;$WS(W_j)$表示顶点V_i的最终排序权重。

[0078] 具体步骤为:对课程语义内容进行切分为字或词的形式;将切分好的字或词构建图模型,即构建的一个字或词与字或词的连接矩阵,选择用滑动窗口的方式对每一个单词取邻居,并假设取一个长度为k的滑动窗口,则$w_1,w_2,\cdots,w_k$;$w_2,w_3,\cdots,w_{k+1}$;$w_3,w_4,\cdots,w_{k+2}$等等,在一个窗口中任意两个单词对应的节点之间村子一个无向无权的边,在邻居上构成图,并计算每一个单词节点的重要性;然后就是计算权重,即设定最大迭代次数,并一次进行迭代;按照连接矩阵,对每一个单词节点更新器排序权重;对连接到自身或连接为空的单词节点不进行计算;对于连接的其他单词节点,则按照模型公式进行逐步更新其权重;同时根据前后两次迭代之间单词的权重变化值来判断是否提前结束循环过程;同时上述算法即可提取到课程语义内容中的语义关键字。

[0079] 分类模块23:用于基于所述课程语义内容对所述语义关键字进行内容层级分类处理,获得语义关键字的内容层级分类结果;

[0080] 在本发明具体实施过程中,所述基于所述课程语义内容对所述语义关键字进行内容层级分类处理,获得语义关键字的内容层级分类结果,包括:获得所述语义关键字所对应的课程语义内容,并确认所述语义关键字所对应的课程语义内容在所述课程文本内容信息中的层级;根据所述层级对所述语义关键字所对应的课程语义内容进行内容层级分类处理,获得语义关键字的内容层级分类结果。

[0081] 具体的,需要获得语义关键字所对应的课程语义内容,并且通过思政课程的教师来确定语义关键字所对应的课程语义内容在课程文本内容信息中的层级;然后根据该语义关键字所对应的课程语义内容在课程文本内容信息中的层级来对语义关键字所对应的课程语义内容进行内容层级分类处理,获得语义关键字的内容层级分类结果。

[0082] 构建模块24:用于基于所述内容层级分类结果对应的语义关键字进行语义关键字对应的树形关系构建处理,获得语义关键字树形结构图;在本发明具体实施过程中,所述基于所述内容层级分类结果对应的语义关键字进行语义关键字对应的树形关系构建处理,获得语义关键字树形结构图,包括:根据所述内容层级分类结果对应的语义关键字确定树形结构中的根节点和其余各个节点的对应位置;将所述内容层级分类结果对应的语义关键字填充至所述树形结构中的根节点和其余各个节点的对应位置中心,形成语义关键字树形结构图。

[0083] 具体的,首先需要根据内容层级分类结果来确定每一个语义关键字在树形结构中的节点位置,即确定树形结构中根节点和其余节点所对应的语义关键字;然后将内容层级分类结果对应的语义关键字填充至树形结构中的根节点和其余各个节点的对应位置中心,形成语义关键字树形结构图。

[0084] 索引模块25:用于将所述语义关键字树形结构图中每一节点的语义关键字与其对应的课程语义内容建立索引,形成思政课程知识图谱,并将每一个节点的语义关键字索引对应的课程语义内容以二进制形式存储至对应的存储空间内。

[0085] 在本发明具体实施过程中,所述将所述关键字树形结构图中每一节点的关键字与其对应的课程语义内容建立索引,形成思政课程知识图谱,并将每一个节点的关键字索引对应的课程语义内容以二进制形式存储至对应的存储空间内,包括:获得所述语义关键字树形结构图中每一个节点的语义关键字所对应的课程语义内容,并为每一个节点的语义关键字分配分配对应的第一元素和语义关键字对应的课程语义内容分配对应的第二元素;基于对应的第一元素和对应的第二元素将每一节点的关键字与其对应的课程语义内容建立索引关联,形成思政课程知识图谱;将一个节点的语义关键字索引对应的课程语义内容转化为二进制数据,并将所述二进制数据索引存储至对应的存储空间内。

[0086] 具体的,需要获得语义关键字树形结构图中每一个节点的语义关键字所对应的课程语义内容,并且为每一个节点的语义关键字分配分配对应的第一元素和语义关键字对应的课程语义内容分配对应的第二元素;并且根据于对应的第一元素和对应的第二元素将每一节点的关键字与其对应的课程语义内容建立索引关联,形成思政课程知识图谱;在建立索引关联时,将第一元素与第二元素进行对应绑定的方式建立索引关联;最后将一个节点的语义关键字索引对应的课程语义内容转化为二进制数据,并将所述二进制数据索引存储至对应的存储空间内。

[0087] 在本发明实施例中,通过将思政课程内容构建成知识图谱,实现将课程内容的串联,提升思政课程的学习效率,提高学习效果。

[0088] 本发明实施例提供的一种计算机可读存储介质,所述计算机可读存储介质上存储有计算机程序,该程序被处理器执行时实现上述实施例中任意一个实施例的思政课程知识图谱构建方法。其中,所述计算机可读存储介质包括但不限于任何类型的盘(包括软盘、硬盘、光盘、CD-ROM、和磁光盘)、ROM(Read-Only Memory,只读存储器)、RAM(Random

AcceSSMemory,随即存储器)、EPROM(EraSable Programmable Read-Only Memory,可擦写可编程只读存储器)、EEPROM(Electrically EraSableProgrammableRead-Only Memory,电可擦可编程只读存储器)、闪存、磁性卡片或光线卡片。也就是,存储设备包括由设备(例如,计算机、手机)以能够读的形式存储或传输信息的任何介质,可以是只读存储器,磁盘或光盘等。

[0089] 本发明实施例还提供了一种计算机应用程序,其运行在计算机上,该计算机应用程序用于执行上述中任意一个实施例的思政课程知识图谱构建方法。

[0090] 此外,图3是本发明实施例中的电子设备的结构组成示意图。

[0091] 本发明实施例还提供了一种电子设备,如图3所示。所述电子设备包括处理器302、存储器303、输入单元304以及显示单元305等器件。本领域技术人员可以理解,图3示出的电子设备结构器件并不构成对所有设备的限定,可以包括比图示更多或更少的部件,或者组合某些部件。存储器303可用于存储应用程序301以及各功能模块,处理器302运行存储在存储器303的应用程序301,从而执行设备的各种功能应用以及数据处理。存储器可以是内存储器或外存储器,或者包括内存储器和外存储器两者。内存储器可以包括只读存储器(ROM)、可编程ROM(PROM)、电可编程ROM(EPROM)、电可擦写可编程ROM(EEPROM)、快闪存储器、或者随机存储器。外存储器可以包括硬盘、软盘、ZIP盘、U盘、磁带等。本发明所公开的存储器包括但不限于这些类型的存储器。本发明所公开的存储器只作为例子而非作为限定。

[0092] 输入单元304用于接收信号的输入,以及接收用户输入的关键字。输入单元304可包括触控面板以及其它输入设备。触控面板可收集用户在其上或附近的触摸操作(比如用户使用手指、触笔等任何适合的物体或附件在触控面板上或在触控面板附近的操作),并根据预先设定的程序驱动相应的连接装置;其它输入设备可以包括但不限于物理键盘、功能键(比如播放控制按键、开关按键等)、轨迹球、鼠标、操作杆等中的一种或多种。显示单元305可用于显示用户输入的信息或提供给用户的信息以及终端设备的各种菜单。显示单元305可采用液晶显示器、有机发光二极管等形式。处理器302是终端设备的控制中心,利用各种接口和线路连接整个设备的各个部分,通过运行或执行存储在存储器303内的软件程序和/或模块,以及调用存储在存储器内的数据,执行各种功能和处理数据。

[0093] 作为一个实施例,所述电子设备包括:一个或多个处理器302,存储器303,一个或多个应用程序301,其中所述一个或多个应用程序301被存储在存储器303中并被配置为由所述一个或多个处理器302执行,所述一个或多个应用程序301配置用于执行上述实施例中的任意一实施例中对的思政课程知识图谱构建方法。

[0094] 在本发明实施例中,通过将思政课程内容构建成知识图谱,实现将课程内容的串联,提升思政课程的学习效率,提高学习效果。

[0095] 另外,以上对本发明实施例所提供的一种基于数字化的思政课程知识图谱构建方法及相关装置进行了详细介绍,本文中应采用了具体个例对本发明的原理及实施方式进行了阐述,以上实施例的说明只是用于帮助理解本发明的方法及其核心思想;同时,对于本领域的一般技术人员,依据本发明的思想,在具体实施方式及应用范围上均会有改变之处,综上所述,本说明书内容不应理解为对本发明的限制。

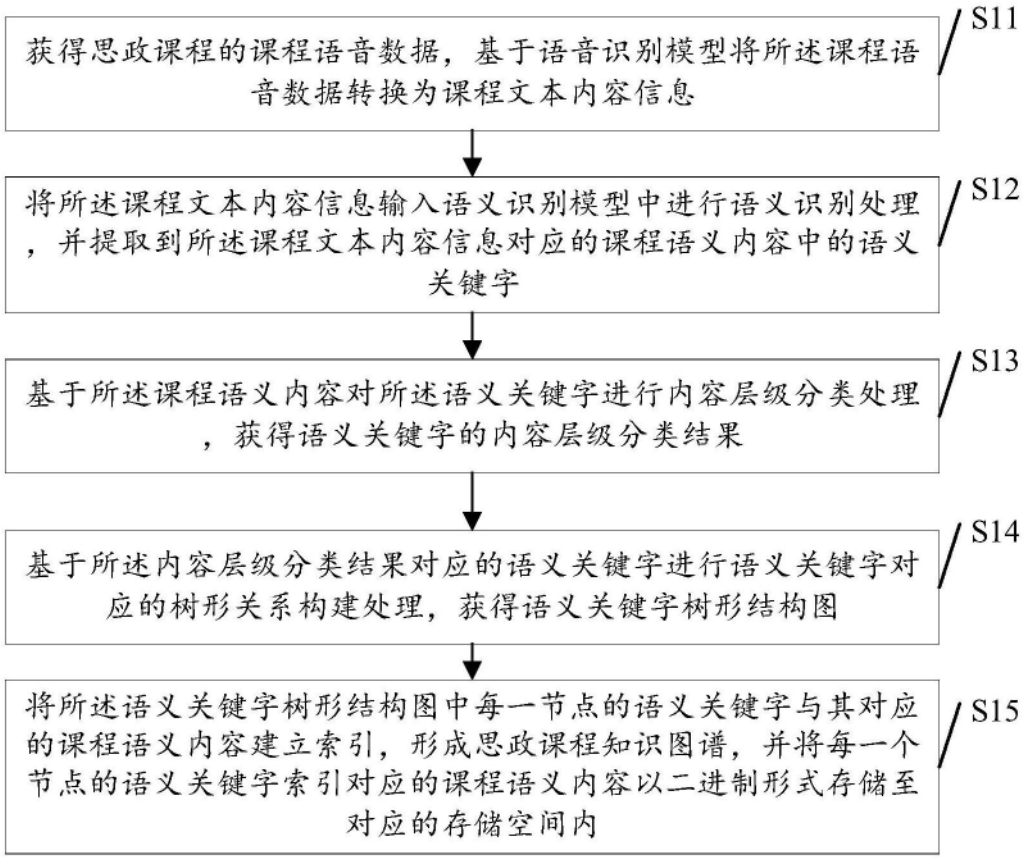

获得思政课程的课程语音数据，基于语音识别模型将所述课程语音数据转换为课程文本内容信息
S11
将所述课程文本内容信息输入语义识别模型中进行语义识别处理，并提取到所述课程文本内容信息对应的课程语义内容中的语义关键字
S12
基于所述课程语义内容对所述语义关键字进行内容层级分类处理，获得语义关键字的内容层级分类结果
S13
基于所述内容层级分类结果对应的语义关键字进行语义关键字对应的树形关系构建处理，获得语义关键字树形结构图
S14
将所述语义关键字树形结构图中每一节点的语义关键字与其对应的课程语义内容建立索引，形成思政课程知识图谱，并将每一个节点的语义关键字索引对应的课程语义内容以二进制形式存储至对应的存储空间内
S15

图1

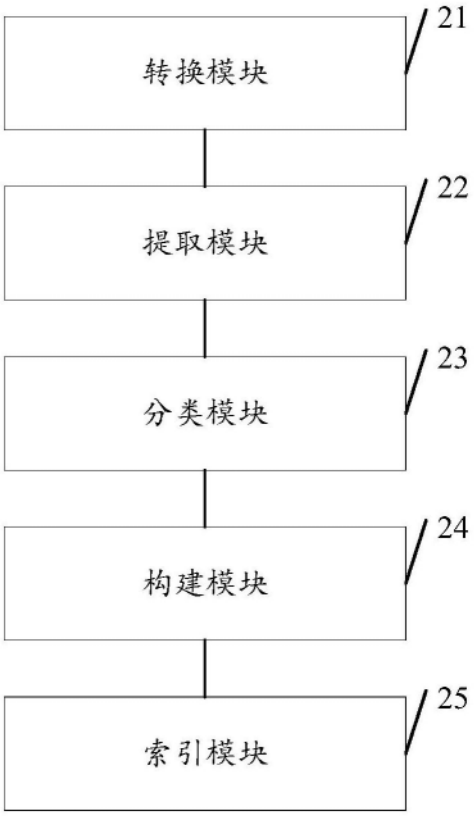

转换模块
21
提取模块
22
分类模块
23
构建模块
24
索引模块
25

图2

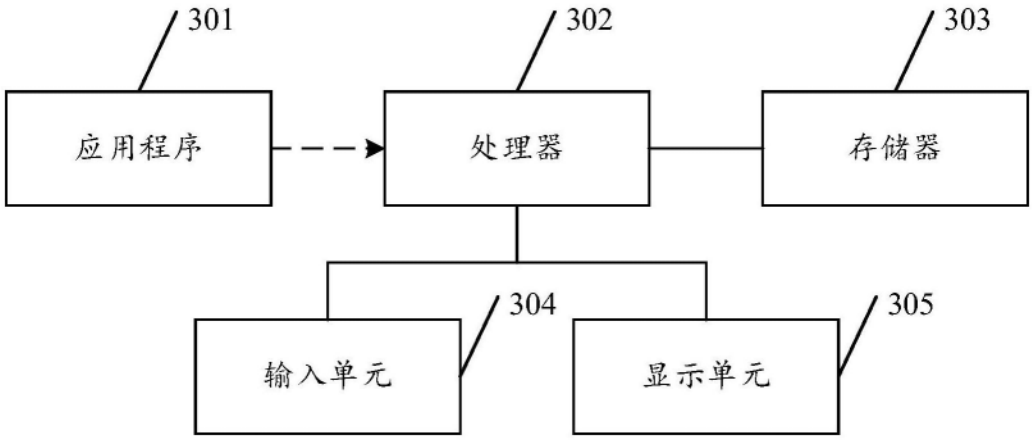

301
应用程序
302
处理器
303
存储器
304
输入单元
305
显示单元

图3