



## (12) 发明专利申请

(10) 申请公布号 CN 112598202 A

(43) 申请公布日 2021.04.02

(21) 申请号 202110236671.8

(22) 申请日 2021.03.03

(71) 申请人 北京世纪好未来教育科技有限公司

地址 100144 北京市海淀区中关村大街32  
号蓝天和盛大厦1702-03室

(72) 发明人 岳祥 王凯夫 孙研 彭守业

(74) 专利代理机构 上海知锦知识产权代理事务  
所(特殊普通合伙) 31327

代理人 王立娜

(51) Int.Cl.

G06Q 10/04 (2012.01)

G06Q 50/20 (2012.01)

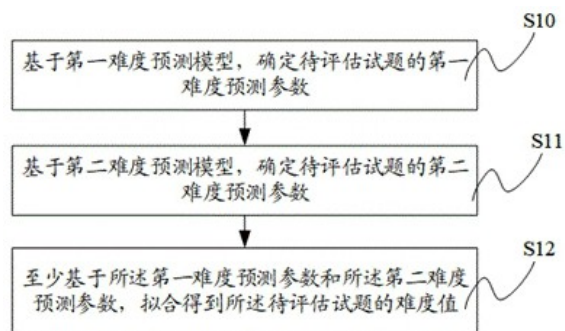
权利要求书3页 说明书17页 附图3页

### (54) 发明名称

试题难度评估方法、装置及存储介质、计算设备

### (57) 摘要

本发明实施例提供了一种试题难度评估方法、装置及存储介质、计算设备,所述方法包括:基于第一难度预测模型,确定待评估试题的第一难度预测参数,所述第一难度预测模型用于在第一群体特征参数下,预测待评估试题的第一难度预测参数;基于第二难度预测模型,确定待评估试题的第二难度预测参数,所述第二难度预测模型用于在第二群体特征参数下,预测待评估试题的第二难度预测参数,所述第二群体特征参数不同于所述第一群体特征参数;至少基于所述第一难度预测参数和所述第二难度预测参数,拟合得到所述待评估试题的难度值,从而提高了试题难度评估的准确度。



1. 一种试题难度评估方法,其特征在于,包括:

基于第一难度预测模型,确定待评估试题的第一难度预测参数,所述第一难度预测模型用于在第一群体特征参数下,预测待评估试题的第一难度预测参数;

基于第二难度预测模型,确定待评估试题的第二难度预测参数,所述第二难度预测模型用于在第二群体特征参数下,预测待评估试题的第二难度预测参数,所述第二群体特征参数不同于所述第一群体特征参数;

至少基于所述第一难度预测参数和所述第二难度预测参数,拟合得到所述待评估试题的难度值。

2. 如权利要求1所述的试题难度评估方法,其特征在于,所述至少基于所述第一难度预测参数和所述第二难度预测参数,拟合得到所述待评估试题的难度值,包括:

确定所述第一难度预测参数和所述第二难度预测参数对应的拟合参数;

基于所述拟合参数,计算得到所述待评估试题的难度值。

3. 如权利要求2所述的试题难度评估方法,其特征在于,采用线性回归、多项式回归或多层神经网络回归确定所述第一难度预测参数和所述第二难度预测参数对应的拟合参数。

4. 如权利要求3所述的试题难度评估方法,其特征在于,采用线性回归的网格搜索方式,确定所述第一难度预测参数和所述第二难度预测参数对应的拟合参数。

5. 如权利要求4所述的试题难度评估方法,其特征在于,所述确定所述第一难度预测参数和所述第二难度预测参数对应的拟合参数,包括:

构建试题样本集,所述试题样本集内的试题至少配置有与之对应的第一难度预测参数和第二难度预测参数;

基于所述试题样本集,寻找所述第一难度预测参数和所述第二难度预测参数对应的最优拟合参数;

以所述第一难度预测参数和所述第二难度预测参数对应的最优拟合参数作为所述第一难度预测参数和所述第二难度预测参数对应的拟合参数。

6. 如权利要求5所述的试题难度评估方法,其特征在于,所述构建试题样本集,包括:

构建试题样本组,所述试题样本组包括至少2个候选试题,所述候选试题至少配置有与之对应的第一难度预测参数和第二难度预测参数;

获取所述试题样本组内,候选试题的相对难度参数;其中,所述相对难度参数用于指示所述试题样本组内,各候选试题的难度排序;

根据所述候选试题的相对难度参数,确定候选试题的第一难度预测参数和第二难度预测参数与所述相对难度参数的相关度和匹配度;

基于所述相关度和匹配度,筛选所述候选试题,得到试题样本集。

7. 如权利要求1所述的试题难度评估方法,其特征在于,所述第二难度预测模型基于所述第一难度预测模型训练得到,所述第二难度预测模型的训练步骤包括:

获取包括第二难度参数的至少一个第二样本数据集;

基于所述第二样本数据集,训练所述第一难度预测模型,得到所述第二难度预测模型。

8. 如权利要求7所述的试题难度评估方法,其特征在于,所述获取包括第二难度参数的至少一个第二样本数据集,包括:

获取至少一个初始数据集,所述初始数据集包括多个试题和与试题相对应的多个作答

数据;

基于试题反应理论模型,确定各所述初始数据集中,各试题的第二难度参数,得到所述第二样本数据集。

9.如权利要求8所述的试题难度评估方法,其特征在于,所述获取至少一个初始数据集之后,所述基于试题反应理论模型,确定各所述初始数据集中,各试题的第二难度参数之前,还包括:

基于测量学指标,筛选所述初始数据集;其中,所述测量学指标包括,初始数据集的样本数量、样本信息缺失率、样本通过率,以及,试题属性参数中的一个或多个;

所述基于试题反应理论模型,确定各所述初始数据集中,各试题的第二难度参数,具体为,基于试题反应理论模型,确定筛选后的所述初始数据集中,各试题的第二难度参数,得到所述第二样本数据集。

10.如权利要求7所述的试题难度评估方法,其特征在于,所述基于所述第二样本数据集,训练所述第一难度预测模型,得到所述第二难度预测模型,包括:

获取所述第二样本数据集中,各试题的特征数据和第二难度参数;所述试题的特征数据包括:试题文本特征、试题公式特征、试题属性标签、试题知识点标签文本和知识点属性中的一个或多个;

根据所述试题的特征数据,拼接所述试题的句子对;

以所述句子对为输入数据,以所述试题的第二难度参数为输出数据,训练所述第一难度预测模型。

11.如权利要求1所述的试题难度评估方法,其特征在于,所述第一难度预测模型基于第一样本数据训练得到,所述第一难度预测模型的训练步骤包括:

获取包括第一难度参数的第一样本数据,所述第一难度参数为基于所述第一群体获取的试题难度数据;

基于所述第一样本数据,训练初始难度预测模型,得到所述第一难度预测模型。

12.如权利要求11所述的试题难度评估方法,其特征在于,所述获取包括第一难度参数的第一样本数据,包括:

获取离散的第一难度初始参数,所述第一难度初始参数基于试题难度的星级评分得到;

归一化所述第一难度初始参数,得到第一样本数据的第一难度参数。

13.如权利要求12所述的试题难度评估方法,其特征在于,所述第一样本数据还包括试题的文本、标签和知识点属性,所述基于所述第一样本数据,训练初始难度预测模型,具体为:

根据所述第一样本数据,确定试题的特征数据,所述试题的特征数据包括:试题文本特征、试题公式特征、试题属性标签、试题知识点标签文本和知识点属性中的一个或多个;

根据所述试题的特征数据,拼接所述试题的句子对;

以所述句子对为输入数据,以所述第一样本数据的第一难度参数为输出数据,训练所述初始难度预测模型。

14.如权利要求3所述的试题难度评估方法,其特征在于,采用多项式回归的网格搜索方式,确定所述第一难度预测参数和所述第二难度预测参数对应的拟合参数。

15. 如权利要求3所述的试题难度评估方法,其特征在于,采用多层神经网络回归的梯度下降方式,确定所述第一难度预测参数和所述第二难度预测参数对应的拟合参数。

16. 一种试题难度评估装置,其特征在于,包括:

第一预测参数确定模块,适于基于第一难度预测模型,确定待评估试题的第一难度预测参数,所述第一难度预测模型用于在第一群体特征参数下,预测待评估试题的第一难度预测参数;

第二预测参数确定模块,适于基于第二难度预测模型,确定待评估试题的第二难度预测参数,所述第二难度预测模型用于在第二群体特征参数下,预测待评估试题的第二难度预测参数,所述第二群体特征参数不同于所述第一群体特征参数;

拟合模块,适于至少基于所述第一难度预测参数和所述第二难度预测参数,拟合得到所述待评估试题的难度值。

17. 一种存储介质,其特征在于,所述存储介质存储有适于试题难度评估的程序,以实现如权利要求1-15任一项所述的试题难度评估方法。

18. 一种计算设备,其特征在于,包括至少一个存储器和至少一个处理器;所述存储器存储有程序,所述处理器调用所述程序,以执行如权利要求1-15任一项所述的试题难度评估方法。

## 试题难度评估方法、装置及存储介质、计算设备

### 技术领域

[0001] 本发明实施例涉及计算机领域,尤其涉及一种试题难度评估方法、装置及存储介质、计算设备。

### 背景技术

[0002] 近年来,互联网对教育和学习过程的影响越来越大,由此推演的学习过程线上化对底层内容资源的要求也不断提高。试题作为底层内容资源中的核心板块,是进行学生个性化资源推荐,实现学生自适应学习和自适应测评的基础。试题题库(Item bank)可以基于试题的各属性值,如难度值等,构建得到,以支持各种场景下试题的有效使用。

[0003] 试题的难度值用于表征试题的难度,以使教学及考试实践中,对试题的难度进行把控,以准确评测学生对试题所反映的知识点的掌控程度。评估试题难度可以基于难度预测模型对试题的难度值进行评估,但目前所用的难度评估方法都存在准确度低的问题。

[0004] 因此,如何提供一种试题难度评估方法,以实现提高试题的难度值的准确度,成为亟需解决的技术问题。

### 发明内容

[0005] 本发明实施例提供一种试题难度评估方法、装置及存储介质、计算设备,以提高试题难度评估的准确度。

[0006] 为解决上述问题,本发明实施例提供一种试题难度评估方法,包括:

基于第一难度预测模型,确定待评估试题的第一难度预测参数,所述第一难度预测模型用于在第一群体特征参数下,预测待评估试题的第一难度预测参数;

基于第二难度预测模型,确定待评估试题的第二难度预测参数,所述第二难度预测模型用于在第二群体特征参数下,预测待评估试题的第二难度预测参数,所述第二群体特征参数不同于所述第一群体特征参数;

至少基于所述第一难度预测参数和所述第二难度预测参数,拟合得到所述待评估试题的难度值。

[0007] 相应的,本发明实施例还提供一种试题难度评估装置,包括:

第一预测参数确定模块,适于基于第一难度预测模型,确定待评估试题的第一难度预测参数,所述第一难度预测模型用于在第一群体特征参数下,预测待评估试题的第一难度预测参数;

第二预测参数确定模块,适于基于第二难度预测模型,确定待评估试题的第二难度预测参数,所述第二难度预测模型用于在第二群体特征参数下,预测待评估试题的第二难度预测参数,所述第二群体特征参数不同于所述第一群体特征参数;

拟合模块,适于至少基于所述第一难度预测参数和所述第二难度预测参数,拟合得到所述待评估试题的难度值。

[0008] 相应的,本发明实施例还提供一种存储介质,包括:所述存储介质存储有适于试题

难度评估的程序,以实现上述试题难度评估方法。

[0009] 相应的,本发明实施例还提供一种计算设备,包括至少一个存储器和至少一个处理器;所述存储器存储有程序,所述处理器调用所述程序,以执行上述试题难度评估方法。

[0010] 与现有技术相比,本发明的技术方案具有以下优点:

本发明实施例所提供的试题难度评估方法、装置及存储介质、计算设备,所述方法包括:基于第一难度预测模型,确定待评估试题的第一难度预测参数,所述第一难度预测模型用于在第一群体特征参数下,预测待评估试题的第一难度预测参数;基于第二难度预测模型,确定待评估试题的第二难度预测参数,所述第二难度预测模型用于在第二群体特征参数下,预测待评估试题的第二难度预测参数,所述第二群体特征参数不同于所述第一群体特征参数;至少基于所述第一难度预测参数和所述第二难度预测参数,拟合得到所述待评估试题的难度值。

[0011] 可以看出,本发明实施例结合第一群体特征参数的第一难度预测模型确定的第一难度预测参数和第二群体特征参数的第二难度预测模型确定的第二难度预测参数,拟合得到待评估试题的难度值,从而可以基于不同群体(例如,学生群体和教师群体)的特征进行难度值的评估,避免了单一群体差异过大可能造成的难度值不准确,使评估得到的试题难度值更贴近试题的实际难度,提高难度值的准确度。

[0012] 在本发明的可选实施例中,还进一步基于第一难度预测模型训练得到第二难度预测模型。由于第一难度预测模型和第二难度预测模型之间具有较大的相关性,采用第一难度预测模型训练得到第二难度预测模型,能够实现模型参数的更快的拟合,从而训练的速度更快,训练效果更好。

## 附图说明

[0013] 图1是本发明实施例所提供的试题难度评估方法的一流程示意图;  
图2是本发明实施例提供的第一难度预测模型的训练步骤的可选流程示意图;  
图3是本发明实施例提供的训练初始难度预测模型的可选流程示意图;  
图4是本发明实施例提供第二难度预测模型的训练步骤的可选流程示意图;  
图5是本发明实施例提供训练得到第二难度预测模型的可选流程图;  
图6是本发明实施例提供确定拟合参数的可选流程示意图;  
图7是本发明实施例提供试题难度评估装置示意图;  
图8是本发明实施例提供的设备的一种可选硬件设备架构图。

## 具体实施方式

[0014] 如背景技术所述,现有的试题难度评估方法的准确度有待提高。

[0015] 具体的,对试题的难度值的预测,可以基于试题本身的特征信息(如题面特征、解析特征、答案特征、所涉及的知识点难易系数等)训练后的预测模型预测获得,也可以基于使用待分析试题的学生群体的群体特征和该群体对试题的解题特征信息训练后的预测模型预测获得。

[0016] 然而,基于试题的特征信息训练的预测模型,没有考虑用户的主观感受,而基于学生群体的特征信息训练的预测模型预测,由于学生群体的解题能力差异过大,使得的其基

于不同的样本训练得到的模型差异较大,进而得到的难度值的准确度不高。

[0017] 基于此,本发明实施例提供了一种试题难度评估方法,包括:基于第一难度预测模型,确定待评估试题的第一难度预测参数,所述第一难度预测模型用于在第一群体特征参数下,预测待评估试题的第一难度预测参数;基于第二难度预测模型,确定待评估试题的第二难度预测参数,所述第二难度预测模型用于在第二群体特征参数下,预测待评估试题的第二难度预测参数,所述第二群体特征参数不同于所述第一群体特征参数;至少基于所述第一难度预测参数和所述第二难度预测参数,拟合得到所述待评估试题的难度值。

[0018] 可以看出,本发明实施例结合第一群体特征参数的第一难度预测模型确定的第一难度预测参数和第二群体特征参数的第二难度预测模型确定的第二难度预测参数,拟合得到待评估试题的难度值,从而可以基于不同群体(例如,学生群体和教师群体)的特征进行难度值的评估,避免了单一群体差异过大可能造成的难度值不准确,使评估得到的试题难度值更贴近试题的实际难度,提高难度值的准确度。

[0019] 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。

[0020] 请参考图1,图1是本发明实施例所提供的试题难度评估方法的一流程示意图。

[0021] 如图中所示,本发明实施例所提供的试题难度评估方法,实现所述待评估试题的难度评估,包括以下步骤:

步骤S10:基于第一难度预测模型,确定待评估试题的第一难度预测参数;

所述第一难度预测模型用于在第一群体特征参数下,预测待评估试题的第一难度预测参数。为实现对待评估试题的难度值的评估,首先可以基于第一群体特征参数下的第一难度预测模型,进行待评估试题的第一难度预测参数的确定。

[0022] 其中,第一群体,可以为学生,也可以为教师,或者,还可以为相关专业的特定人群。第一群体特征参数,指的是基于第一群体的样本数据训练得到的第一难度预测模型的特征参数,在第一群体特征参数下,第一难度预测模型可以基于第一群体的角度,预测待评估试题的第一难度参数,即,第一难度预测参数。

[0023] 在一个可选的示例中,第一群体可以为教师,基于教师对试题知识点的理解更加透彻,从教师的角度预测试题的难度,可以更加贴合实际的试题难度。第一难度预测模型,可以基于教师给出的试题难度数据作为样本数据进行训练得到,从而可以从教师的角度,预测待评估试题的第一难度预测参数。

[0024] 步骤S11:基于第二难度预测模型,确定待评估试题的第二难度预测参数;

所述第二难度预测模型用于在第二群体特征参数下,预测待评估试题的第二难度预测参数,所述第二群体特征参数与所述第一群体特征参数不同;

为提高评估难度值的准确度,在第一难度预测模型确定第一难度预测参数之后,进一步基于第二难度预测模型,确定待评估试题的第二难度预测参数,从而可以结合第一难度预测参数和第二难度预测参数进行多维度的评估,提高难度值的准确度。

[0025] 其中,第二群体,可以为学生,也可以为教师,或者,还可以为相关专业的特定人群。第二群体特征参数,指的是基于第二群体的样本数据训练得到的第二难度预测模型的

特征参数,在第二群体特征参数下,第二难度预测模型可以基于第二群体的角度,预测待评估试题的第二难度预测参数。所述第二群体不同于所述第一群体,相应的,所述第二群体特征参数不同于所述第一群体特征参数。

[0026] 在一个可选的示例中,第二群体可以为学生,基于学生的在解题过程中的特征数据,可以更直观的确定试题的难度。相应的,第二难度预测模型,可以基于学生解题过程中的特征数据作为样本数据进行训练得到,从而可以从学生的角度,预测待评估试题的第二难度预测参数。

[0027] 步骤S12:至少基于所述第一难度预测参数和所述第二难度预测参数,拟合得到所述待评估试题的难度值;

通过拟合的方式,可以将所述第一难度预测参数和所述第二难度预测参数结合,从而从多个维度评估试题的难度值。用于拟合的参数,可以仅为第一难度预测参数和第二难度预测参数,也可以结合其他维度得到的表征难度的参数,同时进行拟合。

[0028] 其中,所述第一难度预测参数和所述第二难度预测参数可以基于不同的拟合参数实现拟合。其中,不同的难度评估参数,对应有不同的拟合参数,例如,第一难度预测参数可以对应第一拟合参数,第二难度预测参数可以对应第二拟合参数。拟合参数可以基于一定的数据计算得到,也可以在已计算得到拟合参数的前提下,通过直接获取的方式得到。拟合参数可以是固定的,也可以是定期更新的,本发明实施例在此不做具体的限定。

[0029] 在拟合得到所述待评估试题的难度值的过程中,可以首先确定所述第一难度预测参数和所述第二难度预测参数对应的拟合参数。具体的,在拟合参数未知的前提下,可以计算得到拟合参数,而在拟合参数已知的前提下,可以直接获取所述拟合参数。

[0030] 可以理解的是,拟合得到所述待评估试题的难度值的过程,可以为线性拟合,也可以为非线性拟合。其中,在线性拟合时,对应的拟合参数可以通过线性回归或多项式回归的方式确定,在非线性拟合时,对应的拟合参数可以通过多项式回归或多层神经网络回归的方式确定。

[0031] 在选取线性拟合的方式进行第一难度预测参数和第二难度预测参数的拟合时,具体可以采用线性回归的网格搜索方式,或,多项式回归的网格搜索方式,确定所述第一难度预测参数和所述第二难度预测参数对应的拟合参数,其中,在线性拟合的方式中,拟合参数用于指示各参数在最终得到的难度值中的权重,拟合参数的值越大,对应的权重越大;在选取非线性拟合的方式进行第一难度预测参数和第二难度预测参数的拟合时,具体可以采用多层神经网络回归的梯度下降方式,确定所述第一难度预测参数和所述第二难度预测参数对应的拟合参数。

[0032] 在确定所述拟合参数后,可以基于所述拟合参数,计算得到所述待评估试题的难度值。

[0033] 可以看出,本发明实施例结合第一群体特征参数的第一难度预测模型确定的第一难度预测参数和第二群体特征参数的第二难度预测模型确定的第二难度预测参数,拟合得到待评估试题的难度值,从而可以基于不同群体(例如,学生群体和教师群体)的特征进行难度值的评估,避免了单一群体差异过大可能造成的难度值不准确,使评估得到的试题难度值更贴近试题的实际难度,提高难度值的准确度。

[0034] 可以理解的是,不同的难度预测模型可以基于不同的样本数据训练得到,下面,针



对不同难度预测模型,说明难度预测模型的训练步骤。

[0035] 具体的,第一难度预测模型可以基于第一样本数据训练得到。参考图2所示的第一难度预测模型的训练步骤的可选流程示意图,第一难度预测模型的训练步骤包括:

步骤S20:获取包括第一难度参数的第一样本数据;

在本示例中,以第一群体为教师为例进行说明,对应的,第一难度预测模型基于教师维度的试题难度数据作为样本数据进行训练得到。其中,所述第一样本数据包括用于训练第一难度预测模型的试题数据,例如,可以包括试题的基础数据和对应试题的难度数据,在一个可选的示例中,所述试题的基础数据可以包括试题的文本、标签和知识点属性等,试题的难度数据可以包括第一难度参数,所述第一难度参数可以为基于所述第一群体获取的试题难度数据。

[0036] 其中,第一难度参数可以理解为第一群体对试题的难度的评估,第一难度预测参数为模型对试题的第一难度参数的预测。基于第一难度参数进行模型的训练得到的第一难度预测模型,可以实现第一难度参数的预测

可以理解的是,对于试题难度的评估,通常采用星级评分的方式得到。例如,难度可以在1~5星之间选择,其中5星为最难,1星为最容易,以便于评估者直观便捷的给出相应的评分。可选的,在试题难度数据基于星级评分得到时,所述第一难度参数可以通过归一化步骤,将离散的参数转化为连续的参数,从而可以加速数据的优化过程,提高数据的处理速度。

[0037] 具体的,以第一难度初始参数作为第一群体基于试题难度的星级评分得到的试题难度数据,本步骤中可以具体包括:

步骤S201:获取离散的第一难度初始参数,所述第一难度初始参数基于试题难度的星级评分得到;

可以理解的是,所述离散的第一难度初始参数,为对应试题的星级评分,例如3星。

[0038] 步骤S202:归一化所述第一难度初始参数,得到第一样本数据的第一难度参数。

[0039] 具体的,归一后的难度值= (归一前的难度值-1.0) / (5.0-1.0);从而将第一难度初始参数转化为0~1之间的数据。其中,数值1表示最难,数值0表示最容易。

[0040] 接着,执行步骤S21:基于所述第一样本数据,训练初始难度预测模型,得到所述第一难度预测模型;

在本示例中,未经训练的模型称为初始难度预测模型,经第一样本数据训练后的模型称为第一难度预测模型。

[0041] 在一个可选的示例中,初始难度预测模型可以为BERT(Bidirectional Encoder Representations from Transformers)模型,BERT是一种语言表示模型,用于在通过联合调节所有层中的左右上下文来预训练深度双向表示。

[0042] 在进行初始难度预测模型训练时,可以首先对第一样本数据进行相应的处理,以使第一样本数据适应初始难度预测模型的数据形式。在初始难度预测模型为BERT模型时,则可以使第一样本数据形成句子对,从而实现对BERT模型的训练。

[0043] 参考图3所示的训练初始难度预测模型的可选流程示意图,在步骤S21中,训练初始难度预测模型的过程可以具体包括:

步骤S211:根据所述第一样本数据,确定试题的特征数据,所述试题的特征数据包

括：试题文本特征、试题公式特征、试题属性标签、试题知识点标签文本和知识点属性中的一个或多个；

具体的，可以通过对数据进行抽取的方式，确定试题的特征数据。具体的，试题的特征数据中，各特征的含义如下：

试题文本特征：包括试题的题干文本信息、选项文本信息、解析文本信息、答案文本信息，该试题文本特征可以从试题的文本中获取；

试题公式特征：包括试题的题干文本、选项文本、解析文本和答案文本中的公式文本信息，该试题公式特征可以从试题的文本中获取；

试题属性标签：试题的可得属性，包括但不限于题干长度、解析长度、答案长度、逻辑题型、书面题型、主知识点标签个数、副知识点标签个数、试题主副知识点覆盖章数（去重）、试题主副知识点覆盖节数（去重）等，其中，试题主副知识点覆盖章数（去重）指的是经过去重处理后的试题主副知识点覆盖章数；试题主副知识点覆盖节数（去重）指的是经过去重处理后的试题主副知识点覆盖节数，其中，该试题属性标签可以从试题的标签中获取。

[0044] 试题的知识点标签：包括主知识点标签文本、副知识点标签文本；其中，该试题的知识点标签可以从试题的标签中获取。

[0045] 知识点属性：可以包括近三年年均主知识点分数占比总和、近三年年均主副知识点分数占比总和、近三年年均主知识点平均难度等级、近三年年均主副知识点的加权难度等级；其中，该知识点属性可以从试题的知识点属性中获取。

[0046] 可以理解的是，上述特征数据用于描述对试题难度具有显著贡献的特征。通过确定上述试题特征，以训练得到试题特征与第一难度参数之间的关联，从而实现第一难度参数的预测。

[0047] 步骤S212：根据所述试题的特征数据，拼接所述试题的句子对；

在确定试题的特征数据后，可以拼接所述试题的句子对。

[0048] 在具体的示例中，所述初始难度预测模型所输入的字段可以为：试题ID、学科名称、年部名称、逻辑题型、书面题型、题干内容、答案、解析、选项（多个选项用；zk；拼接）、主知识点个数、主知识点（多个主知识点用；zk；拼接）、副知识点个数、副知识点（多个副知识点用；zk；拼接）、试题主知识点覆盖章数（去重）、试题副知识点覆盖章数（去重）、试题主知识点覆盖节数（去重）、试题副知识点覆盖节数（去重）、主知识点考频、副知识点考频、主知识点考察难度、副知识点考察难度。

[0049] 相应的，可以在对上述字段进行清洗处理、符号归一化和公式符号归一化后，拼接为句子对。其中，清洗处理可以包括去除html标签，全角转半角；符号归一化可以包括中文冒号转英文冒号，中文问号转英文问号等；公式符号归一化可以将同义的多个符号转化成预设的标准符号，例如 $\leq$ 、 $\leqslant$ 、 $\leqslant$ 、 $\leq$ 都表示小于，可转化为 $\leq$ 。

[0050] 在初始难度预测模型为BERT模型时，参考图4所示的BERT模型可选示意图，输入可以由CLS、句子1（SENT 1）、SEP、句子2（SENT 2）、PAD组成。

[0051] 具体的，输入的各部分的组成元素具体如下：

CLS：BERT模型中与输入token（标志）对应的隐藏表示，根据CLS对应的隐藏表示可以进行分类。

[0052] SENT 1：由逻辑题型，书面题型，主知识点个数，主知识点，副知识点个数，副知识

点,题干,选项等拼接而成。

[0053] SENT 2: 由答案,解析等拼接而成。

[0054] SEP: 用于对句子1和句子2进行分割。

[0055] PAD: 在试题句子对长度不足标准长度时,用于补充试题句子对,以使所有试题的句子对均为标准长度。

[0056] 步骤S213:以所述句子对为输入数据,以所述第一样本数据的第一难度参数为输出数据,训练所述初始难度预测模型;

在拼接得到试题的句子对后,可以以所述第一样本数据的第一难度参数为输出数据,训练得到第一难度预测模型。

[0057] 基于第一样本数据,可以训练得到第一难度预测模型,从而可以基于所述第一难度预测模型,预测得到待评估试题的第一难度预测参数,从而确定待评估试题的第一难度预测参数。

[0058] 可以理解的是,第二难度预测模型可以基于初始难度预测模型训练得到,也可以基于其他难度预测模型训练得到。基于第一难度预测模型和第二难度预测模型之间具有较大的相关性,在本发明实施例中,第二难度预测模型可以基于第一难度预测模型训练得到,从而可以使得采用第一难度预测模型训练得到第二难度预测模型,能够实现模型参数的更快的拟合,训练速度更快,训练效果更好。

[0059] 具体的,参考图4所示的第二难度预测模型的训练步骤的可选流程示意图,第二难度预测模型的训练步骤包括:

步骤S30:获取包括第二难度参数的至少一个第二样本数据集;

在本示例中,以第二群体为学生为例进行说明,对应的,第二难度预测模型基于学生维度的试题难度数据作为样本数据进行训练得到。

[0060] 其中,基于学生个体差异较大,在本示例中,可以基于能够反映较多学生信息的样本数据集进行训练。可选的,用于训练第二难度预测模型的可以为第二样本数据集。在第二样本数据集中,一个试题可以对应多个学生作答该试题的数据信息,从而更全面的反映试题的难度。

[0061] 可选的,一第二样本数据集可以为一个班的学生进行一次测验所产生的数据,该数据可以包括试题的基础数据和试题的难度数据;其中,试题的基础数据可以包括试题的文本、标签、知识点属性,以及试题的作答数据等,试题的难度数据可以包括第二难度参数,所述第二难度参数可以为基于所述第二群体获取的试题难度数据。

[0062] 其中,第二难度参数可以描述第二群体对试题的难度的反映,第二难度预测参数为模型对试题的第二难度参数的预测。基于第二难度参数进行模型的训练得到的第二难度预测模型,可以实现第二难度参数的预测。

[0063] 可以理解的是,在学生进行测验时,通常以班、年级或学校为单位,基于同一套试题(下称试题集)进行,从而在一第二样本数据集中,包括有多个学生对一试题集的作答数据。

[0064] 在第二难度参数的确定过程中,基于第二样本数据集中统计有学生群体的作答数据,例如,作答时间,作答准确率等,可以从统计学的角度反映试题的难度。

[0065] 在本发明的一个可选示例中,可以基于试题反应理论(item response theory,

IRT)模型,确定第二难度参数。具体的,本步骤可以包括:

步骤S301:获取至少一个初始数据集,所述初始数据集包括多个试题和与试题相对应的多个作答数据;

其中,所述初始数据集可以包括多个试题和与试题相对应的多个作答数据。在本示例中,所述初始数据集可以理解为一个测验数据集,具体可以包括试题集(多个试题)和与试题对应的作答数据。其中,所述作答数据可以包括学生答题的作答时间,作答内容,以及作答准确率等。

[0066] 可以理解的是,基于试题,可以确定试题的特征数据,该特征数据的具体描述可以参考前述说明,此处不再赘述。

[0067] 基于初始数据集,可以进一步确定该初始数据集中,各试题的第二难度参数,进而得到第二样本数据集。

[0068] 在一个可选的示例中,在步骤S301后,在进行第二难度参数确定前,还可以进一步对初始数据集进行筛选,以去除影响统计测量的特殊数据集或特殊数据,从而优化模型的训练过程。

[0069] 具体的,可以执行步骤S302:基于测量学指标,筛选所述初始数据集;

其中,所述测量学指标可以包括,初始数据集的样本数量、样本信息缺失率、样本通过率,以及,试题属性参数中的一个或多个;

具体的,样本数量可以体现初始数据集的统计量,统计量过小可能无法体现需要统计的数据特征,因此,可以筛选样本数量大于或等于预设值的初始数据集。其中,样本数量可以表示一个初始数据集中学生的数量。

[0070] 可以理解的是,在进行测验时,试题数量通常也包括多个,在一些可选的示例中,还可以进行试题数量的筛选。

[0071] 样本信息缺失率,用于体现初始数据集中,样本的缺失状况,例如,学生缺考或作答数据不完整等,在样本信息缺失率达到预设值时,则需要删除该学生,并更新样本的数量值。

[0072] 样本通过率,用于体现初始数据集中,样本的综合准确率达到预设值的比例,例如,作答数据与试题答案匹配,则认为作答数据准确,一个样本中的综合准确率可以理解为,作答数据准确的试题的权值(例如试题所占的分值)所占试题总和权值(例如试题的总分)的比例,作答数据的综合准确率达到预设值(可以理解为及格线),即认为作答数据通过,样本数据中,作答数据通过的比例,即为样本通过率。

[0073] 试题属性参数,具体可以包括一致性系数(例如克伦巴赫alpha系数),与试题总和权值相关度等。

[0074] 另外,在进行初始数据集的筛选时,还可以进一步删除初始数据集中的特殊样本,例如,所有试题的作答数据均准确的样本、所有试题的作答数据均错误的样本、所有样本中作答数据均准确的试题、所有样本中作答数据均错误的试题等,从而避免特殊数据对模型的影响。

[0075] 基于初始数据集的筛选过程中,会同时对初始数据集中的样本进行相应的处理,使得初始数据集中的样本数量,使得样本的数量值和试题的数量值可能发生变化,因此,可以将样本数量和试题数量的筛选作为筛选的最后步骤执行,或者,在样本的数量值和试题

的数量值可能发生变化的步骤后均执行样本数量和试题数量的筛选。

[0076] 在一个可选的示例中,可在数据库中进行筛选,其中,以数据库中的数据集为初始数据集,具体的,一个初始数据集为一个测验对应的数据,一个样本作答数据为一个学生的作答数据,具体筛选步骤如下:

步骤S1.筛选目标学段、目标学科的初始数据集;

步骤S2.确定所选初始数据集具有训练模型所需的特征或属性数据;例如,包括试题集(多个试题)和与试题对应的作答数据,具体可以包括初始数据集的样本数量、样本信息缺失率、样本通过率,以及,试题属性参数等。

[0077] 步骤S3.筛选学生数量 $\geq 50$ 、试题数量 $\geq 10$ 的所有测验;

步骤S4.样本信息缺失率筛选:每个测验中,对应试题的学生作答数据的缺失率 $\geq 30\%$ 应删除该试题,对应学生作答数据的试题缺失 $\geq 30\%$ 应删除该学生。若删除部分数据,循环执行步骤S3和S4,直至样本信息缺失率满足筛选条件;

步骤S5.样本通过率筛选:测验中,所有试题的通过率(0~1之间)应从低到高排序,若所有试题都高于预设值(例如90%的试题都 $\geq 0.80$ ),删除该测验;

步骤S6.一致性系数筛选:测验的Cronbach's alpha一致性系数高于或等于预设值(例如0.7或0.6),在测验的Cronbach's alpha一致性系数低于预设值时,删除该测验;

步骤S7.与试题总和权值相关度筛选:计算试题与该测验的总分相关在0.3以上的试题数量,若50%试题 $< 0.2$ 左右,删除该测验。

[0078] 步骤S8.特殊样本筛选:删除在所有试题上作答正确的学生、删除在所有试题上作答错误的学生、删除所有学生均作答正确的试题、删除所有学生均作答错误的试题,并循环执行步骤S3和S8,直至特殊样本筛选完成。

[0079] 通过上述筛选步骤,可以得到有利于模型训练的初始数据集,从而使模型训练的过程中,模型各参数能够迅速优化拟合,降低模型的训练时间。

[0080] 步骤S302:基于试题反应理论模型,确定各所述初始数据集中,各试题的第二难度参数,得到所述第二样本数据集。

[0081] 其中,试题反应理论(Item Response Theory, IRT)模型,是一系列心理统计学模型的总称,是用来分析考试成绩或者问卷调查数据的数学模型,目标在于确定潜在的心理特征(latent trait)是否可以通过测试题被反应出来,以及测试题和被测试者之间的互动关系。基于该模型,可以较为客观的确定试题的难度数据(下称IRT难度参数),在本示例中,以该难度数据作为第二难度参数。

[0082] 在本示例中,以试题的基础数据和所述试题的作答数据作为试题反应理论模型的输入,试题的难度数据作为试题反应理论模型的输出。

[0083] 具体在本示例中,以数据集(例如测验)为单位,分别估计每个初始数据集内试题的IRT难度参数。具体操作可以包括:首先构建每个测验的学生作答矩阵,然后基于恰当的IRT模型(如Rasch模型),以及恰当的估计方法(如极大似然估计法MMLE)进行同时估计,得到学生参数(即学生的能力值)和试题的参数(Rasch模型可得到试题IRT难度参数)。

[0084] 可以理解的是,在进行了初始数据集的筛选处理后,本步骤具体为:基于试题反应理论模型,确定筛选后的所述初始数据集中,各试题的第二难度参数,得到所述第二样本数据集。

[0085] 需要说明的是,在本示例中,基于Rasch模型得到的试题IRT难度参数,为均值为0,标准差为1的高斯分布。为便于后续模型的训练,本步骤中还可以进一步对Rasch模型得到的试题IRT难度参数进行归一化,使之变为0~1之间的参数,其中,数值1表示最难,数值0表示最容易。

[0086] 具体的,归一化的公式为:

$$b\_transform = \{b - [\max(b) + \min(b)]/2 + |\max(b) - \min(b)|/2\} / [\max(b) - \min(b)] * (Upper - Lower) + Lower$$

其中,b\_transform为归一化之后的IRT难度参数,b为原始IRT难度参数,max(b)为b的最大取值,min(b)为b的最小取值,Upper为归一化后的最大取值,Lower为归一化后的最小取值。本示例中,Upper为1,Lower为0。

[0087] 相应的,本示例中所述第二样本数据集中的第二难度参数为归一化后的难度数据。

[0088] 基于第二样本数据集,可以实现对第一难度预测模型的训练,从而得到第二难度预测模型。

[0089] 步骤S31:基于所述第二样本数据集,训练所述第一难度预测模型,得到所述第二难度预测模型

基于第一难度预测模型训练第二难度预测模型,基于第一难度预测模型和第二难度预测模型之间具有较大的相关性,采用第一难度预测模型训练得到第二难度预测模型,能够实现模型参数的更快的拟合,从而训练的速度更快,训练效果更好。

[0090] 具体的,以第一难度预测模型为基础进行训练,并以第二样本数据集训练后的模型称为第二难度预测模型。

[0091] 在进行模型训练时,可以首先对第二样本数据集进行相应的处理,以使第二样本数据集适应模型的数据形式。在以第一难度预测模型为基础进行训练时,则可以使第二样本数据集中的试题数据形成句子对,从而实现对模型的训练。

[0092] 具体的,参考图5所示的训练得到第二难度预测模型的可选流程图,执行步骤S31的过程可以包括:

步骤S311:获取所述第二样本数据集中,各试题的特征数据和第二难度参数;所述试题的特征数据包括:试题文本特征、试题公式特征、试题属性标签、试题知识点标签文本和知识点属性中的一个或多个;

具体的,可以通过对数据集进行抽取的方式,确定试题的特征数据和第二难度参数。

[0093] 其中,特征数据中各参数的含义可以参考前述示例,此处不再赘述。

[0094] 步骤S312:根据所述试题的特征数据,拼接所述试题的句子对;

具体的,步骤S312可以参考步骤S212的描述,此处不再赘述。

[0095] 步骤S313:以所述句子对为输入数据,以所述试题的第二难度参数为输出数据,训练所述初始难度预测模型;

具体的,步骤S313可以参考步骤S213的描述,此处不再赘述。

[0096] 在本示例中,第二难度预测模型基于第一难度预测模型训练得到,由于第一难度预测模型和第二难度预测模型之间具有较大的相关性,采用第一难度预测模型训练得到第

二难度预测模型,能够实现模型参数的更快的拟合,从而训练的速度更快,训练效果更好。

[0097] 具体的,参考表1所示,在本发明实施例中,与基于初始难度预测模型训练得到第二难度预测模型(第一训练方式)相比,基于第一难度预测模型训练得到第二难度预测模型(第二训练方式),可以将模型训练时间降低15%~25%,包括端点值,准确度上可以提高1%~3%,包括端点值。

[0098] 表1:

|      | 第一训练方式 | 第二训练方式 | 差值   |
|------|--------|--------|------|
| 训练时间 | 2.0h   | 1.6h   | 0.4h |
| 准确度  | 72%    | 73%    | 1%   |

其中,需要说明的是,在上述迭代过程中,第一训练方式需要迭代5轮,而第二训练方式则仅迭代4轮,从而降低了训练的时间。

[0099] 在确定试题的第一难度预测参数和第二难度预测参数后,可以进行参数的拟合,从而确定待评估试题的难度值。

[0100] 需要说明的是,在本发明实施例中,用于拟合的难度数据可以仅包括第一难度预测参数和第二难度预测参数,也可以包括其他方式获得的难度数据,例如第一难度参数和第二难度参数,此处并不做具体的限制。

[0101] 具体的,难度值的拟合公式可以如下:

$$\text{que\_diff} = A * \text{star\_diff} + B * \text{star\_predict\_diff} + C * \text{irt\_diff} + D * \text{irt\_predict\_diff}$$

其中A、B、C、D为拟合参数,满足 $A + B + C + D = 1$ 。que\_diff表示拟合后的试题的难度值,star\_diff表示试题的第一难度参数,star\_predict\_diff表示试题的第一难度预测参数,irt\_diff表示试题的第二难度参数,irt\_predict\_diff表示试题的第二难度预测参数。其中,在用于拟合的难度数据仅包括第一难度预测参数和第二难度预测参数时,相应的A和C可以为0。

[0102] 可以理解的是,拟合过程中,确定拟合参数,才能计算所述难度值。基于此,在本示例中,进一步提供了确定所述第一难度预测参数和所述第二难度预测参数对应的拟合参数的可选流程,参考图6所示的确定拟合参数的可选流程示意图,包括:

步骤S40:构建试题样本集,所述试题样本集内的试题至少配置有与之对应的第一难度预测参数和第二难度预测参数;

其中,配置与试题对应的第一难度预测参数和第二难度预测参数,才能实现对第一难度预测参数和第二难度预测参数的拟合参数的确定。

[0103] 可以理解的是,基于前述拟合公式,所述试题样本集内的试题可以配置有第一难度参数和第二难度参数,也可以不包括第一难度参数和第二难度参数。

[0104] 需要说明的是,用于确定拟合参数的试题样本集,原则上应该选择具有精确的难度值的试题,以寻找最优拟合参数。然而,基于试题的绝对难度难以确定,本示例中,基于试题的相对难度与第一难度预测参数和第二难度预测参数的关联程度,确定试题样本集。

[0105] 具体的,所述构建试题样本集的步骤如下:

步骤S401:构建试题样本组,所述试题样本组包括至少2个候选试题,所述候选试题至少配置有与之对应的第一难度预测参数和第二难度预测参数;

其中,候选试题可以为经过第一难度预测模型和第二难度预测模型预测后的试题,或者,还可以为用于训练模型的样本数据中的试题,并经过第一难度预测模型和第二难度预测模型预测后,可以得到第一难度预测参数和第二难度预测参数。

[0106] 试题样本组中包括至少2个候选试题,以使试题样本组内的候选试题可以得到相对于试题样本组内其他候选试题的难度数据,从而可以通过候选试题的相对难度确定其与候选试题的第一难度预测参数和第二难度预测参数的关联程度。

[0107] 步骤S402:获取所述试题样本组内,候选试题的相对难度参数;

其中,所述相对难度参数用于指示所述试题样本组内,各候选试题的难度排序,例如,试题样本组内,候选试题A难度>候选试题B难度>候选试题C难度。

[0108] 所述相对难度参数可以基于特定群体或特定人员给出的相对难度数据获取。例如,可以通过A、B两教师为试题难度进行排序,并在A、B两教师排序一致性大于或等于0.8时,即认为该排序有效。

[0109] 具体的,可以选择30组三题题组,10组五题题组,交由A、B两教师为试题难度进行排序,获取相应的排序数据后,将原三题题组、五题题组分别拆分为两题题组,即每个三题题组拆分为3组试题对、每个五题题组拆分为10组试题对,以每对试题难度对比为结果,AB结果一致的比例大于或等于0.8时,即认为该排序有效,进而,AB结果一致时,则取该结果,AB结果不一致,则可以通过第三方进行确定。

[0110] 其中,需要说明的是,针对同一候选试题集,可以基于不同的分组方式多次循环执行步骤S401和步骤S402,从而更加准确的判断候选试题的相关难度。

[0111] 步骤S403:根据所述候选试题的相对难度参数,确定候选试题的第一难度预测参数和第二难度预测参数与所述相对难度参数的相关度和匹配度;

其中,相关度可以通过相关矩阵确定,匹配度可以通过各参数的排序确定。

[0112] 具体的,在候选试题同时具有第一难度参数和第二难度参数时,基于第二难度参数、第一难度预测参数和第二难度预测参数与所述相对难度参数的相关矩阵确定候选试题内各参数的相关度。

[0113] 例如,可以确定候选试题中各参数(第一难度参数、第二难度参数、第一难度预测参数和第二难度预测参数)与相对难度参数之间的Pearson相关系数,相关系数越大,表明对应的参数越准确。

[0114] 在一个可选的示例中,基于第一难度参数基于教师给出的星级评分得到,所述相对难度参数可以基于第一难度参数确定,进一步的,在确定候选试题中各参数与相对难度参数之间的Pearson相关系数时,可以确定第二难度参数、第一难度预测参数和第二难度预测参数与相对难度参数之间的相关系数,基于该相关系数确定参数的相关度。

[0115] 在候选试题同时具有第一难度参数和第二难度参数时,基于第一难度参数、第二难度参数、第一难度预测参数和第二难度预测参数与所述相对难度参数进行匹配,例如,在相对难度参数表明候选试题A难度>候选试题B难度时,可以确定候选试题A中的第一难度参数是否大于候选试题B的第一难度参数,候选试题A中的第二难度参数是否大于候选试题B的第二难度参数,候选试题A中的第一难度预测参数是否大于候选试题B的第一难度预测参数,候选试题A中的第二难度预测参数是否大于候选试题B的第二难度预测参数,如果结果为是,则相应参数与相对难度参数匹配,如果结果为否,则相应参数与相对难度参数不匹



配。

[0116] 步骤S404:基于所述相关度和匹配度,筛选所述候选试题,得到试题样本集;

具体的,可以选择相关系数大于预设值的候选试题作为试题样本集中的试题,同样的,可以选择匹配度高的候选试题作为试题样本集中的试题,以提升拟合参数的精确度。

[0117] 可选的,为提高拟合参数的精确度,可以选择所有参数均匹配的候选试题作为试题样本集中的试题进行拟合参数的确定。在一些可选的示例中,还可以选择仅部分参数匹配的候选试题作为试题样本集中的试题进行拟合参数的确定。

[0118] 需要说明的是,基于可选示例中的第一难度参数为基于试题难度的星级评分得到,该种方式难易区分难度差别较小的试题,从而不可避免的出现用于比对的候选试题的第一难度参数相等的情况,在本示例中,可以将这种情况认为与相对难度参数不匹配,从而更加严格的筛选候选试题。

[0119] 经过相关度和匹配度的筛选后得到的候选试题的集合,可以作为试题样本集。

[0120] 步骤S41:基于所述试题样本集,寻找所述第一难度预测参数和所述第二难度预测参数对应的最优拟合参数;

其中,可以采用网格搜索方式寻找所述第一难度预测参数和所述第二难度预测参数对应的最优拟合参数,网格搜索方式是指定参数值的一种穷举搜索方式,通过将估计函数的参数通过交叉验证的方法进行优化来得到最优的学习算法。

[0121] 通过网格搜索方式,可以得到第一难度预测参数和所述第二难度预测参数的最优拟合参数。

[0122] 可以理解的是,在试题同时具有第一难度参数和第二难度参数时,还可以基于网格搜索方式同时确定四个难度数据的最优拟合参数。

[0123] 步骤S42:以所述第一难度预测参数和所述第二难度预测参数对应的最优拟合参数作为所述第一难度预测参数和所述第二难度预测参数对应的拟合参数;

在确定最后拟合参数后,可以以所述第一难度预测参数和所述第二难度预测参数对应最优拟合参数作为所述第一难度预测参数和所述第二难度预测参数对应的拟合参数,进而可以基于该拟合参数,进行试题难度值的拟合确定。

[0124] 其中,在试题同时具有第一难度参数和第二难度参数时,以基于网格搜索方式同时确定的四个难度数据的最优拟合参数作为对应难度数据的拟合参数。

[0125] 可以看出,本发明实施例结合第一群体特征参数的第一难度预测模型确定的第一难度预测参数和第二群体特征参数的第二难度预测模型确定的第二难度预测参数,拟合得到待评估试题的难度值,从而可以基于不同群体(例如,学生群体和教师群体)的特征进行难度值的评估,避免了单一群体差异过大可能造成的难度值不准确,使评估得到的试题难度值更贴近试题的实际难度,提高难度值的准确度。

[0126] 为了提高本发明实施例所提供的试题难度评估方法的准确度,参考图7,本发明实施例还提供一种试题难度评估装置,包括:

第一预测参数确定模块400,适于基于第一难度预测模型,确定待评估试题的第一难度预测参数,所述第一难度预测模型用于在第一群体特征参数下,预测待评估试题的第一难度预测参数;

第二预测参数确定模块410,适于基于第二难度预测模型,确定待评估试题的第二

难度预测参数,所述第二难度预测模型用于在第二群体特征参数下,预测待评估试题的第二难度预测参数,所述第二群体特征参数不同于所述第一群体特征参数;

拟合模块420,适于至少基于所述第一难度预测参数和所述第二难度预测参数,拟合得到所述待评估试题的难度值。

[0127] 可选的,所述拟合模块420,适于至少基于所述第一难度预测参数和所述第二难度预测参数,拟合得到所述待评估试题的难度值,包括:

确定所述第一难度预测参数和所述第二难度预测参数对应的拟合参数;

基于所述拟合参数,计算得到所述待评估试题的难度值。

[0128] 可选的,所述拟合模块420,适于采用线性回归、多项式回归或多层神经网络回归确定所述第一难度预测参数和所述第二难度预测参数对应的拟合参数。

[0129] 可选的,所述拟合模块420,适于采用线性回归的网格搜索方式,确定所述第一难度预测参数和所述第二难度预测参数对应的拟合参数。

[0130] 可选的,所述拟合模块420,适于采用多项式回归的网格搜索方式,确定所述第一难度预测参数和所述第二难度预测参数对应的拟合参数。

[0131] 可选的,所述拟合模块420,适于采用多层神经网络回归的梯度下降方式,确定所述第一难度预测参数和所述第二难度预测参数对应的拟合参数。

[0132] 可选的,所述拟合模块420,适于确定所述第一难度预测参数和所述第二难度预测参数对应的拟合参数,包括:

构建试题样本集,所述试题样本集内的试题至少配置有与之对应的第一难度预测参数和第二难度预测参数;

基于所述试题样本集,寻找所述第一难度预测参数和所述第二难度预测参数对应的最优拟合参数;

以所述第一难度预测参数和所述第二难度预测参数对应最优拟合参数作为所述第一难度预测参数和所述第二难度预测参数对应的拟合参数。

[0133] 可选的,所述拟合模块420,适于构建试题样本集,包括:

构建试题样本组,所述试题样本组包括至少2个候选试题,所述候选试题至少配置有与之对应的第一难度预测参数和第二难度预测参数;

获取所述试题样本组内,候选试题的相对难度参数;其中,所述相对难度参数用于指示所述试题样本组内,各候选试题的难度排序;

根据所述候选试题的相对难度参数,确定候选试题的第一难度预测参数和第二难度预测参数与所述相对难度参数的相关度和匹配度;

基于所述相关度和匹配度,筛选所述候选试题,得到试题样本集。

[0134] 可选的,所述第二难度预测模型基于所述第一难度预测模型训练得到,所述第二难度预测模型的训练步骤包括:

获取包括第二难度参数的至少一个第二样本数据集;

基于所述第二样本数据集,训练所述第一难度预测模型,得到所述第二难度预测模型。

[0135] 可选的,所述获取包括第二难度参数的至少一个第二样本数据集,包括:

获取至少一个初始数据集,所述初始数据集包括多个试题和与试题相对应的多个

作答数据；

基于试题反应理论模型，确定各所述初始数据集中，各试题的第二难度参数，得到所述第二样本数据集。

[0136] 可选的，所述获取至少一个初始数据集之后，所述基于试题反应理论模型，确定各所述初始数据集中，各试题的第二难度参数之前，还包括：

基于测量学指标，筛选所述初始数据集；其中，所述测量学指标包括，初始数据集的样本数量、样本信息缺失率、样本通过率，以及，试题属性参数中的一个或多个；

所述基于试题反应理论模型，确定各所述初始数据集中，各试题的第二难度参数，具体为，基于试题反应理论模型，确定筛选后的所述初始数据集中，各试题的第二难度参数，得到所述第二样本数据集。

[0137] 可选的，所述基于所述第二样本数据集，训练所述第一难度预测模型，得到所述第二难度预测模型，包括：

获取所述第二样本数据集中，各试题的特征数据和第二难度参数；所述试题的特征数据包括：试题文本特征、试题公式特征、试题属性标签、试题知识点标签文本和知识点属性中的一个或多个；

根据所述试题的特征数据，拼接所述试题的句子对；

以所述句子对为输入数据，以所述试题的第二难度参数为输出数据，训练所述第一难度预测模型。

[0138] 可选的，所述第一难度预测模型基于第一样本数据训练得到，所述第一难度预测模型的训练步骤包括：

获取包括第一难度参数的第一样本数据，所述第一难度参数为基于所述第一群体获取的试题难度数据；

基于所述第一样本数据，训练初始难度预测模型，得到所述第一难度预测模型。

[0139] 可选的，所述获取包括第一难度参数的第一样本数据，包括：

获取离散的第一难度初始参数，所述第一难度初始参数基于试题难度的星级评分得到；

归一化所述第一难度初始参数，得到第一样本数据的第一难度参数。

[0140] 可选的，所述第一样本数据还包括试题的文本、标签和知识点属性，所述基于所述第一样本数据，训练初始难度预测模型，具体为：

根据所述第一样本数据，确定试题的特征数据，所述试题的特征数据包括：试题文本特征、试题公式特征、试题属性标签、试题知识点标签文本和知识点属性中的一个或多个；

根据所述试题的特征数据，拼接所述试题的句子对；

以所述句子对为输入数据，以所述第一样本数据的第一难度参数为输出数据，训练所述初始难度预测模型。

[0141] 本发明实施例所提供的试题难度评估装置，结合第一群体特征参数的第一难度预测模型确定的第一难度预测参数和第二群体特征参数的第二难度预测模型确定的第二难度预测参数，拟合得到待评估试题的难度值，从而可以基于不同群体（例如，学生群体和教师群体）的特征进行难度值的评估，避免了单一群体差异过大可能造成的难度值不准确，使

评估得到的试题难度值更贴近试题的实际难度,提高难度值的准确度。

[0142] 在本发明的可选实施例中,第二难度预测模型基于第一难度预测模型训练得到,由于第一难度预测模型和第二难度预测模型之间具有较大的相关性,采用第一难度预测模型训练得到第二难度预测模型,能够实现模型参数的更快的拟合,从而训练的速度更快,训练效果更好。

[0143] 当然,本发明实施例还提供一种计算设备,本发明实施例提供的计算设备可以通过程序形式装载上述所述的程序模块架构,以实现本发明实施例提供的试题难度评估方法;该硬件设备可以应用于具体数据处理能力的电子设备,该电子设备可以为:例如终端设备或者服务器设备。

[0144] 可选的,图8示出了本发明实施例提供的设备一种可选硬件设备架构,可以包括:至少一个存储器3和至少一个处理器1;所述存储器存储有程序,所述处理器调用所述程序,以执行前述的试题难度评估方法,另外,至少一个通信接口2和至少一个通信总线4;处理器1和存储器3可以位于同一电子设备,例如处理器1和存储器3可以位于服务器设备或者终端设备;处理器1和存储器3也可以位于不同的电子设备。

[0145] 作为本发明实施例公开内容的一种可选实现,存储器3可以存储程序,处理器1可调用所述程序,以执行本发明上述实施例提供的试题难度评估方法。

[0146] 本发明实施例中,电子设备可以是能够进行试题难度评估的平板电脑、笔记本电脑等设备。

[0147] 在本发明实施例中,处理器1、通信接口2、存储器3、通信总线4的数量为至少一个,且处理器1、通信接口2、存储器3通过通信总线4完成相互间的通信;显然,图7所示的处理器1、通信接口2、存储器3和通信总线4的通信连接示意仅是可选的一种方式;

可选的,通信接口2可以为通信模块的接口,如GSM模块的接口;

处理器1可能是中央处理器CPU,或者是特定集成电路ASIC(Application Specific Integrated Circuit),或者是被配置成实施本发明实施例的一个或多个集成电路。

[0148] 存储器3可能包含高速RAM存储器,也可能还包括非易失性存储器(non-volatile memory),例如至少一个磁盘存储器。

[0149] 需要说明的是,上述的设备还可以包括与本发明实施例公开内容可能并不是必需的其他器件(未示出);鉴于这些其他器件对于理解本发明实施例公开内容可能并不是必需,本发明实施例对此不进行逐一介绍。

[0150] 本发明实施例还提供一种计算机可读存储介质,计算机可读存储介质存储有计算机可执行指令,当该指令被处理器执行时可以实现如上所述试题难度评估方法。

[0151] 上述本发明的实施方式是本发明的元件和特征的组合。除非另外提及,否则所述元件或特征可被视为选择性的。各个元件或特征可在不与其它元件或特征组合的情况下实践。另外,本发明的实施方式可通过组合部分元件和/或特征来构造。本发明的实施方式中所描述的操作顺序可重新排列。任一实施方式的一些构造可被包括在另一实施方式中,并且可用另一实施方式的对应构造代替。对于本领域技术人员而言明显的是,所附权利要求中彼此没有明确引用关系的权利要求可组合成本发明的实施方式,或者可在提交本申请之后的修改中作为新的权利要求包括。

[0152] 本发明的实施方式可通过例如硬件、固件、软件或其组合的各种手段来实现。在硬件配置方式中,根据本发明示例性实施方式的方法可通过一个或多个专用集成电路(ASIC)、数字信号处理器(DSP)、数字信号处理器件(DSPD)、可编程逻辑器件(PLD)、现场可编程门阵列(FPGA)、处理器、控制器、微控制器、微处理器等来实现。

[0153] 在固件或软件配置方式中,本发明的实施方式可以模块、过程、功能等形式实现。软件代码可存储在存储器单元中并由处理器执行。存储器单元位于处理器的内部或外部,并可经由各种已知手段向处理器发送数据以及从处理器接收数据。

[0154] 对所公开的实施例的上述说明,使本领域技术人员能够实现或使用本发明。对这些实施例的多种修改对本领域技术人员来说将是显而易见的,本文中所定义的一般原理可以在不脱离本发明的精神或范围的情况下,在其他实施例中实现。因此,本发明将不会被限制于本文所示的这些实施例,而是符合与本文所公开的原理和新颖特点相一致的最宽的范围。

[0155] 虽然本发明实施例披露如上,但本发明并非限定于此。任何本领域技术人员,在不脱离本发明的精神和范围内,均可作各种变动与修改,因此本发明的保护范围应当以权利要求所限定的范围为准。

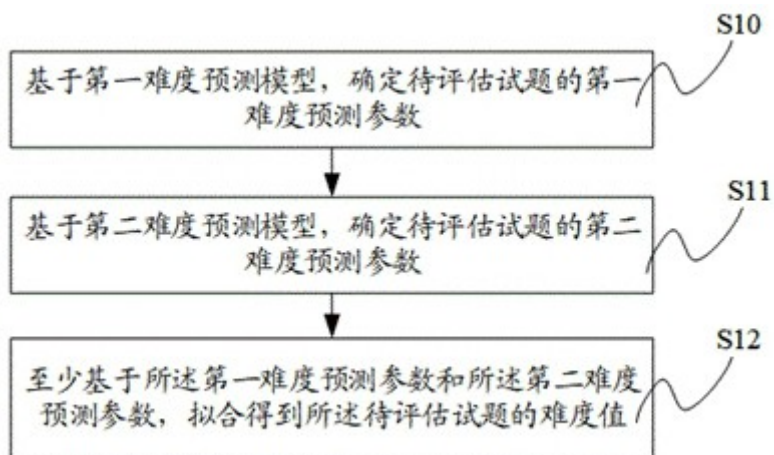


图1

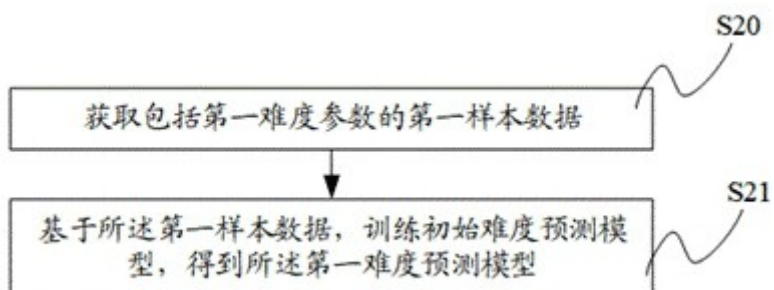


图2

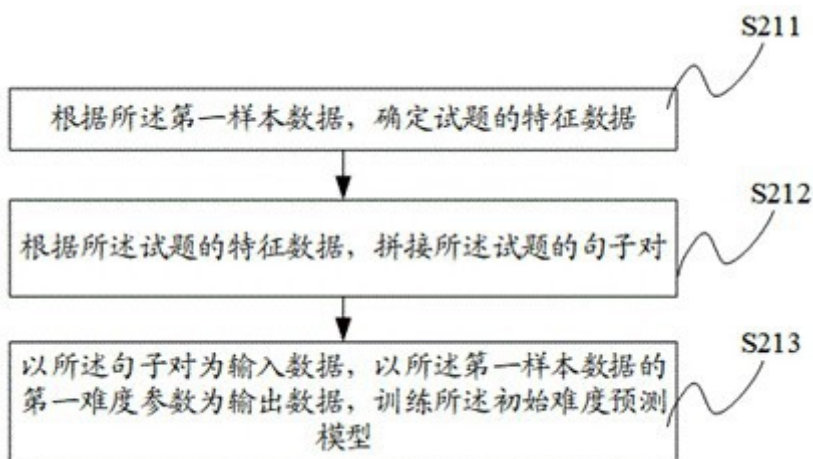


图3

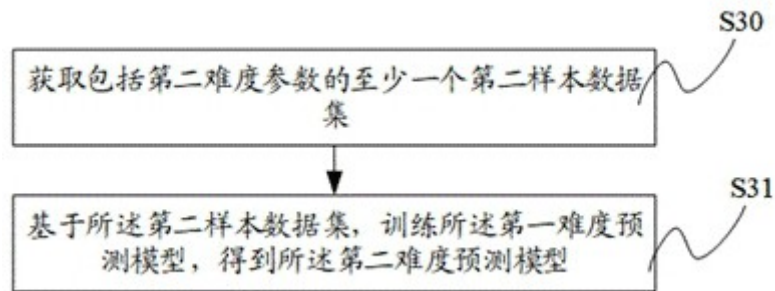


图4

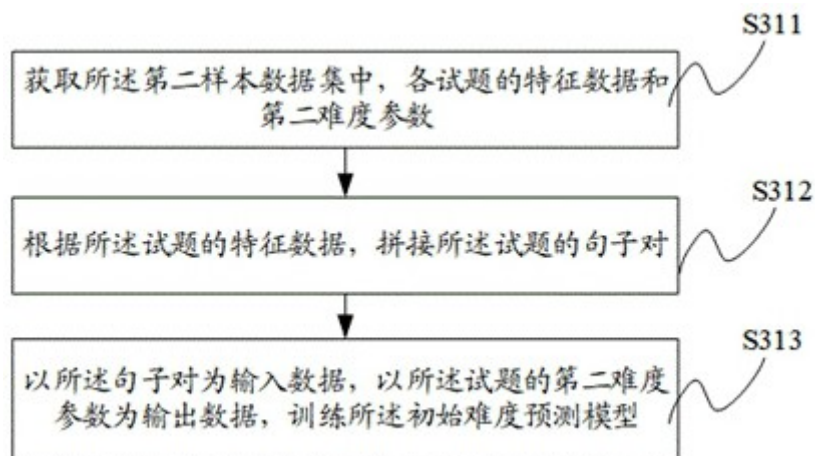


图5

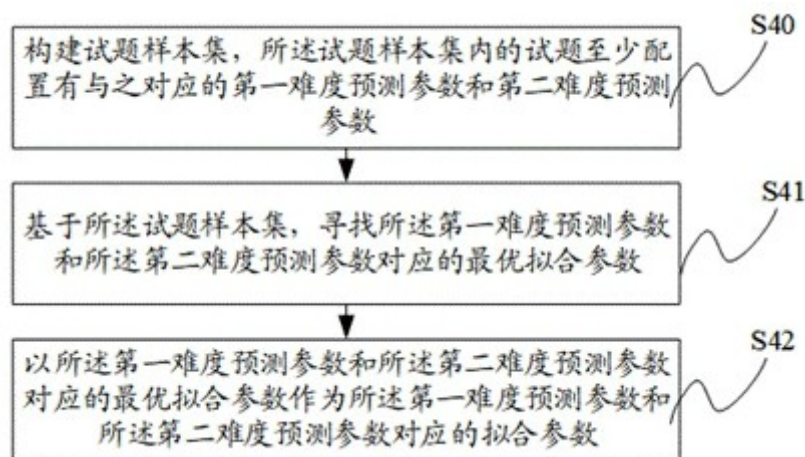


图6



图7

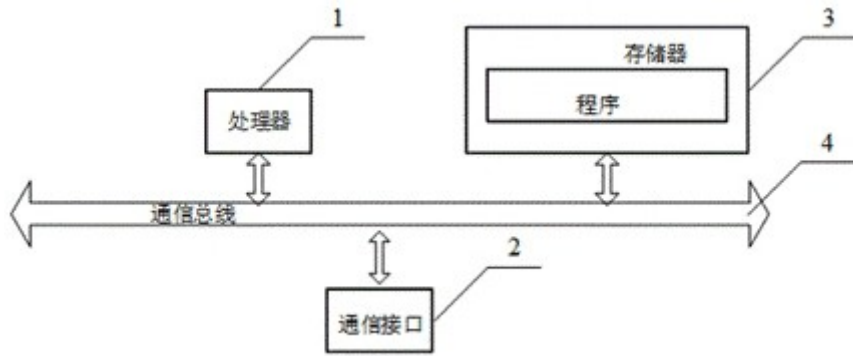


图8