



## (12)发明专利申请

(10)申请公布号 CN 110033402 A

(43)申请公布日 2019.07.19

(21)申请号 201910291996.9

(22)申请日 2019.04.12

(71)申请人 上海义学教育科技有限公司

地址 200233 上海市徐汇区田林东路588号  
B381室

(72)发明人 张丽媛 曹晓晔

(74)专利代理机构 上海科盛知识产权代理有限公司 31225

代理人 赵继明

(51)Int.Cl.

G06Q 50/20(2012.01)

G06K 9/62(2006.01)

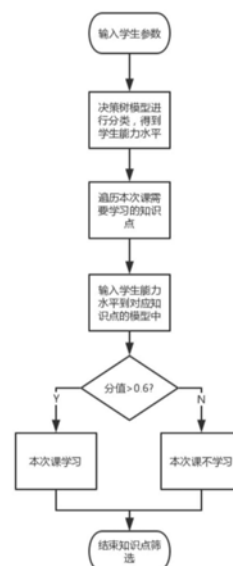
权利要求书1页 说明书5页 附图3页

### (54)发明名称

基于能力分级的思维数学学习方法及计算机学习系统

### (57)摘要

本发明涉及一种基于能力分级的思维数学学习方法及计算机学习系统,所述学习方法包括:获取学生用户的思维数学画像信息;将离散化处理后的思维数学画像信息作为一分类器的输入,获得学生用户对应的等级信息;基于所述等级信息和当前课时的知识点学习模型,获得学生用户对于当前课时内各知识点的筛选信息;基于所述筛选信息推送学习内容。与现有技术相比,本发明具有利于学习效率等优点。



1. 一种基于能力分级的思维数学学习方法,其特征在于,包括:  
获取学生用户的思维数学画像信息;  
将离散化处理后的思维数学画像信息作为一分类器的输入,获得学生用户对应的等级信息;  
基于所述等级信息和当前课时的知识点学习模型,获得学生用户对于当前课时内各知识点的筛选信息;  
基于所述筛选信息推送学习内容。
2. 根据权利要求1所述的基于能力分级的思维数学学习方法,其特征在于,所述思维数学画像信息包括用户作答记录、用户地理位置、用户年龄和用户性别。
3. 根据权利要求1所述的基于能力分级的思维数学学习方法,其特征在于,所述分类器为决策树分类器,通过k-fold方法获得所述决策树分类器的参数。
4. 根据权利要求1所述的基于能力分级的思维数学学习方法,其特征在于,利用动态更新的数据集训练所述分类器,所述数据集包括多组思维数学画像信息及对应等级标签数据。
5. 根据权利要求1所述的基于能力分级的思维数学学习方法,其特征在于,所述知识点学习模型为基于双参数项目反应理论的逻辑回归模型,该逻辑回归模型基于梯度下降算法和k-fold方法获得最优模型参数。
6. 一种基于能力分级的思维数学计算机学习系统,其特征在于,包括:  
采集模块,用于获取学生用户的思维数学画像信息;  
等级分类模块,用于将离散化处理后的思维数学画像信息作为一分类器的输入,获得学生用户对应的等级信息;  
知识点筛选模块,基于所述等级信息和当前课时的知识点学习模型,获得学生用户对于当前课时内各知识点的筛选信息;  
推送模块,基于所述筛选信息推送学习内容。。
7. 根据权利要求6所述的基于能力分级的思维数学计算机学习系统,其特征在于,所述思维数学画像信息包括用户作答记录、用户地理位置、用户年龄和用户性别。
8. 根据权利要求6所述的基于能力分级的思维数学计算机学习系统,其特征在于,所述分类器为决策树分类器,通过k-fold方法获得所述决策树分类器的参数。
9. 根据权利要求6所述的基于能力分级的思维数学计算机学习系统,其特征在于,利用动态更新的数据集训练所述分类器,所述数据集包括多组思维数学画像信息及对应等级标签数据。
10. 根据权利要求6所述的基于能力分级的思维数学计算机学习系统,其特征在于,所述知识点学习模型为基于双参数项目反应理论的逻辑回归模型,该逻辑回归模型基于梯度下降算法和k-fold方法获得最优模型参数。

## 基于能力分级的思维数学学习方法及计算机学习系统

### 技术领域

[0001] 本发明涉及一种在线学习装置,尤其是涉及一种基于能力分级的思维数学学习方法及计算机学习系统。

### 背景技术

[0002] 在现在的学习环境中,家长经常会要求孩子学习思维数学,但是效果往往非常差。传统的学习方式是报名线下的班课,跟着真人老师进行学习。随着在线教育的兴起,也有非常多的学生会选择在网络上报名课程,通过观看视频加做题进行学习。通过在网络上报名上课,能节约学生和家长不少的时间,并且让教育资源得到了更合理的分配。后来又出现了双师课堂的形式,即视频里名师讲解,同时配有班主任,能够督促学生完成作业,并且为学生答疑,一定程度提升了课程的结课率。但是这种学习系统可以说只是线下教学模式的复刻,是在思维数学的学习上效果有限。

[0003] 这种系统存在的缺陷主要有:

[0004] 1、没有考虑学生自身的接受程度和他们的学习目标,通过千篇一律的讲解,学生不一定能够学会知识点,并且固定的练习题也不一定能够符合学生的水平

[0005] 2、在思维数学中有很多知识点都不适合一般的学生,但是系统依旧会让学生学习这部分的知识,反而会让学生得到挫败感,难以提高学习效率。

### 发明内容

[0006] 本发明的目的就是为了解决上述现有技术存在的缺陷而提供一种基于能力分级的思维数学学习方法及计算机学习系统。

[0007] 本发明的目的可以通过以下技术方案来实现:

[0008] 一种基于能力分级的思维数学学习方法,包括:

[0009] 获取学生用户的思维数学画像信息;

[0010] 将离散化处理后的思维数学画像信息作为一分类器的输入,获得学生用户对应的等级信息;

[0011] 基于所述等级信息和当前课时的知识点学习模型,获得学生用户对于当前课时内各知识点的筛选信息;

[0012] 基于所述筛选信息推送学习内容。

[0013] 进一步地,所述思维数学画像信息包括用户作答记录、用户地理位置、用户年龄和用户性别。

[0014] 进一步地,所述分类器为决策树分类器,通过k-fold方法获得所述决策树分类器的参数。

[0015] 进一步地,利用动态更新的数据集训练所述分类器,所述数据集包括多组思维数学画像信息及对应等级标签数据。

[0016] 进一步地,所述知识点学习模型为基于双参数项目反应理论的逻辑回归模型,该

逻辑回归模型基于梯度下降算法和k-fold方法获得最优模型参数。

[0017] 本发明还提供一种基于能力分级的思维数学计算机学习系统,包括:

[0018] 采集模块,用于获取学生用户的思维数学画像信息;

[0019] 等级分类模块,用于将离散化处理后的思维数学画像信息作为一分类器的输入,获得学生用户对应的等级信息;

[0020] 知识点筛选模块,基于所述等级信息和当前课时的知识点学习模型,获得学生用户对于当前课时内各知识点的筛选信息;

[0021] 推送模块,基于所述筛选信息推送学习内容。.

[0022] 进一步地,所述思维数学画像信息包括用户作答记录、用户地理位置、用户年龄和用户性别。

[0023] 进一步地,所述分类器为决策树分类器,通过k-fold方法获得所述决策树分类器的参数。

[0024] 进一步地,利用动态更新的数据集训练所述分类器,所述数据集包括多组思维数学画像信息及对应等级标签数据。

[0025] 进一步地,所述知识点学习模型为基于双参数项目反应理论的逻辑回归模型,该逻辑回归模型基于梯度下降算法和k-fold方法获得最优模型参数。

[0026] 与传统思维数学的教学方式和现有的思维数学网课相比,本发明具备以下改进点及有益效果:

[0027] 第一,本发明首次提出了对用户在思维数学学习上进行能力水平的分级,从而使得根据学生水平挑选合适的知识点和内容,大大提高了用户的学习效率。

[0028] 第二,经过本发明的分级,学生学习适合他们水平的知识点的掌握率达到了70%,有效提高了用户对各知识点的掌握率;

[0029] 第三、本发明根据等级信息获得知识点的筛选信息,进行推送相应知识点的学习内容,用户完课时长更短,知识点的掌握率更高。

## 附图说明

[0030] 图1为本发明分类器的训练过程示意图;

[0031] 图2为本发明知识点模型的训练过程示意图;

[0032] 图3为本发明学习方法的过程示意图。

## 具体实施方式

[0033] 下面结合附图和具体实施例对本发明进行详细说明。本实施例以本发明技术方案为前提进行实施,给出了详细的实施方式和具体的操作过程,但本发明的保护范围不限于下述的实施例。

[0034] 如图3所示,本发明提供一种基于能力分级的思维数学学习方法,包括:

[0035] S1、获取学生用户的思维数学画像信息,包括用户作答记录、用户地理位置、用户年龄和用户性别等,

[0036] 将离散化处理后的思维数学画像信息作为一分类器的输入,获得学生用户对应的等级信息;

[0037] 基于所述等级信息和当前课时的知识点学习模型,获得学生用户对于当前课时内各知识点的筛选信息;

[0038] 基于所述筛选信息推送学习内容。.

[0039] 本发明分类器为决策树分类器,通过k-fold方法获得所述决策树分类器的参数。

[0040] 决策树分类器采用C4.5算法。信息熵 (Information Entropy, 以下用Entropy来简写) 理论被广泛用在了决策树的训练中,它的定义如下:

$$[0041] \quad Entropy(D) = - \sum_{k=1}^{|y|} p_k \log_2 p_k$$

[0042] 其中D代表了样本集,第k类的样本所占比例为 $p_k=1, 2, \dots, |y|$ )。信息熵衡量了信息的不确定性,信息熵越小,表明样本的信息越纯,即分类越准确。基于信息熵,我们还可以定义信息增益 (Gain):

$$[0043] \quad Gain(D, a) = Entropy(D) - \sum_{v=1}^V \frac{|D^v|}{|D|} Entropy(D^v)$$

[0044] 其中a为具体的某一个属性,||表示了数据集合里样本的总体数量, $v = \{1, 2, \dots, V\}$ 表示了a的V个取值。为了使决策树能够进行分类,还需要定义信息增益率 (Gain ratio):

$$[0045] \quad Gain\_ratio = \frac{Gain(D, a)}{IV(a)}, IV(a) = \sum_{v=1}^V \frac{|D^v|}{|D|} \log_2 \frac{|D^v|}{|D|}$$

[0046] IV (a) 为属性a的固有值,a的取值个数越大,IV (a) 的结果越大,信息增益率就会越小。

[0047] 本发明中决策树分类器的训练过程如图1所示,具体为:建立数据集D,该数据集包括多组思维数学画像信息及对应等级标签数据;利用所述数据集对分类器进行训练,遍历数据集中的属性a,并进一步遍历属性a中的可选值,计算数据集中每个a的可选值的信息增益率;选择信息增益率最大的特征进行分类,直至分类结束;通过k-fold (k折交叉验证) 方法挑选精准度最高的模型作为最终的分类器,将参数保存用于模型使用。

[0048] 为了加强分类器的准确度,本发明利用动态更新的数据集训练所述分类器。

[0049] 在每次用户开始一次课学习的时候,需要判断当前课时内的各知识点是否适合用户进行学习,本发明通过知识点学习模型进行判断。

[0050] 本发明中,知识点学习模型为基于双参数项目反应理论的逻辑回归模型,该逻辑回归模型的输入包括归一化为 $[-1, 1]$ 之间的等级信息和知识点难度,基于梯度下降算法和k-fold方法获得最优模型参数。

[0051] 逻辑回归的分类器公式如下:

$$[0052] \quad f(p) = 1 / (1 + e^{-\alpha(p-d)})$$

[0053] 其中e为自然底数, $\alpha$ 为区分度,p为学生的能力,d为知识点的难度。对应的优化目标函数为:

$$[0054] \quad y = J(p) = \frac{1}{m} \sum_{i=1}^m [-y^{(i)}(\log f(p^{(i)})) - (1 - y^{(i)})(1 - \log f(p^{(i)}))]$$

[0055] 其中m为一个知识点中用于训练的样本数; $y^{(i)}$ 为第i个学生是否掌握了知识点的情况,掌握为1,未掌握为0; $p^{(i)}$ 为第i个学生的能力水平。目标函数对于学生能力p进行求导后可得 $\frac{dy}{dp}$ ,每次迭代更新需要优化的参数—区分度( $\alpha$ )。伪代码如下:

```

for i := 1 to 500 do
    y := J(p)
[0056]    gradient :=  $\frac{dy}{dp}$ 
     $\alpha := \alpha - \text{learningRate} * \text{gradient}$ 
end for

```

[0057] 其中learningRate为预先设定好的学习速率。

[0058] 知识点学习模型的训练过程如图2所示,具体为:

[0059] a) 使用学生的知识点学习数据进行模型训练,能力水平和知识点难度作为输入值,优化知识点的区分度(discrimination)

[0060] b) 通过梯度下降算法进行模型训练(gradient descent),通过500次迭代后得到最终的区分度参数。

[0061] c) 使用k-fold(k折交叉验证)算法挑选精准度最高的模型;

[0062] d) 重复上述步骤直到所有知识点的模型训练完毕。

[0063] 本发明的筛选信息包括是否对某一知识点进行学习的信息,筛选信息根据知识点学习模型的输出确定。如知识点学习模型输出的分值超过0.6则认为适合学习,低于0.6则对应知识点不会出现在本次学习中。

[0064] 4. 每次上课前调用该次课的知识点对应的分类器,得到得分,如果学生得分超过0.6则认为适合学习,低于0.6则不会出现在本次学习中

[0065] 本发明还提供一种基于能力分级的思维数学计算机学习系统,包括采集模块、等级分类模块、知识点筛选模块和推送模块,采集模块用于获取学生用户的思维数学画像信息;等级分类模块用于将离散化处理后的思维数学画像信息作为一分类器的输入,获得学生用户对应的等级信息;知识点筛选模块基于所述等级信息和当前课时的知识点学习模型,获得学生用户对于当前课时内各知识点的筛选信息;推送模块基于所述筛选信息推送学习内容。

[0066] 本实施例采集了共372名四年级思维数学的学生学习记录,并且对于这部分的数据进行加工,得到了正答率,知识点掌握情况,能力值等,并将这部分的数据进行离散化,得到学生在各特征维度上的类别。对于这部分的学生,由教研院资深教师进行人工标注,获得分级标签。将标注后的数据用于训练C4.5决策树分类器,训练的过程中每个节点都挑选了信息增益率最大的特征维度和取值进行分类,直到最后没有可以继续分类的属性或者分类后所有的等级的学生都在同一个节点中。最后为了避免出现分类器过拟合的情况,使用了k-fold交叉验证对于决策树进行优化。

[0067] 为了验证本发明的效果,本实施例在同样内容的两个课程上进行了实验,每个课程的学生人数均为50人,课程内容为一般四年级学生学习的思维数学内容,共设有15次课,学生需要依次上完全部的课程。该课程中共设有114个知识点。注意,这里的15次课是我们系统中的15次课,并非传统意义上按照课时来计算的15次课,故根据学生实际学习能力的不同,学习时长会有较大的变化。

[0068] 课程A作为没有采用本发明的控制组,学生在学习的时候会学习课程下全部的114个思维数学知识点;课程B为采用了本发明的对比组,学生在每节课开始学习前,系统会将学生的画像信息和学习数据输入决策树分类器,给出学生对应的思维数学等级水平,然后再结合我们的知识点标签信息为学生挑选知识点。其余学习内容,包括视频、讲义、题目完全不变。

[0069] 为了更好说明本发明的效果,下面对比了如表1所示的两个课程。

[0070] 表1

[0071]

	课程 A (控制组)	课程 B (对比组)
平均课程知识点学习个数 (个)	114	83.4
平均学习总时长 (小时)	27.6	16.3
平均课程知识点掌握率	43%	68%

[0072] 此外,本实施例还分别使用了三年级和五年级的学生数据为他们各自训练分类器,并且在三年级和五年级的思维数学课程中加入了该分级系统。取得了如表2所示的成绩。

[0073] 表2

[0074]

	三年级课程	五年级课程
课程设置知识点个数 (个)	107	121
平均课程知识点学习个数 (个)	82.1	78.7
平均学习总时长 (小时)	17.2	17.9
平均课程知识点掌握率	72%	69%

[0075] 从这两组数据中可以看到,采用了本发明的课程相较于未采用本发明的课程在思维数学的学习效果上更好,具体体现在完课时长更短,知识点的掌握率更高,把大量的学习学不会的内容的时间给节约下来了。同时这个分级方法适用于多个年级的思维数学课程。

[0076] 以上详细描述了本发明的较佳具体实施例。应当理解,本领域的普通技术人员无需创造性劳动就可以根据本发明的构思作出诸多修改和变化。因此,凡本技术领域中技术人员依本发明的构思在现有技术的基础上通过逻辑分析、推理或者有限的实验可以得到的技术方案,皆应在由权利要求书所确定的保护范围内。

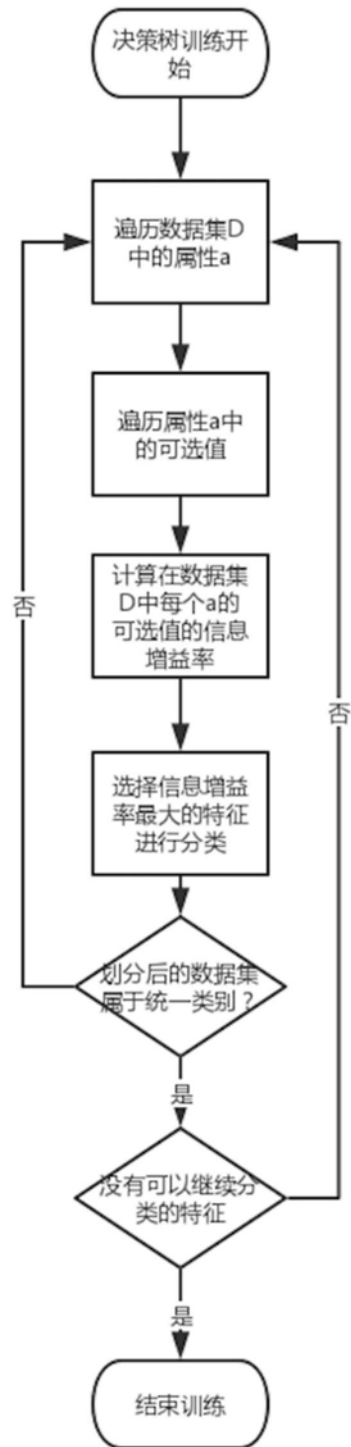


图1



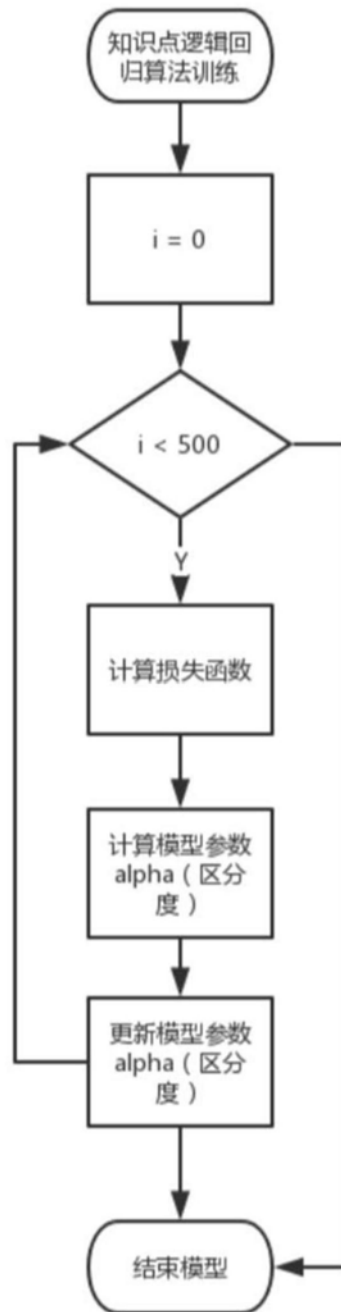


图2

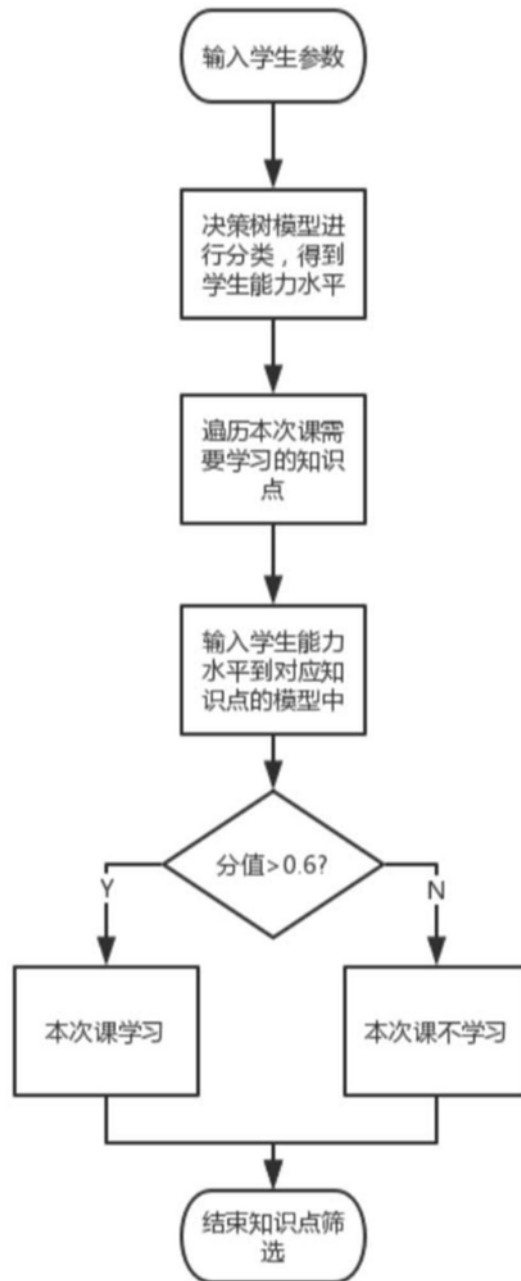


图3