

成语填空题的答案选择方法、装置和计算机设备

[0001] 本申请要求于2020年09月04日提交中国专利局、申请号为202010923909.X，发明名称为“成语填空题的答案选择方法、装置和计算机设备”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

[0002] 本申请涉及人工智能技术领域，特别涉及一种成语填空题的答案选择方法、装置、计算机设备和存储介质。

背景技术

[0003] 成语是汉语言中长期习用的形式简洁、含义丰富而又历史悠久的固定词组，它是中国传统文化的缩影，来源多种多样。成语本身结构复杂，多年以来一直是语文教学中的重点和难点。而成语完形填空是中小学语文学习中一个需要重点掌握的知识点和常见考题。面对成语选择填空这一类题目，现有技术中一直通过使用人脑智力来解答。家长们在辅导孩子作业时，凭借自身知识辅导成语填空这类高难度的题目经常显得力不从心，从而借助很多app软件和互联网搜索来查找答案。然而发明人意识到目前存在的app软件和互联网搜索中，成语填空类型的题目都是依靠人工收集答案，无法通过智能化的方法自动判定答案。经常出现因人工收集的题目数量不全，使得家长或者学生经过长时间的搜索仍找不到答案。

发明概述

技术问题

[0004] 本申请的主要目的为提供一种成语填空题的答案选择方法、装置、计算机设备和存储介质，旨在克服目前人工收集题目数量不全造成无法给出成语填空题答案的缺陷。

问题的解决方案

技术解决方案

[0005] 为实现上述目的，本申请提供了一种成语填空题的答案选择方法，包括以下步

骤：

- [0006] 获取成语填空题的问题文本；其中，所述问题文本包括填空文本以及n个候选成语，所述填空文本中包括m个待填入所述候选成语的填空；
- [0007] 获取所有所述候选成语的解释文本；
- [0008] 将所述填空文本以及所述解释文本输入至预先训练得到的成语选择填空模型中，得到每个候选成语填入每个所述填空的置信度；
- [0009] 从n个所述候选成语中选择出m个成语随机排列填入m个所述填空中，组成多组答案；其中，在每一组答案中，所述候选成语最多只能被选择填入所述填空中一次；
- [0010] 根据每个所述候选成语填入每个所述填空的置信度，基于KM算法，计算每一组答案中的所述候选成语填入所述填空的置信度总和；
- [0011] 获取置信度总和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案。
- [0012] 本申请还提供了一种成语填空题的答案选择装置，包括：
- [0013] 第一获取单元，用于获取成语填空题的问题文本；其中，所述问题文本包括填空文本以及n个候选成语，所述填空文本中包括m个待填入所述候选成语的填空；
- [0014] 第二获取单元，用于获取所有所述候选成语的解释文本；
- [0015] 第一输出单元，用于将所述填空文本以及所述解释文本输入至预先训练得到的成语选择填空模型中，得到每个候选成语填入每个所述填空的置信度；
- [0016] 选择单元，用于从n个所述候选成语中选择出m个成语随机排列填入m个所述填空中，组成多组答案；其中，在每一组答案中，所述候选成语最多只能被选择填入所述填空中一次；
- [0017] 计算单元，用于根据每个所述候选成语填入每个所述填空的置信度，基于KM算法，计算每一组答案中的所述候选成语填入所述填空的置信度总和；
- [0018] 第二输出单元，用于获取置信度总和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案。
- [0019] 本申请还提供一种计算机设备，包括存储器和处理器，所述存储器中存储有计

算机程序，所述处理器执行所述计算机程序时实现一种成语填空题的答案选择方法，包括以下步骤：

- [0020] 获取成语填空题的问题文本；其中，所述问题文本包括填空文本以及n个候选成语，所述填空文本中包括m个待填入所述候选成语的填空；
- [0021] 获取所有所述候选成语的解释文本；
- [0022] 将所述填空文本以及所述解释文本输入至预先训练得到的成语选择填空模型中，得到每个候选成语填入每个所述填空的置信度；
- [0023] 从n个所述候选成语中选择出m个成语随机排列填入m个所述填空中，组成多组答案；其中，在每一组答案中，所述候选成语最多只能被选择填入所述填空中一次；
- [0024] 根据每个所述候选成语填入每个所述填空的置信度，基于KM算法，计算每一组答案中的所述候选成语填入所述填空的置信度总和；
- [0025] 获取置信度总和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案。
- [0026] 本申请还提供一种计算机可读存储介质，其上存储有计算机程序，所述计算机程序被处理器执行时实现一种成语填空题的答案选择方法，包括以下步骤：
- [0027] 获取成语填空题的问题文本；其中，所述问题文本包括填空文本以及n个候选成语，所述填空文本中包括m个待填入所述候选成语的填空；
- [0028] 获取所有所述候选成语的解释文本；
- [0029] 将所述填空文本以及所述解释文本输入至预先训练得到的成语选择填空模型中，得到每个候选成语填入每个所述填空的置信度；
- [0030] 从n个所述候选成语中选择出m个成语随机排列填入m个所述填空中，组成多组答案；其中，在每一组答案中，所述候选成语最多只能被选择填入所述填空中一次；
- [0031] 根据每个所述候选成语填入每个所述填空的置信度，基于KM算法，计算每一组答案中的所述候选成语填入所述填空的置信度总和；
- [0032] 获取置信度总和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案。

发明的有益效果

有益效果

[0033] 本申请提供的成语填空题的答案选择方法、装置、计算机设备和存储介质，获取成语填空题的问题文本；其中，所述问题文本包括填空文本以及n个候选成语，所述填空文本中包括m个待填入所述候选成语的填空；获取所有所述候选成语的解释文本；将所述填空文本以及所述解释文本输入至预先训练得到的成语选择填空模型中，得到每个候选成语填入每个所述填空的置信度；从n个所述候选成语中选择出m个成语随机排列填入m个所述填空中，组成多组答案；根据每个所述候选成语填入每个所述填空的置信度，计算每一组答案中的所述候选成语填入所述填空的置信度总和；获取置信度总和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案。本申请采用人工智能的方式得到每组答案中候选成语填入填空中的置信度总和，从而将置信度总和最高的一组答案作为目标答案；便于高效率、高准确率得出成语填空的答案。

对附图的简要说明

附图说明

[0034] 图1 是本申请一实施例中成语填空题的答案选择方法步骤示意图；
[0035] 图2 是本申请一实施例中成语填空题的答案选择装置结构框图；
[0036] 图3 为本申请一实施例的计算机设备的结构示意图。
[0037] 本申请目的的实现、功能特点及优点将结合实施例，参照附图做进一步说明。

实施该发明的最佳实施例

本发明的最佳实施方式

[0038] 参照图1，本申请一实施例中提供了一种成语填空题的答案选择方法，包括以下步骤：
[0039] 步骤S1，获取成语填空题的问题文本；其中，所述问题文本包括填空文本以及n个候选成语，所述填空文本中包括m个待填入所述候选成语的填空；
[0040] 步骤S2，获取所有所述候选成语的解释文本；
[0041] 步骤S3，将所述填空文本以及所述解释文本输入至预先训练得到的成语选择填

空模型中，得到每个候选成语填入每个所述填空的置信度；

- [0042] 步骤S4，从n个所述候选成语中选择出m个成语随机排列填入m个所述填空中，组成多组答案；其中，在每一组答案中，所述候选成语最多只能被选择填入所述填空中一次；
- [0043] 步骤S5，根据每个所述候选成语填入每个所述填空的置信度，基于KM算法，计算每一组答案中的所述候选成语填入所述填空的置信度总和；
- [0044] 步骤S6，获取置信度总和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案。
- [0045] 在本实施例中，上述方法应用于输出成语填空题的答案的场景中，利用人工智能高效率、高准确率输出成语填空题的答案。本申请中的方案还可以应用于智慧教育的领域中，以推动智慧城市的建设。
- [0046] 如上述步骤S1所述的，上述成语填空题的问题文本通常为学生的试题中的电子文本，上述问题文本中包括由多个语句组成的填空文本，以及多个候选成语（即候选答案），学生可以从上述候选成语中选择填入上述填空文本的填空中。
- [0047] 如上述步骤S2所述的，在数据库（如成语词典）中存储有所有成语的解释文本（即对成语的语义进行解释的文本），从上述数据库中，便可以获取到所有的候选成语的解释文本。
- [0048] 如上述步骤S3所述的，预先训练有一个成语选择填空模型，该成语选择填空模型是基于bert语言模型的自然语言神经网络训练所得，上述成语选择填空模型用于预测每个成语填入在成语填空题中为正确答案的置信度。而判断填入成语填空题中的成语是否为正确答案的方式则是基于上述成语选择填空模型计算上述成语的解释文本与成语填空题中的语句的语义连贯度，从而确定出对应的置信度。
- [0049] 如上述步骤S4所述的，从n个所述候选成语中选择出m个成语随机排列填入m个所述填空中，组成多组答案。在本实施例中，采用数学中的随机排列方式从n个所述候选成语中选择出m个成语随机排列成多组组合，每一组组合作为一种填入上述成语填空题中的预选答案。可以理解的是，上述组合的数量为 $n * (n-1) * (n-2) * \dots * (n-m+1)$ 。

- [0050] 如上述步骤S5所述的，由于上述步骤S3中已经计算出所有候选成语填入每个所述填空的置信度，因此，可以计算出每一组答案中的所述候选成语填入所述填空的置信度总和。在本实施例中，基于KM算法进行上述置信度总和的计算，其中KM算法为通用算法，在此不进行赘述。
- [0051] 如上述步骤S6所述的，获取置信度总和最高的一组答案，并将其作为目标答案，可以理解的是，上述置信度表达的是候选成语填入填空中为正确的置信程度，显然当其置信度越高时，越贴近于正确答案。因此，当上述置信度总和最高时，则对应的目标答案也将最贴近于正确答案，最终将所述目标答案输出为所述成语填空题的答案。在本实施例中，基于人工智能的深度学习训练得到的成语选择填空模型，可以得到每个候选成语填入每个所述填空的置信度，进而根据置信度高效率、高准确率输出成员填空题的目标答案。在此过程中，无需人工进行搜索答案，也无需用户具备专业的成语知识，便可以自动输出对应的目标答案，不仅提高找出答案的效率以及准确率，而且降低人力成本。
- [0052] 在一实施例中，所述将所述填空文本以及所述解释文本输入至预先训练得到的成语选择填空模型中，得到每个候选成语填入每个所述填空的置信度的步骤S3之前，包括：
- [0053] 步骤S01，获取训练样本，其中所述训练样本包括多个已经选择好答案的成语填空题的语料文本以及成语库中的所有成语的解释文本；
- [0054] 步骤S02，将所述训练样本输入至基于bert语言模型的自然语言神经网络中进行训练，得到所述成语选择填空模型；其中，所述自然语言神经网络包括网络输出层以及卷积层，所述卷积层由多个卷积网络依次堆积形成，后一个所述卷积网络的输入层与其之前的所有卷积网络的输出层进行连接，且每一个所述卷积网络的输出层均与所述bert语言模型的输出层连接，最后一层所述卷积网络的输出层与所述网络输出层连接。可以理解的是，上述bert语言模型的输出层输出的特征矩阵分别输出至每一层所述卷积网络，同时，每一层卷积网络的输出层也均连接其之后的所有卷积网络的输入层。在本实施例中，在模型结构上作出上述改进，从而增加神经网络的深度，使其在特征提取上具有更强的提取能力。

- [0055] 在本实施例中，所述将所述训练样本输入至基于bert语言模型的自然语言神经网络中进行训练，得到所述成语选择填空模型的步骤S02，具体包括：
- [0056] a、将所述训练样本输入至所述bert语言模型中，基于所述bert语言模型，提取所述训练样本的特征矩阵；
- [0057] b、将所述特征矩阵输入至所述自然语言神经网络中进行迭代训练，得到所述成语选择填空模型。
- [0058] 在一实施例中，上述bert语言模型是通过在海量的语料的基础上运行自监督学习方法得到，其为词学习到一个好的特征向量表示；其中，上述自监督学习是指在没有人工标注的数据上运行的监督学习。
- [0059] 在另一实施例中，上述bert语言模型使用两个无监督预测任务进行预训练，两个无监督预测任务分别是Masked LM和Next Sentence Prediction。
- [0060] 上述bert语言模型的网络架构使用的是多层Transformer结构，其最大的特点是抛弃了传统的RNN和CNN，通过Attention机制将任意位置的两个单词的距离转换成1，有效的解决了NLP中棘手的长期依赖问题。Transformer的结构在NLP领域中已经得到了广泛应用，在此不再进行赘述。
- [0061] 在NLP（自然语言识别）任务中，可以直接使用bert语言模型输出的特征矩阵的向量表示作为NLP任务的词嵌入特征。bert语言模型提供的是一个供其它任务迁移学习的模型，上述bert语言模型可以根据任务微调或者固定之后作为特征提取器。
- [0062] 在一实施例中，所述根据每个所述候选成语填入每个所述填空的置信度，基于KM算法，计算每一组答案中的所述候选成语填入所述填空的置信度总和的步骤S5之前，包括：
- [0063] 步骤S51，分别判断每一组答案中是否存在至少一个所述候选成语填入所述填空的置信度小于第一阈值；
- [0064] 步骤S52，若存在，则将对应的答案剔除，且不计算剔除掉的所述答案中所述候选成语填入所述填空的置信度总和。
- [0065] 在本实施例中，由于上述组成的答案数量为 $n * (n-1) * (n-2) * \dots * (n-m+1)$ ，当上述n与m较大时，组成的答案数量也比较庞大，若每一组答案均分别计

算置信度总和，则会造成计算量大，影响计算效率。可以理解的是，在上述组成的答案中，显然有一些答案是明显正确的。例如在某一组答案中，填入在填空中的某个候选成语明显不符合时，则可以断定该组答案不是完全正确的答案。因此，只需要分别判断每一种答案中是否具有明显不符合的候选成语，进而将该组答案剔除，不进行置信度总和的计算，则可以显著降低运算量。在本实施例中，在已经获取到每个候选成语填入每个所述填空的置信度时，则可以根据置信度的值是否太低（置信度小于一个阈值，即第一阈值），判断出某个候选成语填入在某个填空是明显错误的。

- [0066] 在一实施例中，所述获取置信度总和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案的步骤S6之后，还包括：
- [0067] 步骤S7，将所述目标答案填入所述填空文本中，得到答案文本；
- [0068] 步骤S8，获取用户对所述答案文本的打分，并判断所述打分是否大于第二阈值；其中，所述用户基于标准答案对所述答案文本进行打分；
- [0069] 步骤S9，若所述打分大于第二阈值，则将所述答案文本以及所述目标答案中所有候选成语的解释文本构成训练集，并将所述训练集输入至所述成语选择填空模型中再训练；
- [0070] 步骤S10，将再训练后的所述成语选择填空模型保存至区块链中。
- [0071] 在本实施例中，由于上述成语填空题的题型在发送变化，其问题文本的具体文本内容也不同于原训练样本中的语料，因此，为了提高上述成语选择填空模型在后续其它场景的计算准确率，提升模型的泛化能力，需要对上述成语选择填空模型继续进行再训练。而在再训练之前，应当确认上述目标答案尽量为正确答案，因此用户（教师等）可以基于标准答案对所述答案文本进行打分，从而获取到对所述答案文本的打分，进而判断所述打分是否大于第二阈值；若大于阈值，可以判定上述成语选择填空模型输出的目标答案正确率高，此时可以采用对应的答案文本以及所述目标答案中所有候选成语的解释文本构成训练集，并将所述训练集输入至所述成语选择填空模型中再训练。在其它实施例中，若可以从教师用户获取到标准答案，则可以利用标准答案进行成语选择填空模型的再训练。

- [0072] 为了加强上述成语选择填空模型的安全性，将上述再训练后的成语选择填空模型保存至区块链中。其中，区块链是分布式数据存储、点对点传输、共识机制、加密算法等计算机技术的新型应用模式。区块链(Blockchain)，本质上是一个去中心化的数据库，是一串使用密码学方法相关联产生的数据块，每一个数据块中包含了一批次网络交易的信息，用于验证其信息的有效性(防伪)和生成下一个区块。区块链可以包括区块链底层平台、平台产品服务层以及应用服务层。
- [0073] 在一实施例中，所述成语填空题用于测试学生；
- [0074] 所述获取置信度总和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案的步骤S6之后，还包括：
- [0075] 步骤S7a，获取学生所在终端上传的测试文本；其中，测试文本为所述学生在所述成语填空题的填空文本中填入学生答案的文本；所述学生所在终端在上传所述测试文本时，将所述学生填入所述成语填空题中的学生答案依次进行拼接得到第一成语组合，并基于区块链对所述第一成语组合进行哈希计算得到对应的哈希值，将所述哈希值作为标识码添加在所述测试文本中；
- [0076] 步骤S8a，提取出所述测试文本中的所有学生答案，并依次进行拼接得到第二成语组合；
- [0077] 步骤S9a，基于区块链对所述第二成语组合进行哈希计算得到对应的验证哈希值；
- [0078] 步骤S10a，识别出所述测试文本中的标识码，并验证所述验证哈希值与所述标识码是否一致；
- [0079] 步骤S11a，若一致，则将所述学生答案与所述目标答案逐一进行对比，以获取所述学生答案的准确率。
- [0080] 在本实施例中，上述方法应用于学生测试场景中，为了测试出学生的真实成语水平，应当保障学生提交的学生答案无误，未被篡改；因此，学生在所述成语填空题的填空文本中填入学生答案之后，通过终端提交具有学生答案的测试文本时，上述终端需要将所述学生填入所述成语填空题中的学生答案依次进行拼接得到第一成语组合，并对第一成语组合进行哈希计算得到对应的哈希值，将

哈希值作为标识码添加在所述测试文本中；哈希计算可以将上述第一成语组合变换成固定长度的散列值，且上述过程是不可逆的，无法将散列值变换为上述第一成语组合，避免其它用户获取到上述哈希值后对其进行倒推获取上述学生答案。上述添加在测试文本中的标识码则作为验证上述测试文本是否被篡改的依据。

[0081] 在获取到上述学生所在终端上传的测试文本时，则提取出学生答案，并按照相同的过程对其进行组合、哈希计算以得到一个验证哈希值，从而判断验证哈希值与上述标识码是否一致，若一致，则表明学生的测试文本未被篡改，其学生答案为有效。此时，再将学生答案与目标答案进行对比，以获取所述学生答案的准确率。

[0082] 参照图2，本申请一实施例中还提供了一种成语填空题的答案选择装置，包括：

[0083] 第一获取单元10，用于获取成语填空题的问题文本；其中，所述问题文本包括填空文本以及n个候选成语，所述填空文本中包括m个待填入所述候选成语的填空；

[0084] 第二获取单元20，用于获取所有所述候选成语的解释文本；

[0085] 第一输出单元30，用于将所述填空文本以及所述解释文本输入至预先训练得到的成语选择填空模型中，得到每个候选成语填入每个所述填空的置信度；

[0086] 选择单元40，用于从n个所述候选成语中选择出m个成语随机排列填入m个所述填空中，组成多组答案；其中，在每一组答案中，所述候选成语最多只能被选择填入所述填空中一次；

[0087] 计算单元50，用于根据每个所述候选成语填入每个所述填空的置信度，基于KM算法，计算每一组答案中的所述候选成语填入所述填空的置信度总和；

[0088] 第二输出单元60，用于获取置信度总和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案。

[0089] 在一实施例中，还包括：

[0090] 第三获取单元，用于获取训练样本，其中所述训练样本包括多个已经选择好答案的成语填空题的语料文本以及成语库中的所有成语的解释文本；

- [0091] 训练单元，用于将所述训练样本输入至基于bert语言模型的自然语言神经网络中进行训练，得到所述成语选择填空模型；其中，所述自然语言神经网络包括网络输出层以及卷积层，所述卷积层由多个卷积网络依次堆积形成，后一个所述卷积网络的输入层与其之前的所有卷积网络的输出层进行连接，且每一个所述卷积网络的输出层均与所述bert语言模型的输出层连接，最后一层所述卷积网络的输出层与所述网络输出层连接。
- [0092] 在一实施例中，所述训练单元具体用于：
- [0093] 将所述训练样本输入至所述bert语言模型中，基于所述bert语言模型，提取所述训练样本的特征矩阵；
- [0094] 将所述特征矩阵输入至所述自然语言神经网络中进行迭代训练，得到所述成语选择填空模型。
- [0095] 在一实施例中，所述装置还包括：
- [0096] 第一判断单元，用于分别判断每一组答案中是否存在至少一个所述候选成语填入所述填空的置信度小于第一阈值；
- [0097] 剔除单元，用于若存在，则将对对应的答案剔除，且不计算剔除掉的所述答案中所述候选成语填入所述填空的置信度总和。
- [0098] 在一实施例中，所述装置还包括：
- [0099] 填入单元，用于将所述目标答案填入所述填空文本中，得到答案文本；
- [0100] 第二判断单元，用于获取用户对所述答案文本的打分，并判断所述打分是否大于第二阈值；其中，所述用户基于标准答案对所述答案文本进行打分；
- [0101] 再训练单元，用于若所述打分大于第二阈值，则将所述答案文本以及所述目标答案中所有候选成语的解释文本构成训练集，并将所述训练集输入至所述成语选择填空模型中再训练；
- [0102] 保存单元，用于将再训练后的所述成语选择填空模型保存至区块链中。
- [0103] 在一实施例中，所述成语填空题用于测试学生；所述装置还包括：
- [0104] 第四获取单元，用于获取学生所在终端上传的测试文本；其中，测试文本为所述学生在所述成语填空题的填空文本中填入学生答案的文本；所述学生所在终端在上传所述测试文本时，将所述学生填入所述成语填空题中的学生答案依次

进行拼接得到第一成语组合，并对所述第一成语组合进行哈希计算得到对应的哈希值，将所述哈希值作为标识码添加在所述测试文本中；

[0105] 提取单元，用于提取出所述测试文本中的所有学生答案，并依次进行拼接得到第二成语组合；

[0106] 哈希计算单元，用于基于区块链对所述第二成语组合进行哈希计算得到对应的验证哈希值；

[0107] 验证单元，用于识别出所述测试文本中的标识码，并验证所述验证哈希值与所述标识码是否一致；

[0108] 对比单元，用于若一致，则将所述学生答案与所述目标答案逐一进行对比，以获取所述学生答案的准确率。

[0109] 在本实施例中，上述装置实施例中各个单元的具体实现请参照上述方法实施例中所述，在此不再进行赘述。

[0110] 参照图3，本申请实施例中还提供一种计算机设备，该计算机设备可以是服务器，其内部结构可以如图3所示。该计算机设备包括通过系统总线连接的处理器、存储器、网络接口和数据库。其中，该计算机设计的处理器用于提供计算和控制能力。该计算机设备的存储器包括非易失性存储介质、内存储器。该非易失性存储介质存储有操作系统、计算机程序和数据库。该内存储器为易失性存储介质中的操作系统和计算机程序的运行提供环境。该计算机设备的数据库用于存储成语填空题等。该计算机设备的网络接口用于与外部的终端通过网络连接通信。该计算机程序被处理器执行时以实现一种成语填空题的答案选择方法，包括以下步骤：

[0111] 获取成语填空题的问题文本；其中，所述问题文本包括填空文本以及n个候选成语，所述填空文本中包括m个待填入所述候选成语的填空；

[0112] 获取所有所述候选成语的解释文本；

[0113] 将所述填空文本以及所述解释文本输入至预先训练得到的成语选择填空模型中，得到每个候选成语填入每个所述填空的置信度；

[0114] 从n个所述候选成语中选择出m个成语随机排列填入m个所述填空中，组成多组答案；其中，在每一组答案中，所述候选成语最多只能被选择填入所述填空中

一次；

[0115] 根据每个所述候选成语填入每个所述填空的置信度，基于KM算法，计算每一组答案中的所述候选成语填入所述填空的置信度总和；

[0116] 获取置信度总和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案。

[0117] 本领域技术人员可以理解，图3中示出的结构，仅仅是与本申请方案相关的部分结构的框图，并不构成对本申请方案所应用于其上的计算机设备的限定。

[0118] 本申请一实施例还提供一种计算机可读存储介质，其上存储有计算机程序，计算机程序被处理器执行时实现一种成语填空题的答案选择方法，包括以下步骤：

[0119] 获取成语填空题的问题文本；其中，所述问题文本包括填空文本以及n个候选成语，所述填空文本中包括m个待填入所述候选成语的填空；

[0120] 获取所有所述候选成语的解释文本；

[0121] 将所述填空文本以及所述解释文本输入至预先训练得到的成语选择填空模型中，得到每个候选成语填入每个所述填空的置信度；

[0122] 从n个所述候选成语中选择出m个成语随机排列填入m个所述填空中，组成多组答案；其中，在每一组答案中，所述候选成语最多只能被选择填入所述填空中一次；

[0123] 根据每个所述候选成语填入每个所述填空的置信度，基于KM算法，计算每一组答案中的所述候选成语填入所述填空的置信度总和；

[0124] 获取置信度总和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案。

[0125] 可以理解的是，本实施例中的计算机可读存储介质可以是易失性可读存储介质，也可以为非易失性可读存储介质。

[0126] 综上所述，为本申请实施例中提供的成语填空题的答案选择方法、装置、计算机设备和存储介质，包括：获取成语填空题的问题文本；其中，所述问题文本包括填空文本以及n个候选成语，所述填空文本中包括m个待填入所述候选成语的填空；获取所有所述候选成语的解释文本；将所述填空文本以及所述解释文

本输入至预先训练得到的成语选择填空模型中，得到每个候选成语填入每个所述填空的置信度；从n个所述候选成语中选择出m个成语随机排列填入m个所述填空中，组成多组答案；根据每个所述候选成语填入每个所述填空的置信度，基于KM算法，计算每一组答案中的所述候选成语填入所述填空的置信度总和；获取置信度总和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案。本申请采用人工智能的方式得到每组答案中候选成语填入填空中的置信度总和，从而将置信度总和最高的一组答案作为目标答案；便于高效率、高准确率得出成语填空的答案。

权利要求书

- [权利要求 1] 一种成语填空题的答案选择方法，其中，包括以下步骤：
- 获取成语填空题的问题文本；其中，所述问题文本包括填空文本以及n个候选成语，所述填空文本中包括m个待填入所述候选成语的填空；
- 获取所有所述候选成语的解释文本；
- 将所述填空文本以及所述解释文本输入至预先训练得到的成语选择填空模型中，得到每个候选成语填入每个所述填空的置信度；
- 从n个所述候选成语中选择出m个成语随机排列填入m个所述填空中，组成多组答案；其中，在每一组答案中，所述候选成语最多只能被选择填入所述填空中一次；
- 根据每个所述候选成语填入每个所述填空的置信度，基于KM算法，计算每一组答案中的所述候选成语填入所述填空的置信度总和；
- 获取置信度总和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案。
- [权利要求 2] 根据权利要求1所述的成语填空题的答案选择方法，其中，所述将所述填空文本以及所述解释文本输入至预先训练得到的成语选择填空模型中，得到每个候选成语填入每个所述填空的置信度的步骤之前，包括：
- 获取训练样本，其中所述训练样本包括多个已经选择好答案的成语填空题的语料文本以及成语库中的所有成语的解释文本；
- 将所述训练样本输入至基于bert语言模型的自然语言神经网络中进行训练，得到所述成语选择填空模型；其中，所述自然语言神经网络包括网络输出层以及卷积层，所述卷积层由多个卷积网络依次堆积形成，后一个所述卷积网络的输入层与其之前的所有卷积网络的输出层进行连接，且每一个所述卷积网络的输出层均与所述bert语言模型的输出层连接，最后一层所述卷积网络的输出层与所述网络输出层连接。
- [权利要求 3] 根据权利要求2所述的成语填空题的答案选择方法，其中，所述将所述训练样本输入至基于bert语言模型的自然语言神经网络中进行训练

，得到所述成语选择填空模型的步骤，包括：

将所述训练样本输入至所述bert语言模型中，基于所述bert语言模型，提取所述训练样本的特征矩阵；

将所述特征矩阵输入至所述自然语言神经网络中进行迭代训练，得到所述成语选择填空模型。

[权利要求 4] 根据权利要求1所述的成语填空题的答案选择方法，其中，所述根据每个所述候选成语填入每个所述填空的置信度，基于KM算法，计算每一组答案中的所述候选成语填入所述填空的置信度总和的步骤之前，包括：

分别判断每一组答案中是否存在至少一个所述候选成语填入所述填空的置信度小于第一阈值；

若存在，则将对应的答案剔除，且不计算剔除掉的所述答案中所述候选成语填入所述填空的置信度总和。

[权利要求 5] 根据权利要求1所述的成语填空题的答案选择方法，其中，所述获取置信度总和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案的步骤之后，还包括：

将所述目标答案填入所述填空文本中，得到答案文本；

获取用户对所述答案文本的打分，并判断所述打分是否大于第二阈值；其中，所述用户基于标准答案对所述答案文本进行打分；

若所述打分大于第二阈值，则将所述答案文本以及所述目标答案中所有候选成语的解释文本构成训练集，并将所述训练集输入至所述成语选择填空模型中再训练；

将再训练后的所述成语选择填空模型保存至区块链中。

[权利要求 6] 根据权利要求1所述的成语填空题的答案选择方法，其中，所述成语填空题用于测试学生；

所述获取置信度总和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案的步骤之后，还包括：

获取学生所在终端上传的测试文本；其中，测试文本为所述学生在所

述成语填空题的填空文本中填入学生答案的文本；所述学生所在终端在上传所述测试文本时，将所述学生填入所述成语填空题中的学生答案依次进行拼接得到第一成语组合，并基于区块链对所述第一成语组合进行哈希计算得到对应的哈希值，将所述哈希值作为标识码添加在所述测试文本中；

提取出所述测试文本中的所有学生答案，并依次进行拼接得到第二成语组合；

基于区块链对所述第二成语组合进行哈希计算得到对应的验证哈希值；

识别出所述测试文本中的标识码，并验证所述验证哈希值与所述标识码是否一致；

若一致，则将所述学生答案与所述目标答案逐一进行对比，以获取所述学生答案的准确率。

[权利要求 7]

一种成语填空题的答案选择装置，其中，包括：

第一获取单元，用于获取成语填空题的问题文本；其中，所述问题文本包括填空文本以及n个候选成语，所述填空文本中包括m个待填入所述候选成语的填空；

第二获取单元，用于获取所有所述候选成语的解释文本；

第一输出单元，用于将所述填空文本以及所述解释文本输入至预先训练得到的成语选择填空模型中，得到每个候选成语填入每个所述填空的置信度；

选择单元，用于从n个所述候选成语中选择出m个成语随机排列填入m个所述填空中，组成多组答案；其中，在每一组答案中，所述候选成语最多只能被选择填入所述填空中一次；

计算单元，用于根据每个所述候选成语填入每个所述填空的置信度，基于KM算法，计算每一组答案中的所述候选成语填入所述填空的置信度总和；

第二输出单元，用于获取置信度总和最高的一组答案作为目标答案，

将所述目标答案输出为所述成语填空题的答案。

[权利要求 8] 根据权利要求7所述的成语填空题的答案选择装置，其中，还包括：
第三获取单元，用于获取训练样本，其中所述训练样本包括多个已经选择好答案的成语填空题的语料文本以及成语库中的所有成语的解释文本；

训练单元，用于将所述训练样本输入至基于bert语言模型的自然语言神经网络中进行训练，得到所述成语选择填空模型；其中，所述自然语言神经网络包括网络输出层以及卷积层，所述卷积层由多个卷积网络依次堆积形成，后一个所述卷积网络的输入层与其之前的所有卷积网络的输出层进行连接，且每一个所述卷积网络的输出层均与所述bert语言模型的输出层连接，最后一层所述卷积网络的输出层与所述网络输出层连接。

[权利要求 9] 一种计算机设备，包括存储器和处理器，所述存储器中存储有计算机程序，其中，所述处理器执行所述计算机程序时实现一种成语填空题的答案选择方法，其中，包括以下步骤：

获取成语填空题的问题文本；其中，所述问题文本包括填空文本以及n个候选成语，所述填空文本中包括m个待填入所述候选成语的填空；
获取所有所述候选成语的解释文本；

将所述填空文本以及所述解释文本输入至预先训练得到的成语选择填空模型中，得到每个候选成语填入每个所述填空的置信度；

从n个所述候选成语中选择出m个成语随机排列填入m个所述填空中，组成多组答案；其中，在每一组答案中，所述候选成语最多只能被选择填入所述填空中一次；

根据每个所述候选成语填入每个所述填空的置信度，基于KM算法，计算每一组答案中的所述候选成语填入所述填空的置信度总和；

获取置信度总和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案。

[权利要求 10] 根据权利要求9所述的计算机设备，其中，所述将所述填空文本以及

所述解释文本输入至预先训练得到的成语选择填空模型中，得到每个候选成语填入每个所述填空的置信度的步骤之前，包括：

获取训练样本，其中所述训练样本包括多个已经选择好答案的成语填空题的语料文本以及成语库中的所有成语的解释文本；

将所述训练样本输入至基于bert语言模型的自然语言神经网络中进行训练，得到所述成语选择填空模型；其中，所述自然语言神经网络包括网络输出层以及卷积层，所述卷积层由多个卷积网络依次堆积形成，后一个所述卷积网络的输入层与其之前的所有卷积网络的输出层进行连接，且每一个所述卷积网络的输出层均与所述bert语言模型的输出层连接，最后一层所述卷积网络的输出层与所述网络输出层连接。

[权利要求 11] 根据权利要求10所述的计算机设备，其中，所述将所述训练样本输入至基于bert语言模型的自然语言神经网络中进行训练，得到所述成语选择填空模型的步骤，包括：

将所述训练样本输入至所述bert语言模型中，基于所述bert语言模型，提取所述训练样本的特征矩阵；

将所述特征矩阵输入至所述自然语言神经网络中进行迭代训练，得到所述成语选择填空模型。

[权利要求 12] 根据权利要求9所述的计算机设备，其中，所述根据每个所述候选成语填入每个所述填空的置信度，基于KM算法，计算每一组答案中的所述候选成语填入所述填空的置信度总和的步骤之前，包括：

分别判断每一组答案中是否存在至少一个所述候选成语填入所述填空的置信度小于第一阈值；

若存在，则将对答案剔除，且不计算剔除掉的所述答案中所述候选成语填入所述填空的置信度总和。

[权利要求 13] 根据权利要求9所述的计算机设备，其中，所述获取置信度总和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案的步骤之后，还包括：

将所述目标答案填入所述填空文本中，得到答案文本；

获取用户对所述答案文本的打分，并判断所述打分是否大于第二阈值；其中，所述用户基于标准答案对所述答案文本进行打分；

若所述打分大于第二阈值，则将所述答案文本以及所述目标答案中所有候选成语的解释文本构成训练集，并将所述训练集输入至所述成语选择填空模型中再训练；

将再训练后的所述成语选择填空模型保存至区块链中。

- [权利要求 14] 根据权利要求9所述的计算机设备，其中，所述成语填空题用于测试学生；
- 所述获取置信度总和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案的步骤之后，还包括：
- 获取学生所在终端上传的测试文本；其中，测试文本为所述学生在所述成语填空题的填空文本中填入学生答案的文本；所述学生所在终端在上传所述测试文本时，将所述学生填入所述成语填空题中的学生答案依次进行拼接得到第一成语组合，并基于区块链对所述第一成语组合进行哈希计算得到对应的哈希值，将所述哈希值作为标识码添加在所述测试文本中；
- 提取出所述测试文本中的所有学生答案，并依次进行拼接得到第二成语组合；
- 基于区块链对所述第二成语组合进行哈希计算得到对应的验证哈希值；
- 识别出所述测试文本中的标识码，并验证所述验证哈希值与所述标识码是否一致；
- 若一致，则将所述学生答案与所述目标答案逐一进行对比，以获取所述学生答案的准确率。

- [权利要求 15] 一种计算机可读存储介质，其上存储有计算机程序，其中，所述计算机程序被处理器执行时实现一种成语填空题的答案选择方法，其中，包括以下步骤：
- 获取成语填空题的问题文本；其中，所述问题文本包括填空文本以及

n个候选成语，所述填空文本中包括m个待填入所述候选成语的填空；
获取所有所述候选成语的解释文本；
将所述填空文本以及所述解释文本输入至预先训练得到的成语选择填空模型中，得到每个候选成语填入每个所述填空的置信度；
从n个所述候选成语中选择出m个成语随机排列填入m个所述填空中，组成多组答案；其中，在每一组答案中，所述候选成语最多只能被选择填入所述填空中一次；
根据每个所述候选成语填入每个所述填空的置信度，基于KM算法，计算每一组答案中的所述候选成语填入所述填空的置信度总和；
获取置信度总和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案。

[权利要求 16] 根据权利要求15所述的计算机可读存储介质，其中，所述将所述填空文本以及所述解释文本输入至预先训练得到的成语选择填空模型中，得到每个候选成语填入每个所述填空的置信度的步骤之前，包括：
获取训练样本，其中所述训练样本包括多个已经选择好答案的成语填空题的语料文本以及成语库中的所有成语的解释文本；
将所述训练样本输入至基于bert语言模型的自然语言神经网络中进行训练，得到所述成语选择填空模型；其中，所述自然语言神经网络包括网络输出层以及卷积层，所述卷积层由多个卷积网络依次堆积形成，后一个所述卷积网络的输入层与其之前的所有卷积网络的输出层进行连接，且每一个所述卷积网络的输出层均与所述bert语言模型的输出层连接，最后一层所述卷积网络的输出层与所述网络输出层连接。

[权利要求 17] 根据权利要求16所述的计算机可读存储介质，其中，所述将所述训练样本输入至基于bert语言模型的自然语言神经网络中进行训练，得到所述成语选择填空模型的步骤，包括：
将所述训练样本输入至所述bert语言模型中，基于所述bert语言模型，提取所述训练样本的特征矩阵；
将所述特征矩阵输入至所述自然语言神经网络中进行迭代训练，得到

所述成语选择填空模型。

- [权利要求 18] 根据权利要求15所述的计算机可读存储介质，其中，所述根据每个所述候选成语填入每个所述填空的置信度，基于KM算法，计算每一组答案中的所述候选成语填入所述填空的置信度总和的步骤之前，包括：分别判断每一组答案中是否存在至少一个所述候选成语填入所述填空的置信度小于第一阈值；若存在，则将对应的答案剔除，且不计算剔除掉的所述答案中所述候选成语填入所述填空的置信度总和。
- [权利要求 19] 根据权利要求15所述的计算机可读存储介质，其中，所述获取置信度总和和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案的步骤之后，还包括：将所述目标答案填入所述填空文本中，得到答案文本；获取用户对所述答案文本的打分，并判断所述打分是否大于第二阈值；其中，所述用户基于标准答案对所述答案文本进行打分；若所述打分大于第二阈值，则将所述答案文本以及所述目标答案中所有候选成语的解释文本构成训练集，并将所述训练集输入至所述成语选择填空模型中再训练；将再训练后的所述成语选择填空模型保存至区块链中。
- [权利要求 20] 根据权利要求15所述的计算机可读存储介质，其中，所述成语填空题用于测试学生；所述获取置信度总和和最高的一组答案作为目标答案，将所述目标答案输出为所述成语填空题的答案的步骤之后，还包括：获取学生所在终端上传的测试文本；其中，测试文本为所述学生在所述成语填空题的填空文本中填入学生答案的文本；所述学生所在终端在上传所述测试文本时，将所述学生填入所述成语填空题中的学生答案依次进行拼接得到第一成语组合，并基于区块链对所述第一成语组合进行哈希计算得到对应的哈希值，将所述哈希值作为标识码添加在所述测试文本中；

提取出所述测试文本中的所有学生答案，并依次进行拼接得到第二成语组合；

基于区块链对所述第二成语组合进行哈希计算得到对应的验证哈希值；

识别出所述测试文本中的标识码，并验证所述验证哈希值与所述标识码是否一致；

若一致，则将所述学生答案与所述目标答案逐一进行对比，以获取所述学生答案的准确率。