



(12)发明专利

(10)授权公告号 CN 106776523 B

(45)授权公告日 2020.04.07

(21)申请号 201710047783.2

G06F 16/34(2019.01)

(22)申请日 2017.01.22

(56)对比文件

(65)同一申请的已公布的文献号

申请公布号 CN 106776523 A

CN 106021389 A, 2016.10.12,

CN 102929930 A, 2013.02.13,

US 2004102958 A1, 2004.05.27,

(43)申请公布日 2017.05.31

叶娜等.基于相似计算的信息抽取模板自动获取方法.《第二届全国学生计算语言学研讨会论文集》.2004,

(73)专利权人 百度在线网络技术(北京)有限公司

Maslennikov等. Combining relations for information extraction from free text.《ACM Transactions on Information Systems》.2010,第28卷(第3期),

地址 100085 北京市海淀区上地十街10号
百度大厦三层

(72)发明人 袁洁

审查员 张子瑜

(74)专利代理机构 北京清亦华知识产权代理事务所(普通合伙) 11201

代理人 宋合成

(51)Int.Cl.

G06F 40/186(2020.01)

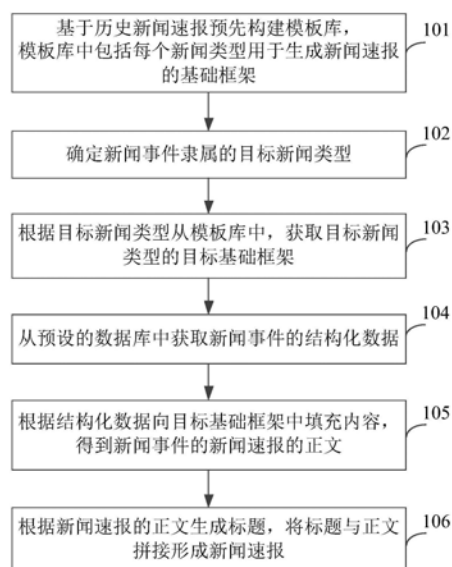
权利要求书3页 说明书10页 附图4页

(54)发明名称

基于人工智能的新闻速报生成方法及装置

(57)摘要

本发明提出一种基于人工智能的新闻速报生成方法及装置,其中,方法包括:基于历史新闻速报预先构建模板库,其中模板库中包括每个新闻类型用于生成新闻速报的基础框架;确定新闻事件隶属的目标新闻类型,根据目标新闻类型从模板库中,获取目标新闻类型的目标基础框架,从预设的数据库中获取新闻事件的结构化数据,根据结构化数据向目标基础框架中填充内容,得到新闻事件的新闻速报的正文,根据正文生成新闻速报的标题,将标题与正文拼接形成新闻速报。本实施例中,通过对历史新闻速报的学习,可以为不同的新闻类型构建基础框架,然后基于基础框架智能地生成新闻速报,提高新闻事件的报道的时效性。



1. 一种基于人工智能的新闻速报生成方法,其特征在于,包括:

基于历史新闻速报预先构建模板库,所述模板库中包括每个新闻类型用于生成新闻速报的基础框架;

确定新闻事件隶属的目标新闻类型;

根据所述目标新闻类型从所述模板库中,获取所述目标新闻类型的目标基础框架;

从预设的数据库中获取所述新闻事件的结构化数据;

根据所述结构化数据向所述目标基础框架中填充内容,得到所述新闻事件的新闻速报的正文;

根据所述正文生成所述新闻速报的标题,将所述标题与所述正文拼接形成所述新闻速报;

其中,所述基于历史新闻速报,预先构建新闻速报的模板库,所述模板库中包括每个新闻类型用于生成新闻速报的基础框架,包括:

针对同一新闻类型,对所有历史新闻速报进行提取,得到所述历史新闻速报中的结构化数据与新闻语料之间的对应关系;

利用所述历史新闻速报中的结构化数据、新闻语料以及所述对应关系形成训练集;

对所述训练集进行文本训练,挖掘各历史新闻速报中可替换的结构化数据,形成所述基础框架。

2. 根据权利要求1所述的基于人工智能的新闻速报生成方法,其特征在于,所述形成所述基础框架之后,还包括:

挖掘预设数量的所述历史新闻速报中每两个描述性短语的语义相似度;

根据所述语义相似度获取表示同一语义的所有描述性短语;

利用表示同一语义的所有描述性短语生成该同一语义的润色语句集。

3. 根据权利要求2所述的基于人工智能的新闻速报生成方法,所述根据所述结构化数据向所述目标基础框架中填充内容,得到所述新闻事件的新闻速报的正文之后,还包括:

利用所述润色语句集中对所述正文中的同一语义的描述性短语进行随机替换。

4. 根据权利要求1所述的基于人工智能的新闻速报生成方法,其特征在于,所述形成所述基础框架之后,还包括:

对预设数量的所述历史新闻速报中的结构化数据进行挖掘,得到表示对象的结构化数据;其中,所述对象的结构化数据为名词性的结构化数据;

识别用于表示同一对象的所有结构化数据;

利用同一对象的所有结构化数据生成该同一对象的同义词表。

5. 根据权利要求4所述的基于人工智能的新闻速报生成方法,所述根据所述结构化数据向所述目标基础框架中填充内容,得到所述新闻事件的新闻速报的正文之后,还包括:

利用所述同义词表中对所述正文中的同一对象的所述结构化数据进行随机替换。

6. 根据权利要求1-5任一项所述的基于人工智能的新闻速报生成方法,其特征在于,所述确定新闻事件隶属的目标新闻类型,包括:

检测是否有所述新闻事件发生;

当检测到所述新闻事件后,确定所述新闻事件隶属的所述目标新闻类型。

7. 根据权利要求1-5任一项所述的基于人工智能的新闻速报生成方法,其特征在于,所

述根据所述结构化数据向所述目标基础框架中填充内容,得到所述新闻事件的新闻速报的正文,包括:

基于知识图谱从所述数据库中获取与所述结构化数据相关的内容;

向所述目标基础框架中填充所述结构化数据以及与所述结构化数据相关的内容,得到所述新闻速报的正文。

8. 一种基于人工智能的新闻速报生成装置,其特征在于,包括:

模板库构建模块,用于基于历史新闻速报预先构建模板库,所述模板库中包括每个新闻类型用于生成新闻速报的基础框架;

类型确定模块,用于确定新闻事件隶属的目标新闻类型;

框架获取模块,用于根据所述目标新闻类型从所述模板库中,获取所述目标新闻类型的目标基础框架;

数据获取模块,用于从预设的数据库中获取所述新闻事件的结构化数据;

填充模块,用于根据所述结构化数据向所述目标基础框架中填充内容,得到所述新闻事件的新闻速报的正文;

生成模块,用于根据所述正文生成所述新闻速报的标题,将所述标题与所述正文拼接形成所述新闻速报;

其中,所述模板库构建模块,具体用于针对同一新闻类型,对所有历史新闻速报进行提取,得到所述历史新闻速报中的结构化数据与新闻语料之间的对应关系,利用所述历史新闻速报中的结构化数据、新闻语料以及所述对应关系形成训练集,对所述训练集进行文本训练,挖掘各历史新闻速报中可替换的结构化数据,形成所述基础框架。

9. 根据权利要求8所述的基于人工智能的新闻速报生成装置,其特征在于,所述模板库构建模块,还用于挖掘预设数量的所述历史新闻速报中每两个描述性短语的语义相似度,根据所述语义相似度获取表示同一语义的所有描述性短语,利用表示同一语义的所有描述性短语生成该同一语义的润色语句集。

10. 根据权利要求9所述的基于人工智能的新闻速报生成装置,还包括:润色模块,用于利用所述润色语句集中对所述正文中的同一语义的描述性短语进行随机替换。

11. 根据权利要求8所述的基于人工智能的新闻速报生成装置,其特征在于,所述模板库构建模块,还用于对预设数量的所述历史新闻速报中的结构化数据进行挖掘,得到表示对象的结构化数据;其中,所述对象的结构化数据为名词性的结构化数据,识别用于表示同一对象的所有结构化数据,利用同一对象的所有结构化数据生成该同一对象的同义词表。

12. 根据权利要求11所述的基于人工智能的新闻速报生成装置,还包括:优化模块,用于利用所述同义词表中对所述正文中的同一对象的所述结构化数据进行随机替换。

13. 根据权利要求8-12任一项所述的基于人工智能的新闻速报生成装置,其特征在于,所述类型确定模块,具体用于检测是否有所述新闻事件发生,当检测到所述新闻事件后,确定所述新闻事件隶属的所述目标新闻类型。

14. 根据权利要求8-12任一项所述的基于人工智能的新闻速报生成装置,其特征在于,所述填充模块,具体用于基于知识图谱从所述数据库中获取与所述结构化数据相关的内容,向所述目标基础框架中填充所述结构化数据以及与所述结构化数据相关的内容,得到所述新闻速报的正文。

15. 一种基于人工智能的新闻速报生成装置,其特征在于,包括:

存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,其特征在于,所述处理器执行所述程序时实现如权利要求1-7中任一所述的基于人工智能的新闻速报生成方法。

16. 一种计算机可读存储介质,其上存储有计算机程序,其特征在于,该程序被处理器执行时实现如权利要求1-7中任一所述的基于人工智能的新闻速报生成方法。

基于人工智能的新闻速报生成方法及装置

技术领域

[0001] 本发明涉及信息处理技术领域,尤其涉及一种基于人工智能的新闻速报生成方法及装置。

背景技术

[0002] 人工智能(Artificial Intelligence),英文缩写为AI。它是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学。人工智能是计算机科学的一个分支,它企图了解智能的实质,并生产出一种新的能以人类智能相似的方式做出反应的智能机器,该领域的研究包括机器人、语音识别、图像识别、自然语言处理和专家系统等。

[0003] 目前,传统的新闻速报多采用人工编辑方式产生,使得新闻速报的时效性较差。即使已经有人工编辑好的模板,从获取到用于新闻速报的数据,到校验完成最终生成新闻速报,仍然需要十几分钟的延时。

发明内容

[0004] 本发明旨在至少在一定程度上解决相关技术中的技术问题之一。

[0005] 为此,本发明的第一个目的在于提出一种基于人工智能的新闻速报生成方法,用于解决现有的通过人工编辑生成新闻速报的失效性较差的问题。

[0006] 本发明的第二个目的在于提出一种基于人工智能的新闻速报生成装置。

[0007] 本发明的第三个目的在于提出另一种基于人工智能的新闻速报生成装置。

[0008] 本发明的第四个目的在于提出一种非临时性计算机可读存储介质。

[0009] 本发明的第五个目的在于提出一种计算机程序产品。

[0010] 为达上述目的,本发明第一方面实施例提出了一种基于人工智能的新闻速报生成方法,包括:

[0011] 基于历史新闻速报预先构建模板库,所述模板库中包括每个新闻类型用于生成新闻速报的基础框架;

[0012] 确定新闻事件隶属的目标新闻类型;

[0013] 根据所述目标新闻类型从所述模板库中,获取所述目标新闻类型的目标基础框架;

[0014] 从预设的数据库中获取所述新闻事件的结构化数据;

[0015] 根据所述结构化数据向所述目标基础框架中填充内容,得到所述新闻事件的新闻速报的正文;

[0016] 根据所述正文生成所述新闻速报的标题,将所述标题与所述正文拼接形成所述新闻速报。

[0017] 本发明实施例的基于人工智能的新闻速报生成方法,通过基于历史新闻速报预先构建模板库,其中模板库中包括每个新闻类型用于生成新闻速报的基础框架;确定新闻事

件隶属的目标新闻类型,根据目标新闻类型从模板库中,获取目标新闻类型的目标基础框架,从预设的数据库中获取新闻事件的结构化数据,根据结构化数据向目标基础框架中填充内容,得到新闻事件的新闻速报的正文,根据正文生成新闻速报的标题,将标题与正文拼接形成新闻速报。本实施例中,通过对历史新闻速报的学习,可以为不同的新闻类型构建基础框架,然后基于基础框架智能地生成新闻速报,提高新闻事件的报道的时效性。

[0018] 为达上述目的,本发明第二方面实施例提出了一种基于人工智能的新闻速报生成装置,包括:

[0019] 模板库构建模块,用于基于历史新闻速报预先构建模板库,所述模板库中包括每个新闻类型用于生成新闻速报的基础框架;

[0020] 类型确定模块,用于确定新闻事件隶属的目标新闻类型;

[0021] 框架获取模块,用于根据所述目标新闻类型从所述模板库中,获取所述目标新闻类型的目标基础框架;

[0022] 数据获取模块,用于从预设的数据库中获取所述新闻事件的结构化数据;

[0023] 填充模块,用于根据所述结构化数据向所述目标基础框架中填充内容,得到所述新闻事件的新闻速报的正文;

[0024] 生成模块,用于根据所述正文生成所述新闻速报的标题,将所述标题与所述正文拼接形成所述新闻速报。

[0025] 本发明实施例的基于人工智能的新闻速报生成装置,通过基于历史新闻速报预先构建模板库,其中模板库中包括每个新闻类型用于生成新闻速报的基础框架;确定新闻事件隶属的目标新闻类型,根据目标新闻类型从模板库中,获取目标新闻类型的目标基础框架,从预设的数据库中获取新闻事件的结构化数据,根据结构化数据向目标基础框架中填充内容,得到新闻事件的新闻速报的正文,根据正文生成新闻速报的标题,将标题与正文拼接形成新闻速报。本实施例中,通过对历史新闻速报的学习,可以为不同的新闻类型构建基础框架,然后基于基础框架智能地生成新闻速报,提高新闻事件的报道的时效性。

[0026] 为达上述目的,本发明第三方面实施例提出了另一种基于人工智能的新闻速报生成装置,包括:处理器;用于存储所述处理器可执行指令的存储器;其中,所述处理器被配置为:

[0027] 基于历史新闻速报预先构建模板库,所述模板库中包括每个新闻类型用于生成新闻速报的基础框架;

[0028] 确定新闻事件隶属的目标新闻类型;

[0029] 根据所述目标新闻类型从所述模板库中,获取所述目标新闻类型的目标基础框架;

[0030] 从预设的数据库中获取所述新闻事件的结构化数据;

[0031] 根据所述结构化数据向所述目标基础框架中填充内容,得到所述新闻事件的新闻速报的正文;

[0032] 根据所述正文生成所述新闻速报的标题,将所述标题与所述正文拼接形成所述新闻速报。

[0033] 为了实现上述目的,本发明第四方面实施例提出了一种非临时性计算机可读存储介质,当所述存储介质中的指令由服务器端的处理器被执行时,使得服务器端能够执行一

种基于人工智能的新闻速报生成方法,所述方法包括:

[0034] 基于历史新闻速报预先构建模板库,所述模板库中包括每个新闻类型用于生成新闻速报的基础框架;

[0035] 确定新闻事件隶属的目标新闻类型;

[0036] 根据所述目标新闻类型从所述模板库中,获取所述目标新闻类型的目标基础框架;

[0037] 从预设的数据库中获取所述新闻事件的结构化数据;

[0038] 根据所述结构化数据向所述目标基础框架中填充内容,得到所述新闻事件的新闻速报的正文;

[0039] 根据所述正文生成所述新闻速报的标题,将所述标题与所述正文拼接形成所述新闻速报。

[0040] 为了实现上述目的,本发明第五方面实施例提出了一种计算机程序产品,当所述计算机程序产品中的指令处理器执行时,执行一种基于人工智能的新闻速报生成方法,所述方法包括:

[0041] 基于历史新闻速报预先构建模板库,所述模板库中包括每个新闻类型用于生成新闻速报的基础框架;

[0042] 确定新闻事件隶属的目标新闻类型;

[0043] 根据所述目标新闻类型从所述模板库中,获取所述目标新闻类型的目标基础框架;

[0044] 从预设的数据库中获取所述新闻事件的结构化数据;

[0045] 根据所述结构化数据向所述目标基础框架中填充内容,得到所述新闻事件的新闻速报的正文;

[0046] 根据所述正文生成所述新闻速报的标题,将所述标题与所述正文拼接形成所述新闻速报。

[0047] 本发明附加的方面和优点将在下面的描述中部分给出,部分将从下面的描述中变得明显,或通过本发明的实践了解到。

附图说明

[0048] 本发明上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解,其中:

[0049] 图1为本发明实施例提供的一种基于人工智能的新闻速报生成方法的流程示意图;

[0050] 图2为本发明实施例提供的一种预先构建模板库的流程示意图;

[0051] 图3为本发明实施例提供的一种基于人工智能的新闻速报生成方法的应用示意图;

[0052] 图4为本发明实施例提供的另一种基于人工智能的新闻速报生成方法的流程示意图;

[0053] 图5为本发明实施例提供的另一种基于人工智能的新闻速报生成装置的结构示意图;

[0054] 图6为本发明实施例提供的另一种基于人工智能的新闻速报生成装置的结构示意图；

[0055] 图7为本发明实施例提供的另一种基于人工智能的新闻速报生成装置的结构示意图。

具体实施方式

[0056] 下面详细描述本发明的实施例，所述实施例的示例在附图中示出，其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的，旨在用于解释本发明，而不能理解为对本发明的限制。

[0057] 下面参考附图描述本发明实施例的基于人工智能的新闻速报生成方法及装置。

[0058] 图1为本发明实施例所提供的一种基于人工智能的新闻速报生成方法的流程示意图。该基于人工智能的新闻速报生成方法包括以下步骤：

[0059] S101、基于历史新闻速报预先构建模板库，模板库中包括每个新闻类型用于生成新闻速报的基础框架。

[0060] 本实施例中，可以从开发数据平台中获取到历史新闻速报，然后对历史新闻速报进行提取和训练，得到一个模板库，在该模板库中包括每个新闻类型用于生成新闻速报的基础框架，即包括每个新闻类别用于生成新闻速报的模板。在基础框架包括生成新闻速报的规则，在新闻事件发生后，根据该规则来生成该新闻事件的速报。本实施例中，新闻速报可以涉及体育赛事、彩票分析、股票看盘等新闻类型。

[0061] 图2为本发明实施例提供的一种预先构建模板库的流程示意图。该预先构建模板库具体包括以下步骤：

[0062] S201、针对同一新闻类型，对所有历史新闻速报进行提取，得到历史新闻速报中结构化数据与新闻语料之间的对应关系。

[0063] 具体地，采集同一新闻类型的所有历史新闻速报，从所有的历史新闻速报中提取，得到历史新闻速报中的结构化数据，该结构化数据可以为一个数据，例如比赛的比分。该结构化数据也可以为一个数据块，例如双色球的开奖结果。该结构化数据也可以为一个可控集合，例如比赛中球队的球员名单等。

[0064] 新闻速报中包括对结构化数据的描述内容，这些描述内容就是新闻语料。本实施例中，从历史新闻速报中可以提取出结构化数据与新闻语料之间的对应关系，即结构化数据与新闻速报中各句段之间的对应关系。

[0065] S202、利用历史新闻速报中的结构化数据、新闻语料以及所述对应关系形成训练集。

[0066] 在获取到历史新闻速报中的结构化数据与新闻语料之间的对应关系后，可以利用历史新闻速报中的结构化数据、新闻语料以及所述对应关系，形成一个用于文本训练的训练集。

[0067] S203、对训练集进行文本训练，挖掘各历史新闻速报中可替换的结构化数据形成基础框架。

[0068] 进一步地，可以利用得到的训练集进行文本训练，挖掘历史新闻速报中可替换的结构化数据，然后形成新闻类型的基础框架。例如，可以从新闻速报中提取出新闻速报中可

替换的结构化数据,例如比分、球队名称等,队员名称。

[0069] 在该基础框架中包括形成一个新闻类型的新闻速报的策略。本实施例中,不同的新闻类型形成的基础框架不同。例如,在比赛赛事的新闻速报中可以包括总比分、队员、每节的比分以及对比赛的评价等。在彩票分析的新闻速报中可以包括最近一期的开奖结果、累计奖金、选号分析等内容。

[0070] 基础框架所包括的形成一个新闻类型的新闻速报的策略,可以包括每段主要介绍的对象或者主干元素,以及段落之间的关系等。例如,在比赛赛事的新闻速报中,可以在不同的段落分别介绍总比分、各节的比分情况,以及队员分布、单个球员所贡献的分值或者进球数,以及球场上个最佳球员等。

[0071] 本实施例中,通过对历史新闻速报的学习,可以为不同的新闻类型构建基础框架,然后基于基础框架智能地生成新闻速报,提高新闻事件的报道的时效性。

[0072] S102、确定新闻事件隶属的目标新闻类型。

[0073] 具体地,可以实时地或者周期性检测是否有新闻事件发生,当检测到有新闻事件发生后,说明需要对新闻事件进行报道,则需要确定该新闻事件所隶属的目标新闻类型。本实施例中,可以为不同的新闻类型设置不同的触发字段,通过这些触发字段来检测是否有相应类型的新闻事件发生。

[0074] S103、根据目标新闻类型从模板库中,获取目标新闻类型的目标基础框架。

[0075] 由于预先构建的模板库中存储有多个新闻类型的基础矿建,在确定出新闻事件的目标新闻类型后,可以根据该目标新闻类型从模块库中,获取与该目标新闻类型匹配的目标基础框架。本实施例中,可以为一个新闻类型设置多个基础框架。例如,针对同一新闻类型,不同的基础框架所包括的策略中可以设定段落的个数,以及每个段落介绍的对象,以及各个段落之间的位置关系。当针对目标新闻类型有多个基础框架时,可以从该目标新闻类型的所有基础框架中,随机地选取一个基础框架作为目标基础框架,由于每个新闻类型设置有多个基础框架,可以使针对同一新闻类型的新闻速报多样化。

[0076] S104、从预设的数据库中获取新闻事件的结构化数据。

[0077] 在选取了目标基础框架后,为了完成新闻事件的报道,需要从预设的数据库中获取该新闻事件的结构化数据。例如,一个篮球体育赛事刚刚结束,可以从预设的数据库中提取到该体育赛事的结构化数据,包括:两支球队的名称、比分、单节比分、两个球队的球员名称等。

[0078] S105、根据结构化数据向目标基础框架中填充内容,得到新闻事件的新闻速报的正文。

[0079] 为了能够形成一个新闻报道,可以基于知识图谱从数据库中获取与结构化数据相关的内容,然后向目标基础框架中填充结构化数据以及与结构化数据相关的内容,得到新闻速报的正文。例如,根据结构化数据“A队和B队”和“100:78”,基于知识图谱技术,可以得到一个关于两个球队比赛情况介绍的相关内容。A队与B队进行了激烈的角逐,最后比分锁定在100:78,A队大胜B队。然后可以根据单节比分或者队形排布等结构化数据,进一步地获取到相关内容,通过这些相关内容来丰富新闻速报,从而为互联网用户提供更多的关于该篮球赛事的信息。

[0080] S106、根据新闻速报的正文生成标题,将标题与正文拼接形成新闻速报。

[0081] 在获取到新闻速报的正文后,可以根据新闻速报的正文为该新闻速报生成一个标题,然后将标题与新闻速报的正文进行拼接,生成一个新闻速报。例如,该新闻事件为A队与B队的篮球比赛,生成的新闻速报的正文是针对该赛事的介绍,在形成正文后,就可以根据正文内容,智能生成一个标题,如“A队轻取B队,率先冲进八强”。

[0082] 图3为本实施例提供的基于人工智能的新闻速报生成方法的应用示意图。图3中为一个彩票的新闻速报,该新闻速报为通过基础框架智能生成。其中,基础框架中包括:开奖详情和选号分析两个模块。其中,开奖详情中包括:中奖情况以及奖金和累计奖金等介绍,然后可以在号码分析中对红球表现进行描述。当彩票开奖这一新闻事件发生后,就可以从预设的数据库中获取到新闻事件的结构化数据,例如,多少期、开奖号码、一等奖的注数、二等奖的注数、每注的奖金设置,以及累计的奖金。进一步地还可以获取到号码分析模块所需的结构化数据,例如冷号、温号、热号等。在获取到这些结果化数据后,就可以进一步地获取到与结果话数据相关的内容,将结构化数据以及相关内容填充到基础框架中,形成一个关于彩票开奖的新闻速报的正文,然后添加上题目就可以进行报道。

[0083] 本实施例提供的基于人工智能的新闻速报生成方法,通过基于历史新闻速报预先构建模板库,其中模板库中包括每个新闻类型用于生成新闻速报的基础框架;确定新闻事件隶属的目标新闻类型,根据目标新闻类型从模板库中,获取目标新闻类型的目标基础框架,从预设的数据库中获取新闻事件的结构化数据,根据结构化数据向目标基础框架中填充内容,得到新闻事件的新闻速报的正文,根据正文生成新闻速报的标题,将标题与正文拼接形成新闻速报。本实施例中,通过对历史新闻速报的学习,可以为不同的新闻类型构建基础框架,然后基于基础框架智能地生成新闻速报,提高新闻事件的报道的时效性。

[0084] 图4为本发明实施例提供的另一种基于人工智能的新闻速报生成方法的流程示意图。在上述实施例的基础之上,在形成所述该基于人工智能的新闻速报生成方法包括以下步骤:

[0085] S301、利用润色语句集中对正文中的同一语义的描述性短语进行随机替换。

[0086] 本实施例中,针对同一新闻类型,可以在形成基础框架之后,对预设数量的历史新闻速报中每两个描述性短语的语音相似度进行挖掘,然后根据获取到的语义相似度可以确定出表示同一语义的所有描述性短语,例如可以设置一个阈值,将两个短语之间的语义相似度与阈值进行比较,当两个短语之间的语义相似度高于阈值时,可以认定两个短语表示同一语义。进一步地,利用表示同一语义的所有描述性短语生成该同一语义的润色语句集,不同的语义生成不同的润色语句集。进一步地,可以通过所有的润色语句集对生成的新闻速报的内容进行优化和润色。例如,表示胜利这一语义的润色语句集中可以包括“大胜”、“完胜”、“轻取”、“险胜”、“逆袭”、“碾压”等。

[0087] 本实施例中,为了使智能生成的新闻速报更加多样性,避免针对同一类型的新闻事件的速报出现统一的报道样式,因此,可以在形成新闻速报的正文之后,利用润色语句集对正文中的同一语义的描述性短语进行随机替换,以实现正文的优化。

[0088] S302、利用同义词表中对正文中的同一对象的结构化数据进行随机替换。

[0089] 进一步地,针对同一新闻类型,可以在形成基础框架之后,对预设数量的历史新闻速报中结构化数据进行挖掘,得到表示对象的结构化数据,然后可以识别出用于同一对象的所有结构化数据,然后利用同一对象的所有结构化数据生成该对象的同义词表。例如,某

一球星,该球星的名字就是一个结构化数据,而且该结构化数据为一个名词性结构化数据,该球星可能还有其他的昵称,昵称也可以用来上表示该球星,从而球星的名字、昵称都可以表示该球星,则名字、昵称则可以形成一个与该球星相关的同义词表,可以通过同义词表对生成的新闻速报的内容进行优化和润色。

[0090] 本实施例中,避免针对同一类型的新闻事件的速报出现统一的报道样式,新闻速报机械生成化的痕迹较重,因此,可以在形成新闻速报的正文之后,利用同一词表对正文中的同一对象的结构化数据进行随机替换,以实现正文的优化,使得为了使智能生成的新闻速报更加多样性。例如将球星的名字替换成球星的昵称等。

[0091] 实际应用中,通过人工编辑生成新闻速报时,可能因为人工编辑的成本过高,为了保证新闻速报的超高时效性,人工编辑时只能承担热点事件的跟踪,例如热点赛色、热点股票或者热点彩票的报道,缺少对长尾比赛、股票、彩票等内容的覆盖,使得无法满足长尾用户的需求。而本实施例中,由于新闻速报是基于基础框架智能生成的,不再依赖人工进行编辑,从而可以对长尾比赛、股票、彩票等新闻事件的覆盖,能够满足长尾用户的需求。

[0092] 此处需要说明,可以单独执行S301也可以单独执行302,也可以先执行S302,再执行S301,来实现对新闻速报的润色和完善,以提高智能生成新闻速报的质量。

[0093] 图5为本发明实施例提供的一种基于人工智能的新闻速报生成装置的结构示意图。该基于人工智能的新闻速报生成装置包括:模板库构建模块11、类型确定模块12、框架获取模块13、数据获取模块14、填充模块15和生成模块16。

[0094] 其中,模板库构建模块11,用于基于历史新闻速报预先构建模板库,所述模板库中包括每个新闻类型用于生成新闻速报的基础框架。

[0095] 类型确定模块12,用于确定新闻事件隶属的目标新闻类型。

[0096] 框架获取模块13,用于根据所述目标新闻类型从所述模板库中,获取所述目标新闻类型的目标基础框架。

[0097] 数据获取模块14,用于从预设的数据库中获取所述新闻事件的结构化数据。

[0098] 填充模块15,用于根据所述结构化数据向所述目标基础框架中填充内容,得到所述新闻事件的新闻速报的正文。

[0099] 生成模块16,用于根据所述正文生成所述新闻速报的标题,将所述标题与所述正文拼接形成所述新闻速报。

[0100] 进一步地,模板库构建模块11,具体用于针对同一新闻类型,对所有历史新闻速报进行提取,得到所述历史新闻速报中的结构化数据与新闻语料之间的对应关系,利用所述历史新闻速报中的结构化数据、新闻语料以及所述对应关系形成训练集,对所述训练集进行文本训练,挖掘各历史新闻速报中可替换的结构化数据,形成所述基础框架。

[0101] 本实施例提供的基于人工智能的新闻速报生成装置,通过基于历史新闻速报预先构建模板库,其中模板库中包括每个新闻类型用于生成新闻速报的基础框架;确定新闻事件隶属的目标新闻类型,根据目标新闻类型从模板库中,获取目标新闻类型的目标基础框架,从预设的数据库中获取新闻事件的结构化数据,根据结构化数据向目标基础框架中填充内容,得到新闻事件的新闻速报的正文,根据正文生成新闻速报的标题,将标题与正文拼接形成新闻速报。本实施例中,通过对历史新闻速报的学习,可以为不同的新闻类型构建基础框架,然后基于基础框架智能地生成新闻速报,提高新闻事件的报道的时效性。

[0102] 图6为本发明实施例提供的另一种基于人工智能的新闻速报生成装置的结构示意图。在上述实施例的基础之上,该基于人工智能的新闻速报生成装置还包括:润色模块17和优化模块18。

[0103] 进一步地,模板库构建模块11,还用于挖掘预设数量的所述历史新闻速报中每两个描述性短语的语义相似度,根据所述语义相似度获取表示同一语义的所有描述性短语,利用表示同一语义的所有描述性短语生成该同一语义的润色语句集。

[0104] 润色模块17,用于利用所述润色语句集中对所述正文中的同一语义的描述性短语进行随机替换。

[0105] 进一步地,模板库构建模块11,还用于对预设数量的所述历史新闻速报中的结构化数据进行挖掘,得到表示对象的结构化数据;其中,所述对象的结构化数据为名词性的结构化数据,识别用于表示同一对象的所有结构化数据,利用同一对象的所有结构化数据生成该同一对象的同义词表。

[0106] 优化模块18,用于利用所述同义词表中对所述正文中的同一对象的所述结构化数据进行随机替换。

[0107] 进一步地,类型确定模块12,具体用于检测是否有所述新闻事件发生,当检测到所述新闻事件后,确定所述新闻事件隶属的所述目标新闻类型。

[0108] 进一步地,填充模块15,具体用于基于知识图谱从所述数据库中获取与所述结构化数据相关的内容,向所述目标基础框架中填充所述结构化数据以及与所述结构化数据相关的内容,得到所述新闻速报的正文。

[0109] 本实施例中,通过对历史新闻速报的学习,可以为不同的新闻类型构建基础框架,然后基于基础框架智能地生成新闻速报,提高新闻事件的报道的时效性。进一步地,可以通过生成的润色语句集和同义词表,对生成的新闻速报进行润色和优化,提高新闻速报的可读性。进一步地,由于新闻速报是基于基础框架智能生成的,不再依赖人工进行编辑,从而可以对长尾比赛、股票、彩票等新闻事件的覆盖,能够满足长尾用户的需求。

[0110] 图7为本发明实施例提供的另一种基于人工智能的新闻速报生成装置的结构示意图。该基于人工智能的新闻速报生成装置包括:

[0111] 存储器21、处理器22及存储在存储器21上并可在处理器22上运行的计算机程序。

[0112] 处理器22执行所述程序时实现上述实施例中提供的基于人工智能的新闻速报生成方法。

[0113] 进一步地,基于人工智能的新闻速报生成装置还包括:

[0114] 通信接口23,用于存储器21和处理器22之间的通信。

[0115] 存储器21,用于存放可在处理器22上运行的计算机程序。

[0116] 存储器21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。

[0117] 处理器22,用于执行所述程序时实现上述实施例所述的业务管理方法。

[0118] 如果存储器21、处理器22和通信接口23独立实现,则通信接口21、存储器21和处理器22可以通过总线相互连接并完成相互间的通信。所述总线可以是工业标准体系结构(Industry Standard Architecture,简称为ISA)总线、外部设备互连(Peripheral Component,简称为PCI)总线或扩展工业标准体系结构(Extended Industry Standard

Architecture, 简称为EISA) 总线等。所述总线可以分为地址总线、数据总线、控制总线等。为便于表示, 图7中仅用一条粗线表示, 但并不表示仅有一根总线或一种类型的总线。

[0119] 可选的, 在具体实现上, 如果存储器21、处理器22及通信接口23, 集成在一块芯片上实现, 则存储器21、处理器22及通信接口23可以通过内部接口完成相互间的通信。

[0120] 处理器22可能是一个中央处理器 (Central Processing Unit, 简称为CPU), 或者是特定集成电路 (Application Specific Integrated Circuit, 简称为ASIC), 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。

[0121] 在本说明书的描述中, 参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本发明的至少一个实施例或示例中。在本说明书中, 对上述术语的示意性表述不必须针对的是相同的实施例或示例。而且, 描述的具体特征、结构、材料或者特点可以在任一个或多个实施例或示例中以合适的方式结合。此外, 在不相互矛盾的情况下, 本领域的技术人员可以将本说明书中描述的不同实施例或示例以及不同实施例或示例的特征进行结合和组合。

[0122] 此外, 术语“第一”、“第二”仅用于描述目的, 而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此, 限定有“第一”、“第二”的特征可以明示或者隐含地包括至少一个该特征。在本发明的描述中, “多个”的含义是至少两个, 例如两个, 三个等, 除非另有明确具体的限定。

[0123] 流程图中或在此以其他方式描述的任何过程或方法描述可以被理解为, 表示包括一个或更多个用于实现定制逻辑功能或过程的步骤的可执行指令的代码的模块、片段或部分, 并且本发明的优选实施方式的范围包括另外的实现, 其中可以不按所示出或讨论的顺序, 包括根据所涉及的功能按基本同时的方式或按相反的顺序, 来执行功能, 这应被本发明的实施例所属技术领域的技术人员所理解。

[0124] 在流程图中表示或在此以其他方式描述的逻辑和/或步骤, 例如, 可以被认为用于实现逻辑功能的可执行指令的定序列表, 可以具体实现在任何计算机可读介质中, 以供指令执行系统、装置或设备 (如基于计算机的系统、包括处理器的系统或其他可以从指令执行系统、装置或设备取指令并执行指令的系统) 使用, 或结合这些指令执行系统、装置或设备而使用。就本说明书而言, “计算机可读介质”可以是任何可以包含、存储、通信、传播或传输程序以供指令执行系统、装置或设备或结合这些指令执行系统、装置或设备而使用的装置。计算机可读介质的更具体的示例 (非穷尽性列表) 包括以下: 具有一个或多个布线的电连接部 (电子装置), 便携式计算机盘盒 (磁装置), 随机存取存储器 (RAM), 只读存储器 (ROM), 可擦除可编程只读存储器 (EPROM或闪速存储器), 光纤装置, 以及便携式光盘只读存储器 (CDROM)。另外, 计算机可读介质甚至可以是可在其上打印所述程序的纸或其他合适的介质, 因为可以例如通过对纸或其他介质进行光学扫描, 接着进行编辑、解译或必要时以其他合适方式进行处理来以电子方式获得所述程序, 然后将其存储在计算机存储器中。

[0125] 应当理解, 本发明的各部分可以用硬件、软件、固件或它们的组合来实现。在上述实施方式中, 多个步骤或方法可以用存储在存储器中且由合适的指令执行系统执行的软件或固件来实现。如, 如果用硬件来实现和在另一实施方式中一样, 可用本领域公知的下列技术中的任一项或他们的组合来实现: 具有用于对数据信号实现逻辑功能的逻辑门电路的离

散逻辑电路,具有合适的组合逻辑门电路的专用集成电路,可编程门阵列(PGA),现场可编程门阵列(FPGA)等。

[0126] 本技术领域的普通技术人员可以理解实现上述实施例方法携带的全部或部分步骤是可以通过程序来指令相关的硬件完成,所述的程序可以存储于一种计算机可读存储介质中,该程序在执行时,包括方法实施例的步骤之一或其组合。

[0127] 此外,在本发明各个实施例中的各功能单元可以集成在一个处理模块中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个模块中。上述集成的模块既可以采用硬件的形式实现,也可以采用软件功能模块的形式实现。所述集成的模块如果以软件功能模块的形式实现并作为独立的产品销售或使用,也可以存储在一个计算机可读取存储介质中。

[0128] 上述提到的存储介质可以是只读存储器,磁盘或光盘等。尽管上面已经示出和描述了本发明的实施例,可以理解的是,上述实施例是示例性的,不能理解为对本发明的限制,本领域的普通技术人员在本发明的范围内可以对上述实施例进行变化、修改、替换和变型。

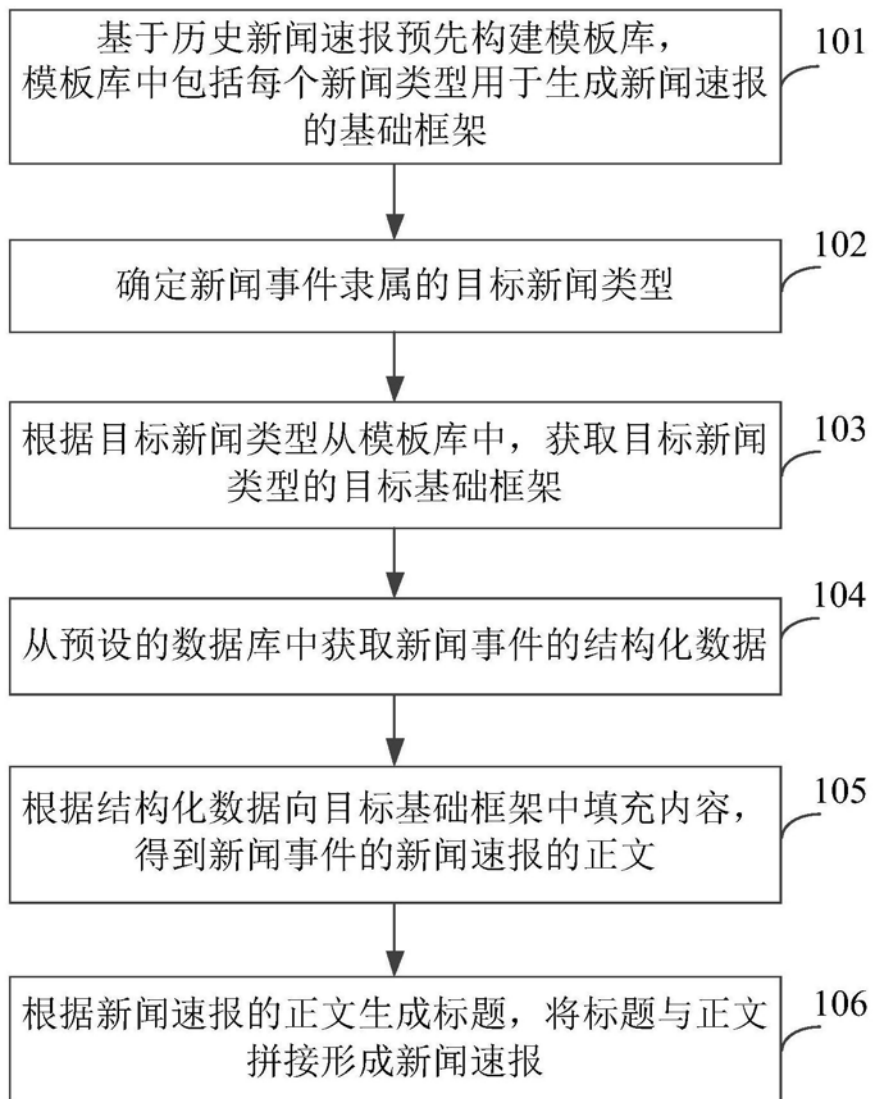


图1

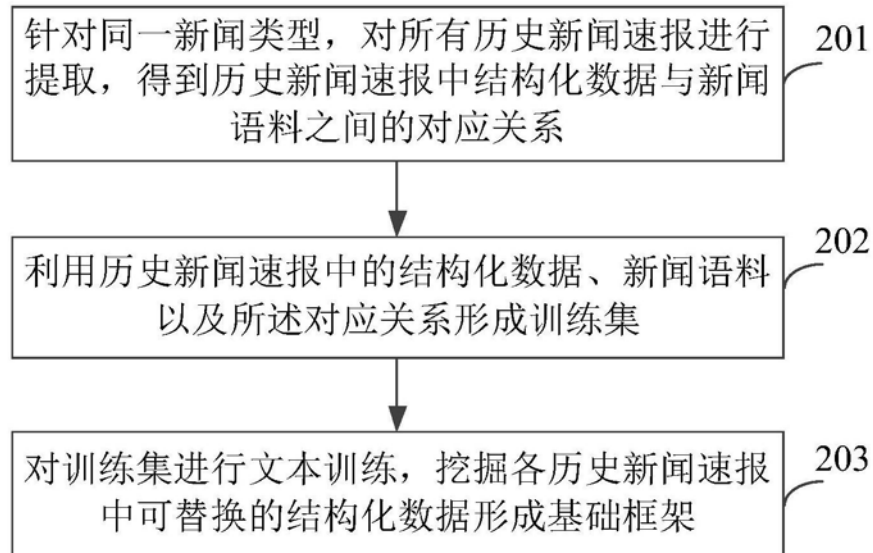


图2

开奖详情

双色球第 *** 期 *** 开奖，本期中奖号码为：

红球：***；篮球：***

本期一等奖：*注，每注 *** 元

本期二等奖：*注，每注 *** 元

目前累积奖池奖金 * 亿 * 万元。

选号分析

最近5期以来

红球表现

热号分析（出现5次、4次）：从5期走势可以看出，热码选号范围不足，目前暂无热号

温号分析（出现3次、2次）：近5期出号频率来看，温码选号范围充裕，目前温号有 *

冷号分析（出现1次）：冷号近期出号频率来看，冷码选号范围充裕，目前冷号有 **

***** 近5期没有出现，值得关注

图3

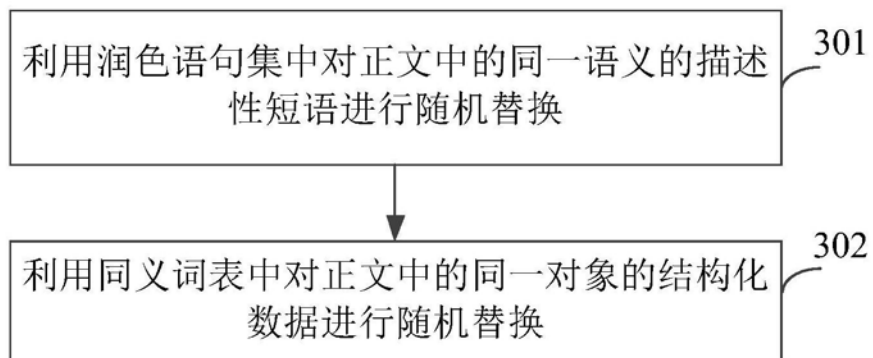


图4

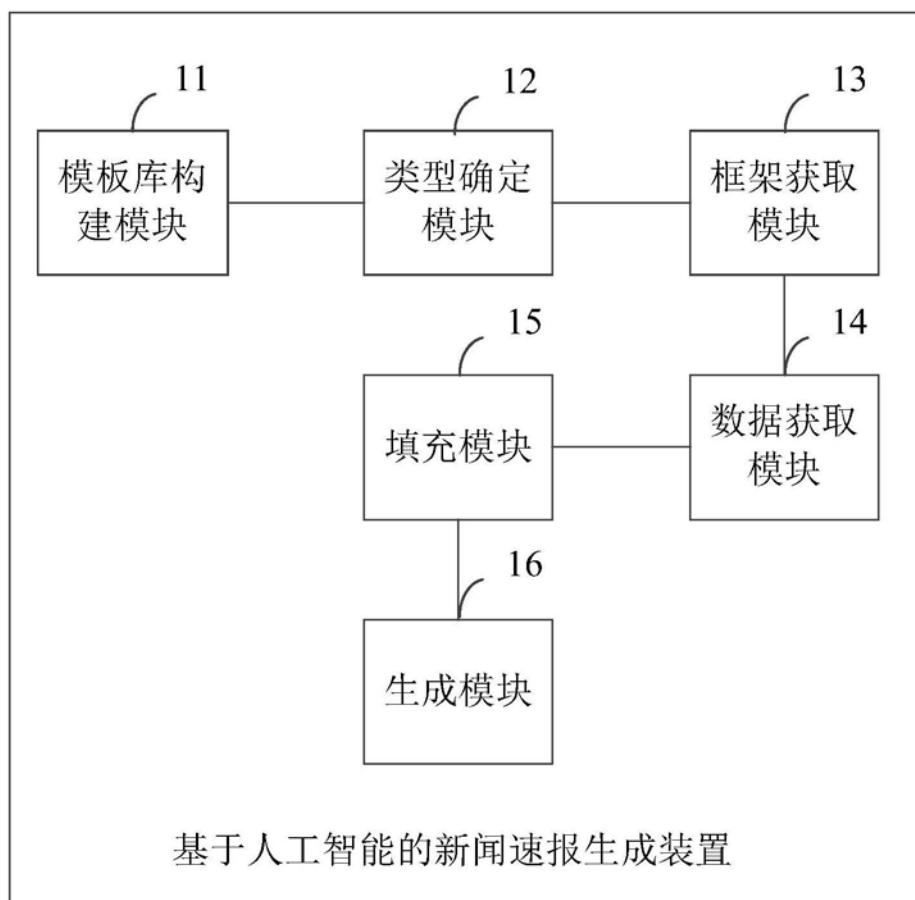


图5

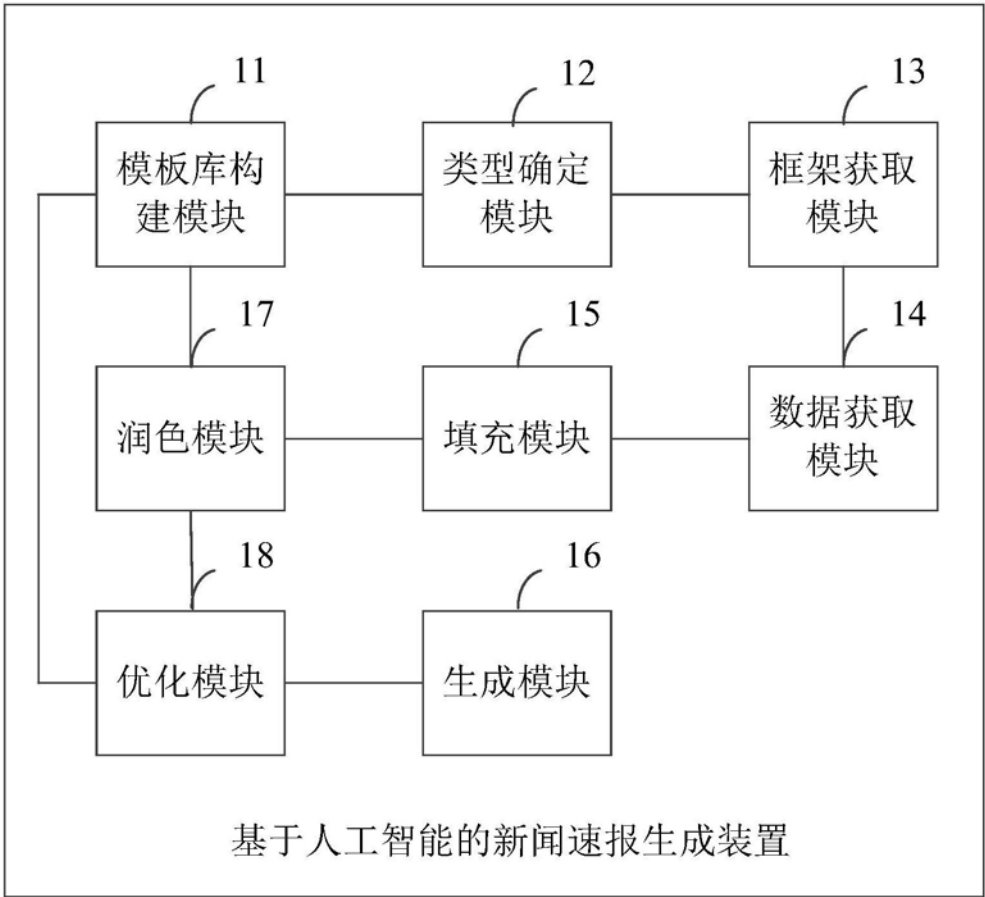


图6

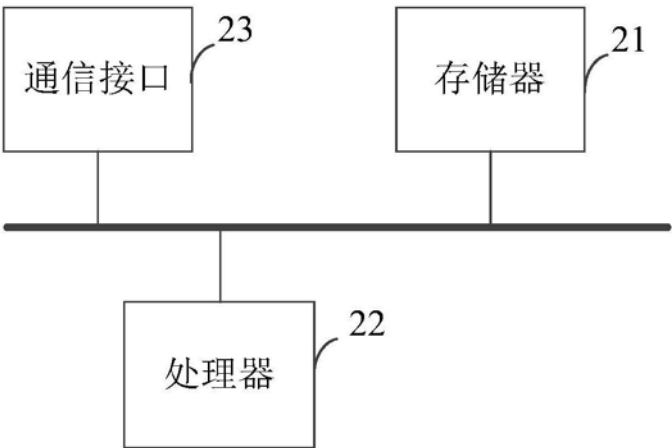


图7