



(12) 发明专利申请

(10) 申请公布号 CN 112446214 A

(43) 申请公布日 2021.03.05

(21) 申请号 202011450409.5

(22) 申请日 2020.12.09

(71) 申请人 北京有竹居网络技术有限公司
地址 101299 北京市平谷区林荫北街13号
信息大厦802室

(72) 发明人 陈陌信

(74) 专利代理机构 北京远智汇知识产权代理有限公司 11659
代理人 范坤坤

(51) Int. Cl.
G06F 40/289 (2020.01)
G06Q 30/02 (2012.01)

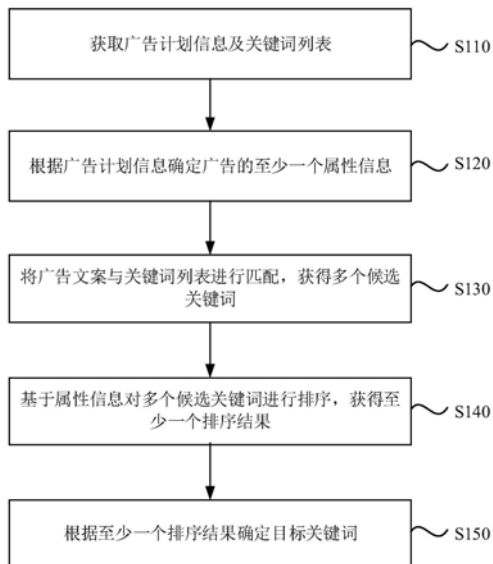
权利要求书2页 说明书9页 附图2页

(54) 发明名称

广告关键词的生成方法、装置、设备及存储介质

(57) 摘要

本公开实施例公开了一种广告关键词的生成方法、装置、设备及存储介质。包括：获取广告计划信息及关键词列表；所述广告计划信息包括广告文案；根据所述广告计划信息确定所述广告的至少一个属性信息；将所述广告文案与所述关键词列表进行匹配，获得多个候选关键词；基于所述属性信息对所述多个候选关键词进行排序，获得至少一个排序结果；根据所述至少一个排序结果确定目标关键词。本公开实施例提供的广告关键词的生成方法，基于广告的属性信息对候选关键词进行排序，并根据至少一个排序结果确定目标关键词，可以提高广告关键词的生成效率及质量。



1. 一种广告关键词的生成方法,其特征在于,包括:
获取广告计划信息及关键词列表;所述广告计划信息包括广告文案;
根据所述广告计划信息确定所述广告的至少一个属性信息;
将所述广告文案与所述关键词列表进行匹配,获得多个候选关键词;
基于所述属性信息对所述多个候选关键词进行排序,获得至少一个排序结果;
根据所述至少一个排序结果确定目标关键词。
2. 根据权利要求1所述的方法,其特征在于,将所述广告文案与所述关键词列表进行匹配,获得多个候选关键词,包括:
遍历所述关键词列表,将遍历到的关键词与所述广告文案进行比对;
若广告文案中存在遍历到的关键词或者存在与遍历到的关键词语义相近的词,则将遍历到的关键词确定为候选关键词。
3. 根据权利要求1所述的方法,其特征在于,基于所述属性信息对所述多个候选关键词进行排序,获得多个排序结果,包括:
对于每个候选关键词,确定所述候选关键词分别与各属性信息对应的置信度;
对于每个属性信息,根据所述置信度对所述多个候选关键词进行排序,获得至少一个排序结果。
4. 根据权利要求3所述的方法,其特征在于,确定所述候选关键词分别与各属性信息对应的置信度,包括:
将所述候选关键词分别输入各属性信息分别对应的识别模型中,获得所述候选关键词分别与各属性信息对应的置信度;其中,属性信息包括行业类别和商品类别;识别模型包括行业类别识别模型和商品类别识别模型。
5. 根据权利要求4所述的方法,其特征在于,所述识别模型的训练方式为:
获取训练数据;所述训练数据包括广告关键词以及所述广告关键词对应的真实属性信息;
将所述广告关键词输入所述识别模型中,获得预测属性信息;
根据所述真实属性信息和所述预测属性信息确定损失函数;
基于所述损失函数对所述识别模型进行训练。
6. 根据权利要求1所述的方法,其特征在于,根据所述至少一个排序结果确定目标关键词,包括:
对所述至少一个排序结果进行融合,获得目标排序结果;
从所述目标排序结果中提取设定数量的目标关键词。
7. 根据权利要求6所述的方法,其特征在于,对所述至少一个排序结果进行融合,获得目标排序结果,包括:
对候选关键词在各排序结果中的序号进行加权求和,获得各候选关键词的目标序号;
基于所述目标序号对所述候选关键词进行重新排序,获得目标排序结果;
相应的,从所述目标排序结果中提取设定数量的目标关键词,包括:
若是按照置信度由大到小排序,则从所述目标排序结果中提取排序靠前设定数量的目标关键词;若是按照置信度由小到大排序,则从所述目标排序结果中提取排序靠后设定数量的目标关键词。

8. 一种广告关键词的生成装置,其特征在于,包括:

关键词列表获取模块,用于获取广告计划信息及关键词列表;所述广告计划信息包括广告文案;

属性信息确定模块,用于根据所述广告计划信息确定所述广告的至少一个属性信息;

候选关键词获取模块,用于将所述广告文案与所述关键词列表进行匹配,获得多个候选关键词;

排序结果获取模块,用于基于所述属性信息对所述多个候选关键词进行排序,获得至少一个排序结果;

目标关键词确定模块,用于根据所述至少一个排序结果确定目标关键词。

9. 一种电子设备,其特征在于,所述电子设备包括:

一个或多个处理装置;

存储装置,用于存储一个或多个程序;

当所述一个或多个程序被所述一个或多个处理装置执行,使得所述一个或多个处理装置实现如权利要求1-7中任一所述的广告关键词的生成方法。

10. 一种计算机可读介质,其上存储有计算机程序,其特征在于,该程序被处理装置执行时实现如权利要求1-7中任一所述的广告关键词的生成方法。

广告关键词的生成方法、装置、设备及存储介质

技术领域

[0001] 本公开实施例涉及广告技术领域,尤其涉及一种广告关键词的生成方法、装置、设备及存储介质。

背景技术

[0002] 广告关键词是广告内容的高度提炼和概括,在广告投放环节中,通过关键词进行检索和匹配具有很强的针对性,能够大幅提高广告的投放效果。获得高质量广告关键词的方法之一是人工对每个广告打上关键词标签,但是人工标注的成本高昂,扩展到多语言的业务场景中标注成本更是急剧增加。另一种方法是通过文本匹配的方法从关键词候选集中为每个广告匹配出若干个关键词。匹配的方法虽然能够大幅缩减成本,但是无法保证匹配结果的质量。

发明内容

[0003] 本公开实施例提供一种广告关键词的生成方法、装置、设备及存储介质,可以提高广告关键词的生成效率及质量。

[0004] 第一方面,本公开实施例提供了一种广告关键词的生成方法,包括:

[0005] 获取广告计划信息及关键词列表;所述广告计划信息包括广告文案;

[0006] 根据所述广告计划信息确定所述广告的至少一个属性信息;

[0007] 将所述广告文案与所述关键词列表进行匹配,获得多个候选关键词;

[0008] 基于所述属性信息对所述多个候选关键词进行排序,获得至少一个排序结果;

[0009] 根据所述至少一个排序结果确定目标关键词。

[0010] 第二方面,本公开实施例还提供了一种广告关键词的生成装置,包括:

[0011] 关键词列表获取模块,用于获取广告计划信息及关键词列表;所述广告计划信息包括广告文案;

[0012] 属性信息确定模块,用于根据所述广告计划信息确定所述广告的至少一个属性信息;

[0013] 候选关键词获取模块,用于将所述广告文案与所述关键词列表进行匹配,获得多个候选关键词;

[0014] 排序结果获取模块,用于基于所述属性信息对所述多个候选关键词进行排序,获得至少一个排序结果;

[0015] 目标关键词确定模块,用于根据所述至少一个排序结果确定目标关键词。

[0016] 第三方面,本公开实施例还提供了一种电子设备,所述电子设备包括:

[0017] 一个或多个处理装置;

[0018] 存储装置,用于存储一个或多个程序;

[0019] 当所述一个或多个程序被所述一个或多个处理装置执行,使得所述一个或多个处理装置实现如本公开实施例所述的广告关键词的生成方法。

[0020] 第四方面,本公开实施例还提供了一种计算机可读介质,其上存储有计算机程序,该程序被处理装置执行时实现如本公开实施例所述的广告关键词的生成方法。

[0021] 本公开实施例公开了一种广告关键词的生成方法、装置、设备及存储介质。首先获取广告计划信息及关键词列表,然后根据广告计划信息确定广告的至少一个属性信息,再然后将广告文案与关键词列表进行匹配,获得多个候选关键词,再然后基于属性信息对多个候选关键词进行排序,获得至少一个排序结果,最后根据至少一个排序结果确定目标关键词。本公开实施例提供的广告关键词的生成方法,基于广告的属性信息对候选关键词进行排序,并根据至少一个排序结果确定目标关键词,可以提高广告关键词的生成效率及质量。

附图说明

[0022] 图1是本公开实施例中的一种广告关键词的生成方法的流程图;

[0023] 图2是本公开实施例中的一种广告关键词的生成装置的结构示意图;

[0024] 图3是本公开实施例中的一种电子设备的结构示意图。

具体实施方式

[0025] 下面将参照附图更详细地描述本公开的实施例。虽然附图中显示了本公开的某些实施例,然而应当理解的是,本公开可以通过各种形式来实现,而且不应该被解释为限于这里阐述的实施例,相反提供这些实施例是为了更加透彻和完整地理解本公开。应当理解的是,本公开的附图及实施例仅用于示例性作用,并非用于限制本公开的保护范围。

[0026] 应当理解,本公开的方法实施方式中记载的各个步骤可以按照不同的顺序执行,和/或并行执行。此外,方法实施方式可以包括附加的步骤和/或省略执行示出的步骤。本公开的范围在此方面不受限制。

[0027] 本文使用的术语“包括”及其变形是开放性包括,即“包括但不限于”。术语“基于”是“至少部分地基于”。术语“一个实施例”表示“至少一个实施例”;术语“另一实施例”表示“至少一个另外的实施例”;术语“一些实施例”表示“至少一些实施例”。其他术语的相关定义将在下文描述中给出。

[0028] 需要注意,本公开中提及的“第一”、“第二”等概念仅用于对不同的装置、模块或单元进行区分,并非用于限定这些装置、模块或单元所执行的功能的顺序或者相互依存关系。

[0029] 需要注意,本公开中提及的“一个”、“多个”的修饰是示意性而非限制性的,本领域技术人员应当理解,除非在上下文另有明确指出,否则应该理解为“一个或多个”。

[0030] 本公开实施方式中的多个装置之间所交互的消息或者信息的名称仅用于说明性的目的,而并不是用于对这些消息或信息的范围进行限制。

[0031] 图1为本公开实施例提供的一种广告关键词的生成方法的流程图,本实施例可适用于提取广告中关键词的情况,该方法可以由广告关键词的生成装置来执行,该装置可由硬件和/或软件组成,并一般可集成在具有广告关键词的生成功能的设备中,该设备可以是服务器、移动终端或服务器集群等电子设备。如图1所示,该方法具体包括如下步骤:

[0032] 步骤110,获取广告计划信息及关键词列表。

[0033] 其中,广告计划信息包括广告文案。广告计划信息还可以包括广告的音视频、图片

等信息。关键词列表可以预先建立的关键词库,可以通过收集已发布广告的关键词形成。

[0034] 步骤120,根据广告计划信息确定广告的至少一个属性信息。

[0035] 其中,属性信息可以包括行业类别及商品类别等。例如:行业类别可以是美妆类、食品类、服装类、家电类及生活用品类;商品类别可以是广告涉及的商品的具体类别,例如:洗发水、口红、衬衫、冰箱等。本实施例中,根据广告计划信息确定广告的至少一个属性信息的方式可以是,通过人工的方式对广告计划信息进行分析,获得广告的至少一个属性信息;或者将广告计划信息输入属性识别模型中进行识别,获得广告的至少一个属性信息。示例性的,某个广告的属性信息为行业类别是美妆类、商品类别为防晒霜。

[0036] 步骤130,将广告文案与关键词列表进行匹配,获得多个候选关键词。

[0037] 其中,广告文案与关键词列表进行匹配的方式可以是将关键词列表中的每个关键词与广告文案进行逐行比对,判断广告文案中是否出现关键词列表中的关键词,若出现,则将该关键词确定为候选关键词。

[0038] 具体的,将广告文案与关键词列表进行匹配,获得多个候选关键词的方式可以是:遍历关键词列表,将遍历到的关键词与广告文案进行比对;若广告文案中存在遍历到的关键词或者存在与遍历到的关键词语义相近的词,则将遍历到的关键词确定为候选关键词。

[0039] 其中,语义相近可以理解为两个词在特定语言环境下,表达的意思相同。例如:防晒乳与防晒霜,就可以理解为语义相近的词。

[0040] 步骤140,基于属性信息对多个候选关键词进行排序,获得至少一个排序结果。

[0041] 其中,基于属性信息对多个候选关键词进行排序可以理解为:将多个候选关键词分别按照各个属性信息进行排序。

[0042] 具体的,基于属性信息对多个候选关键词进行排序,获得多个排序结果的方式可以是:对于每个候选关键词,确定候选关键词分别与各属性信息对应的置信度;对于每个属性信息,根据置信度对候选关键词进行排序,获得多个排序结果。

[0043] 其中,确定候选关键词分别与各属性信息对应的置信度的方式可以是:将候选关键词输入各属性信息分别对应的识别模型中,获得候选关键词与各属性信息对应的置信度。

[0044] 本实施例中,属性信息包括行业类别和商品类别;识别模型包括行业类别识别模型和商品类别识别模型。行业类别识别模型的输入为广告关键词,输出为该广告关键词对应的各行业类别的置信度;商品类别识别模型的输入为广告关键词,输出为该广告关键词对应的各商品类别的置信度。

[0045] 将各候选关键词分别输入行业类别识别模型和商品类别识别模型,则获得各候选关键词的行业类别的置信度以及商品类别的置信度。假设本实施例中的广告的属性信息为行业类别是服装类及商品类别是男士衬衫,则提取各候选关键词属于服装类的置信度以及属于男士衬衫的置信度。示例性的,假设候选关键词包括5个,分别为关键词a、关键词b、关键词c、关键词d和关键词e,这5个候选关键词属于服装类的置信度分别为50%、30%、60%、45%及80%。属于男士衬衫的置信度分别为75%、55%、20%、80%及30%。则按照行业类别的置信度由大到小进行排序的排序结果为:关键词e、关键词c、关键词a、关键词d和关键词b;按照商品类别的置信度由大到小进行排序的排序结果为:关键词d、关键词a、关键词b、关键词e和关键词c。

[0046] 可选的,属性信息识别模型是基于训练样本对设定机器学习算法训练获得的。属性识别模型的训练方式为:获取训练数据;将广告关键词输入识别模型中,获得预测属性信息;根据真实属性信息和预测属性信息确定损失函数;基于损失函数对识别模型进行训练。

[0047] 其中,训练数据包括广告关键词以及广告关键词对应的真实属性信息。具体的,对于行业类别识别模型的训练方式为:将广告关键词与真实行业类别组成训练数据,将广告关键词输入行业类别识别模型中,获得预测行业类别,根据真实行业类别和预测行业类别确定损失函数,基于损失函数对行业类别识别模型进行训练。对于商品类别识别模型的训练方式为:将广告关键词与真实商品类别组成训练数据,将广告关键词输入商品类别识别模型中,获得预测商品类别,根据真实商品类别和预测商品类别确定损失函数,基于损失函数对商品类别识别模型进行训练。

[0048] 步骤150,根据至少一个排序结果确定目标关键词。

[0049] 具体的,在获得各属性信息分别对应的排序结果后,将各排序结果进行融合,获得目标排序结果,最后从目标排序结果中提取设定数量的目标关键词。

[0050] 对多个排序结果进行融合,获得目标排序结果的方式可以是:对候选关键词在各排序结果中的序号进行加权求和,获得各候选关键词的目标序号;基于目标序号对候选关键词进行重新排序,获得目标排序结果。

[0051] 本实施例中,可以先确定每个属性信息的权重,然后对序号按照确定的权重进行加权求和。可选的,可以直接对序号求取平均值。示例性的,假设按照行业类别的置信度由大到小进行排序的排序结果为:关键词e、关键词c、关键词a、关键词d和关键词b;按照商品类别的置信度由大到小进行排序的排序结果为:关键词d、关键词a、关键词b、关键词e和关键词c。对于关键词a,在行业类别的排序结果中序号为第3,在商品类别的排序结果中的序号为第2,那么目标序号为2.5。同样,对于关键词b,计算出的目标序号为4;关键词c的目标序号为3.5;关键词d的目标序号为2.5;关键词e的目标序号为2.5。那么目标排序结果为:关键词b、关键词c、关键词a、关键词d、关键词e。

[0052] 具体的,从目标排序结果中提取设定数量的目标关键词的方式可以是:若是按照置信度由大到小排序,则从目标排序结果中提取排序靠前设定数量的目标关键词;若是按照置信度由小到大排序,则从目标排序结果中提取排序靠后设定数量的目标关键词。

[0053] 其中,设定数量可以设置为5-10之间的任意值。

[0054] 本公开实施例的技术方案,首先获取广告计划信息及关键词列表,然后根据广告计划信息确定广告的至少一个属性信息,再然后将广告文案与关键词列表进行匹配,获得多个候选关键词,再然后基于属性信息对多个候选关键词进行排序,获得至少一个排序结果,最后根据至少一个排序结果确定目标关键词。本公开实施例提供的广告关键词的生成方法,基于广告的属性信息对候选关键词进行排序,并根据至少一个排序结果确定目标关键词,可以提高广告关键词的生成效率及质量。

[0055] 下面以一个具体的广告实例解释本实施例的广告关键词的生成方法:

[0056] 广告文案为:「韩国直邮」毛孔隐形妆前乳,360°美到无死角!闺蜜力荐!该广告的行业类别为“美妆类”、商品类别为“妆前乳”。将广告文案与关键词类别进行匹配,获得的候选关键词依次为:韩国直邮、毛孔、妆前乳、无死角、闺蜜。对于每个候选关键词,确定其与行业类别及商品类别对应的置信度。对于广告的行业类别“美妆类”,这些候选关键词与美妆

类对应的置信度依次为:0.91、0.94、0.98、0.82、0.71,并根据置信度对候选关键词进行排序:妆前乳、毛孔、韩国直邮、无死角、闺蜜;对于广告的商品类别“妆前乳”,这些候选关键词与妆前乳对应的置信度依次为:0.93、0.87、0.99、0.90、0.78,并根据执行度对候选关键词进行排序:妆前乳、韩国直邮、无死角、毛孔、闺蜜。对两个排序结果进行融合,获得最终的排序结果为:妆前乳、韩国直邮、毛孔、无死角、闺蜜。最后提取排序靠前的两个候选关键词作为目标关键词,即最终结果为:妆前乳、韩国直邮。

[0057] 图2是本公开实施例公开的一种广告关键词的生成装置的结构示意图。如图2所示,该装置包括:

[0058] 关键词列表获取模块210,用于获取广告计划信息及关键词列表;广告计划信息包括广告文案;

[0059] 属性信息确定模块220,用于根据广告计划信息确定广告的至少一个属性信息;

[0060] 候选关键词获取模块230,用于将广告文案与关键词列表进行匹配,获得多个候选关键词;

[0061] 排序结果获取模块240,用于基于属性信息对多个候选关键词进行排序,获得至少一个排序结果;

[0062] 目标关键词确定模块250,用于根据至少一个排序结果确定目标关键词。

[0063] 可选的,候选关键词获取模块230,还用于:

[0064] 遍历关键词列表,将遍历到的关键词与广告文案进行比对;

[0065] 若广告文案中存在遍历到的关键词或者存在与遍历到的关键词语义相近的词,则将遍历到的关键词确定为候选关键词。

[0066] 可选的,排序结果获取模块240,还用于:

[0067] 对于每个候选关键词,确定所述候选关键词分别与各属性信息对应的置信度;

[0068] 对于每个属性信息,根据置信度对候选关键词进行排序,获得至少一个排序结果。

[0069] 可选的,排序结果获取模块240,还用于:

[0070] 将候选关键词分别输入各属性信息分别对应的识别模型中,获得候选关键词分别与各属性信息对应的置信度;其中,属性信息包括行业类别和商品类别;识别模型包括行业类别识别模型和商品类别识别模型。

[0071] 可选的,还包括:识别模型训练模块,用于:

[0072] 获取训练数据;所述训练数据包括广告关键词以及所述广告关键词对应的真实属性信息;

[0073] 将所述广告关键词输入所述识别模型中,获得预测属性信息;

[0074] 根据所述真实属性信息和所述预测属性类别确定损失函数;

[0075] 基于所述损失函数对所述识别模型进行训练。

[0076] 可选的,目标关键词确定模块250,还用于:

[0077] 对至少一个排序结果进行融合,获得目标排序结果;

[0078] 从目标排序结果中提取设定数量的目标关键词。

[0079] 可选的,目标关键词确定模块250,还用于:

[0080] 对候选关键词在各排序结果中的序号进行加权求和,获得各候选关键词的目标序号;

[0081] 基于目标序号对候选关键词进行重新排序,获得目标排序结果;

[0082] 可选的,目标关键词确定模块250,还用于:

[0083] 若是按照置信度由大到小排序,则从目标排序结果中提取排序靠前设定数量的目标关键词;若是按照置信度由小到大排序,则从目标排序结果中提取排序靠后设定数量的目标关键词。

[0084] 上述装置可执行本公开前述所有实施例所提供的方法,具备执行上述方法相应的功能模块和有益效果。未在本实施例中详尽描述的技术细节,可参见本公开前述所有实施例所提供的方法。

[0085] 下面参考图3,其示出了适于用来实现本公开实施例的电子设备300的结构示意图。本公开实施例中的电子设备可以包括但不限于诸如移动电话、笔记本电脑、数字广播接收器、PDA(个人数字助理)、PAD(平板电脑)、PMP(便携式多媒体播放器)、车载终端(例如车载导航终端)等等的移动终端以及诸如数字TV、台式计算机等等的固定终端,或者各种形式的服务器,如独立服务器或者服务器集群。图3示出的电子设备仅仅是一个示例,不应对本公开实施例的功能和使用范围带来任何限制。

[0086] 如图3所示,电子设备300可以包括处理装置(例如中央处理器、图形处理器等)301,其可以根据存储在只读存储装置(ROM)302中的程序或者从存储装置305加载到随机访问存储装置(RAM)303中的程序而执行各种适当的动作和处理。在RAM 303中,还存储有电子设备300操作所需的各种程序和数据。处理装置301、ROM 302以及RAM 303通过总线304彼此相连。输入/输出(I/O)接口305也连接至总线304。

[0087] 通常,以下装置可以连接至I/O接口305:包括例如触摸屏、触摸板、键盘、鼠标、摄像头、麦克风、加速度计、陀螺仪等的输入装置306;包括例如液晶显示器(LCD)、扬声器、振动器等的输出装置307;包括例如磁带、硬盘等的存储装置308;以及通信装置309。通信装置309可以允许电子设备300与其他设备进行无线或有线通信以交换数据。虽然图3示出了具有各种装置的电子设备300,但是应理解的是,并不要求实施或具备所有示出的装置。可以替代地实施或具备更多或更少的装置。

[0088] 特别地,根据本公开的实施例,上文参考流程图描述的过程可以被实现为计算机软件程序。例如,本公开的实施例包括一种计算机程序产品,其包括承载在计算机可读介质上的计算机程序,该计算机程序包含用于执行词语的推荐方法的程序代码。在这样的实施例中,该计算机程序可以通过通信装置309从网络上被下载和安装,或者从存储装置305被安装,或者从ROM 302被安装。在该计算机程序被处理装置301执行时,执行本公开实施例的方法中限定的上述功能。

[0089] 需要说明的是,本公开上述的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本

公开中,计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号,其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式,包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读信号介质还可以是计算机可读存储介质以外的任何计算机可读介质,该计算机可读信号介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输,包括但不限于:电线、光缆、RF(射频)等等,或者上述的任意合适的组合。

[0090] 在一些实施方式中,客户端、服务器可以利用诸如HTTP(HyperText Transfer Protocol,超文本传输协议)之类的任何当前已知或未来研发的网络协议进行通信,并且可以与任意形式或介质的数字数据通信(例如,通信网络)互连。通信网络的示例包括局域网(“LAN”),广域网(“WAN”),网际网(例如,互联网)以及端对端网络(例如,ad hoc端对端网络),以及任何当前已知或未来研发的网络。

[0091] 上述计算机可读介质可以是上述电子设备中所包含的;也可以是单独存在,而未装配入该电子设备中。

[0092] 上述计算机可读介质承载有一个或者多个程序,当上述一个或者多个程序被该电子设备执行时,使得该电子设备:获取广告计划信息及关键词列表;所述广告计划信息包括广告文案;根据所述广告计划信息确定所述广告的多个属性信息;将所述广告文案与所述关键词列表进行匹配,获得多个候选关键词;基于所述属性信息对所述多个候选关键词进行排序,获得多个排序结果;根据所述多个排序结果确定目标关键词。

[0093] 可以以一种或多种程序设计语言或其组合来编写用于执行本公开的操作的计算机程序代码,上述程序设计语言包括但不限于面向对象的程序设计语言—诸如Java、Smalltalk、C++,还包括常规的过程式程序设计语言—诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中,远程计算机可以通过任意种类的网络—包括局域网(LAN)或广域网(WAN)—连接到用户计算机,或者,可以连接到外部计算机(例如利用因特网服务提供商来通过因特网连接)。

[0094] 附图中的流程图和框图,图示了按照本公开各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分,该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为替换的实现中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个接连地表示的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意,框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合,可以用执行规定的功能或操作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0095] 描述于本公开实施例中所涉及到的单元可以通过软件的方式实现,也可以通过硬件的方式来实现。其中,单元的名称在某种情况下并不构成对该单元本身的限定。

[0096] 本文中以上描述的功能可以至少部分地由一个或多个硬件逻辑部件来执行。例

如,非限制性地,可以使用的示范类型的硬件逻辑部件包括:现场可编程门阵列(FPGA)、专用集成电路(ASIC)、专用标准产品(ASSP)、片上系统(SOC)、复杂可编程逻辑设备(CPLD)等等。

[0097] 在本公开的上下文中,机器可读介质可以是有形的介质,其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备,或者上述内容的任何合适组合。机器可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、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或快闪存储器)、光纤、便捷式紧凑盘只读存储器(CD-ROM)、光学储存设备、磁储存设备、或上述内容的任何合适组合。

[0098] 根据本公开实施例的一个或多个实施例,本公开实施例公开了一种广告关键词的生成方法,包括:

[0099] 获取广告计划信息及关键词列表;所述广告计划信息包括广告文案;

[0100] 根据所述广告计划信息确定所述广告的至少一个属性信息;

[0101] 将所述广告文案与所述关键词列表进行匹配,获得多个候选关键词;

[0102] 基于所述属性信息对所述多个候选关键词进行排序,获得至少一个排序结果;

[0103] 根据所述至少一个排序结果确定目标关键词。

[0104] 进一步地,将所述广告文案与所述关键词列表进行匹配,获得多个候选关键词,包括:

[0105] 遍历所述关键词列表,将遍历到的关键词与所述广告文案进行比对;

[0106] 若广告文案中存在遍历到的关键词或者存在与遍历到的关键词语义相近的词,则将遍历到的关键词确定为候选关键词。

[0107] 进一步地,基于所述属性信息对所述多个候选关键词进行排序,获得多个排序结果,包括:

[0108] 对于每个候选关键词,确定所述候选关键词分别与各属性信息对应的置信度;

[0109] 对于每个属性信息,根据所述置信度对所述候选关键词进行排序,获得至少一个排序结果。

[0110] 进一步地,确定所述候选关键词分别与各属性信息对应的置信度,包括:

[0111] 将候选关键词分别输入各属性信息分别对应的识别模型中,获得所述候选关键词分别与各属性信息对应的置信度;其中,属性信息包括行业类别和商品类别;识别模型包括行业类别识别模型和商品类别识别模型。

[0112] 进一步地,所述识别模型的训练方式为:

[0113] 获取训练数据;所述训练数据包括广告关键词以及所述广告关键词对应的真实属性信息;

[0114] 将所述广告关键词输入所述识别模型中,获得预测属性信息;

[0115] 根据所述真实属性信息和所述预测属性类别确定损失函数;

[0116] 基于所述损失函数对所述识别模型进行训练。

[0117] 进一步地,根据所述多个排序结果确定目标关键词,包括:

- [0118] 对所述至少一个排序结果进行融合,获得目标排序结果;
- [0119] 从所述目标排序结果中提取设定数量的目标关键词。
- [0120] 进一步地,对所述多个排序结果进行融合,获得目标排序结果,包括:
- [0121] 对候选关键词在各排序结果中的序号进行加权求和,获得各候选关键词的目标序号;
- [0122] 基于所述目标序号对所述候选关键词进行重新排序,获得目标排序结果;
- [0123] 相应的,从所述目标排序结果中提取设定数量的目标关键词,包括:
- [0124] 若是按照置信度由大到小排序,则从所述目标排序结果中提取排序靠前设定数量的目标关键词;若是按照置信度由小到大排序,则从所述目标排序结果中提取排序靠后设定数量的目标关键词。
- [0125] 注意,上述仅为本公开的较佳实施例及所运用技术原理。本领域技术人员会理解,本公开不限于这里所述的特定实施例,对本领域技术人员来说能够进行各种明显的变化、重新调整和替代而不会脱离本公开的保护范围。因此,虽然通过以上实施例对本公开进行了较为详细的说明,但是本公开不仅仅限于以上实施例,在不脱离本公开构思的情况下,还可以包括更多其他等效实施例,而本公开的范围由所附的权利要求范围决定。

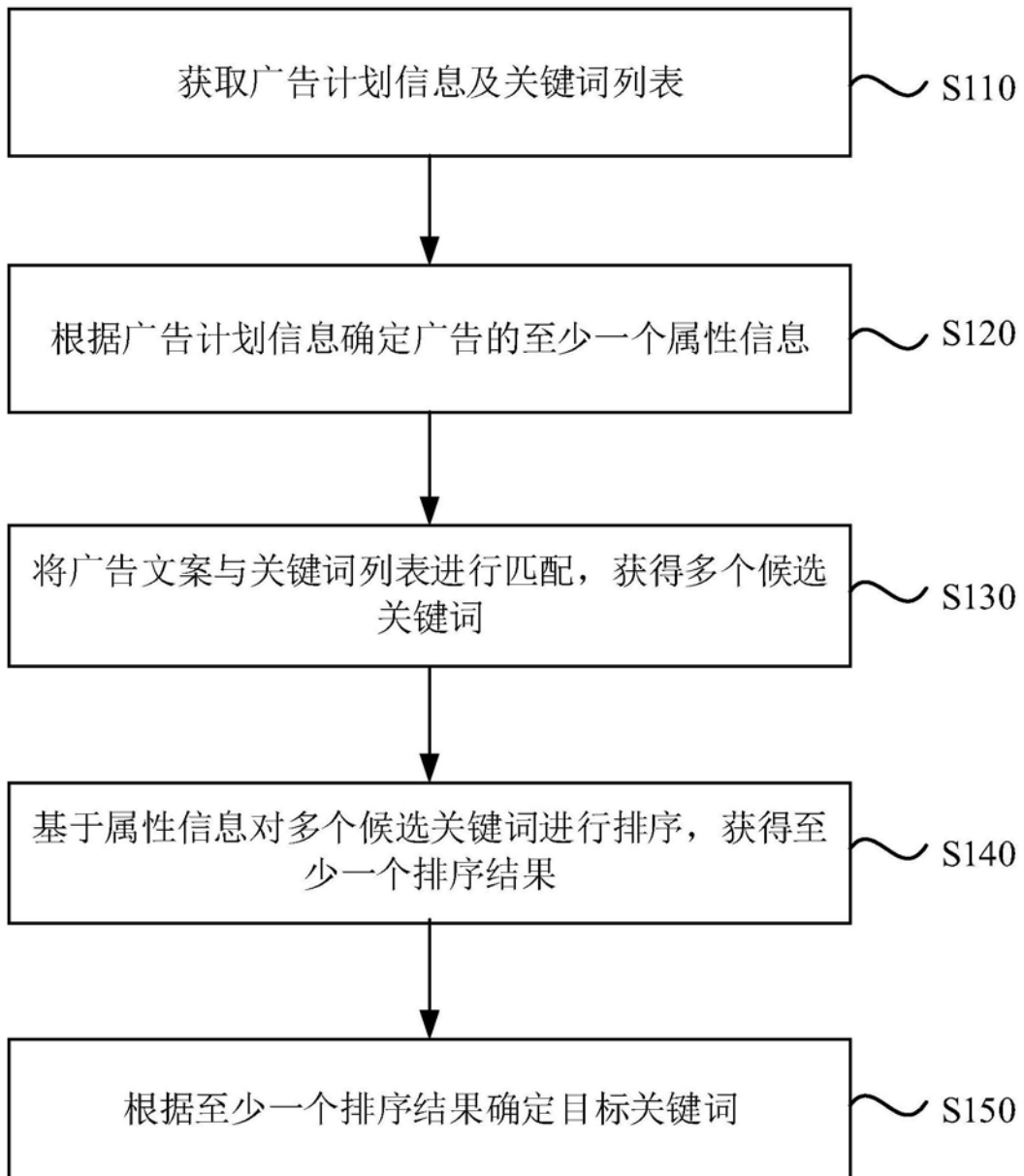


图1

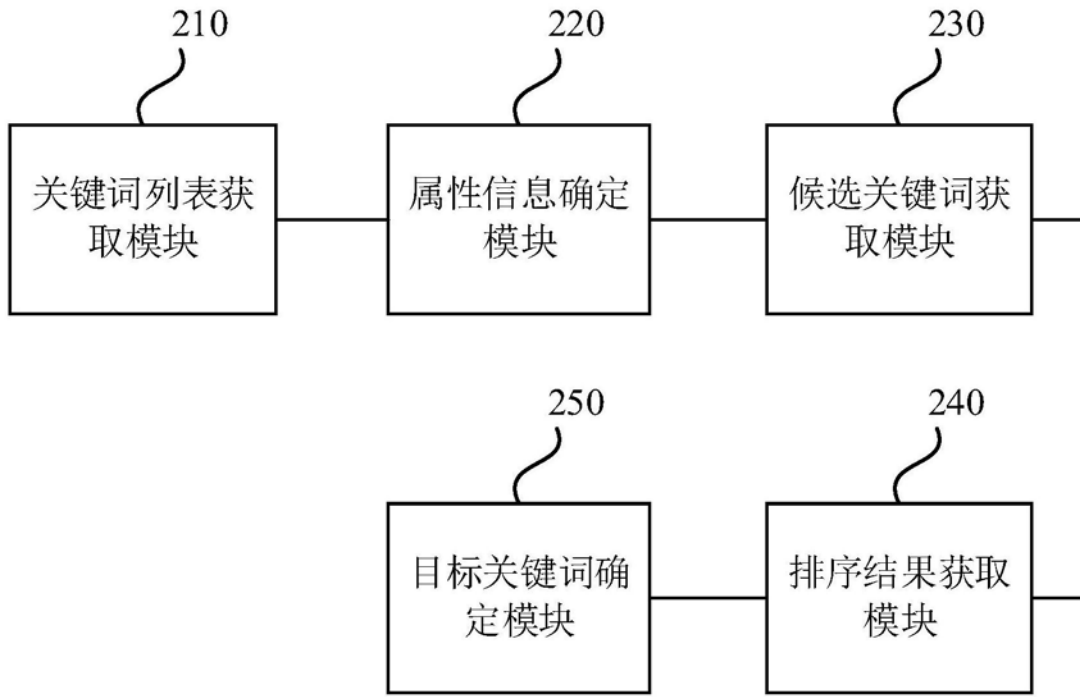


图2

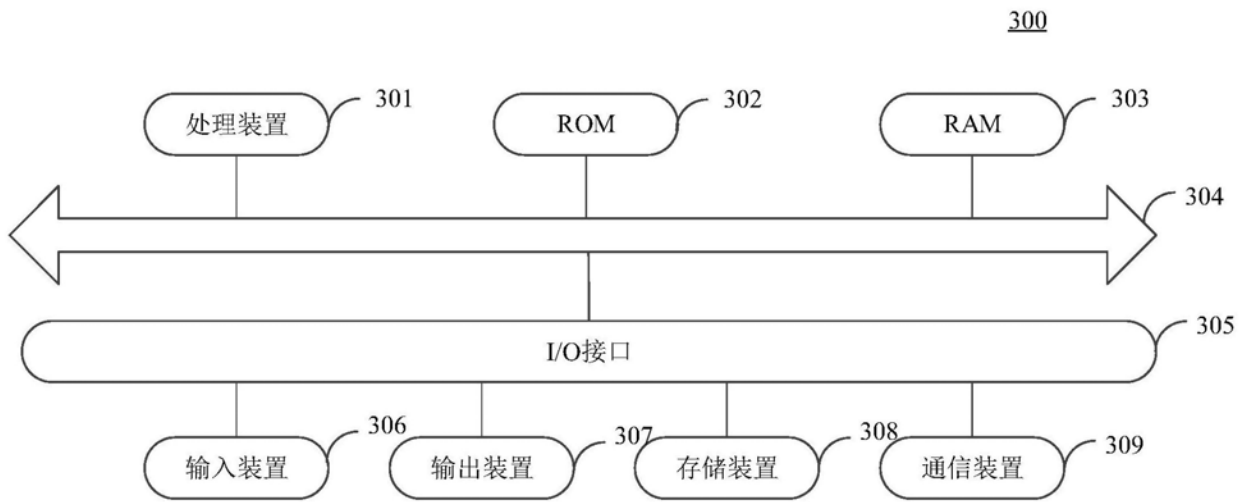


图3