



(12)发明专利

(10)授权公告号 CN 105335391 B

(45)授权公告日 2019.02.15

(21)申请号 201410326142.7

(22)申请日 2014.07.09

(65)同一申请的已公布的文献号
申请公布号 CN 105335391 A

(43)申请公布日 2016.02.17

(73)专利权人 阿里巴巴集团控股有限公司
地址 英属开曼群岛大开曼资本大厦一座四
层847号邮箱

(72)发明人 崔保良

(74)专利代理机构 北京博浩百睿知识产权代理
有限责任公司 11134

代理人 宋子良

(51)Int.Cl.
G06F 16/9535(2019.01)

(56)对比文件

CN 103412882 A,2013.11.27,
CN 103049474 A,2013.04.17,
CN 103729351 A,2014.04.16,
EP 2680172 A2,2014.01.01,
US 2013282750 A1,2013.10.24,
US 2011191311 A1,2011.08.04,
CN 102609433 A,2012.07.25,

审查员 唐文俊

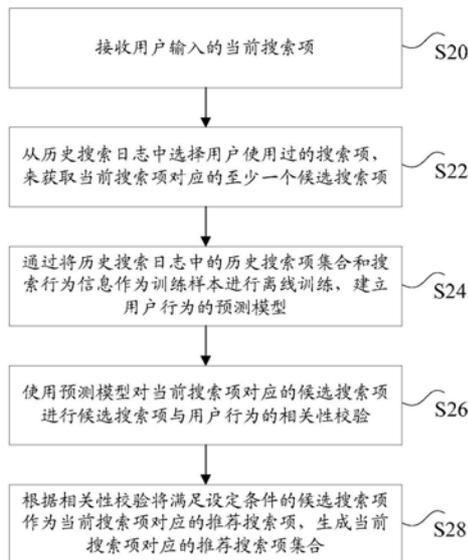
权利要求书3页 说明书13页 附图4页

(54)发明名称

基于搜索引擎的搜索请求的处理方法和装置

(57)摘要

本发明公开了一种基于搜索引擎的搜索请求的处理方法和装置。其中,该方法包括:基于搜索引擎的搜索请求的处理方法,该方法包括:接收用户输入的当前搜索项;从历史搜索日志中选择用户使用过的搜索项,获取当前搜索项对应的至少一个候选搜索项;通过将历史搜索日志中的历史搜索项集合和搜索行为信息作为训练样本进行离线训练,建立用户行为的预测模型;使用预测模型对当前搜索项对应的候选搜索项进行候选搜索项与用户行为的相关性校验;根据相关性校验将满足设定条件的候选搜索项作为当前搜索项对应的推荐搜索项,生成当前搜索项对应的推荐搜索项集合。本发明解决了搜索结果不准确的技术问题。



1. 一种基于搜索引擎的搜索请求的处理方法,其特征在于,包括:

接收用户输入的当前搜索项;

从历史搜索日志中选择用户使用过的搜索项,获取所述当前搜索项对应的至少一个候选搜索项;

通过将所述历史搜索日志中的历史搜索项集合和搜索行为信息作为训练样本进行离线训练,建立用户行为的预测模型;

使用所述预测模型对所述当前搜索项对应的候选搜索项进行所述候选搜索项与用户行为的相关性校验;

根据所述相关性校验将满足设定条件的候选搜索项作为当前搜索项对应的推荐搜索项,生成所述当前搜索项对应的推荐搜索项集合;

使用所述预测模型对所述当前搜索项对应的候选搜索项进行所述候选搜索项与用户行为的相关性校验包括:将候选搜索项与预测模型所确定的搜索行为的反馈信息进行匹配,提取匹配度大于或等于阈值的候选搜索项作为推荐搜索项。

2. 根据权利要求1所述的方法,其特征在于,通过将所述历史搜索日志中的历史搜索项集合和搜索行为信息作为训练样本进行离线训练,建立用户行为的预测模型的步骤包括:

读取所述历史搜索日志,获取预定时间段内的所述历史搜索项集合和所述历史搜索项集合中每个历史搜索项的搜索行为信息;

在所述历史搜索项集合中提取至少一个历史搜索项匹配对<历史搜索项A,历史搜索项B>,其中,所述历史搜索项B包含的分词集合是所述历史搜索项A包含的分词集合的真子集;

抽取所述历史搜索项匹配对<历史搜索项A,历史搜索项B>的特征和所述历史搜索项匹配对中的所述历史搜索项B的特征,组合生成特征集合;

从所述历史搜索项集合中每个历史搜索项的搜索行为信息中提取所述历史搜索项B的搜索行为信息;

根据所述特征集合和所述历史搜索项B的搜索行为信息生成所述训练样本;

采用逻辑回归LR模型对所述训练样本进行模型训练处理,生成所述预测模型。

3. 根据权利要求2所述的方法,其特征在于,所述历史搜索项匹配对<历史搜索项A,历史搜索项B>的特征包括如下任意一个或多个特征:行为特征、文本特征、知识库属性特征和统计特征;所述历史搜索项B的特征包括如下任意一个或多个特征:历史统计特征、分词组合特征、文本特征、知识库属性特征和词性特征。

4. 根据权利要求1所述的方法,其特征在于,在接收用户输入的当前搜索项之后,所述方法还包括:

从历史搜索日志中查询得到所述当前搜索项对应的历史行为信息;

在所述当前搜索项对应的历史行为信息满足推荐条件的情况下,执行获取所述当前搜索项对应的至少一个候选搜索项的步骤,其中,所述推荐条件包括以下任意一个或多个条件:所述当前搜索项对应的所述历史行为信息中的点击行为次数小于预定点击次数;当前搜索项搜索到的商品数少于预定个数;当前搜索项检索到的商品的相关性均小于预定值。

5. 根据权利要求1所述的方法,其特征在于,从历史搜索日志中选择用户使用过的搜索项,获取所述当前搜索项对应的至少一个候选搜索项的步骤包括:

从所述历史搜索日志中提取包含于所述当前搜索项的搜索项,得到初始搜索项集合;

根据至少一个初始候选搜索项生成所述至少一个候选搜索项,得到候选搜索项集合。

6. 根据权利要求5所述的方法,其特征在于,根据所述至少一个初始候选搜索项生成所述至少一个候选搜索项,得到候选搜索项集合的步骤包括:

直接保存所述至少一个初始候选搜索项,构成所述至少一个候选搜索项;或者,

对每个所述初始候选项进行特征筛选处理,其中,对每个所述初始候选项进行特征筛选处理的步骤包括:获取对应的候选特征集合,其中,筛选特征集合包括:中心词识别值、分词权重值和知识库属性丰富度;分别对每个所述初始候选搜索项的所述候选特征集合中规定多个筛选特征进行线性加权计算,得到每个所述初始候选搜索项的排序值;按照每个所述初始候选搜索项的排序值进行排序,选择预定数目的初始候选搜索项得到所述候选搜索项集合。

7. 一种基于搜索引擎的搜索请求的处理装置,其特征在于,包括:

接收模块,用于接收用户输入的当前搜索项;

获取模块,用于从历史搜索日志中选择用户使用过的搜索项,获取所述当前搜索项对应的至少一个候选搜索项;

模型建立模块,用于通过将所述历史搜索日志中的历史搜索项集合和搜索行为信息作为训练样本进行离线训练,建立用户行为的预测模型;

校验模块,用于使用所述预测模型对所述当前搜索项对应的候选搜索项进行所述候选搜索项与用户行为的相关性校验;

生成模块,用于根据所述相关性校验将满足设定条件的候选搜索项作为当前搜索项对应的推荐搜索项,生成所述当前搜索项对应的推荐搜索项集合;

所述校验模块还用于将候选搜索项与预测模型所确定的搜索行为的反馈信息进行匹配,提取匹配度大于或等于阈值的候选搜索项作为推荐搜索项。

8. 根据权利要求7所述的装置,其特征在于,所述模型建立模块包括:

读取模块,用于读取所述历史搜索日志,获取预定时间段内的所述历史搜索项集合和所述历史搜索项集合中每个历史搜索项的搜索行为信息;

搜索项提取模块,用于在所述历史搜索项集合中提取至少一个历史搜索项匹配对<历史搜索项A,历史搜索项B>,其中,所述历史搜索项B包含的分词集合是所述历史搜索项A包含的分词集合的真子集;

抽取特征模块,用于抽取所述历史搜索项匹配对<历史搜索项A,历史搜索项B>的特征和所述历史搜索项匹配对中的所述历史搜索项B的特征,组合生成特征集合;

搜索行为提取模块,用于从所述历史搜索项集合中每个历史搜索项的搜索行为信息中提取所述历史搜索项B的搜索行为信息;

组和模块,用于根据所述特征集合和所述历史搜索项B的搜索行为信息生成所述训练样本;

训练模块,用于采用逻辑回归LR模型对所述训练样本进行模型训练处理,生成所述预测模型。

9. 根据权利要求8所述的装置,其特征在于,所述历史搜索项匹配对<历史搜索项A,历史搜索项B>的特征包括如下任意一个或多个特征:行为特征、文本特征、知识库属性特征和统计特征;所述历史搜索项B的特征包括如下任意一个或多个特征:历史统计特征、分词组

合特征、文本特征、知识库属性特征和词性特征。

10. 根据权利要求7所述的装置,其特征在於,所述装置还包括:

查询模块,用于从历史搜索日志中查询得到所述当前搜索项对应的历史行为信息;

运行模块,用于在所述当前搜索项对应的历史行为信息满足推荐条件的情况下,执行获取所述当前搜索项对应的至少一个候选搜索项的步骤,其中,所述推荐条件包括以下任意一个或多个条件:所述当前搜索项对应的所述历史行为信息中的点击行为次数小于预定点击次数;当前搜索项搜索到的商品数少于预定个数;当前搜索项检索到的商品的相关性均小于预定值。

11. 根据权利要求7所述的装置,其特征在於,所述获取模块包括:

提取模块,用于从历史搜索日志中提取包含所述当前搜索项的搜索项,得到初始搜索项集合;

得到模块,用于根据至少一个初始候选搜索项生成所述至少一个候选搜索项,得到候选搜索项集合。

12. 根据权利要求11所述的装置,其特征在於,所述得到模块包括:

保存模块,用于直接保存所述至少一个初始候选搜索项,构成所述至少一个候选搜索项;或者,

筛选处理模块,用于对每个所述初始候选项进行特征筛选处理,其中,所述筛选处理模块包括:子获取模块,用于获取每个所述初始候选项对应的候选特征集合,其中,筛选特征集合包括:中心词识别值、分词权重值和知识库属性丰富度;计算模块,用于分别对每个所述初始候选搜索项的所述候选特征集合中规定多个筛选特征进行线性加权计算,得到每个所述初始候选搜索项的排序值;选择模块,用于按照每个所述初始候选搜索项的排序值进行排序,选择预定数目的初始候选搜索项得到所述候选搜索项集合。

基于搜索引擎的搜索请求的处理方法和装置

技术领域

[0001] 本发明涉及计算机互联网领域,具体而言,涉及一种基于搜索引擎的搜索请求的处理方法和装置。

背景技术

[0002] 在电子商务搜索中,用户为了尽量快的搜索到自己需要的商品,输入的搜索项(query)可能非常详细,通过构造多个分词(term)组合来表达需求。然而对于商务搜索引擎而言,召回的方式一般认为用户输入的多个分词之间是“AND”关系,很容易发生无结果和少结果的情形。针对这种情形,最常用的方式就是通过搜索项省略技术对用户输入的搜索项做改写来召回更多满足用户意图的商品。

[0003] 搜索项省略就是通过搜索项变换技术将搜索项中的一些分词删除,从而得到一个新的较短的发生改写后的改写搜索项(sub_query),使得发生改写后的改写搜索项(sub_query)尽量保留原始搜索项的重要信息,并且改写后搜索得到的商品尽量多且满足用户的原始购物意图。

[0004] 现有技术提供的搜索项省略技术的方案主要包括以下两种方案:

[0005] 方案一:计算搜索项中包含的各个分词的重要性,然后通过对分词进行排序来进行选择丢词。具体步骤如下:首先,给定一个用户输入的搜索项(包含若干分词);然后,使用逻辑算法计算各个分词的重要性;最后,保留重要性最大的分词,其他分词按照重要性从小到大丢弃,生成子搜索项。

[0006] 方案二:先丢出发生改写后的改写搜索项,然后对这些发生改写后的改写搜索项进行排序。具体步骤如下:首先,给定一个用户输入的搜索项(包含若干分词);然后,使用枚举方式或其他方式丢掉一些词,生成可能的候选分词子集;接着,使用在文本分类中通过特征选择方法确定分子之间的相关性的互信息方法对上述子集进行评估,最后,用最大生成树条件生成最优的发生改写后的改写搜索项。

[0007] 在系统对用户输入的搜索项进行改写的应用中,通过上述搜索项省略技术生成一个候选的发生改写后的改写搜索项集合的使用的方式主要有下面两种:

[0008] 方式一:直接用改写后的最优发生改写后的改写搜索项进行查询或最优发生改写后的改写搜索项集合组合起来进行筛选式的查询,将得到的搜索结果集合按照相关性排序后展示给用户。

[0009] 方式二:以提示条的方式把改写得到的最优候选集合展示给用户,让用户根据自己的意图来决策点击某一个提示的搜索项获取搜索结果。

[0010] 分析可知,上述现有的应用所采用的改写方法存在以下缺陷:

[0011] 一方面,不能预知改写之后的召回商品数。一方面多次的在线尝试会严重影响查询的性能和效率;另一方面为了精确度虽然使得发生改写后的改写搜索项丢失信息较少,但是不能召回更多质量好的搜索结果,也会影响用户下一步的决策。

[0012] 另一方面,无法获知搜索项改写后是否也满足用户的搜索要求。确定是否满足用

户搜索要求的衡量方式可以用是否有利于用户接下来的操作行为。

[0013] 而且,上述现有技术无法把改写后的发生改写后的改写搜索项当成一个有意义的整体来考虑,有可能认为重要的分词的组合毫无意义,甚至发生转义。

[0014] 针对上述现有技术对用户输入的搜索项的处理结果不完善,导致搜索结果不准确的问题,目前尚未提出有效的解决方案。

发明内容

[0015] 本发明实施例提供了一种基于搜索引擎的搜索请求的处理方法和装置,以提升搜索结果的准确性。

[0016] 根据本发明实施例的一个方面,提供了一种基于搜索引擎的搜索请求的处理方法,该方法包括:接收用户输入的当前搜索项;从历史搜索日志中选择用户使用过的搜索项,获取当前搜索项对应的至少一个候选搜索项;通过将历史搜索日志中的历史搜索项集合和搜索行为信息作为训练样本进行离线训练,建立用户行为的预测模型;使用预测模型对当前搜索项对应的候选搜索项进行候选搜索项与用户行为的相关性校验;根据相关性校验将满足设定条件的候选搜索项作为当前搜索项对应的推荐搜索项,生成当前搜索项对应的推荐搜索项集合。

[0017] 根据本发明实施例的另一方面,还提供了一种基于搜索引擎的搜索请求的处理装置,该装置包括:接收模块,用于接收用户输入的当前搜索项;获取模块,用于从历史搜索日志中选择用户使用过的搜索项,获取当前搜索项对应的至少一个候选搜索项;模型建立模块,用于通过将历史搜索日志中的历史搜索项集合和搜索行为信息作为训练样本进行离线训练,建立用户行为的预测模型;校验模块,用于使用预测模型对当前搜索项对应的候选搜索项进行候选搜索项与用户行为的相关性校验;生成模块,用于根据相关性校验将满足设定条件的候选搜索项作为当前搜索项对应的推荐搜索项,生成当前搜索项对应的推荐搜索项集合。

[0018] 在本发明实施例中,采用接收用户输入的当前搜索项;从历史搜索日志中选择用户使用过的搜索项,获取当前搜索项对应的至少一个候选搜索项;通过将历史搜索日志中的历史搜索项集合和搜索行为信息作为训练样本进行离线训练,建立用户行为的预测模型;使用预测模型对当前搜索项对应的候选搜索项进行候选搜索项与用户行为的相关性校验;根据相关性校验将满足设定条件的候选搜索项作为当前搜索项对应的推荐搜索项,生成当前搜索项对应的推荐搜索项集合的方式,提供了一个基于历史搜索日志中记录的历史搜索项以及每个历史搜索项的历史行为,对用户输入的当前搜索项进行改写的预测模型,通过学习历史搜索日志session中用户主动改写搜索项的模式,并抽取包括有效特征的训练样本进行建模,使得当前用户输入的搜索项可以使用建模得到的预测模型进行处理,实现了为当前输入的搜索项确定一个推荐搜索项集合。由于历史搜索日志中还提供了历史搜索项的搜索行为信息,即预测模型中融合了搜索行为的反馈信息,因此,可以使得对当前搜索项的改写做出更好的决策,从而提高了用户的查询速度和准确率,并能最大程度的满足用户的原始意图。进而解决了搜索结果不准确的技术问题。

附图说明

[0019] 此处所说明的附图用来提供对本发明的进一步理解,构成本申请的一部分,本发明的示意性实施例及其说明用于解释本发明,并不构成对本发明的不当限定。在附图中:

[0020] 图1是本发明实施例的一种运行基于搜索引擎的搜索请求的处理方法的硬件结构框图;

[0021] 图2是根据本发明实施例一的基于搜索引擎的搜索请求的处理方法的流程图;

[0022] 图3是根据本发明实施例一的基于搜索引擎的搜索请求的处理方法的详细流程图;

[0023] 图4是根据本发明实施例二的基于搜索引擎的搜索请求的处理装置的示意图;

[0024] 图5是根据本发明实施例二的一种可选的基于搜索引擎的搜索请求的处理装置的示意图;

[0025] 图6是根据本发明实施例二的一种可选的基于搜索引擎的搜索请求的处理装置的示意图;

[0026] 图7是根据本发明实施例二的一种可选的基于搜索引擎的搜索请求的处理装置的示意图。

具体实施方式

[0027] 为了使本技术领域的人员更好地理解本发明方案,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的实施例,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,都应当属于本发明保护的范围。

[0028] 需要说明的是,本发明的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本发明的实施例能够以除了在这里图示或描述的那些以外的顺序实施。此外,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不排他的包含,例如,包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0029] 实施例1

[0030] 根据本发明实施例,提供了一种基于搜索引擎的搜索请求的处理方法的方法实施例,需要说明的是,在附图的流程图示出的步骤可以在诸如一组计算机可执行指令的计算机系统中执行,并且,虽然在流程图中示出了逻辑顺序,但是在某些情况下,可以以不同于此处的顺序执行所示出或描述的步骤。

[0031] 本申请实施例一所提供的方法实施例可以在计算机终端、移动终端或者类似的用户运算设备中执行。以运行在计算机终端上为例,图1是本发明实施例的一种运行基于搜索引擎的搜索请求的处理方法的计算机终端10以及后台搜索服务器30的硬件结构框图。如图1所示,计算机终端10可以包括一个或多个(图中仅示出一个)处理器102(处理器102可以包括但不限于微处理器MCU或可编程逻辑器件FPGA等的处理装置)、用于存储数据的存储器

104、以及用于通信功能的传输装置106,后台搜索服务器30可以包括一个或多个(图中仅示出一个)搜索引擎处理器301(搜索引擎处理器301可以包括但不限于微处理器MCU或可编程逻辑器件FPGA等的处理装置)和通信装置303。本领域普通技术人员可以理解,图1所示的结构仅为示意,其并不对上述电子装置的结构造成限定。例如,计算机终端10还可包括比图1中所示更多或者更少的组件,或者具有与图1所示不同的配置。

[0032] 存储器104可用于存储应用程序的软件程序以及模块,如本发明实施例中的基于搜索引擎的搜索请求的处理方法对应的程序指令/模块,处理器102通过运行存储在存储器104内的软件程序以及模块,从而执行各种功能应用以及数据处理,即实现上述的升级应用程序的处理方法。存储器104可包括高速随机存储器,还可包括非易失性存储器,如一个或者多个磁性存储装置、闪存、或者其他非易失性固态存储器。在一些实例中,存储器104可进一步包括相对于处理器102远程设置的存储器,这些远程存储器可以通过网络连接至计算机终端10。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0033] 传输装置106用于经由一个网络接收或者发送数据。上述的网络具体实例可包括计算机终端10的通信供应商提供的无线网络。

[0034] 在上述运行环境下,本申请提供了如图2所示的基于搜索引擎的搜索请求的处理方法。图2是根据本发明实施例一的基于搜索引擎的搜索请求的处理方法的流程图。

[0035] 如图2所示,该基于搜索引擎的搜索请求的处理方法可以包括如下步骤:

[0036] 步骤S20,接收用户输入的当前搜索项。

[0037] 本申请上述步骤S20中的搜索项可以由用户通过终端设备的输入设备进行输入。用户输入的搜索项的信息可以发送至图1所示的处理器102。

[0038] 步骤S22,从历史搜索日志中选择用户使用过的搜索项,来获取当前搜索项对应的至少一个候选搜索项。

[0039] 本申请上述步骤S22中,可以通过图1所示的处理器102来执行根据用户输入的搜索项,从历史搜索日志中选取用户使用过的搜索项作为样本。从用户使用过的搜索项中选择至少一个搜索词作为候选搜索项。候选搜索项为用户历史输入的搜索项集合中的一个子集。所有的候选搜索项都有可能作为推荐搜索项显示在搜索栏中供用户选择。

[0040] 步骤S24,通过将历史搜索日志中的历史搜索项集合和搜索行为信息作为训练样本进行离线训练,建立用户行为的预测模型。其中,该步骤中的搜索行为信息为对应用户历史使用搜索项时发生的行为信息。

[0041] 上述步骤S24可以通过图1所示的后台搜索服务器30的搜索引擎处理器301或图1所示的处理器102来实现通过将历史搜索日志中的历史搜索项集合和用户基于搜索项发生的行为信息(比如在使用搜索项搜索获得搜索结果后,对搜索结果的点击行为)作为训练样本进行离线训练,获得用户行为的预测模型。

[0042] 上述预测模型用于根据当前输入的搜索项预测一个用户可能感兴趣的推荐搜索项集合。另外,上述利用预测模型对候选搜索项进行校验处理的过程可以是进行相关性校验,即实现将候选搜索项与预测模型所确定的搜索行为的反馈信息进行匹配,提取匹配度大于或等于阈值的候选搜索项作为推荐搜索项。

[0043] 上述历史搜索项集合为用户历史使用过的搜索项,可以是用户在一定时间内已经

输入的搜索项,在搜索的过程中,用户在输入一个搜索项进行搜索后可以紧随着更换一个相似的搜索项进行搜索。即如果用户对输入的某一个搜索项对应的搜索结果不满意,则会修改该搜索项的分词(term)(可以包括增加、删除、更新分词(term)的方式)来得到一个新的搜索项,系统会将该新的搜索项作为一个输入搜索项保存至所述历史搜索项集合中,上述过程是一个用户主动改写搜索项的模式。

[0044] 步骤S26,使用预测模型对当前搜索项对应的候选搜索项进行候选搜索项与用户行为的相关性校验。该步骤实现了,使用基于用户历史使用过的搜索项及用户基于所述搜索项发生的行为信息所建立的用户行为的预测模型,执行对候选搜索项与用户行为的相关性校验。

[0045] 步骤S28,根据相关性校验将满足设定条件的候选搜索项作为当前搜索项对应的推荐搜索项,生成当前搜索项对应的推荐搜索项集合。

[0046] 分析可知,本申请上述实施例,提供了历史搜索日志中记录的历史搜索项以及每个历史搜索项的历史行为。通过机器学习历史搜索日志中用户主动改写搜索项的模式,并抽取包括有效特征的训练样本进行搜索项改写模式的预测模型的建模,使得当前用户输入的搜索项可以使用建模得到的预测模型来预测用户当前输入的搜索项对应的改写模式,即预测当前输入的搜索项对应的改写的搜索项。由于历史搜索日志中还提供了历史搜索项的搜索行为信息,即预测模型中融合了搜索行为的反馈信息,因此,可以使得对当前搜索项的改写做出更好的决策,从而提高了用户的查询速度和准确率,并能最大程度的满足用户的原始意图。进而解决了搜索结果不准确的技术问题。

[0047] 本申请上述实施例所提供的步骤S20至步骤S28可以在计算机客户端上运行,实施过程中,在上述实施例中的计算机客户端可以是安装了用于搜索的浏览器或搜索客计算机终端。分析可知,上述方法所使用的搜索项的省略技术,其应用范围可以不局限于搜索项省略方向,改写方式可以扩展为对分词(term)的增加、删除、替换的其中一种或混合方式,目的是为了改善当前搜索项的质量,提升用户搜索的效率和体验。

[0048] 下面就结合图3所示的流程图对上述实施例进行详细描述。

[0049] 结合图3可知,本申请实施例上述基于搜索引擎的搜索请求的处理方法是一个对当前用户输入的搜索项进行自动改写的方案,可以包括线下的模型训练流程和线上的实时预测流程两个部分。

[0050] 其中,线下的模型训练的功能实现了通过将历史搜索日志中的历史搜索项集合和搜索行为信息作为训练样本进行离线训练,获得预测模型的流程,该部分功能可以包括如下三部分实施步骤:a)从历史搜索日志中进行样本选取;b)对选取得到的样本进行特征抽取;c)根据抽取到的特征生成训练样本,并对训练样本进行模型训练。

[0051] 线上的实时预测的功能主要实现了,基于已经得到的训练模块对当前接收到的搜索项进行相关性校验处理,得到系统为当前搜索项推荐的搜索项集合,该部分功能主要包括以下两部分实施步骤:a)可以基于历史搜索日志对当前搜索词进行初选;b)可以基于已经得到的训练模型对初选得到的候选搜索项进行预测处理,得到当前搜索词对应的推荐搜索项。

[0052] 下面就结合图3对上述线下的模型训练的功能和线上的实时预测的功能进行详细描述。

[0053] 本申请上述实施例中,步骤S24执行了的线下的模型训练的功能,此处需要说明的是,通过将历史搜索日志中的历史搜索项集合和搜索行为信息作为训练样本进行离线训练,建立用户行为的预测模型的步骤可以包括如下实施方案:

[0054] 步骤S241,读取历史搜索日志,获取预定时间段内的历史搜索项集合和历史搜索项集合中每个历史搜索项的搜索行为信息。

[0055] 结合图3可知,该步骤S241实现了样本选取的过程。可以保存在图1所示的存储器104中的历史搜索日志session指的是预定时间段的间隔内用户连续做出的一系列动作的集合(历史搜索日志session中每一条目志通常会包括:搜索用户cookie、用户使用搜索引擎的搜索行为信息、搜索行为产生的时间、搜索行为产生的位置、搜索行为对应的搜索项),此处一系列动作包括输入的历史搜索项以及对每个历史搜索项执行的搜索行为信息(诸如搜索行为、搜索结果的点击行为、翻页行为、关闭页面行为等),由于历史搜索日志session的内容海量的,因此,可以仅抽取出预定时间段内的一定数量的session作为样本即可。

[0056] 本申请该步骤S241中关键获取的是一定时间间隔内用户使用过的搜索词构成的搜索项序列(即用户预定时间内在搜索框页面的输入框中输入的搜索项,例如:电脑、电脑键盘、电脑鼠标、修理电脑等)以及每个搜索项对应的搜索后的搜索行为信息(例如点击反馈信息)。

[0057] 步骤S243,在历史搜索项集合中提取至少一个历史搜索项匹配对<历史搜索项A,历史搜索项B>,其中,历史搜索项B包含的分词(term)集合是历史搜索项A包含的分词(term)集合的真子集。

[0058] 上述步骤S243实现了,在获取到历史搜索日志中的历史搜索项集合之后,从历史搜索项集合中选取满足以下特征的历史搜索项匹配对<A,B>,其中,B的分词(term)集合是A的分词(term)集合的真子集。上述历史搜索项集合中的历史搜索项可以按照队列的形式进行保存,从而历史搜索项集合包含了一组历史搜索项序列,因此,可以采用正向扫描的方式从历史搜索项序列中提取得到历史搜索项匹配对。

[0059] 步骤S245,抽取历史搜索项匹配对<历史搜索项A,历史搜索项B>的特征和历史搜索项匹配对中的历史搜索项B的特征,组合生成特征集合。

[0060] 优选地,上述步骤S245中的历史搜索项匹配对<历史搜索项A,历史搜索项B>的特征包括如下任意一个或多个特征:行为特征、文本特征、知识库属性特征和统计特征;历史搜索项B的特征包括如下任意一个或多个特征:历史统计特征、分词(term)组合特征、文本特征、知识库属性特征和词性特征。

[0061] 此处需要说明的是,上述历史搜索项匹配对中的历史搜索项A和历史搜索项B都是用户在一定时间内已经输入的搜索项,其中,历史搜索项B可以是用户在输入历史搜索项A之后,对历史搜索项A的分词(term)进行修改得到的搜索项,因此,结合图3可知,历史搜索项A和历史搜索项B可以看做是对一个历史搜索项进行改写后的一个改写搜索项,由此,历史搜索项匹配对也可以表示为query-sub_query对。由此,步骤S245抽取特征的过程是针对query-sub_query对的特征抽取和针对改写搜索项sub_query本身的特征抽取。

[0062] 下面就对历史搜索项匹配对query-sub_query对和其中的改写搜索项sub_query的各个特征进行详细说明:

[0063] 1、针对历史搜索项匹配对(query-sub_query对)的特征:

[0064] 行为特征:改写之后引导的点击率ctr,成交。

[0065] 上述行为特征表示的是用户从历史搜索项A修改为搜索项B这个过程中的行为特征:点击率ctr和是否成交,以表征这个改写行为的有效性。引入这个特征可以把改写后用户发生的行为信息作为经验特征有效融入到模型中。这个特征不一定总是可以获取到值,获取不到值的情况用均值填充。

[0066] 文本特征:前缀重合比例、后缀重合比例、丢词位置;丢掉分词(term)占比、丢掉长度占比;丢掉的词的id、保留词的id等。

[0067] 上述文本特征表示历史搜索项A和历史搜索项B在文本上的差别。其中,前缀重合比例、后缀重合比例和丢词位置用来学习用户倾向于对重要信息的输入位置;丢掉分词(term)占比和丢掉长度占比用来学习用户是否倾向于保留更多的词;丢掉词的id、保留词的id是一个很重要的特征,可以直接学习到用户认为重要的词。

[0068] 知识库属性特征:列举丢掉、保留分词(term)的知识库属性类型以及占比信息。

[0069] 由于商业搜索一般是用产品(如手机)、品牌(如华为)、修饰词(如红色)等知识库属性特征来表征,把这些特征加进来可以学习到什么知识库属性用户更为倾向于保留或丢弃。

[0070] 词性特征:列举丢掉分词(term)的词性特征以及占比信息。

[0071] 词性特征和知识库属性类似,在商业搜索中词性也是有区分度的,比如名词,人名,专名一般比较重要,而语气词,助词等可以忽略。

[0072] 统计特征:类目预测相似度;性别意图相似度以及其他重要属性相似度。

[0073] 搜索项的重要属性匹配程度是衡量搜索项相似度的重要指标。如类目预测,性别意图,品牌等其他重要属性。

[0074] 2、针对改写搜索项本身的特征:

[0075] 历史统计特征:召回结果集的类目熵;历史点击商品的类目熵;点击率ctr和成交。

[0076] 这些历史统计特征可以通过改写搜索项sub_query的历史统计信息来反映改写搜索项sub_query的质量。类目:商品分类名称,每个商品挂在一个唯一类目下。结果集类目熵:获取一个搜索项在历史30天搜索展现的商品结果集的类目以及占比信息,计算出熵,可以衡量query结果集的意图明确程度。点击商品的类目熵:获取一个搜索项在历史30天点击的商品集的类目以及占比信息,计算熵,可以反映这个搜索项下用户点击意图的明确程度。点击率ctr和成交表示这个改写搜索项引导点击和成交的能力。

[0077] 分词(term)组合特征:1) .首先获取分词(term)级别的相关信息:类目熵,自由度,分词权重等;2)对分词(term)的相关信息组合,方式为:互信息、和、标准差、算数平均数等。

[0078] 分词(term)的类目熵:包括结果熵和点击熵,和的类目熵计算方式相同,区别在于统计的粒度到了分词(term)级别。分词权重用于标记一个分词(term)在搜索项中的重要程度。获取了分词(term)级别的信息,可以通过一些运算把分词(term)级别的信息映射到级别,方式主要有互信息、和、标准差、算数平均数等。

[0079] 文本特征:分词(term)个数;文本长度;数字字母串的长度和比例等。

[0080] 这些文本特征用于表示改写搜索项的文本特征,用来学习用户倾向于改写的文本状态。

[0081] 知识库属性特征:包含的各种知识库属性类型的个数以及比例。该特征用来学习用户倾向于改写的改写搜索项的知识库属性的分布情况。

[0082] 词性特征:包含的各种词性类型的个数以及比例。该用来学习用户倾向于改写的改写搜索项的词性的分布情况。

[0083] 步骤S247,从历史搜索项集合中每个历史搜索项的搜索行为信息中提取历史搜索项B的搜索行为信息。

[0084] 步骤S249,根据特征集合和历史搜索项B的搜索行为信息生成训练样本。该步骤中的搜索行为信息可以是模型训练的目标,目标是指在针对历史搜索项B的搜索结果是否存在点击行为。

[0085] 此处需要说明的是,可以通过特征集合和历史搜索项B的搜索行为信息的结合生成训练样本,此处的结合方式为:训练样本由特征和目标构成,如果历史搜索项B发生了点击行为,则认为目标是1(正样本),否则为0(负样本)。比如用户搜索了“红色华为手机”之后做了两次改写:“红色手机”:没有点击;“华为手机”:发生了点击。我们则认为(红色华为手机,红色手机)为负样本,(红色华为手机,华为手机)为正样本。

[0086] 由此可知,上述步骤S247至步骤S249实现了训练样本的生成过程,其中,为了更细致的描述现实场景,可以对搜索项的分词(term)个数,文本长度等分段后抽取上述特征进行组合,可以使模型更加准确。

[0087] 此处还需要说明的是,本申请训练样本还可以采取如下方式的样本:(1)用simrank等搜索项相似性算法计算得到的相似的搜索项对以及相似程度,以相似度为目标进行训练。这种方式也可以从用户的点击行为中学习到搜索项的相似性,但是训练的query-sub_query在时间顺序上没有很强的关联,不利于模拟现实的改写场景。(2)当改写系统运行一段时间后,可以拿到改写系统做出改写之后用户实际的行为反馈来构造样本,更新模型。这种方式可以直接学习到用户对于当前改写系统的反馈,有利于系统的自适应更新。

[0088] 步骤S251,采用逻辑回归LR模型对训练样本进行模型训练处理,生成预测模型。

[0089] 本申请上述步骤S251中选择的LR(Logistic Regression)模型具有很好的描述能力,模型训练和预测的时间复杂度和空间复杂度较低。离线训练会生成每个特征id对应的权重值,在预先计算好可以离线计算的特征之后,存储上述计算得到的特征;然后,在线预测过程中对于离线计算的特征进行直接查表,使得特征id对应的权重值可以直接查表获得,接着通过简单计算即可得到预测值。发明还可以选用其他机器学习模型,如决策树,SVM等。

[0090] 在上述线下的模型训练生成的预测模型的基础上,用户通过将该离线训练获得的预测模型在线上使用,从而实时预测的功能,对当前输入的当前搜索项进行系统自动改写的结果做出实时预测。下面就对该线上的实时预测的功能进行详细描述。

[0091] 本申请上述实施例中,结合图3可知,在步骤S20接收用户输入的当前搜索项之后,还可以获取用户输入的当前搜索项进行是否具有改写的必要的判断。可以通过执行如下实施方案实现:

[0092] 步骤S301,从历史搜索日志中查询得到当前搜索项对应的历史行为信息。

[0093] 步骤S303,在当前搜索项对应的历史行为信息满足推荐条件的情况下,执行获取当前搜索项对应的至少一个候选搜索项的步骤,其中,推荐条件包括以下任意一个或多个

条件:当前搜索项对应的历史行为信息中的点击行为次数小于预定点击次数。

[0094] 此处的推荐条件还可以是如下条件:当前搜索项搜索到的商品数少于预定个数;当前搜索项检索到的商品的相关性均小于预定值。

[0095] 本申请上述一种可选的实施例中,步骤S22实现的从历史搜索日志中选择用户使用过的搜索项,获取当前搜索项对应的至少一个候选搜索项的方案,实现了对当前搜索项对应的候选搜索项进行初步筛选,获取改写搜索项sub_query的候选项。可以包括如下实施步骤:

[0096] 步骤S221,从历史搜索日志中提取包含当前搜索项的搜索项,得到初始搜索项集合。

[0097] 步骤S223,根据至少一个初始候选搜索项生成至少一个候选搜索项,得到候选搜索项集合。

[0098] 优选地,上述步骤S223实现的根据至少一个初始候选搜索项生成至少一个候选搜索项,得到候选搜索项集合的步骤可以包括如下两种方案:

[0099] 方案一:可以直接保存至少一个初始候选搜索项,构成至少一个候选搜索项。

[0100] 上述方案一提供的方法适用于针对比较大的电子商务网站进行搜索行为的应用场景。在该应用场景中,用户的历史行为非常丰富,因此,可以从搜索日志中选取历史上用户曾经搜索过的历史搜索项作为改写的sub_候选,此处的历史搜索项的分词(term)集合包含于当前输入的当前搜索项。这样可以方便的离线获取候选的各种行为信息:召回商品质量、点击率ctr、转化率等以供模型训练使用。

[0101] 为了节省方案一的性能消耗,可以对方案一生成的候选搜索项作为初步筛选结果,做进一步的筛选特征,详细可见方案二。

[0102] 方案二:

[0103] 首先,获取每个初始候选项对应的候选特征集合,其中,筛选特征集合包括:中心词识别、分词权重值和知识库属性丰富度。

[0104] 此处需要说明的是,上述中心词识别是指需要对初始选择的初始候选项进行核心词挖掘,检测当前搜索项中最核心的词是否被删除,如果中心词如果丢了,则该初始候选项不符合要求,有可能发生了转义;分词(term)的权重值,是指可以根据知识库属性,词性,自由度,类目熵进行分析;知识库属性丰富度:包含的知识库属性种类个数。这一步骤的主要作用在于尽量保证质量好的改写搜索项能通过海选,在保证精确性的前提下尽量节省后续相关性校验的计算工作量。

[0105] 然后,分别对每个初始候选搜索项的候选特征集合中规定多个筛选特征进行线性加权计算,得到每个初始候选搜索项的排序值。

[0106] 最后,按照每个初始候选搜索项的排序值进行排序,选择预定数目的初始候选搜索项得到候选搜索项集合。

[0107] 上述两个步骤实现了对上述筛选特征进行线性加权组合,算出每个初始候选搜索项query的规则分,然后使用规则分对候选搜索项进行排序,并选取指定个数的最优集合进入执行推荐候选搜索项的模型预测阶段。从而实现了按照模型算分进行排序,输出当前输入的当前搜索项对应的最终的改写结果。

[0108] 此处需要说明的是,上述线下的模型训练实现了将搜索项query的改写过程采用

一个模型来描述,这个模型基于用户的历史行为来训练,在线可以做到实时预测,并可自适应更新。而在模型目标,样本选取和特征抽取中,尽量多的考虑了用户的历史行为反馈信息和统计信息,使得模型可以更加精确的描述和预测现实场景。

[0109] 此处还需要说明的是,关于候选搜索项的获取,对于一些非常长尾的查询,当历史搜索项集合中找不到合适的候选搜索项的时候,可以使用其他算法来计算改写的候选,使得这些候选在概率上可以有较好的表现和获取较好的召回率,后续仍可以继续用模型来预测,不过特征方面要做些调整。另外,候选集的初步筛选可以采取更加详尽的方式,尽量使得进入到模型训练的候选集合是最优的

[0110] 本申请上述实施例中进行搜索结果的数据统计及模型训练都是在Hadoop集群中进行,程序可以使用Java及C++语言实现。线上部分可以使用C++语言实现。整体流程图如下:

[0111] 需要说明的是,对于前述的各方法实施例,为了简单描述,故将其都表述为一系列的动作组合,但是本领域技术人员应该知悉,本发明并不受所描述的动作顺序的限制,因为依据本发明,某些步骤可以采用其他顺序或者同时进行。其次,本领域技术人员也应该知悉,说明书中所描述的实施例均属于优选实施例,所涉及的动作和模块并不一定是本发明所必须的。

[0112] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到根据上述实施例的方法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质(如ROM/RAM、磁碟、光盘)中,包括若干指令用以使得一台终端设备(可以是手机,计算机,服务器,或者网络设备等)执行本发明各个实施例所述的方法。

[0113] 实施例2

[0114] 根据本发明实施例,还提供了一种用于实施上述方法实施例的装置。图4是根据本发明实施例二的基于搜索引擎的搜索请求的处理装置的示意图。本申请上述实施例所提供的装置可以在网络游戏客户端上运行。

[0115] 如图4所示,该装置可以包括:接收模块40、获取模块42、模型建立模块44、校验模块46和生成模块48。

[0116] 其中,接收模块40,用于接收用户输入的当前搜索项;获取模块42,用于从历史搜索日志中选择用户使用过的搜索项,获取当前搜索项对应的至少一个候选搜索项;模型建立模块44,用于通过将历史搜索日志中的历史搜索项集合和搜索行为信息作为训练样本进行离线训练,建立用户行为的预测模型;校验模块46,用于使用预测模型对当前搜索项对应的候选搜索项进行候选搜索项与用户行为的相关性校验;生成模块48,用于根据相关性校验将满足设定条件的候选搜索项作为当前搜索项对应的推荐搜索项,生成当前搜索项对应的推荐搜索项集合。

[0117] 本申请上述实施例,提供了一个基于历史搜索日志中记录的历史搜索项以及每个历史搜索项的历史行为,对用户输入的当前搜索项进行改写的预测模型,通过学习历史搜索日志session中用户主动改写搜索项的模式,并抽取包括有效特征的训练样本进行建模,使得当前用户输入的搜索项可以使用建模得到的预测模型进行处理,实现了为当前输入的

搜索项确定一个推荐搜索项集合。由于历史搜索日志中还提供了历史搜索项的搜索行为信息,即预测模型中融合了搜索行为的反馈信息,因此,可以使得对当前搜索项的改写做出更好的决策,从而提高了用户的查询速度和准确率,并能最大程度的满足用户的原始意图。进而解决了搜索结果不准确的技术问题。

[0118] 本申请上述实施例所提供的接收模块40、获取模块42、模型建立模块44、校验模块46和生成模块48可以在计算机客户端上运行,实施过程中,在上述实施例中的计算机客户端可以是安装了用于搜索的浏览器或搜索客计算机终端。分析可知,上述方法所使用的搜索项的省略技术,其应用范围可以不局限于搜索项省略方向,搜索项的改写方式可以扩展为对搜索项中的分词(term)的增加、删除、替换的其中一种或混合方式,目的是为了改善当前搜索项的质量,提升用户搜索的效率和体验。

[0119] 此处需要说明的是,上述接收模块40、获取模块42、模型建立模块44、校验模块46和生成模块48对应于实施例一中的步骤S20至步骤S28,五个模块与对应的步骤所实现的示例和应用场景相同,但不限于上述实施例一所公开的内容。需要说明的是,上述模块作为装置的一部分可以运行在实施例一提供的计算机终端10中,可以通过软件实现,也可以通过硬件实现。

[0120] 优选地,如图5所示,上述模型建立模块44可以包括:读取模块441、搜索项提取模块443、抽取特征模块445、搜索行为提取模块447、组和模块449和训练模块451。

[0121] 其中,读取模块441,用于读取历史搜索日志,获取预定时间段内的历史搜索项集合和历史搜索项集合中每个历史搜索项的搜索行为信息;搜索项提取模块443,用于在历史搜索项集合中提取至少一个历史搜索项匹配对<历史搜索项A,历史搜索项B>,其中,历史搜索项B包含的分词(term)集合是历史搜索项A包含的分词(term)集合的真子集;抽取特征模块445,用于抽取历史搜索项匹配对<历史搜索项A,历史搜索项B>的特征和历史搜索项匹配对中的历史搜索项B的特征,组合生成特征集合;搜索行为提取模块447,用于从历史搜索项集合中每个历史搜索项的搜索行为信息中提取历史搜索项B的搜索行为信息;组和模块449,用于根据特征集合和历史搜索项B的搜索行为信息生成训练样本;训练模块451,用于采用逻辑回归LR模型对训练样本进行模型训练处理,生成预测模型。

[0122] 此处需要说明的是,上述读取模块441、搜索项提取模块443、抽取特征模块445、搜索行为提取模块447、组和模块449和训练模块451,六个模块与实施例一中对应的步骤S247至步骤S249所实现的示例和应用场景相同,但不限于上述实施例一所公开的内容。需要说明的是,上述模块作为装置的一部分可以运行在实施例一提供的计算机终端10中,可以通过软件实现,也可以通过硬件实现。

[0123] 优选地,上述历史搜索项匹配对<历史搜索项A,历史搜索项B>的特征包括如下任意一个或多个特征:行为特征、文本特征、知识库属性特征和统计特征;历史搜索项B的特征包括如下任意一个或多个特征:历史统计特征、分词(term)组合特征、文本特征、知识库属性特征和词性特征。

[0124] 优选地,如图6所示,在执行接收模块40的功能之后,本申请上述装置还可以执行如下功能模块:查询模块411和运行模块413。

[0125] 其中,查询模块411,用于从历史搜索日志中查询得到当前搜索项对应的历史行为信息;运行模块413,用于在当前搜索项对应的历史行为信息满足推荐条件的情况下,执行

获取当前搜索项对应的至少一个候选搜索项的步骤,其中,推荐条件包括以下任意一个或多个条件:当前搜索项对应的历史行为信息中的点击行为次数小于预定点击次数;当前搜索项搜索到的商品数少于预定个数;当前搜索项检索到的商品的相关性均小于预定值。

[0126] 此处需要说明的是,上述查询模块411和运行模块413,两个模块与对应的步骤S301至步骤S303所实现的示例和应用场景相同,但不限于上述实施例一所公开的内容。需要说明的是,上述模块作为装置的一部分可以运行在实施例一提供的计算机终端10中,可以通过软件实现,也可以通过硬件实现。

[0127] 如图7所示,本申请上述获取模块42可以包括:提取模块421和得到模块423。

[0128] 其中,提取模块421,用于从历史搜索日志中提取包含当前搜索项的搜索项,得到初始搜索项集合;得到模块423,用于根据至少一个初始候选搜索项生成至少一个候选搜索项,得到候选搜索项集合。

[0129] 此处需要说明的是,上述提取模块421和得到模块423,两个模块与对应的步骤S221至步骤S223所实现的示例和应用场景相同,但不限于上述实施例一所公开的内容。需要说明的是,上述模块作为装置的一部分可以运行在实施例一提供的计算机终端10中,可以通过软件实现,也可以通过硬件实现。

[0130] 优选地,上述得到模块423可以包括:保存模块或者筛选处理模块。

[0131] 上述保存模块,用于直接保存至少一个初始候选搜索项,构成至少一个候选搜索项。

[0132] 筛选处理模块,用于对每个初始候选项进行特征筛选处理,其中,该筛选处理模块可以包括:子获取模块,用于获取每个初始候选项对应的候选特征集合,其中,筛选特征集合包括:中心词识别值、分词权重值和知识库属性丰富度;计算模块,用于分别对每个初始候选搜索项的候选特征集合中规定多个筛选特征进行线性加权计算,得到每个初始候选搜索项的排序值;选择模块,用于按照每个初始候选搜索项的排序值进行排序,选择预定数目的初始候选搜索项得到候选搜索项集合。

[0133] 此处需要说明的是,上述得到模块423包括的子模块与实施例一中对应的方案一和方案二所实现的步骤示例和应用场景相同,但不限于上述实施例一所公开的内容。需要说明的是,上述模块作为装置的一部分可以运行在实施例一提供的计算机终端10中,可以通过软件实现,也可以通过硬件实现。

[0134] 上述本发明实施例序号仅仅为了描述,不代表实施例的优劣。

[0135] 在本发明的上述实施例中,对各个实施例的描述都各有侧重,某个实施例中沒有详述的部分,可以参见其他实施例的相关描述。

[0136] 在本申请所提供的几个实施例中,应该理解到,所揭露的客户端,可通过其它的方式实现。其中,以上所描述的装置实施例仅仅是示意性的,例如单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,单元或模块的间接耦合或通信连接,可以是电性或其它的形式。

[0137] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个

网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0138] 另外,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单元中,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单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0139] 所述集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可为个人计算机、服务器或者网络设备)执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、移动硬盘、磁碟或者光盘等各种可以存储程序代码的介质。

[0140] 以上所述仅是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。

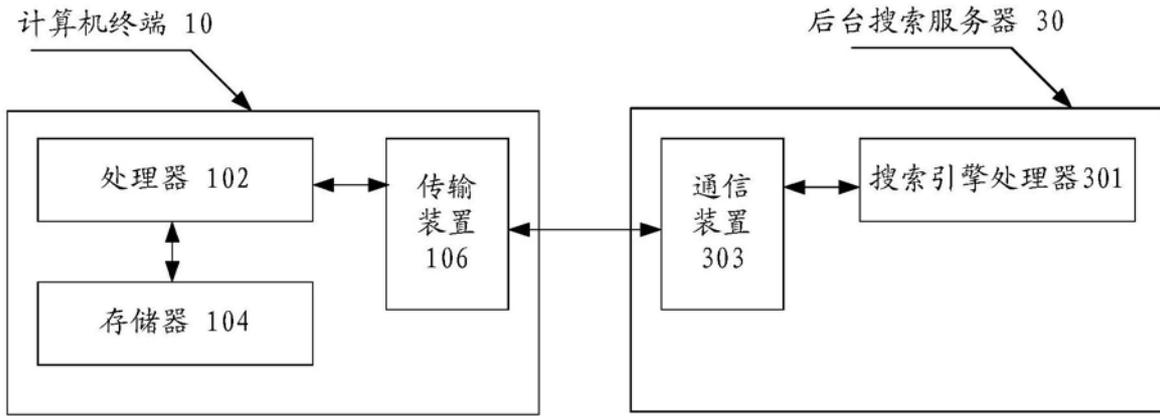


图1

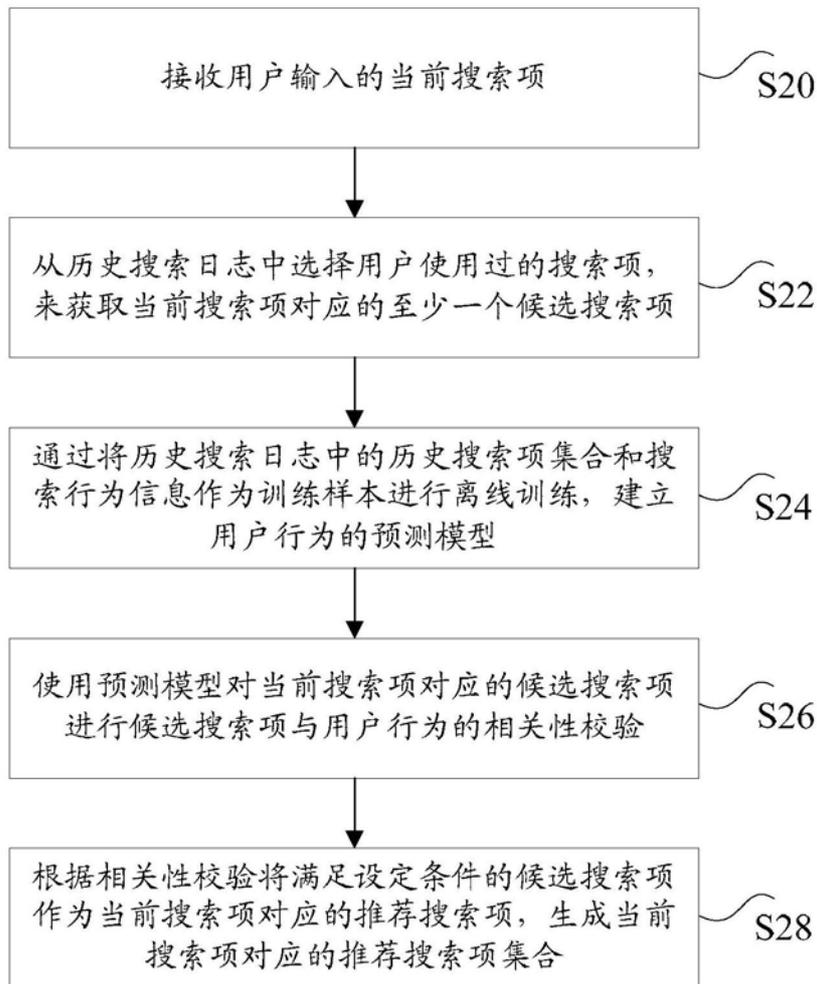


图2

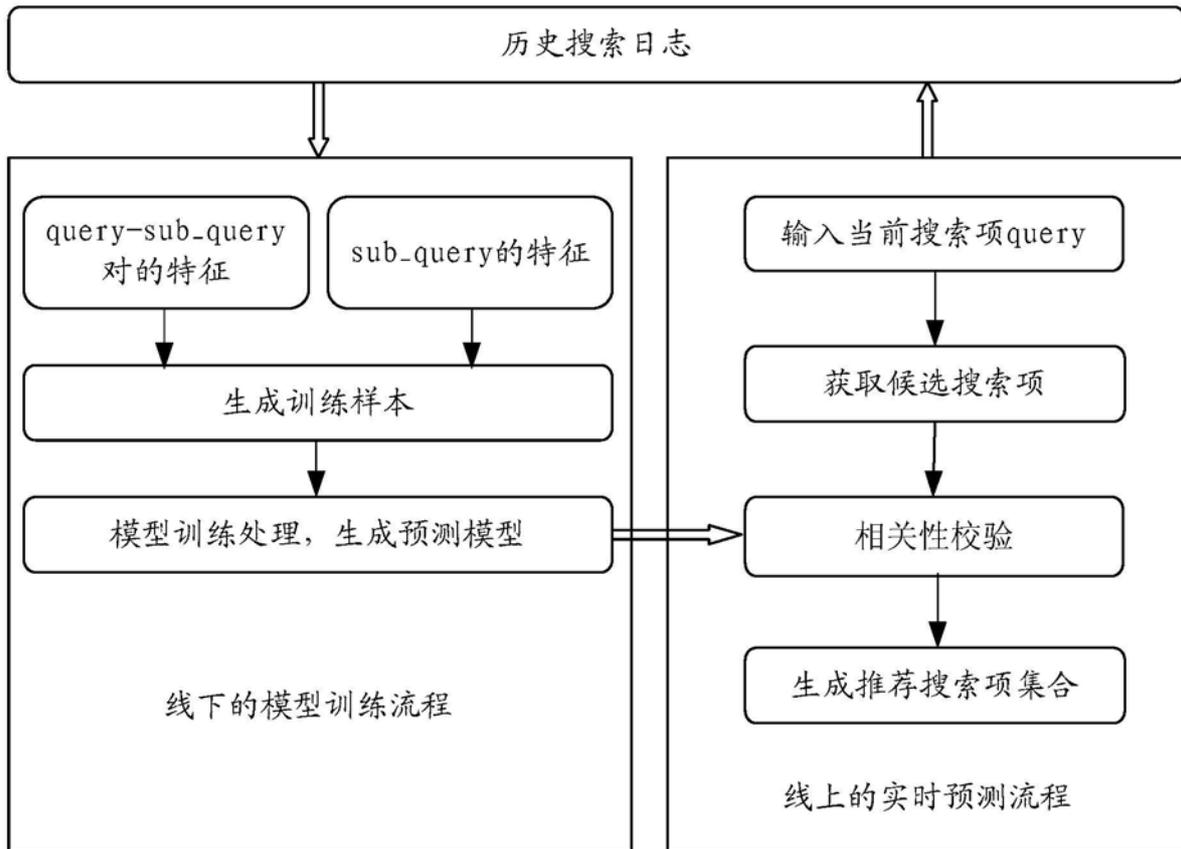


图3



图4

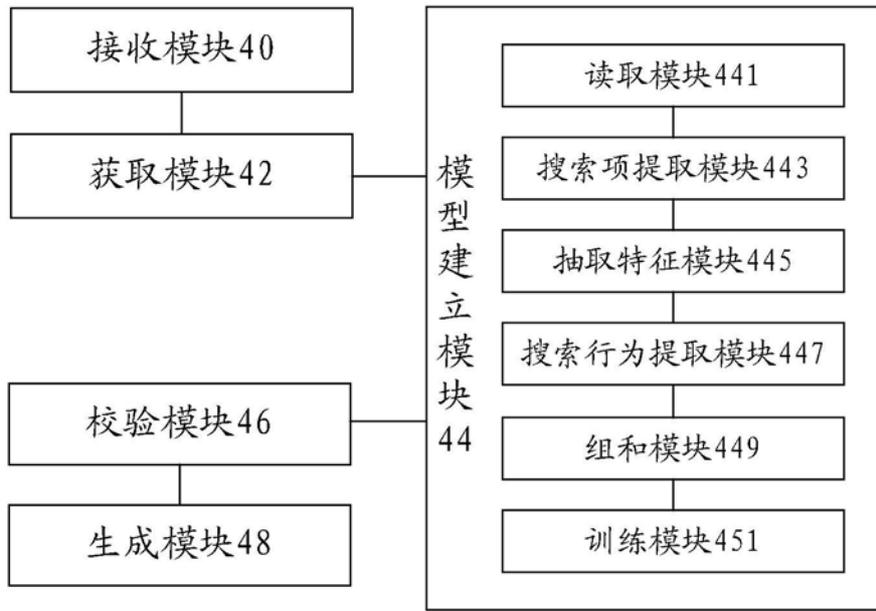


图5



图6

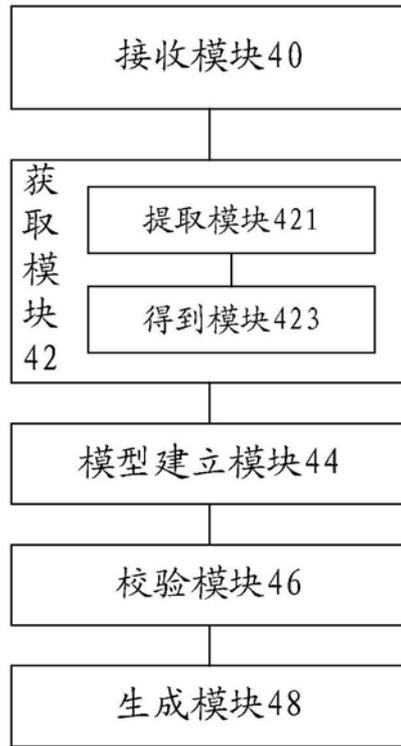


图7