



(12)发明专利申请

(10)申请公布号 CN 110851651 A

(43)申请公布日 2020.02.28

(21)申请号 201911088257.6

(22)申请日 2019.11.08

(71)申请人 杭州趣维科技有限公司

地址 310012 浙江省杭州市西湖区天目山路294号杭钢、冶金科技大厦16层

(72)发明人 李文杰 范俊 张智伟 顾湘余

(74)专利代理机构 杭州橙知果专利代理事务所  
(特殊普通合伙) 33261

代理人 贺龙萍

(51)Int.Cl.

G06F 16/735(2019.01)

G06F 16/783(2019.01)

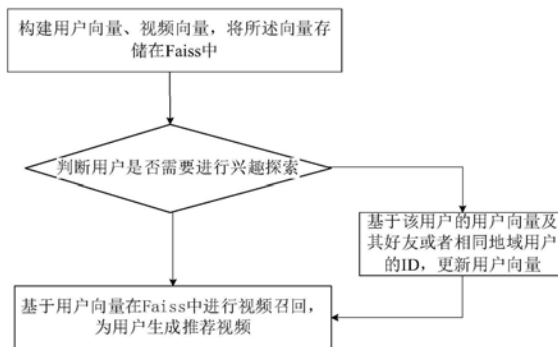
权利要求书1页 说明书8页 附图1页

(54)发明名称

一种个性化视频推荐方法及系统

(57)摘要

本发明公开了一种个性化视频推荐方法及系统,其中方法包括:S1、构建用户向量、视频向量,将所述向量存储在Faiss中;S2、判断用户是否需要进行兴趣探索,若是,执行步骤S3、若否,执行步骤S4;S3、基于该用户的用户向量及其好友或者相同地域用户的ID,更新用户向量;S4、基于用户向量在Faiss中进行视频召回,为用户生成推荐视频。本发明在避免视频推荐过程中推荐的内容类型固定问题的同时,仍能维持高的推荐效率。同时,结合社交关系网及自身特性进行推荐,推荐效果好。



1. 一种个性化视频推荐方法,其特征在于,包括:
  - S1、构建用户向量、视频向量,将所述向量存储在Faiss中;
  - S2、判断用户是否需要兴趣探索,若是,执行步骤S3、若否,执行步骤S4;
  - S3、基于该用户的用户向量及其好友或者相同地域用户的ID,更新用户向量;
  - S4、基于用户向量在Faiss中进行视频召回,为用户生成推荐视频。
2. 根据权利要求1所述的个性化视频推荐方法,其特征在于,所述步骤S1具体为:
  - S1.1、采用特征表示用户、视频的相关属性;
  - S1.2、基于因子分解机训练生成用户、视频的特征向量;
  - S1.3、对所述特征向量求和生成用户向量、视频向量。
3. 根据权利要求1所述的个性化视频推荐方法,其特征在于,通过计算概率随机数判断是否需要兴趣探索。
4. 根据权利要求3所述的个性化视频推荐方法,其特征在于,生成一个在0-1之间均匀分布的随机数,当生成的随机数大于等于0.4就进行兴趣探索,否则不进行兴趣探索。
5. 根据权利要求3所述的个性化视频推荐方法,其特征在于,所述步骤S4具体为:

采用Faiss中自带的向量函数进行运算,基于用户向量在Faiss中进行视频召回获取视频候选集;对所述视频候选集进行排序;基于策略调整排序后的视频候选集,返回推荐视频。
6. 一种个性化视频推荐系统,其特征在于,包括:

向量构建模块,用于构建用户向量、视频向量,将所述向量存储在Faiss中;

判断模块,用于判断用户是否需要兴趣探索,若是,调用向量更新模块,若否,调用推荐模块;

向量更新模块,用于基于该用户的用户向量及其好友或者相同地域用户的ID,更新用户向量;

推荐模块,用于基于用户向量在Faiss中进行视频召回,为用户生成推荐视频。
7. 根据权利要求6所述的个性化视频推荐系统,其特征在于,所述向量构建模块包括:

特征表示模块,用于采用特征表示用户、视频的相关属性;

特征向量生成模块,用于基于因子分解机训练生成用户、视频的特征向量;

特征向量求和模块,用于对所述特征向量求和生成用户向量、视频向量。
8. 根据权利要求6所述的个性化视频推荐系统,其特征在于,通过计算概率随机数判断是否需要兴趣探索。
9. 根据权利要求8所述的个性化视频推荐系统,其特征在于,生成一个在0-1之间均匀分布的随机数,当生成的随机数大于等于0.4就进行兴趣探索,否则不进行兴趣探索。
10. 根据权利要求8所述的个性化视频推荐方法,其特征在于,所述推荐模块包括:

采用Faiss中自带的向量函数进行运算,基于用户向量在Faiss中进行视频召回获取视频候选集;对所述视频候选集进行排序;基于策略调整排序后的视频候选集,返回推荐视频。

## 一种个性化视频推荐方法及系统

### 技术领域

[0001] 本发明涉及内容推荐领域,具体涉及一种个性化视频推荐方法及系统。

### 背景技术

[0002] 随着各种应用的流行,企业能收集更多更全的用户数据,如何利用这些数据提高收入是各企业都会面临的问题。最常见的方式就是个性化推荐,特别是在电商、视频网站或其它的内容平台。基于社交关系网的推荐是常用的推荐方法。相同或相近志趣的人喜欢的内容越有可能相同,因此,基于社交关系网的推荐是基于自己关注的朋友所喜欢内容去推荐。然而,在基于社交关系网的推荐过程中,仅考虑了用户间的社交关系,而没有考虑用户自身的特征,完全依赖社交关系网人物的喜好。

[0003] 此外,针对现有的个性化推荐进行推荐的内容类型趋向不变,不能不断探索用户的其他兴趣的问题,现有的推荐系统提出Multi Armed Bandit (简记为bandit)算法,Bandit算法的效果通常采用累积遗憾进行衡量,具体为:

$$[0004] \quad R_T = \sum_{i=1}^T (w_{opt} - w_{B(i)}) = Tw^* - \sum_{i=1}^T w_{B(i)}$$

[0005] 推荐的收益用户点击和没有点击,即非0即1,也就是伯努利收益,其中, $w_B(i)$ 是第*i*次试验时被选中臂的期望收益, $w^*$ 是所有臂中的最佳那个, $T$ 为试验次数。

[0006] 该方法通过累计遗憾,来选择最佳的方案,具体步骤如下:

[0007] 1) 假设概率*p*的概率分布符合beat (wins,loses)分布,有2个参赛wins,loses;

[0008] 2) 对应的每一类维护beta分布参数,每次实验,有收益wins加一,没有收益loses加一;

[0009] 3) 推荐的时候每次根据每个类的beta分布产生一个随机数*b*,选择所有类中随机数最大的去推荐。

[0010] 该算法赋予商品(视频/文章)一个Beta分布,而不是单一的值。每次排序时通过采样随机获得一个排序值,通过这种随机性来加大商品排序的变动性。但同时beta分布的均值会跟着商品的表现而改变,这样就能对商品进行区分,让表现好的商品有更大的概率获得高排序值,而不会像均匀分布那样每个商品的曝光机率永远保持不变。

[0011] 然而,Bandit算法在通过累计用户表现来选择推荐的结果,当开始的时候用完对该类数据少的时候,结果可信度不高,因此在算法初期效果不佳;此外,当探索的类别过多的时候,每次需要计算每个类的随机数,耗时多;每个用户的每个类别都需要维护一个beta分布参数,存储消耗大。

[0012] 因此,如何克服现有个性化推荐的缺点,实现高效、低耗的个性化视频推荐是本领域亟待解决的问题。

### 发明内容

[0013] 本发明的目的是针对现有技术的缺陷,提供了一种个性化视频推荐方法及系统。

通过对用户向量的随机更新,实现视频推荐过程中推荐的内容类型多样化,同时仍能维持高的推荐效率。此外,与现有的召回算法结合,系统开销小。

[0014] 为了实现以上目的,本发明采用以下技术方案:

[0015] 一种个性化视频推荐方法,包括:

[0016] S1、构建用户向量、视频向量,将所述向量存储在Faiss中;

[0017] S2、判断用户是否需要进行搜索,若是,执行步骤S3、若否,执行步骤S4;

[0018] S3、基于该用户的用户向量及其好友或者相同地域用户的ID,更新用户向量;

[0019] S4、基于用户向量在Faiss中进行视频召回,为用户生成推荐视频。

[0020] 进一步地,所述步骤S1具体为:

[0021] S1.1、采用特征表示用户、视频的相关属性;

[0022] S1.2、基于因子分解机训练生成用户、视频的特征向量;

[0023] S1.3、对所述特征向量求和生成用户向量、视频向量。

[0024] 进一步地,通过计算概率随机数判断是否需要进行搜索。

[0025] 进一步地,生成一个在0-1之间均匀分布的随机数,当生成的随机数大于等于0.4就进行搜索,否则不进行搜索。

[0026] 进一步地,所述步骤S4具体为:

[0027] 采用Faiss中自带的向量函数进行运算,基于用户向量在Faiss中进行视频召回获取视频候选集;对所述视频候选集进行排序;基于策略调整排序后的视频候选集,返回推荐视频。

[0028] 本发明还提出一种个性化视频推荐系统,包括:

[0029] 向量构建模块,用于构建用户向量、视频向量,将所述向量存储在Faiss中;

[0030] 判断模块,用于判断用户是否需要进行搜索,若是,调用向量更新模块,若否,调用推荐模块;

[0031] 向量更新模块,用于基于该用户的用户向量及其好友或者相同地域用户的ID,更新用户向量;

[0032] 推荐模块,用于基于用户向量在Faiss中进行视频召回,为用户生成推荐视频。

[0033] 进一步地,所述向量构建模块包括:

[0034] 特征表示模块,用于采用特征表示用户、视频的相关属性;

[0035] 特征向量生成模块,用于基于因子分解机训练生成用户、视频的特征向量;

[0036] 特征向量求和模块,用于对所述特征向量求和生成用户向量、视频向量。

[0037] 进一步地,通过计算概率随机数判断是否需要进行搜索。

[0038] 进一步地,生成一个在0-1之间均匀分布的随机数,当生成的随机数大于等于0.4就进行搜索,否则不进行搜索。

[0039] 进一步地,所述推荐模块包括:

[0040] 采用Faiss中自带的向量函数进行运算,基于用户向量在Faiss中进行视频召回获取视频候选集;对所述视频候选集进行排序;基于策略调整排序后的视频候选集,返回推荐视频。

[0041] 与现有技术相比,本发明具有如下优点:

[0042] (1) 本发明并不为每次用户的请求进行搜索。而是通过计算概率随机数的方

式进行随机探索。在避免视频推荐过程中推荐的内容类型固定问题的同时，仍能维持高的推荐效率；

[0043] (2) 本发明结合了用户本身特征和关系网好友特征去推荐，克服了现有的基于社交关系网推荐的方法中，仅依赖社交关系网，而没有考虑到自身特性的问题；

[0044] (3) 本发明可以与现有的日常召回算法结合，不需要额外增加系统开销，实现视频的快速召回；

[0045] (4) 本发明基于FM进行特征组合，构建用户、视频的特征向量，计算复杂度低，但效果有大大提升；同时适用于在大规模稀疏特征应用环境下，泛化能力强；

[0046] (5) 本发明将用户向量、视频向量存储在Faiss中，基于Faiss中自带的向量函数进行运算，实现视频的快速召回。

## 附图说明

[0047] 图1是实施例一提供的一种个性化视频推荐的方法流程图；

[0048] 图2是实施例二提供的一种个性化视频推荐的系统结构图。

## 具体实施方式

[0049] 以下通过特定的具体实例说明本发明的实施方式，本领域技术人员可由本说明书所揭露的内容轻易地了解本发明的其他优点与功效。本发明还可以通过另外不同的具体实施方式加以实施或应用，本说明书中的各项细节也可以基于不同观点与应用，在没有背离本发明的精神下进行各种修饰或改变。需说明的是，在不冲突的情况下，以下实施例及实施例中的特征可以相互组合。

[0050] 需要说明的是，以下实施例中所提供的图示仅以示意方式说明本发明的基本构想，遂图式中仅显示与本发明中有关的组件而非按照实际实施时的组件数目、形状及尺寸绘制，其实际实施时各组件的型态、数量及比例可为一种随意的改变，且其组件布局型态也可能更为复杂。

[0051] 下面结合附图和具体实施例对本发明作进一步说明，但不作为本发明的限定。

[0052] 实施例一

[0053] 如图1所示，本实施例提出了一种个性化视频推荐方法，包括：

[0054] S1、构建用户向量、视频向量，将所述向量存储在Faiss中；

[0055] 为了向用户推荐视频，本发明首先构建用户向量、视频向量，基于用户向量、视频向量进行视频推荐。具体为：

[0056] S1.1、采用特征表示用户、视频的相关属性；

[0057] 用户的属性包括基础属性，行为属性等，因此本发明可以抽取表征用户属性的用户特征，如ID类特征（用户名字等）、基础属性类特征（如性别、年龄、城市、爱好等）、统计类特征（如一周上网时间、浏览视频类型等）。相应地，用于表征视频属性的视频特征包括ID类特征（视频地址等）、基础属性类特征（如内容、展示位置等）、统计类特征（视频点击率、观看时间等）。用户、视频的特征表示可以直接从现有的画像系统中提取，也可以根据用户及视频的特征抽取代表性的属性进行特征表示，在此不作限定。

[0058] S1.2、基于因子分解机训练生成用户、视频的特征向量；

[0059] 用户、视频包括多个特征,不同的特征之间存在关联,例如,男性用户更喜欢球类比赛视频,女性更喜欢娱乐类视频。因此,特征组合在推荐排序过程中极为重要,然而,推荐过程中,特征组合量大,如果进行人工特征组合,其工作量大。因此,Rendle在2010年提出的因子分解机(Factorization Machine,FM),被各大厂大规模在CTR(click-through rate)预估和推荐领域广泛使用,FM这一思路解决大量人工特征组合的工作量。

[0060] 二阶FM模型具体为:

$$[0061] \quad \hat{y}(X) := w_0 + \sum_{i=1}^n w_i x_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^n \langle v_i, v_j \rangle x_i x_j$$

$$[0062] \quad \langle v_i, v_j \rangle := \sum_{f=1}^k v_{i,f} \cdot v_{j,f}$$

[0063] 其中, $n$ 为样本特征数量, $x_i$ 、 $x_j$ 分别为第 $i$ 、 $j$ 维特征的值, $w_0$ 、 $w_i$ 为模型参数。

$w_0 + \sum_{i=1}^n w_i x_i$ 为普通的线性模型, $\sum_{i=1}^{n-1} \sum_{j=i+1}^n \langle v_i, v_j \rangle x_i x_j$ 为交叉项即特征的组合。 $v_i$ 、 $v_j$ 是分别是第 $i$ 维、第 $j$ 维特征的隐向量, $\langle \cdot, \cdot \rangle$ 为向量的点积,隐向量的长度为 $k \ll n$ ,包含 $k$ 个描述特征的因子。

[0064] FM模型引入任意两个特征的二阶特征组合,FM对于每个特征,学习一个大小为 $k$ 的一维向量,于是,两个特征 $x_i$ 和 $x_j$ 的特征组合的权重值,通过特征对应的向量 $v_i$ 和 $v_j$ 内积来表示。而且两个特征并未在训练实例中同时出现,也可以通过这两个特征向量内积得到新的特征组合,在大规模稀疏特征应用环境下比较好用,泛化能力强。

$$[0065] \quad \text{具体地, } \sum_{i=1}^n \sum_{j=i+1}^n \langle v_i, v_j \rangle x_i x_j$$

$$= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \langle v_i, v_j \rangle x_i x_j - \frac{1}{2} \sum_{i=1}^n \langle v_i, v_i \rangle x_i x_i$$

$$[0066] \quad = \frac{1}{2} \left( \sum_{i=1}^n \sum_{j=1}^n \sum_{f=1}^k v_{i,f} v_{j,f} x_i x_j - \sum_{i=1}^n \sum_{f=1}^k v_{i,f} v_{i,f} x_i x_i \right)$$

$$= \frac{1}{2} \left( \sum_{f=1}^k \sum_{i=1}^n v_{i,f} x_i \sum_{j=1}^n v_{j,f} x_j - \sum_{i=1}^n \sum_{f=1}^k v_{i,f} v_{i,f} x_i x_i \right)$$

[0067] 由此可知,FM的二次项化简为只与 $v_{i,f}$ 有关,FM可以在线性时间对新样本做出预测,复杂度和普通的线性模型一样,但效果有大大的提升。

[0068] FM模型训练时学习模型参数和每一个特征的权重向量。因此,本发明对于用户和视频,基于因子分解机训练生成用户、视频的一批特征权重值,即特征向量。FM模型能有效地学习模型参数, $w_i$ 和 $w_j$ 的更新不需要特征 $x_i$ 和 $x_j$ 同时出现在一个样本中。本发明FM模型通过对稀疏特征组合进行低维连续空间的转换,特征 $x_i$ 和 $x_j$ 不出现在一个样本的情况下,模

型参数依旧可以更新,因此,基于FM模型的特征组合方式能适用于特征子集 $\{x_i, x_j\}$ 稀疏的情况。

[0069] S1.3、对所述特征向量求和生成用户向量、视频向量。

[0070] 用户向量、视频向量包括多个特征向量集合。假设特征集合包括 $n$ 个特征,则用户向量、视频向量可表示为 $\{x_1, x_2, \dots, x_n\}$ ,其中, $x_i$ 表示 $n$ 个特征向量中的各个特征向量。

[0071] 为用户进行视频推荐过程中,主要包括召回和排序阶段,召回是从视频集中选取一部分作为候选集,基于用户向量和视频向量进行计算。然而,在视频召回过程中,可能需要召回大量的候选集,因此,对于单个用户的推荐,都可能需要进行大量的处理。例如,当推荐系统在召回阶段需要召回5000个视频候选集时,如果用户和视频的向量表示都是32维,则对于单个用户的单次推荐需要进行的运算处理为 $5000 \times 32 \times 32 = 512$ 万。如此大的向量计算量对于常规的索引不能承受。因此,本发明将用户向量、视频向量存储在Faiss中,基于Faiss中自带的向量函数进行运算,实现视频的快速召回。

[0072] S2、判断用户是否需要进行搜索,若是,执行步骤S3、若否,执行步骤S4;

[0073] 当用户发送请求时,为用户推荐相应的视频,发送的请求可以为访问视频网站等。视频推荐过程中,如果基于固定的信息进行推荐,如用户的喜好等,所推荐出的内容类型相对固定,不能不断探索用户的其他兴趣的问题,但是兴趣的探索需要对用户数据进行进一步处理、更新,增加了处理花销,降低了推荐的效率。因此,为了在避免视频推荐过程中推荐的内容类型固定问题的同时,仍能维持高的推荐效率,本发明并不为每次用户的请求进行兴趣探索。而是通过计算概率随机数的方式进行随机探索。

[0074] 例如,生成一个在0-1之间均匀分布的随机数,当生成的随机数大于等于0.4就进行兴趣探索,否则不进行兴趣探索。

[0075] S3、基于该用户的用户向量及其好友或者相同地域用户的ID,更新用户向量;

[0076] 当进行兴趣探索时,需要对用户的向量进行更新。具体地,当获取到用户请求时,首先根据用户ID等信息构建用户向量,其具体步骤与步骤S1一致,在此不再赘述。当进行兴趣探索的时候,将该用户的ID类特征替换成好友或者相同地域用户的ID,求出新的用户向量。因此,本发明进行兴趣探索时,结合了好友及相同地域用户的信息,根据好友及相同地域用户的兴趣去拓展该用户的兴趣。此外,兴趣探索过程中,仅对ID类特征进行替换,保留了用户自身的基础属性类、统计类特征等,因此,本发明结合了用户本身特征和关系网好友特征去推荐,克服了现有的基于社交关系网推荐的方法中,仅依赖社交关系网,而没有考虑到自身特性的问题。

[0077] S4、基于用户向量在Faiss中进行视频召回,为用户生成推荐视频。

[0078] 个性化视频推荐中主要包括三部分:视频召回、排序、以及策略调整。其中,视频召回是指海量的视频中选择一部分视频作为候选集。排序主要是调整召回完的候选集中视频的展现顺序,例如,将候选集中的视频基于点击率大小由高到低进行排序。排序完之后还有一些策略的调整,比如控制相同上传作者的视频数目等。

[0079] 本发明将样本集中的用户向量、视频向量存储在Faiss中,因此,当用户发出请求后,如果用户需要进行探索,则基于更新后的用户向量在Faiss中进行视频召回,召回的计算过程基于Faiss中自带的向量函数进行运算,实现视频的快速召回。如果用户不需要进行探索,则不对用户向量进行更新,直接基于用户自身信息构建的向量进行召回。



[0080] 本发明不对具体的召回算法进行限定,可以与现有的日常召回算法结合,不需要额外增加系统开销,实现视频的快速召回。

[0081] 实施例二

[0082] 如图2所示,本实施例提出了一种个性化视频推荐系统,包括:

[0083] 向量构建模块,用于构建用户向量、视频向量,将所述向量存储在Faiss中;

[0084] 为了向用户推荐视频,本发明首先构建用户向量、视频向量,基于用户向量、视频向量进行视频推荐。具体包括:

[0085] 特征表示模块,用于采用特征表示用户、视频的相关属性;

[0086] 用户的属性包括基础属性,行为属性等,因此本发明可以抽取表征用户属性的用户特征,如ID类特征(用户名字等)、基础属性类特征(如性别、年龄、城市、爱好等)、统计类特征(如一周上网时间、浏览视频类型等)。相应地,用于表征视频属性的视频特征包括ID类特征(视频地址等)、基础属性类特征(如内容、展示位置等)、统计类特征(视频点击率、观看时间等)。用户、视频的特征表示可以直接从现有的画像系统中提取,也可以根据用户及视频的特征抽取代表性的属性进行特征表示,在此不作限定。

[0087] 特征向量生成模块,用于基于因子分解机训练生成用户、视频的特征向量;

[0088] 用户、视频包括多个特征,不同的特征之间存在关联,例如,男性用户更喜欢球类比赛视频,女性更喜欢娱乐类视频。因此,特征组合在推荐排序过程中极为重要,然而,推荐过程中,特征组合量大,如果进行人工特征组合,其工作量大。因此,Rendle在2010年提出的因子分解机(Factorization Machine, FM),被各大厂大规模在CTR(click-through rate)预估和推荐领域广泛使用,FM这一思路解决大量人工特征组合的工作量。

[0089] 二阶FM模型具体为:

$$[0090] \quad \hat{y}(X) := w_0 + \sum_{i=1}^n w_i x_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^n \langle v_i, v_j \rangle x_i x_j$$

$$[0091] \quad \langle v_i, v_j \rangle := \sum_{f=1}^k v_{i,f} \cdot v_{j,f}$$

[0092] 其中, $n$ 为样本特征数量, $x_i$ 、 $x_j$ 分别为第 $i$ 、 $j$ 维特征的值, $w_0$ 、 $w_i$ 为模型参数。

$w_0 + \sum_{i=1}^n w_i x_i$ 为普通的线性模型, $\sum_{i=1}^{n-1} \sum_{j=i+1}^n \langle v_i, v_j \rangle x_i x_j$ 为交叉项即特征的组合。 $v_i$ 、 $v_j$ 是分别是

第 $i$ 维、第 $j$ 维特征的隐向量, $\langle \cdot, \cdot \rangle$ 为向量的点积,隐向量的长度为 $k \ll n$ ,包含 $k$ 个描述特征的因子。

[0093] FM模型引入任意两个特征的二阶特征组合,FM对于每个特征,学习一个大小为 $k$ 的一维向量,于是,两个特征 $x_i$ 和 $x_j$ 的特征组合的权重值,通过特征对应的向量 $v_i$ 和 $v_j$ 内积来表示。而且两个特征并未在训练实例中同时出现,也可以通过这两个特征向量内积得到新的特征组合,在大规模稀疏特征应用环境下比较好用,泛化能力强。

$$[0094] \quad \text{具体地, } \sum_{i=1}^n \sum_{j=i+1}^n \langle v_i, v_j \rangle x_i x_j$$



$$\begin{aligned}
&= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \langle v_i, v_j \rangle x_i x_j - \frac{1}{2} \sum_{i=1}^n \langle v_i, v_i \rangle x_i x_i \\
[0095] \quad &= \frac{1}{2} \left( \sum_{i=1}^n \sum_{j=1}^n \sum_{f=1}^k v_{if} v_{jf} x_i x_j - \sum_{i=1}^n \sum_{f=1}^k v_{if} v_{if} x_i x_i \right) \\
&= \frac{1}{2} \left( \sum_{f=1}^k \sum_{i=1}^n v_{if} x_i \sum_{j=1}^n v_{jf} x_j - \sum_{i=1}^n \sum_{f=1}^k v_{if} v_{if} x_i x_i \right)
\end{aligned}$$

[0096] 由此可知,FM的二次项化简为只与 $v_{i,f}$ 有关,FM可以在线性时间对新样本做出预测,复杂度和普通的线性模型一样,但效果有大的提升。

[0097] FM模型训练时学习模型参数和每一个特征的权重向量。因此,本发明对于用户和视频,基于因子分解机训练生成用户、视频的一批特征权重值,即特征向量。FM模型能有效地学习模型参数, $\omega_i$ 和 $\omega_j$ 的更新不需要特征 $x_i$ 和 $x_j$ 同时出现在一个样本中。本发明FM模型通过对稀疏特征组合进行低维连续空间的转换,特征 $x_i$ 和 $x_j$ 不出现在一个样本的情况下,模型参数依旧可以更新,因此,基于FM模型的特征组合方式能适用于特征子集 $\{x_i, x_j\}$ 稀疏的情况。

[0098] 特征向量求和模块,用于对所述特征向量求和生成用户向量、视频向量。

[0099] 用户向量、视频向量包括多个特征向量集合。假设特征集合包括 $n$ 个特征,则用户向量、视频向量可表示为 $\{x_1, x_2, \dots, x_n\}$ ,其中, $x_i$ 表示 $n$ 个特征向量中的各个特征向量。

[0100] 为用户进行视频推荐过程中,主要包括召回和排序阶段,召回是从视频集中选取一部分作为候选集,基于用户向量和视频向量进行计算。然而,在视频召回过程中,可能需要召回大量的候选集,因此,对于单个用户的推荐,都可能需要进行大量的处理。例如,当推荐系统在召回阶段需要召回5000个视频候选集时,如果用户和视频的向量表示都是32维,则对于单个用户的单次推荐需要进行的运算处理为 $5000 \times 32 \times 32 = 512$ 万。如此大的向量计算量对于常规的索引不能承受。因此,本发明将用户向量、视频向量存储在Faiss中,基于Faiss中自带的向量函数进行运算,实现视频的快速召回。

[0101] 判断模块,用于判断用户是否需要继续进行兴趣探索,若是,调用向量更新模块,若否,调用推荐模块;

[0102] 当用户发送请求时,为用户推荐相应的视频,发送的请求可以为访问视频网站等。视频推荐过程中,如果基于固定的信息进行推荐,如用户的喜好等,所推荐出的内容类型相对固定,不能不断探索用户的其他兴趣的问题,但是兴趣的探索需要对用户数据进行进一步处理、更新,增加了处理花销,降低了推荐的效率。因此,为了在避免视频推荐过程中推荐的内容类型固定问题的同时,仍能维持高的推荐效率,本发明并不为每次用户的请求进行兴趣探索。而是通过计算概率随机数的方式进行随机探索。

[0103] 例如,生成一个在0-1之间均匀分布的随机数,当生成的随机数大于等于0.4就进行兴趣探索,否则不进行兴趣探索。

[0104] 向量更新模块,用于基于该用户的用户向量及其好友或者相同地域用户的ID,更新用户向量;

[0105] 当进行兴趣探索时,需要对用户的向量进行更新。具体地,当获取到用户请求时,

首先根据用户ID等信息构建用户向量,其具体步骤与步骤S1一致,在此不再赘述。当进行兴趣探索的时候,将该用户的ID类特征替换成好友或者相同地域用户的ID,求出新的用户向量。因此,本发明进行兴趣探索时,结合了好友及相同地域用户的信息,根据好友及相同地域用户的兴趣去拓展该用户的兴趣。此外,兴趣探索过程中,仅对ID类特征进行替换,保留了用户自身的基础属性类、统计类特征等,因此,本发明结合了用户本身特征和关系网好友特征去推荐,克服了现有的基于社交关系网推荐的方法中,仅依赖社交关系网,而没有考虑到自身特性的问题。

[0106] 推荐模块,用于基于用户向量在Faiss中进行视频召回,为用户生成推荐视频。

[0107] 个性化视频推荐中主要包括三部分:视频召回、排序、以及策略调整。其中,视频召回是指海量的视频中选择一部分视频作为候选集。排序主要是调整召回完的候选集中视频的展现顺序,例如,将候选集中的视频基于点击率大小由高到低进行排序。排序完之后还有一些策略的调整,比如控制相同上传作者的视频数目等。

[0108] 本发明将样本集中的用户向量、视频向量存储在Faiss中,因此,当用户发出请求后,如果用户需要进行探索,则基于更新后的用户向量在Faiss中进行视频召回,召回的计算过程基于Faiss中自带的向量函数进行运算,实现视频的快速召回。如果用户不需要进行探索,则不对用户向量进行更新,直接基于用户自身信息构建的向量进行召回。

[0109] 本发明不对具体的召回算法进行限定,可以与现有的日常召回算法结合,不需要额外增加系统开销,实现视频的快速召回。

[0110] 由此可知,本发明提出的个性化推荐方法及系统,并不为每次用户的请求进行兴趣探索。而是通过计算概率随机数的方式进行随机探索。在避免视频推荐过程中推荐的内容类型固定问题的同时,仍能维持高的推荐效率;结合了用户本身特征和关系网好友特征去推荐,克服了现有的基于社交关系网推荐的方法中,仅依赖社交关系网,而没有考虑到自身特性的问题;同时可以与现有的日常召回算法结合,不需要额外增加系统开销,实现视频的快速召回;此外基于FM进行特征组合,构建用户、视频的特征向量,计算复杂度低,但效果有大大的提升;同时适用于在大规模稀疏特征应用环境下,泛化能力强;最后,本发明将用户向量、视频向量存储在Faiss中,基于Faiss中自带的向量函数进行运算,实现视频的快速召回。

[0111] 注意,上述仅为本发明的较佳实施例及所运用技术原理。本领域技术人员会理解,本发明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发明的保护范围。因此,虽然通过以上实施例对本发明进行了较为详细的说明,但是本发明不仅仅限于以上实施例,在不脱离本发明构思的情况下,还可以包括更多其他等效实施例,而本发明的范围由所附的权利要求范围决定。

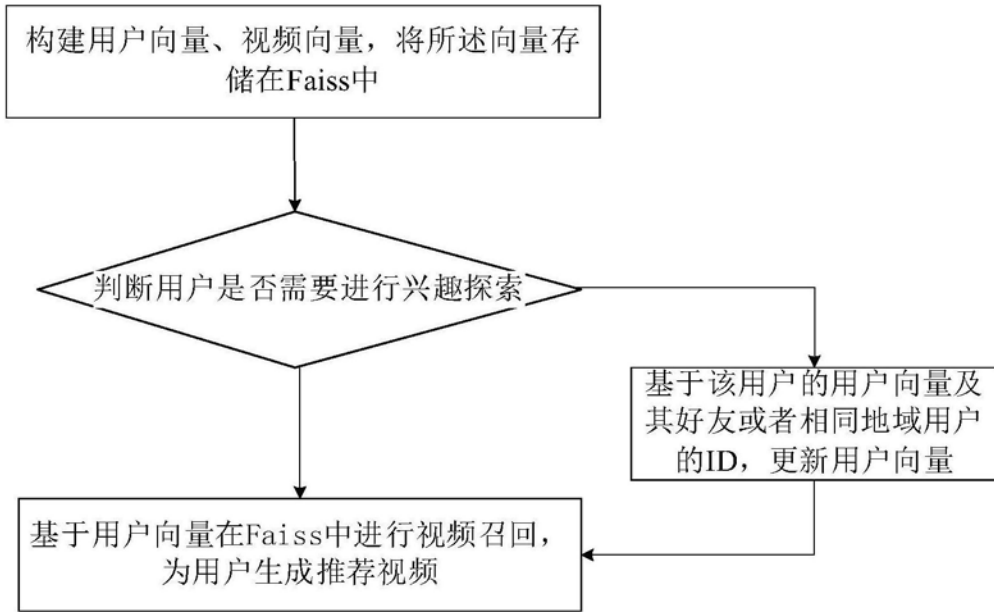


图1

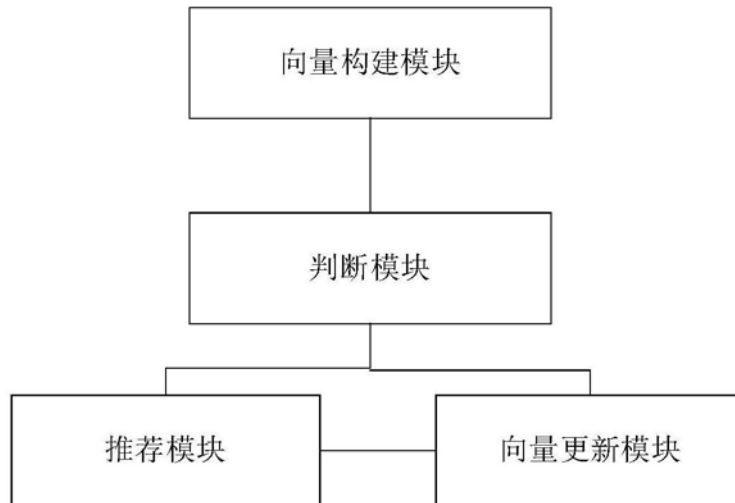


图2