



(12) 发明专利申请

(10) 申请公布号 CN 117973535 A

(43) 申请公布日 2024. 05. 03

(21) 申请号 202410045363.0

G06N 3/082 (2023.01)

(22) 申请日 2024.01.12

(71) 申请人 大连理工大学

地址 116024 辽宁省大连市甘井子区凌工
路2号

(72) 发明人 张童 孔雨秋 张立和 尹宝才

(74) 专利代理机构 辽宁鸿文知识产权代理有限
公司 21102

专利代理师 王海波

(51) Int. Cl.

G06N 5/04 (2023.01)

G06N 5/025 (2023.01)

G06N 3/045 (2023.01)

G06N 3/042 (2023.01)

G06N 3/0985 (2023.01)

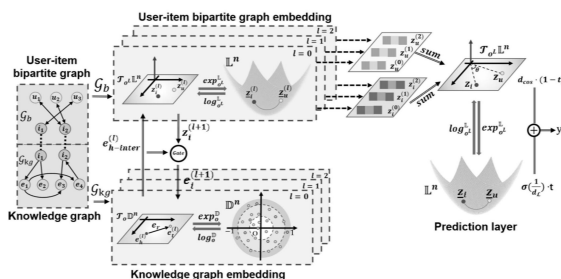
权利要求书2页 说明书7页 附图3页

(54) 发明名称

一种多双曲空间的知识感知推荐方法

(57) 摘要

本发明属于大数据挖掘技术领域,具体涉及一种多双曲空间的知识感知推荐方法。本发明提出了一种基于多个双曲空间的网络模型,并设计了不同双曲空间之间的特征交互,从而有效地学习到不同分布的数据。并且因为是基于双曲空间的知识感知推荐方法,继承了双曲嵌入的相关优势,例如:避免了嵌入失真的问题,获得更有层次性的嵌入结果等优点,这些都是现有的基于欧式空间的知识感知推荐方法不具备的。本发明突破了现有的基于知识感知的推荐方法的无法判断用户对商品的知识属性存在敏感性枷锁,将双曲距离信息整合到嵌入学习中,挖掘用户商品二部图的潜在层次,并对每个用户对商品的知识属性的敏感程度做出打分。



1. 一种多双曲空间的知识感知推荐方法,其特征在于,步骤如下:

第一步:对历史交互与知识图谱数据进行预处理

第二步:构建基于多双曲空间的知识感知推荐系DHN模型,该模型包括知识图谱编码器、用户商品二部图编码器、编码融合模块以及预测模块;

2.1 知识图谱编码器

在DHN中,利用经典的知识图谱嵌入模型TransE的庞加莱空间版本来学习嵌入;使用双曲空间中的平行位移的方法,使得能区分不同的头节点对关系与尾节点的影响;对商品节点 h ,具体的公式如下:

$$\underline{e}_{h-inter}^{(l)} = \frac{1}{|N_h|} \sum_{(r,t) \in N_h} \log_O^D \left(\exp_{\underline{e}_h^{(l)}}^D \left(P_{O \rightarrow \underline{e}_h^{(l)}}^D (e_r + e_t^{(l)}) \right) \right),$$

其中,带下划线的字母表示双曲空间中的表示,不带下划线的字母表示在对应原点切空间中的表示;对于非商品节点 h ,则有 $\underline{e}_h^{(l)} = e_{h-inter}^{(l)}$ 即这就是在第1层的最终表示;

2.2 用户商品二部图编码器

DHN采用一阶邻域进行特征聚合,并通过递归聚合获得用户与商品的高阶表示;对于二部图中的用户与商品,如下:

$$\underline{z}_u^{(l+1)} = \exp_{O^L}^L \left(\frac{1}{|N_u|} \sum_{i \in N_u} \log_{O^L}^L (z_i^{(l)}) \right),$$

$$\underline{z}_i^{(l+1)} = \exp_{O^L}^L \left(\frac{1}{|N_i|} \sum_{u \in N_i} \log_{O^L}^L (\underline{z}_u^{(l)}) \cdot a_{u \rightarrow i}^{(l)} \right),$$

利用用户嵌入距离原点的位置来判断用户是否对商品的知识属性敏感,并且设计了基于嵌入到原点距离的注意力分数;这里 $a_{u \rightarrow i}^{(l)}$ 的计算公式为:

$$a_{u \rightarrow i}^{(l)} = \frac{\|z_i^{(l)}\|^2}{\|z_u^{(l)}\|^2}$$

更加靠近原点的用户将在二部图的嵌入中获得更高的权重;

2.3 编码融合模块

DHN通过增加维数来实现KG与二部图信息之间的交互,设计了一个可学习的门控模块来控制知识图谱中项目嵌入的更新:

$$g_h^{(l)} = \sigma \left(W_0 z_h^{(l+1)} [1:] + W_1 e_{h-inter}^{(l)} \right),$$

$$e_h^{(l+1)} = g_h^{(l)} \cdot z_h^{(l+1)} [1:] + (1 - g_h^{(l)}) \cdot e_{h-inter}^{(l)},$$

2.4 预测模块

在L层之后,得到了用户和物品在不同尺度上的表示,利用求和来获得用户与商品的最终表示:

$$z_i = z_i^{(0)} + \dots + z_i^{(L)},$$

$$z_u = z_u^{(0)} + \dots + z_u^{(L)},$$

DHN通过对双曲距离和余弦相似度加权求和来计算匹配分数:

$$y_{ui} = t \cdot \sigma\left(\frac{1}{d_L(\underline{z}_i, \underline{z}_u)}\right) + (1 - t) \cdot \cos(z_i, z_u),$$

其中 t 为超参数, σ 为sigmoid激活函数;

第三步:模型训练与性能评估。

2.如权利要求1所述的一种多双曲空间的知识感知推荐方法,其特征在于,所述的第三步,具体操作如下:

3.1模型训练:将批量的用户交互信息与知识图谱输入到DHN模型当中后获得模型预测结果,选择损失函数并根据模型预测结果与标签值计算损失值,选择优化算法计算模型参数的最优解;在模型训练期间,根据验证集拟合情况调整模型超参数;

3.2性能评估:确定评价指标,用于衡量测试集中模型预测结果与标签值的拟合情况,进而评估模型整体预测效果。

3.如权利要求2所述的一种多双曲空间的知识感知推荐方法,其特征在于,所述的第3.1步,具体操作如下:将批量用户历史交互数据输入到模型当中后获得模型预测结果,选择双曲距离以及切空间中的余弦相似度作为损失函数并根据模型预测结果与标签值计算损失值,选择Adam优化算法计算模型参数的最优解;在模型训练期间,根据验证集拟合情况调整模型超参数;

其中有用户商品二部图中嵌入的双曲距离损失函数:

$$L_{dis-cf} = \text{mean}_{u \in U} [\text{mean}_{i \in N_u} (d_L(\underline{z}_u, \underline{z}_i)) - \text{mean}_{j \in M_u} (d_L(\underline{z}_u, \underline{z}_j))].$$

4.如权利要求1或2或3所述的一种多双曲空间的知识感知推荐方法,其特征在于,所述的第一步,具体操作如下:

1.1数据清洗:先对知识图谱进行数据清洗,以保证每个实体至少出现在10个三元组中;而对用户商品历史交互图,则采用与神经协同过滤方法中相同的处理方法即可;

1.2数据集划分:对用户与商品的历史交互数据按比例划分为训练集、验证集和测试集,分别用于模型的训练、验证和测试。

一种多双曲空间的知识感知推荐方法

技术领域

[0001] 本发明属于大数据挖掘技术领域,具体涉及一种多双曲空间的知识感知推荐方法。

背景技术

[0002] 推荐作为互联网数据挖掘系统的重要组成部分,能够有效帮助用户从互联网海量的数据中获取有效的信息,从而提升用户体验。然而,由于传统推荐方法存在稀疏和冷启动等难题,知识图谱作为一种静态数据库可用于增强商品表示,增强推荐效果。但显然用户与商品发生交互并不完全依赖于商品的知识属性,我们尝试使用多个双曲空间作为嵌入空间,在利用知识图谱完成推荐任务的同时,获取每个用户对商品知识属性依赖度的评分,这也是本文的发明动机。接下来详细介绍这一领域中相关的背景技术。

[0003] (一) 基于知识感知的推荐方法

[0004] 用户与商品的历史交互记为 $G_b = \{(u, i) | u \in U, i \in I\}$ 表示用户 u 与商品 i 之间存在交互历史, U 与 I 表示用户与商品的集合;知识图谱记为 $G_{kg} = \{(h, r, t) | h, t \in E, r \in R\}$ 以三元组的方式存在,其中 h, t 表示三元组中的头节点与尾节点, r 表示两者之间的有向关系,其中 E 与 R 表示知识图谱中的实体与关系的集合。一般情况下有 $I \subseteq E$,即商品作为知识图谱中的部分实体节点存在。基于知识感知的推荐方法旨在寻找一个映射函数 f 使得能够以知识图谱与用户与商品的历史交互作为输入来预测这个用户与其他商品发生交互的可能性,并进行排序。将上述内容进行公式化如下所示:

$$[0005] \quad \{s_1, s_2, \dots, s_n\} = f_u(G_b, G_{kg})$$

[0006] 其中 s_k 是指用户 u 对第 k 商品的打分。根据图中节点之间的消息传递机制,现有的基于知识感知的推荐方法可以分为根据路径的方法与根据图神经网络的方法。其中主要以基于图神经网络为主,如KGAT, KGIN等。然而,现有的方法主要有两个问题:(1)都使用欧氏空间作为嵌入空间,并没有考虑到节点的自然增长速度会大过空间的表示能力,这将会导致嵌入失真,影响最终的推荐效果。同时,这些方法也并没有对用户对知识属性的依赖程度进行分析,直接假设所有的用户的交互行为都会与商品的知识属性挂钩。

[0007] (二) 双曲空间

[0008] 双曲空间是只那些曲率设定为常负数的非欧几里得空间。与欧几里得空间相比,双曲空间具有优越的表征能力。双曲空间的表示能力随着空间半径呈指数增长,使其善于捕捉数据的潜在层次属性,而欧式空间的表示能力随空间半径只呈现多项式级别的增长。常见的双曲空间有洛伦兹空间、洛伦兹空间等。其中常负曲率为-1的 n 维洛伦兹流形被定义为黎曼流形 (L^n, g^L) :

$$[0009] \quad L^n = \{x \in \mathbb{R}^{n+1} | \langle x, x \rangle_L = -1, x_0 > 0\},$$

$$[0010] \quad g^L = \text{diag}\{-1, 1, \dots, 1\} \in \mathbb{R}^{n+1},$$

$$[0011] \quad \langle x, y \rangle_L = x^T g^L y,$$

[0012] 这里 $\langle x, x \rangle_L$ 指洛伦兹空间的内积, g^L 是洛伦兹空间的黎曼度量形如双曲面的上半

平面, x_0 指的是 x 的第一个维度。而常负曲率为-1的 n 维庞加莱流形被定义为黎曼流形 (D^n, g_x^D) :

$$[0013] \quad D^n = \{x \in \mathbb{R}^n \mid \langle x, x \rangle_2 < 1\},$$

$$[0014] \quad g_x^D = \left(\frac{2}{1 - \|x\|_2^2} \right)^2 g^E,$$

[0015] 这里 g^E 是 n 维的标准单位阵。注意到,在欧式空间中的四则运算对双曲空间并不满足封闭性,为了能够将双曲空间应用于深度学习中,我们将利用到双曲空间的切空间。切空间是一个局部的欧式空间,我们将这个局部的性质假设其在全局成立,这样可以在切空间中完成运算后将切空间中的结果投影回到双曲空间。我们将数据从双曲空间投影到切空间的函数称为对数函数,将数据重新投影回到双曲空间的函数称为指数函数。这一对函数在洛伦兹空间与庞加莱空间中都有明确的定义,分别记做:

$$[0016] \quad \exp_x^L(z) = \cosh(\|z\|)x + \sinh(\|z\|) \frac{z}{\|z\|},$$

$$[0017] \quad \log_x^L(y) = d_L(x, y) \frac{y + \langle x, y \rangle_L x}{\|y + \langle x, y \rangle_L x\|},$$

$$[0018] \quad \exp_x^D(z) = x \oplus \tanh\left(\frac{\|z\|}{1 - \|z\|}\right) \frac{z}{\|z\|},$$

$$[0019] \quad \log_x^D(y) = (1 - \|x\|^2) \tanh^{-1}(-x \oplus y) \frac{-x \oplus y}{\|-x \oplus y\|},$$

[0020] 其中 $d_L(x, y)$ 与 \oplus 分别代表了洛伦兹空间中的距离与莫比乌斯加法。

[0021] (三) 基于双曲空间的知识感知推荐方法

[0022] 受2019年Chami等人提出的基于双曲空间的图神经网络影响,双曲空间逐步被应用于图神经网络中,并因为其优秀的表示能力而受到关注。当使用双曲空间作为嵌入空间,对原有的欧式空间进行替代时,嵌入失真的问题得到解决。此外,Facebook的AI实验室的实验发现,由于双曲空间的距离定义与其在空间中的位置相关,对远离原点位置的数据双曲距离会产生更加严格的惩罚,利用双曲空间作为嵌入空间可以在一定程度上揭示数据之间的层次关系。双曲空间中的距离设置如下所示:

$$[0023] \quad d_L(x, y) = \operatorname{arcosh}(-\langle x, y \rangle_L),$$

$$[0024] \quad d_D(x, y) = \operatorname{arcosh}\left(1 + 2 \frac{\|x - y\|^2}{(1 - \|x\|^2)(1 - \|y\|^2)}\right),$$

[0025] 此前,LKGR与HAKG都就将双曲空间应用于基于知识感知的推荐任务进行了尝试,并且都提升了推荐的效果。其中LKGR将KGAT的方法在洛伦兹空间中实现,利用切空间计算不同节点之间的关系的权值,进行迭代计算,其主要迭代公式为:

$$[0026] \quad s_{N(u)} = \prod_u^{\exp, c} \left(\sum_{i \in N(u)} \hat{\pi}(u, r^*, i) \prod_u^{\log, c}(i) \right),$$

$$[0027] \quad s_{N(i)} = \prod_u^{\exp, c} \left(\sum_{u \in N(i)} \hat{\pi}(u, r^*, i) \prod_u^{\log, c}(u) + \sum_{e \in N_{kg}(i)} \hat{\pi}(u, r, e) \prod_u^{\log, c}(e) \right),$$

[0028] 而HAKG使用了庞加莱空间,学习到了数据之间的层次关系,文章的主要方程式为:

$$[0029] \quad e_i^{(l)} = \exp_0 \left(\frac{1}{|N_i|} \sum_{(r,t) \in N_i} \log_{e_i^{(l-1)}} (e_t^{(l-1)} \oplus e_r) \right),$$

[0030] 然而,现有的基于双曲空间的知识感知推荐方法使用用户与商品嵌入之间的余弦相似度来衡量用户对商品的感兴趣的程度,并没有建立基于双曲距离的损失。这将在一定程度上损害建模的层次属性。同时,他们都仅仅使用了相同的双曲空间,没有考虑到知识图谱与交互图之间的数据分布存在区别。

[0031] 此外,目前的方法因为完全建立在用户的交互行为都会与商品的知识属性挂钩的假设上,没有考虑到不同用户对商品知识属性的不同的接受程度。

发明内容

[0032] 为了突破现有的基于知识感知的推荐方法的缺陷,本发明通过使用多个双曲空间来处理不同的数据分布来提升模型的预测效果。同时,建立了一种基于双曲距离的注意力机制,利用双曲距离来进一步的增强建模的层次属性。本发明将提出了模型称为DHN即基于多双曲空间的知识感知推荐方法。此外,DHN还设计了方案获取每个用户对知识属性的敏感程度的指标。

[0033] 本发明的技术方案:

[0034] 一种多双曲空间的知识感知推荐方法,步骤如下:

[0035] 第一步:对历史交互与知识图谱数据进行预处理

[0036] 1.1数据清洗:对于原始数据,特别是知识图谱中,存在部分的节点出现次数过少的情况。在大部分情况下,这些节点反而会对最终的学习产生干扰。为了避免这种情况的出现,我们应该先对知识图谱进行数据清洗,以保证每个实体至少出现在10个三元组中。而对用户商品历史交互图,则采用与神经协同过滤方法中相同的处理方法即可。

[0037] 1.2数据集划分:对用户与商品的历史交互数据按比例划分为训练集、验证集和测试集,分别用于模型的训练、验证和测试。

[0038] 第二步:构建基于多双曲空间的知识感知推荐系DHN模型,该模型由知识图谱编码器、用户商品二部图编码器、编码融合模块以及预测模块这四部分构成。

[0039] 2.1知识图谱编码器

[0040] 知识图谱编码器可以在有效地将实体和关系参数化为矢量表示,同时保留图结构。在DHN中,我们利用经典的知识图谱嵌入模型TransE的庞加莱空间版本来学习嵌入。TransE作为一种经典的知识图谱嵌入方法,使得具有相似知识属性的实体在嵌入空间中足够的接近。在HAKG中使用基于莫比乌斯加法的方式(\oplus)来聚合关系与节点的信息,但是这种方式忽视了不同的头节点对模型学习的影响。我们使用双曲空间中的平行位移的方法,使得可以区分不同的头节点对关系与尾节点的影响。对商品节点h,具体的公式如下:

$$[0041] \quad e_{h-inte}^{(l)} = \frac{1}{|N_h|} \sum_{(r,t) \in N_h} \log_{e_h^{(l)}}^D \left(\exp_{e_h^{(l)}}^D \left(P_{O \rightarrow e_h^{(l)}}^D (e_r + e_t^{(l)}) \right) \right),$$

[0042] 其中,我们使用带下划线的字母表示双曲空间中的表示,用不带下划线的字母表示在对应原点切空间中的表示。我们用 $e_{h-inter}^{(l)}$ 表示第1层的切空间的图谱中间结果的表

示, \log_o^D 与 $\exp_{e_h}^D$ 分别是庞加莱空间对应点上的对数运算与指数运算。平行位移是指将对应双曲空间中某个点对应的切空间平移到另一个点对应的切空间中, 式子中 $P_{o \rightarrow e_h}^D$ 是指将庞加莱原点切空间中的点平移到头节点对应的切空间中。我们采用这种中间结果表示商品节点在知识图谱中的表示, 对于非商品节点 h , 则有 $e_h^{(l)} = e_{h-inter}^{(l)}$ 即这就是在第1层的最终表示。

[0043] 2.2 用户商品二部图编码器

[0044] 先前的研究发现, 用户商品二部图中数据经常呈现为幂律分布。为了处理二部图中的幂律分布并获取层次信息, 本发明在洛伦兹空间中学习用户商品二部图嵌入。此前 LightGCN 已经证明, 自信息聚合和特征转换对于推荐任务来说是不必要的。因此, DHN 采用一阶邻域进行特征聚合, 并通过递归聚合获得用户与商品的高阶表示。对于二部图中的用户与商品, 我们有如下的表示:

$$[0045] \quad z_u^{(l+1)} = \exp_{o^L}^L \left(\frac{1}{|N_u|} \sum_{i \in N_u} \log_{o^L}^L \left(z_i^{(l)} \right) \right),$$

$$[0046] \quad z_i^{(l+1)} = \exp_{o^L}^L \left(\frac{1}{|N_i|} \sum_{u \in N_i} \log_{o^L}^L \left(z_u^{(l)} \right) \cdot a_{u \rightarrow i}^{(l)} \right),$$

[0047] 其中, $a_{u \rightarrow i}^{(l)}$ 为设计出的与双曲距离相关的注意力权值。与以往的方法不同, DHN 引入了与双曲距离相关的约束。它要求用户和物品之间的双曲距离之和最小化。我们假设具有相同知识属性的商品在嵌入空间中的位置也相对靠近, 对于对商品的知识属性不敏感的用户, 他们交互的商品大多分散在许多不同的集群中。这些用户的嵌入应该位于原点附近, 因为远离原点的点的双曲距离惩罚更大。对于其他用户, 他们与之交互的商品的嵌入形成了空间上接近的集群, 导致这些用户的嵌入主要位于这些集群的中心附近。所以我们这里利用用户嵌入距离原点的位置来判断用户是否对商品的知识属性敏感, 并且设计了基于嵌入到原点距离的注意力分数。这里 $a_{u \rightarrow i}^{(l)}$ 的计算公式为:

$$[0048] \quad a_{u \rightarrow i}^{(l)} = \frac{\|z_i^{(l)}\|^2}{\|z_u^{(l)}\|^2}$$

[0049] 更加靠近原点的用户将在二部图的嵌入中获得更高的权重。

[0050] 2.3 编码融合模块

[0051] 在学习节点嵌入的过程中, 知识图谱和用户商品二部图的知识应该相互增强, 以获得更高质量的嵌入。为了实现这一目标, 我们的目标是在庞加莱空间和洛伦兹空间之间构建一个映射, 使不同双曲空间之间的信息交互成为可能。具体来说, DHN 通过增加维数来实现 KG 与二部图信息之间的交互。公式可以看作 $z_h^{(l)} = (0, e_{h-int}^{(l)})$, 利用庞加莱空间与洛伦兹空间在切空间上的构造的相似性完成了信息的传递。我们设计了一个可学习的门控模块来控制知识图谱中项目嵌入的更新:

$$[0052] \quad g_h^{(l)} = \sigma \left(W_0 z_h^{(l+1)} [1:] + W_1 e_{h-inter}^{(l)} \right),$$

$$[0053] \quad e_h^{(l+1)} = g_h^{(l)} \cdot z_h^{(l+1)} [1:] + (1 - g_h^{(l)}) \cdot e_{h-inter}^{(l)},$$

[0054] 这里 $g_h^{(l)} \in R^n$, W_0 和 $W_1 \in R^{n \times n}$ 都是可学习的参数, σ 表示激活函数。

[0055] 2.4 预测模块

[0056] 在L层之后,我们得到了用户和物品在不同尺度上的表示,利用求和来获得用户与商品的最终表示:

$$[0057] \quad z_i = z_i^{(0)} + \dots + z_i^{(L)},$$

$$[0058] \quad z_u = z_u^{(0)} + \dots + z_u^{(L)},$$

[0059] DHN通过对双曲距离和余弦相似度加权求和来计算匹配分数:

$$[0060] \quad y_{ui} = t \cdot \sigma \left(\frac{1}{d_L(z_i, z_u)} \right) + (1 - t) \cdot \cos(z_i, z_u),$$

[0061] 其中t为超参数, σ 为sigmoid激活函数。

[0062] 第三步:模型训练与性能评估

[0063] 3.1 模型训练:将批量的用户交互信息与知识图谱输入到DHN模型当中后获得模型预测结果,选择合适的损失函数并根据模型预测结果与标签值计算损失值,选择合适的优化算法计算模型参数的最优解。在模型训练期间,可根据验证集拟合情况调整模型超参数。

[0064] 3.2 性能评估:确定评价指标,用于衡量测试集中模型预测结果与标签值的拟合情况,进而评估模型整体预测效果。

[0065] 本发明的有益效果:

[0066] (1) 本发明提出了一种基于多个双曲空间的网络模型,并设计了不同双曲空间之间的特征交互,从而有效地学习到不同分布的数据。并且因为是基于双曲空间的知识感知推荐方法,继承了双曲嵌入的相关优势,例如:避免了嵌入失真的问题,获得更有层次性的嵌入结果等优点,这些都是现有的基于欧式空间的知识感知推荐方法不具备的。

[0067] (2) 本发明突破了现有的基于知识感知的推荐方法的无法判断用户对商品的知识属性存在敏感性枷锁,将双曲距离信息整合到嵌入学习中,挖掘用户商品二部图的潜在层次,并对每个用户对商品的知识属性的敏感程度做出打分。

附图说明

[0068] 图1为常见双曲空间庞加莱空间与洛伦兹空间对常见树状结构图的编码,展示欧式空间中嵌入失真的情况。

[0069] 图2为DHN模型的总体框架。

[0070] 图3为注意力权重计算的示意图。

[0071] 图4为DHN对嵌入位置不同的用户的评分情况图。

具体实施方式

[0072] 如图1所示,树状结构是最常见的图结构之一,随着树状结构的深度加深,节点的

数量呈现出指数级别的增长。以图中的二维空间为例,欧式空间的表示面积的增长呈现为平方级,这将远远低于节点的增长速度,这导致欧式空间中嵌入的位置非常的靠近,无法对正负例进行区分,产生了嵌入失真。双曲空间中因为表示能力更强,则不会有这种情况出现。同时,使用不同的双曲空间可以对不同分布的数据采用不同的空间度量进行约束。

[0073] 下面将结合设计意图、附图以及具体实例对本发明的具体实施方式做详细说明。

[0074] 一种多双曲空间的知识感知推荐方法,步骤如下:

[0075] 第一步:将数据进行预处理并且划分数据

[0076] (1) 数据集:使用Alibaba-iFashion数据集对本发明内容做进一步说明,该数据集在现有的基于知识感知的推荐方法研究中被广泛使用。这是一个从阿里巴巴在线购物系统收集的时尚数据集。对知识图谱中的数据采用10Core策略进行清洗。

[0077] (2) 数据划分:对数据以8:1:1的比例划分为训练集、验证集和测试集。

[0078] 第二步:前向学习过程

[0079] (1) 初始的嵌入表示:对所有的节点赋予一个独立的编号并且赋予随机的64维度的嵌入表示,用于最终的优化。

[0080] (2) 分批次输入模型:将获得的表示按批次输入图2所示的DHN模型中,不同批次的计算过程,为了使得模型更加具有鲁棒性并且避免过拟合风险,都会对知识图谱做出随机的边dropout,即不同批次的输入对应的是随机过滤掉一半数量边的知识图谱。

[0081] (3) 注意力机制:如图3所示,我们用二维的庞加莱圆盘作为例子,进一步说明了提出的注意力机制的合理性。与以往的方法不同,DHN引入了与双曲距离相关的约束。它要求用户和商品之间的双曲距离之和最小化。当该条件满足时,商品的嵌入应该位于以用户嵌入为中心的双曲圆内。由于双曲距离的几何特性,如图3所示,商品嵌入应该位于以用户为中心的液滴形状区域内。与用户的双曲嵌入相比,当项目嵌入距离原点较远时,商品与用户的余弦相似度趋于较高(如图中的 i_0)。这样我们可以用最少的计算成本估计不同尺度下用户和商品之间的余弦相似度。

[0082] (4) 预测模块:在经过所有的前向传播后,最终获得预测的结果。图4为DHN对嵌入位置不同的用户的评分情况,说明DHN得到的嵌入结果的层次性能更加明显。

[0083] 第三步:模型训练与性能评估

[0084] (1) 模型训练:将批量用户历史交互数据输入到模型当中后获得模型预测结果,选择双曲距离以及切空间中的余弦相似度作为损失函数并根据模型预测结果与标签值计算损失值,选择Adam优化算法计算模型参数的最优解。在模型训练期间,根据验证集拟合情况调整模型超参数。

[0085] 其中有用户商品二部图中嵌入的双曲距离损失函数:

$$[0086] \quad L_{dis-c} = \text{mean}_{u \in U} [\text{mean}_{i \in N_u} (d_L(\underline{z}_u, \underline{z}_i)) - \text{mean}_{j \in M_u} (d_L(\underline{z}_u, \underline{z}_j))],$$

[0087] 其中 M_u 是随机选取的负样本集合。知识图谱中的双曲损失函数与TransE的损失函数大致相似,记作 L_{dis-kg} 。还有二部图中的余弦相似度的损失函数:

$$[0088] \quad L_{ang-c} = \text{mean}_{u \in U} [1 - \text{mean}_{i \in N_u} (\cos(z_u, z_i)) - \text{mean}_{j \in M_u} (\cos(z_u, z_j) - m)_+],$$

[0090] 这里 m 是一个阈值,以超参数的形式呈现,当数值超过这个超参数时,才会作为负

例出现,以避免将潜在的正例错当成了负例的情况出现。最终将上述的几个损失函数加权使得数量级大致一致后求和获得最终的损失函数。

[0091] (2) 性能评估:推荐方法的评价指标一般为评分前20的召回率 (Recall@20) 与归一化折损累计增益 (NDCG@20)。

[0092] Recall计算公式如下所示:

[0093]
$$Recall = \frac{TP}{TP + FN}$$

[0094] 其中,真正例 (True Positive, TP):真实类别为正例,预测类别为正例;假负例 (False Negative, FN):真实类别为正例,预测类别为负例。Recall@20主要检测评分前20的商品的查全率。

[0095] NDCG计算公式如下所示:

[0096]
$$NDCG_i = \frac{DCG_i}{1 + \sum_{k=2}^{n_i^+} \frac{1}{\log_2 k}}$$

[0097] NDCG@20主要是查看前20个排序中的顺序是否合理。

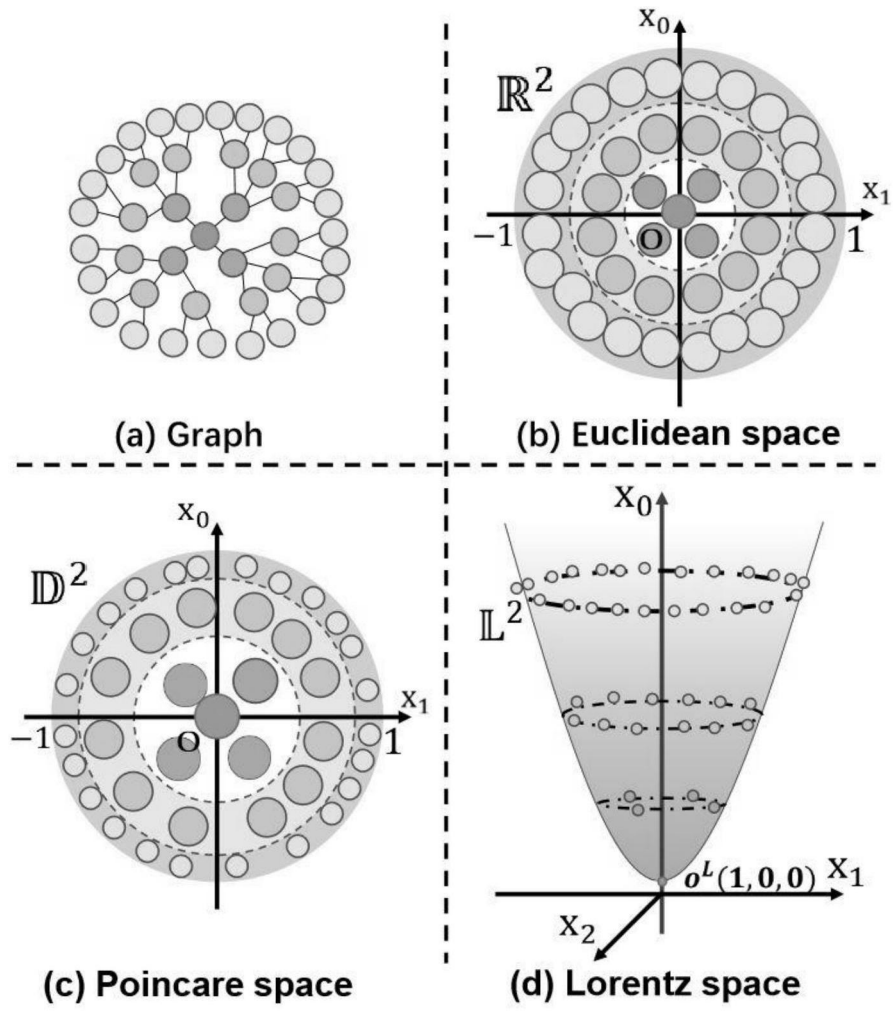


图1

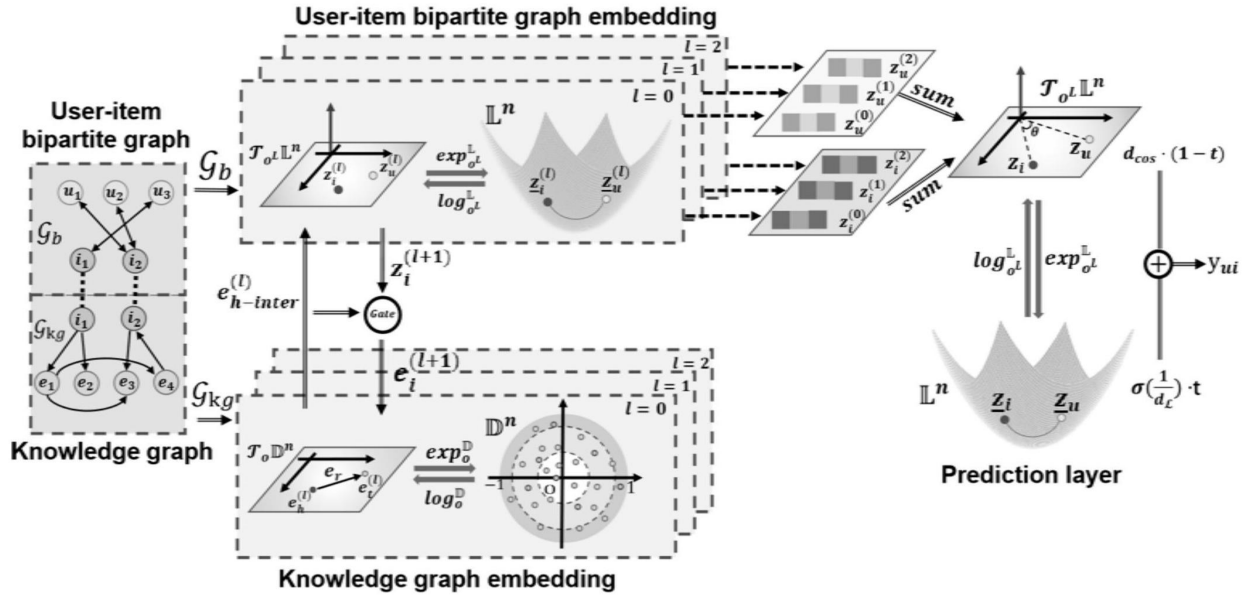


图2

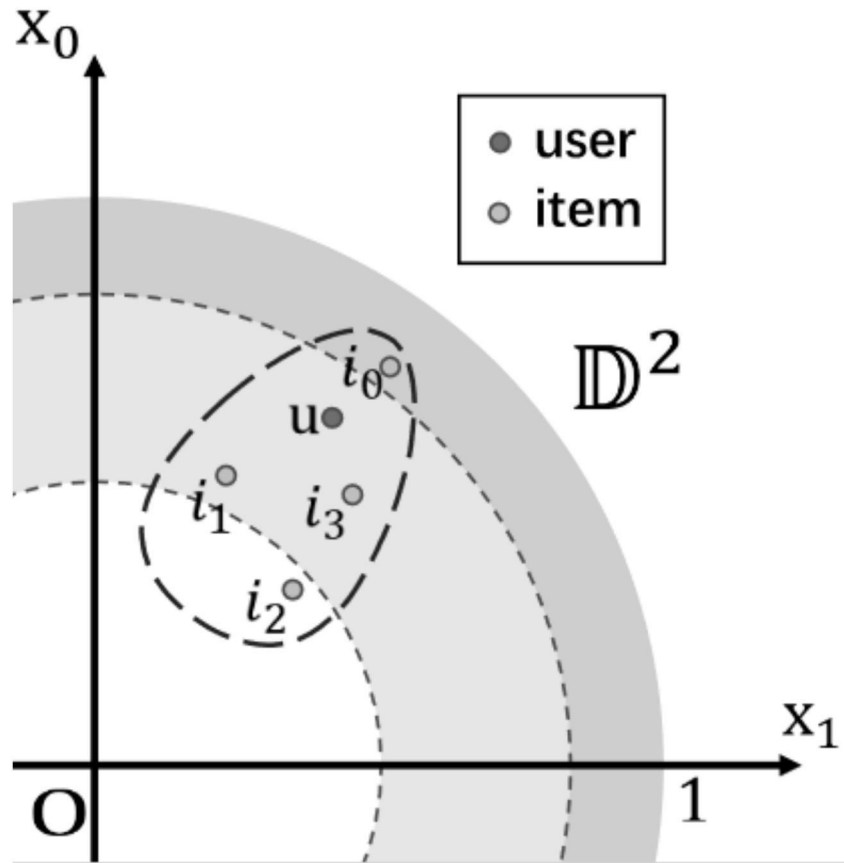


图3

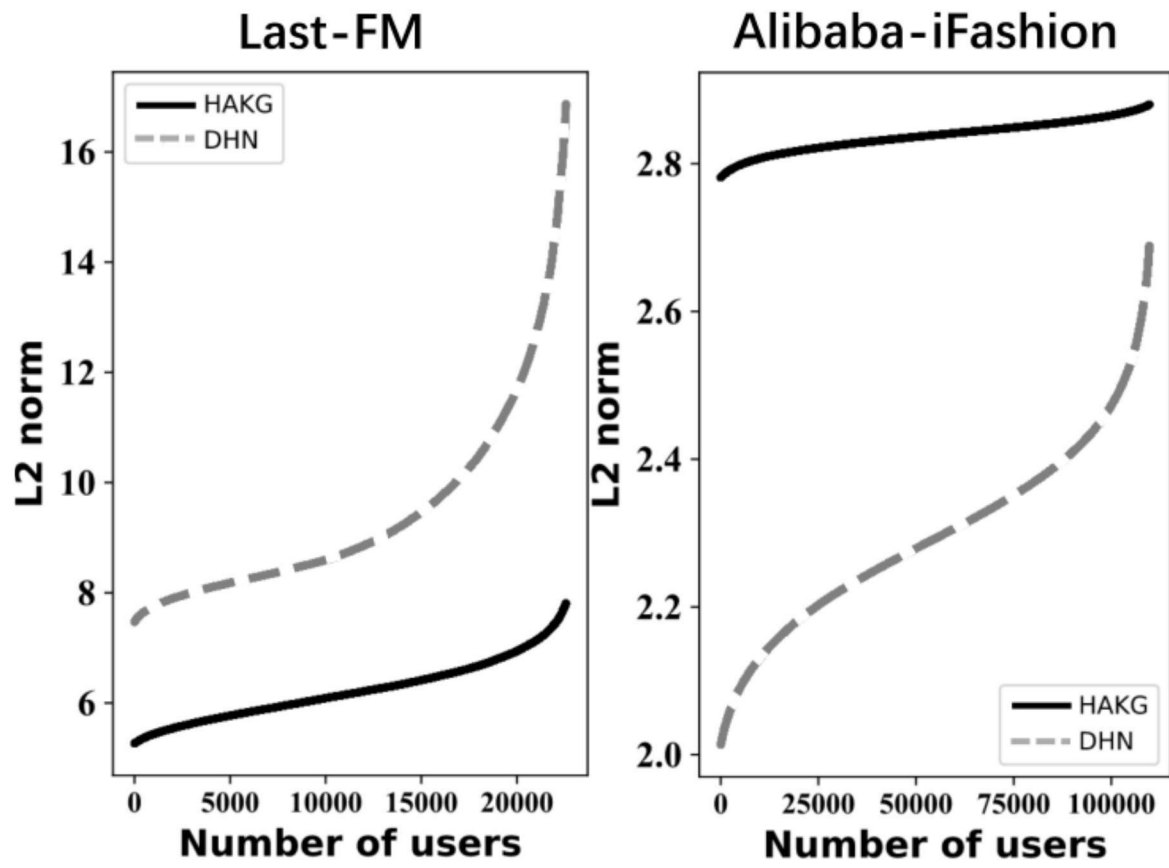


图4