

基于知识图谱和图嵌入的个性化学习资源推荐^①



张栩翔, 汤玉祺, 赵文, 马华, 唐文胜

(湖南师范大学 信息科学与工程学院, 长沙 410081)

通信作者: 马华, E-mail: huama@hunnu.edu.cn

摘要: 面对海量的在线学习资源, 学习者往往面临“信息过载”和“信息迷航”等问题, 帮助学习者高效准确地获取适合自己的学习资源来提升学习效果, 已成为研究热点. 针对现有方法存在的可解释性差、推荐效率和准确度不足等问题, 提出了一种基于知识图谱和图嵌入的个性化学习资源推荐方法, 它基于在线学习通用本体模型构建在线学习环境知识图谱, 利用图嵌入算法对知识图谱进行训练, 以优化学习资源推荐中的图计算效率. 基于学习者的学习风格特征进行聚类来优化学习者的资源兴趣度, 以获得排序后的学习资源推荐结果. 实验结果表明, 相对于现有方法, 所提方法能在大规模图数据场景下显著提升计算效率和个性化学习资源推荐的准确度.

关键词: 知识图谱; 图嵌入; 个性化推荐; 学习资源; 推荐系统

引用格式: 张栩翔, 汤玉祺, 赵文, 马华, 唐文胜. 基于知识图谱和图嵌入的个性化学习资源推荐. 计算机系统应用, 2023, 32(5): 180-187. <http://www.c-s-a.org.cn/1003-3254/9068.html>

Personalized Learning Resource Recommendation Based on Knowledge Graph and Graph Embedding

ZHANG Xu-Xiang, TANG Yu-Qi, ZHAO Wen, MA Hua, TANG Wen-Sheng

(College of Information Science and Engineering, Hunan Normal University, Changsha 410081, China)

Abstract: Faced with numerous online learning resources, learners often suffer from information overload and information disorientation problems. It has become a hotspot to help learners efficiently and accurately obtain suitable learning resources to improve their learning effects. Considering the deficiencies of existing approaches, such as the poor interpretability as well as the limited efficiency and accuracy of recommendation, a new recommendation approach of personalized learning resources is proposed on the basis of knowledge graphs and graph embeddings. In this approach, a knowledge graph of the online learning environment is established through a generic ontology model, and the graph embedding algorithm is applied to train the knowledge graph for optimized efficiency of graph computation in learning resource recommendation. Then, the learners' interest in learning resources is optimized via clustering based on the learning style features of learners. Finally, the ranked recommendation results of learning resources are obtained. The experiments demonstrate that the proposed approach significantly improves the computational efficiency and the accuracy of personalized learning resource recommendations compared with existing methods in large-scale graph data scenarios.

Key words: knowledge graph; graph embedding; personalized recommendation; learning resource; recommendation system

近年来, 全球范围内在线学习发展迅速, 为学习者的自主学习带来了极大的帮助, 但新的问题也随之出

现. 面对海量的学习资源, 学习者需要不断地根据自身学习需求来筛选信息, 在这个过程中, 学习者将耗费大

^① 基金项目: 国家自然科学基金 (62077014, 71971221); 湖南省自然科学基金 (2021JJ30886)

收稿时间: 2022-10-11; 修改时间: 2022-11-14; 采用时间: 2022-11-18; csa 在线出版时间: 2023-02-17

CNKI 网络首发时间: 2023-02-17

量精力,由于无用信息过载,学习者容易迷失学习目标和方向,深陷于海量的网络资源之中^[1],导致学习效率低下,进而影响学习者的学习积极性。因此,针对学习者在线学习常面临的“信息超载”和“信息迷航”等问题,为学习者推荐个性化的学习资源来提升学习效果,已成为当前的研究热点^[2]。

现有研究通常从挖掘学习者的学习兴趣入手,研究个性化学习资源推荐^[2]。但是,学习者的学习过程是循序渐进的,学习的知识点是由浅入深。在评估学习者对学习资源的需求时,不仅要考虑其学习兴趣,也要分析知识点的内在语义联系。知识图谱(knowledge graph)直接表示现实世界中各类实体间的关系,可建模概念与知识的语义关系并支持可视化的交互方式,近年来已在问答与推荐系统等领域得到广泛的应用,然而,至今在学习资源推荐方面尚处于探索阶段,尤其是,大规模应用场景中知识图谱在计算上存在明显的效率问题。而图嵌入(graph embedding)算法通过把知识图谱的图网络映射到低维空间,进而将图运算转化为数值计算,使得计算效率明显提升。目前,尚未见到公开发表将图嵌入算法与知识图谱相结合进行学习资源推荐的成果^[3]。

因此,为缓解学习者在学习过程中面临的“信息过载”和“信息迷航”等问题,本文提出了一种基于知识图谱和图嵌入的个性化学习资源推荐方法 PLRec-KGGE (personalized learning resource recommendation based on knowledge graph and graph embedding)。该方法使用在线学习通用本体模型构建在线学生知识图谱,对知识图谱进行图嵌入训练来提高图数据库的计算效率,以图匹配方式来快速获取相似子图,根据约束条件和实体类别过滤子图中的不相关实体与关系以获得相似学习资源集合。同时,基于学习风格特征对学习者的聚类,由此优化计算学习者对学习资源的兴趣度,计算学习者对图匹配后的各个资源的兴趣度,通过排序获得学习资源的推荐结果。实验表明本文方法在大规模图数据场景下能显著提升计算效率和学习资源推荐的准确率、召回率和 $F1$ 值。

1 相关工作

1.1 基于传统推荐算法的个性化学习资源推荐

传统推荐算法可分为3类:(1)基于内容的推荐方法,它通过提取项目属性特征计算项目间的相似性,再根据用户偏好进行项目推荐。(2)基于协同过滤的推荐

方法,它利用与使用者相似的用户找出其可能感兴趣的项目,传统的基于协同过滤的推荐算法可分为基于用户的协同过滤和基于项目的协同过滤,前者根据用户历史记录计算用户间的相似度,为当前用户推荐项目,后者根据用户历史记录计算项目间的相似度,给用户推荐与其历史偏好类似的新项目。(3)混合推荐方法,它组合和发挥多种推荐方法的优势进行推荐。

学习资源的内容通常包括教学视频和课件、试题和习题、实训和实验项目、辅导资料等。Kong 等人^[2]提出一种用户上下文感知的 MOOC 资源推荐方法,它优化了用户相似度的计算方法,从兴趣信息、搜索痕迹、停留痕迹、学习层次等多角度挖掘用户的个性化特征。针对现有试题推荐研究难以处理知识点本身可能存在的复杂性和学习者认知状态的不确定性, Ma 等人^[4]提出一种新的个性化试题推荐方法,该方法引入中智集理论,从知识点的掌握程度、误解程度及不确定程度等3个侧角度对学习者的认知状态进行综合诊断。针对评分矩阵稀疏和冷启动的问题,王根生等人^[5]提出一种基于改进型协同过滤的学习资源推荐算法,它将用户对资源的学习行为转化为用户对资源的评分,并引入用户初始化标签改进用户的相似度计算。通过识别学习者的学习风格并根据学习风格对学习者的分类, Chen 等人^[6]基于传统的协同过滤方法提出了 AROLS 模型,以用于学习资源推荐。

传统的推荐方法通过融合用户行为信息、个性化需求信息、文本、图像等多源异构信息,在一定程度上可缓解冷启动与数据稀疏等问题。但在数据量庞大、数据类型多样、数据结构复杂的场景中,仍呈现明显的局限性。近年来深度学习应用广泛,基于深度学习的个性化学习资源推荐^[7]可一定程度上突破传统模型的不足,获得更准确和高效的推荐结果,但也存在训练成本较高、推荐结果缺乏可解释性等不足。

1.2 基于知识图谱的个性化学习资源

传统推荐算法未考虑学习资源间的语义关系,难以保证获得令人满意的推荐结果,因此, Xu 等人^[8]提出一个新的个性化课程推荐系统,它使用知识图谱将学习资源的语义信息嵌入到低维空间,结合资源间的语义相似度,使用协同过滤算法实现更准确的推荐效果。针对现有研究未考虑学习过程中知识点的顺序关系, Lv 等人^[9]提出一个基于加权知识图谱的试题推荐方法,该方法将知识点建模为一个节点,将学习者对于

知识点的掌握程度作为权重分配到相应节点. 以上研究构建知识图谱的数据量较小, 均未在真实数据集中讨论图数据的计算效率问题^[3]. 在融合知识图谱信息与深度协同过滤算法方面, 邱玥^[10]提出基于知识图谱增强的在线课程推荐方法, 它引入共现实体网络捕获用户的显性兴趣特征. 但是, 该方法将共现网络中的课程建模为知识图谱中的实体, 获得的关系权重可能存在局限性.

学习路径是一类特殊的学习资源, 近年来受到研究者的关注^[3]. 假设同一学习场景下学习路径是相似的, Zhu 等人^[11]提出一种基于知识图谱的多约束学习路径推荐算法, 但它仅考虑了4种基本的学习情境. 针对现有方法主要从单一维度构建知识图谱, Shi 等人^[12]提出一种基于多维度知识图谱框架的学习路径推荐模型, 它将学习对象间的关系划分为6种类型. 但该模型对知识图谱的数据规模与结构有较强依赖, 难以保障小数据规模或复杂语义关系时的效果.

实际应用中, 知识图谱的数据规模通常较大, 在图网络上直接进行图运算, 会存在明显的效率问题. 尤其是在大规模数据的应用场景下, 上述方法的执行效率会明显降低.

2 知识图谱建模与图嵌入训练

2.1 在线学习环境的知识图谱建模

知识图谱由实体-关系-实体的三元组构成, 为构建知识图谱, 需要对在线学习系统中的非结构化信息进行结构化操作, 并采取人工方式对结构化数据进行实体识别和关系类型梳理. 在对国内主流的在线学习平台, 如超星学习通、EduCoder 和中国大学 MOOC 等, 进行调研分析的基础上, 本文使用课程、成绩单、考试、资源站点等4种实体类型来定义交互资源的核心概念, 并对学习平台的注册类信息与学生的个人背景信息等辅助概念进行实体类型梳理, 再根据实体之间的关联关系与现实语义, 抽象出知识图谱中数据间的各类关系, 形成本体关系模型. 本文设计的在线学习本体模型包含17种概念与18种关系, 它涵盖了学生的个人背景信息、系统注册信息、课程学习信息、在线测试信息、资源交互信息等5类在线学习信息, 这些信息体现了当前主流在线学习平台中的核心要素. 本体结构如图1所示.

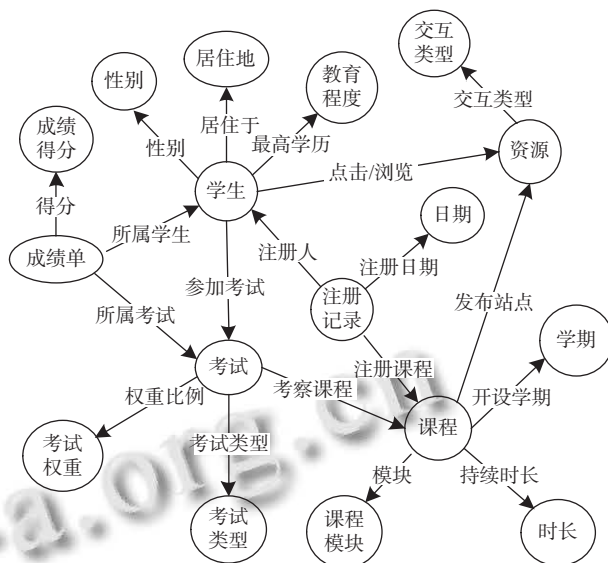


图1 在线学习环境本体模型

本文对 OULAD 数据集^[13]进行了数据抽取与知识存储. OULAD 是包含 32 593 名学习者、7 门课程以及相关考评信息的真实数据集, 共有 10 655 280 条学习者与在线学习环境交互的日志信息. 本文依据本体概念抽取所有实体数据, 再以这些实体抽取关系数据, 将知识图谱中的各个实体关联起来, 得到了包含 242 402 个实体与 2 865 114 条关系的在线学习环境的知识图谱. 图谱的数据存储在 Neo4j 图数据库.

2.2 图嵌入模型训练

基于在线学习环境本体生成的知识图谱能体现学习者之间的社会关联性和学习资源之间的关联度^[3]. 因此, 本文进行图嵌入时重点关注知识图谱的整体网络拓扑结构. 为完整表现原图的节点类型、网络结构与其丰富的语义关系, 且在一定程度上保留网络的隐性特征, 采用 Node2vec 算法^[14]进行图嵌入训练. 该算法生成随机游走, 其领域采样策略分为广度优先采样 (breadth-first sampling, BFS) 和深度优先采样 (depth-first sampling, DFS), 如图2所示.

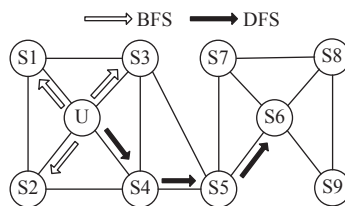


图2 随机游走示例

BFS的邻域仅限于与源节点直接相邻的节点,而DFS的邻域覆盖与源节点的距离不断增加的节点. 随机游走采样获得节点路径组合后,使用词向量方法进行建模,以获得网络节点的代表形式. Node2vec在随机游走过程中加入规则约束,采用支持BFS和DFS动态调整的邻域采样策略,使得生成的随机游走可同时反映BFS和DFS的特性,从而提高网络嵌入的效果.

Node2vec算法中, G 表示图、 V 为图的顶点集合、 E 为图的边集、 d 为向量维度、每个节点生成的游走个数为 r 、游走长度默认为1、上下文的窗口尺寸为 k . 首先计算一个节点到它的邻居的转移概率,再将这个转移概率加到 G 中形成 G' ,使用walks存储随机游走,先初始化为空,循环 r 次,即每个节点作为初始节点要生成 r 个随机游走,然后对图中每个节点,生成一条随机游走walk. 最后,使用随机梯度递减方法来训练walks.

3 一种新的个性化学习资源推荐流程

3.1 个性化学习资源推荐流程

结合知识图谱和图嵌入算法,本文提出了一种新的个性化学习资源推荐方法PLRec-KGGE,其推荐流程如下.

(1) 预处理阶段:先抽取系统中的交互信息,以构建在线学生知识图谱.再引入图嵌入算法训练构建好的知识图谱得到嵌入模型,以嵌入模型上的数值计算代替原有的图数据库计算,以提高计算效率.

(2) 基于知识图谱的相似资源匹配阶段:为获得个性化的学习资源集合,采用图匹配快速获取相似子图,过滤子图中的元素,得到相似学习资源集合.

(3) 基于学习风格特征的学生聚类阶段:考虑到学习者在学习过程中所带有的群体倾向性,从学习风格特征角度对学习进行聚类分组,并以此为基础优化用户对资源的兴趣度计算.

(4) 基于兴趣度的资源推荐阶段:计算学习者对相似资源集合中各资源的兴趣度并排序,最终获得学习资源的推荐结果集合.

接下来,将介绍以上的3个关键步骤.

3.2 基于知识图谱的相似资源匹配

将学生的注册信息、学习记录等使用数据,按通用本体模型转换成信息图谱,作为输入,如图3所示.在完整的在线学习环境图谱中进行图匹配,快速获取与输入图最相似的匹配子图.

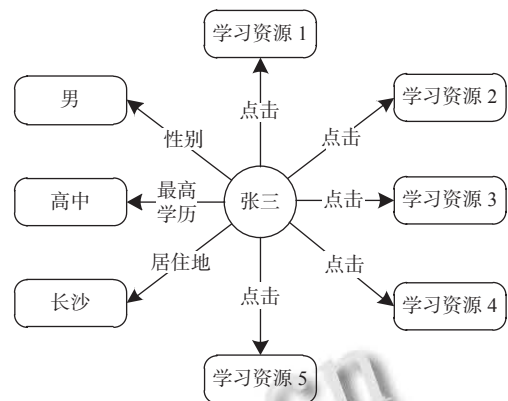


图3 输入的学生信息图谱

由于随机游走只能在单个图上进行,需将两张待匹配的图转化为单个图,即伴随图(association graph).如图4所示,考虑两个节点(A, B)匹配3个节点(C, D, E)的情况.图4(a)中两个图结构代表原始的图匹配问题,图4(b)中的图为伴随图.伴随图中的节点代表原图匹配问题中节点间的对应关系(即点虚线双箭头),例如,伴随图中节点AC代表原图中节点A与节点C的匹配关系.边代表原图匹配问题中边与边的相似度信息,例如,伴随图中的有权边AC-BD代表原图边a与b的相似度(即 $K_{AC,BD}$).因此图匹配问题的伴随图是一个无向权值图.通过随机游走算法可求得伴随图中节点的权重,这样图匹配问题转化为寻找伴随图中具有最大权重的若干节点的问题.

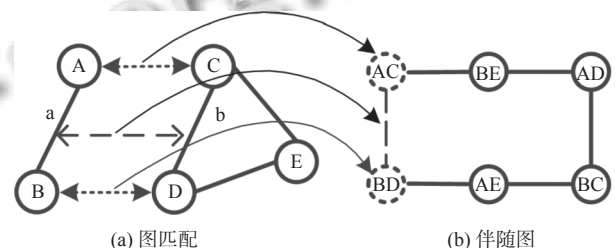


图4 图匹配过程

获得相似用户子图后,汇总子图中的元素类型,滤除不属于学习资源类型以及不满足限制约束条件的元素,从而获得最终的相似学习资源的集合.

3.3 基于学习风格特征的学生聚类

学习者在学习过程中各有不同的学习习惯、偏好和学习方法,在心理学领域,学习者间的这些个体差异被称为学习风格^[15].引入学习风格理论,识别学习者的学习风格,并根据学习风格将学习者分类,有利于提升

学习资源推荐系统的推荐效率^[6]. 本文在个性化学习资源推荐的过程中, 依据目前应用最广泛的学习风格理论——费尔德-西尔弗曼模型^[15]对学习者的聚类. 该模型从情感、社会学、生理学和心理学4个方面、8个特征来描述在线学习者. 如表1所示.

表1 在线学习风格模型

类别	特征	学习行为
情感	动力型	学习时间长, 做大量习题
社会学	活跃型	在论坛中表现活跃等
生理学	视觉型	容易记住看到的信息, 如视频、图片
	语言型	容易记住听到、读到和说到的信息
心理学	感觉型	喜欢具体事实、数据及按既定程序解决问题
	直觉型	喜欢学习原理和理论
	序列型	解决问题时会伴随一大串逻辑推理过程
	综合型	针对提供的信息, 由整体构想获得正确结论

通过分析学生的在线学习数据, 并计算不同活动类型的点击次数, 可以表格的形式呈现学生学习数据和交互数据的分析结果. 表格中的每一行代表一个学习者, 每一列代表他对该活动类型的偏好水平, 并将交互类型的组合作为聚集数据的特征. 例如, 论坛上的点击次数能反映学生在学习过程中的交流活跃程度. 因OULAD数据集缺少每个页面中的详细交互细节数据, 故无法建模与直觉型和感觉型特征对应的交互行为类型, 只能对动力型、活跃型、视觉型、语言型、序列型、综合型等6个特征建模. 现有研究^[6]表明, 使用这6个特征也可获得令人满意的结果.

本文采用K-means聚类算法对学习者的分组. 给定N个学习者, 每个学习者 x_n 可表示为一个向量 $f(d_{n1})$, d_{nj} 代表学习者 x_n 的第j个特征值. 使用欧几里得距离度量计算两个学习者之间的相似度, 并根据相似度值, 将N个学习者分成K个不相交的聚类, C_k 代表第k个聚类, 可以用均值(或质心)来描述第k个聚类中的学习者.

本文根据预处理结果对学习者的聚类, 每一个类代表学习者的一种在线学习风格特征. 为评估K的最优取值, 本文选用轮廓系数(SC)和Calinski-Harabasz(CH)指数评估聚类模型的效果. 其中, 轮廓系数SC的值是介于[-1, 1], 高SC分数代表聚类结果的内聚度和分离度都相对较优, 高CH分数说明模型的聚类效果良好. 实验结果证明, 当K=5时, 模型聚类效果最佳.

3.4 基于兴趣度的资源推荐

基于学习风格分组情况, 本文优化了学生对学习

资源兴趣度的计算. 具体计算步骤如下.

本文用余弦相似度计算 C_k 组中学习资源m与n的相似度 $sim(i_m, i_n | C_k)$, 即 C_k 中既点击了m又点击了n的学生数与点击资源m的总人数 $S(i_m | C_k)$ 和点击资源n的总人数 $S(i_n | C_k)$ 的几何平均的比值.

C_k 组中学习资源I的子集表示为 I_j , C_k 对于 I_j 的支撑度为 C_k 中点击 I_j 中资源的学生比, 即:

$$sup(I_j | C_k) = \frac{|S(I_j | C_k)|}{N_{C_k}} \quad (1)$$

将关联规则“ $I_i \rightarrow I_j$ ”的支撑度定义为 C_k 中同时点击了资源i与j的学生占比, 即:

$$sup(I_i \rightarrow I_j | C_k) = \frac{|N(I_i | C_k) \cap N(I_j | C_k)|}{N_{C_k}} \quad (2)$$

将关联规则“ $I_i \rightarrow I_j$ ”的可信度定义为:

$$conf(I_i \rightarrow I_j | C_k) = \frac{sup(I_i \rightarrow I_j | C_k)}{sup(I_i | C_k)} \quad (3)$$

将学习者x对学习资源 i_m 的兴趣度定义为:

$$P(x, i_m | C_k) = \sum_{i_n \in S(x) \cap S(i_m, L)} sim(i_m, i_n | C_k) \quad (4)$$

其中, $S(i_m, L)$ 表示与学习资源 i_m 最相似的L个推荐结果. 即, 兴趣度表示为特定的学习资源(即被x浏览过并且属于与 i_m 资源最相似的L个学习资源), 与 i_m 的基于 C_k 的相似度之和.

基于学生对关联资源的兴趣度, 对它们进行排序过滤, 最终得到“学习资源推荐结果集合”.

4 实验分析

4.1 图嵌入效率对比

本文对比图数据库直接运算(简记为graph_D)与图嵌入训练后模型计算(简记为embed_M)的响应时间与资源占用情况. 对抽取的200万条边规模的图数据进行随机数扩充, 测试2000万条边规模的图数据.

时间效率对比结果由图5中的圆点折线和星形折线所示, 在图数据规模较小时, 二者的平均响应时间差距较小. 但随着图规模的增长, embed_M与graph_D的平均响应时间差距越来越大. 在 10^7 规模时, embed_M的响应时间为701 ms, 而graph_D的响应时间为14785 ms. 可见, 图嵌入模型在大规模图运算时能有效提高计算效率.

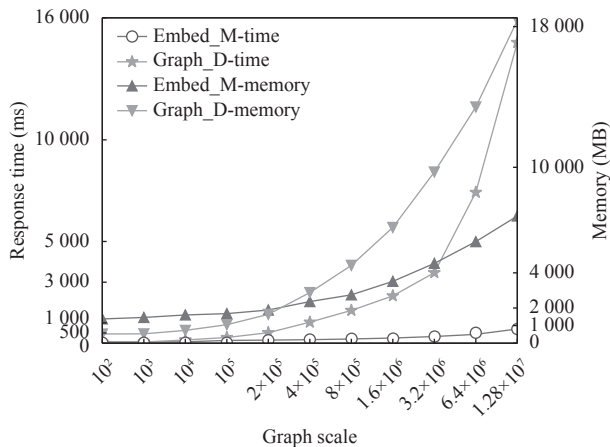


图5 图运算时间空间效率对比

空间效率对比结果如图5中正三角折线和倒三角折线所示,图数据规模较小时,两者差异不明显.当图数据规模增大时,embed_M体现明显优势.例如,当图数据规模为 1.28×10^7 时,graph_D占用系统资源18357 MB,而embed_M占用资源仅为7231 MB.整体而言,图嵌入模型占用的资源没有随着图数据规模的增大而产生明显增长,在大规模图数据场景下可显著节约系统资源.

4.2 学习风格聚类数K值分析

在OULAD数据集上分析K-means算法中学习风格聚类数K值的变化,结果如图6所示.图6给出了SC和CH分数随聚类数K的变化.其中,圆点折线表示SC分数,三角形折线表示CH分数.可见, $K=6$ 时,SC分数的峰值为0.285, $K=5$ 时,CH分数的峰值为5864.但是,CH分数在 $K=6$ 时下降为5431,SC在 $K=5$ 时为0.282时,综合对比,取 $K=5$ 作为最佳聚类效果.

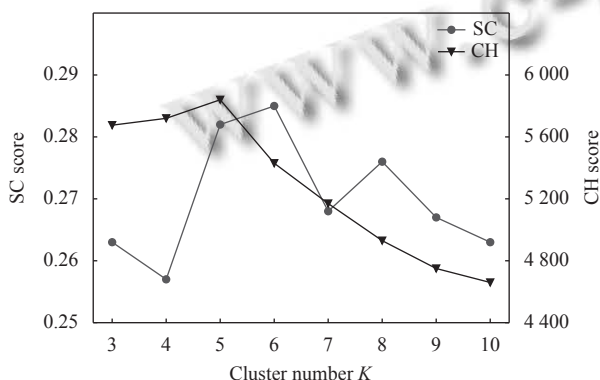


图6 学习风格聚类分组数评估

4.3 对比试验

本实验对比方法包括:结合学习风格模型和协同

过滤的AROLS方法^[6]、基于深度神经网络的DNN方法^[7]、基于用户的协同过滤(简记为UserCF)方法、基于项目的协同过滤(简记为ItemCF)方法.为了确保数据的可靠性和准确性,本文抽取OULAD中平均评估结果等于或大于60的19263名学习者.同时,K-means算法计算欧式距离时可能出现较大的异常值,为避免其影响到结果的准确性,本文删除至少一个特征的标准偏差超过10的记录.

本文采用准确率(*precision*)、召回率(*recall*)以及F1值(*F1*)作为评估指标,计算方法如下所示.

$$precision = \frac{TP}{TP + FP} \quad (5)$$

$$recall = \frac{TP}{TP + FN} \quad (6)$$

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \quad (7)$$

准确率对比的实验结果如图7所示,图中横坐标的N表示协同过滤算法中top-N推荐的N取值.PLRec-KGGE方法相对于传统的ItemCF方法、UserCF方法、基于深度神经网络的DNN方法,在准确率上有大幅度提升.如在N为20时,PLRec-KGGE平均准确率为0.306,而准确率最低的UserCF方法0.164,但PLRec-KGGE与利用学习风格优化的协同过滤推荐方法AROLS相比,准确率平均提升8.6%.其原因在于,PLRec-KGGE方法使用了知识图谱,它包含了各物品的背景信息以及物品间的关系,同时能集成由用户行为数据组成的用户物品网络,扩展用户与物品之间的隐藏关联关系,补充用户与物品之间的交互数据,从而进一步提升推荐效果.

召回率对比的实验结果如图8所示.随着N值的增加,各对比方法的召回率均逐渐升高.结果显示,ItemCF与UserCF在召回率上不及DNN方法.由于知识图谱特殊的图网络结构,能发现图中的弱关联关系,挖掘可能的关联路径,从而使基于知识图谱的推荐结果相对于现有方法在召回率上有比较大的提升.所以,PLRec-KGGE方法在召回率上较DNN方法平均有68.9%的提升.

F1值对比的实验结果如图9所示,其中,ItemCF与UserCF在综合性能上无明显区别.而DNN方法在F1值上优于现有的其他3个方法,这可能是由于DNN能更好挖掘多特征非线性数据中各元素间的隐藏特征,

并利用多层网络来拟合这些隐藏映射,表现出较强的综合性能.知识图谱是通过整合各类异构辅助信息形成的一个大图,同时利用图网络结构在关联匹配与路径分析领域的优势,因此,PLRec-KGGE方法相对于其他方法,在F1值上具有较为明显的提升,相对于DNN方法有35.7%的提升.

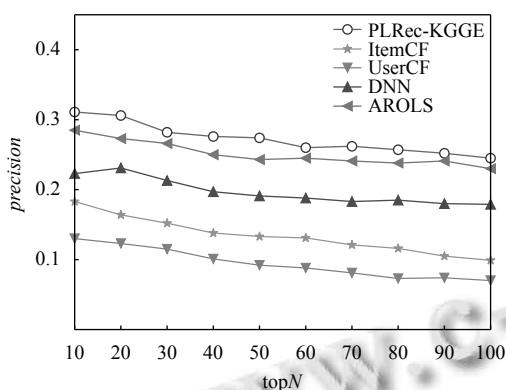


图7 准确率对比

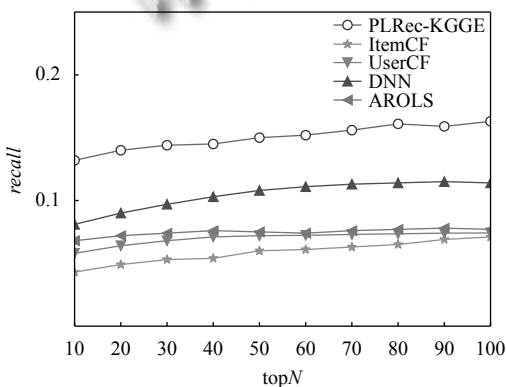


图8 召回率对比

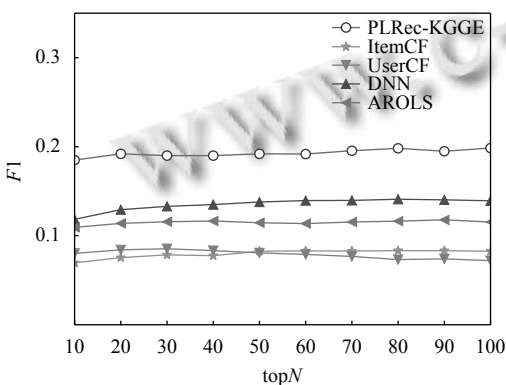


图9 F1值对比

由于在线学习环境复杂,图数据规模通常较大,在实际应用场景中,需考虑图数据的存储和部署设备的配置需求.本文使用Neo4j图数据库存储数据,同时,

知识图谱嵌入模型的训练对服务器的内存配置等要求较高,如本文使用的服务器内存为256GB.在效率对比实验中,当数据规模为 1.28×10^7 时,已用内存为7231MB,而实际应用中的数据规模将更庞大.

5 结语

在线学习环境下,如何高效准确地挖掘学习者与资源之间潜在的关联关系并为学习者推荐个性化的学习资源,对实现个性化教育具有重要意义.针对现有个性化学习资源推荐方法的局限性,本文设计了通用本体模型,构建了在线学习环境的知识图谱,考虑到传统知识图谱推荐系统在大规模图计算场景中存在的效率低下问题,采用图嵌入算法提高在学生信息知识图谱上的图运算效率,结合图匹配原理能快速获取相似资源集合,并通过引入学习风格理论对学生进行聚类,进一步优化了学生对资源的兴趣度计算,最终保障了本文所提方法在大规模图数据场景下显著提升的计算效率和个性化学习资源推荐的准确度.在后续研究中,我们将进一步探讨基于图嵌入与图神经网络算法的个性化学习资源推荐,并利用知识图谱在呈现教师、学习者、学习资源和知识点之间的关联关系上的优势,研究其在知识检索、自动问答等领域中的应用.

参考文献

- 郭磊,郭华.网络环境下学习者信息超载和迷航的成因及化解.远程教育杂志,2004,(4):31-32,27.[doi:10.3969/j.issn.1672-0008.2004.04.007]
- Kong LY, Ma H, Peng Y, et al. Hybrid recommendation of personalized MOOC resources: A user context-aware approach. Proceedings of the 2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design. Hangzhou: IEEE, 2022. 537-542.
- 张翔翔,马华.知识图谱与图嵌入在个性化教育中的应用综述.计算机系统应用,2022,31(3):48-55.[doi:10.15888/j.cnki.csa.008377]
- Ma H, Huang ZX, Tang WS, et al. Exercise recommendation based on cognitive diagnosis and neutrosophic set. Proceedings of the 2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design. Hangzhou: IEEE, 2022. 1467-1472.
- 王根生,袁红林,黄学坚,等.基于改进型协同过滤的网络学习资源推荐算法.小型微型计算机系统,2021,42(5):940-945.[doi:10.3969/j.issn.1000-1220.2021.05.007]

- 6 Chen H, Yin CT, Li RM, *et al.* Enhanced learning resource recommendation based on online learning style model. *Tsinghua Science and Technology*, 2020, 25(3): 348–356. [doi: [10.14177/j.cnki.32-1397n.2022.46.02.009](https://doi.org/10.14177/j.cnki.32-1397n.2022.46.02.009)]
- 7 宋菲菲, 隋栋, 周湘贞. 基于深度学习的智能学习资源推荐算法. *南京理工大学学报*, 2022, 46(2): 185–191.
- 8 Xu GW, Jia GY, Shi L, *et al.* Personalized course recommendation system fusing with knowledge graph and collaborative filtering. *Computational Intelligence and Neuroscience*, 2021, 2021: 9590502.
- 9 Lv P, Wang XX, Xu J, *et al.* Intelligent personalised exercise recommendation: A weighted knowledge graph-based approach. *Computer Applications in Engineering Education*, 2021, 29(5): 1403–1419. [doi: [10.1002/cae.22395](https://doi.org/10.1002/cae.22395)]
- 10 邱玥. 知识图谱增强的在线课程推荐方法研究 [硕士学位论文]. 武汉: 华中师范大学, 2020.
- 11 Zhu HP, Tian F, Wu K, *et al.* A multi-constraint learning path recommendation algorithm based on knowledge map. *Knowledge-based Systems*, 2018, 143: 102–114. [doi: [10.1016/j.knosys.2017.12.011](https://doi.org/10.1016/j.knosys.2017.12.011)]
- 12 Shi DQ, Wang T, Xing H, *et al.* A learning path recommendation model based on a multidimensional knowledge graph framework for e-learning. *Knowledge-based Systems*, 2020, 195: 105618. [doi: [10.1016/j.knosys.2020.105618](https://doi.org/10.1016/j.knosys.2020.105618)]
- 13 Kuzilek J, Hlosta M, Zdrahal Z. Open university learning analytics dataset. *Scientific Data*, 2017, 4: 170171. [doi: [10.1038/sdata.2017.171](https://doi.org/10.1038/sdata.2017.171)]
- 14 Grover A, Leskovec J. Node2vec: Scalable feature learning for networks. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco: ACM, 2016. 855–864.
- 15 Truong HM. Integrating learning styles and adaptive e-learning system: Current developments, problems and opportunities. *Computers in Human Behavior*, 2016, 55: 1185–1193. [doi: [10.1016/j.chb.2015.02.014](https://doi.org/10.1016/j.chb.2015.02.014)]

(校对责编: 牛欣悦)