

基于人才知识图谱推理的强化学习可解释推荐研究^{*}

阮小芸 廖健斌 李 祥 杨 阳 李岱峰

(中山大学信息管理学院 广州 510006)

摘要:【目的】为解决现有工作推荐存在的难以大规模应用、冷启动、缺乏新颖性和解释性等问题,提出基于人才知识图谱推理的强化学习可解释推荐方法。【方法】基于真实的简历数据集构建人才社会经历知识图谱,依据强化学习的理论在知识图谱上训练一个策略智能体,将一次推理过程分解为选择方向、选择节点两个子过程,使其能够在知识图谱上寻找潜在的优质推荐目标。【结果】相比于LR、BPR、JRL-int、JRL-rep及PGPR模型,基于人才知识图谱推理的强化学习可解释推荐模型在MRR@20(81.7%)、Hit@1(74.8%)、Hit@5(92.2%)以及Hit@10(97.0%)均表现最优。【局限】实验数据集规模和任务类型相对有限。【结论】该模型有效结合人才历史工作经历、相似人才工作经历进行推荐,结合知识图谱工作岗位的属性关联,在给出推荐结果的同时,提供推理路径,能够有效应对冷启动和缺乏新颖性、可解释性问题。

关键词: 工作推荐 知识图谱推理 强化学习 可解释推荐

分类号: TP393

DOI: 10.11925/infotech.2096-3467.2020.1218

引用本文: 阮小芸, 廖健斌, 李祥等. 基于人才知识图谱推理的强化学习可解释推荐研究[J]. 数据分析与知识发现, 2021, 5(6): 36-50. (Ruan Xiaoyun, Liao Jianbin, Li Xiang, et al. Interpretable Recommendation of Reinforcement Learning Based on Talent Knowledge Graph Reasoning[J]. Data Analysis and Knowledge Discovery, 2021, 5(6): 36-50.)

1 引 言

人才工作岗位匹配一直是人们关注的热点问题,无论是刚毕业的学生还是已经有丰富工作经验的从业者都面临着工作岗位与自身发展的匹配问题。受新冠肺炎疫情等多重因素影响,根据中华人民共和国人力资源和社会保障部最新数据显示,2020年前7个月全国城镇新增就业674万人,同比减少196万人,而城镇失业人员再就业人数为259万人,就业困难人员就业人数84万人,由此可见不论是新增就业还是失业再就业的形势均较为严峻^[1]。同时,网络招聘服务的广泛使用导致了就业市场的信息爆炸,就业信息繁多冗杂,无法简单依靠人工进

行信息筛选,越来越多的招聘网站开始为求职者进行工作推荐,然而现有的推荐效果一般且不具有可解释性,不能够为求职者提供最精准可解释的工作推荐。因此,在竞争激烈的就业市场上如何为求职者选择最合适自己的工作成为各方关注的焦点,如何结合求职者自身的社会经历与目标公司和岗位信息进行精准的匹配已成为亟待解决的问题。可解释推荐旨在向寻求工作的人提供更加精准的工作推荐,以大量前人的社会经历作为参考来帮助他们快速找到合适自己的工作。然而,现有的工作推荐存在难以大规模应用^[2-3]、冷启动^[4-5]、缺乏新颖性^[6]和解释性等问题。虽然已有研究^[7-8]将知识图谱和强化

通讯作者(Corresponding author): 李岱峰(Li Daifeng), ORCID: 0000-0002-5457-9324, E-mail: lidaifeng@mail.sysu.edu.cn。

^{*}本文系国家自然科学基金青年项目(项目编号: 61702564)、广东省软科学面上项目(项目编号: 2019A101002020)和国家自然科学基金面上项目(项目编号: 72074231)的研究成果之一。

The work is supported by the National Natural Science Foundation of China (Grant No. 61702564, 72074231), the Soft Science General Program of Guangdong Province, China (Grant No. 2019A101002020).

学习理论相结合并运用到问答、推荐等任务中,但是仍面临着奖励稀疏、动作空间过大、探索与利用难以平衡等挑战。

为应对上述挑战,本研究基于真实的简历数据集构建人才社会经历知识图谱,依据强化学习的理论在知识图谱上训练一个策略智能体,使其能够在知识图谱上寻找潜在的优质推荐目标。训练好的智能体能够根据用户在知识图谱中所处的环境推理出有效路径并找到合适的工作候选集,这些推理路径将作为推荐的理由为用户提供推荐解释,不仅避免了决策的绝对化,同时增加了推荐候选集的新颖性,而且推理路径的展示增强了推荐的可解释性。推荐示例如图 1 所示。

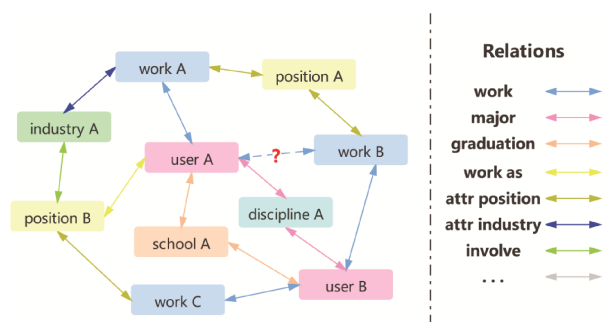


图 1 基于人才社会经历知识图谱的工作推荐示例
Fig.1 Examples of Job Recommendations Based on the Talent Social Experience Knowledge Graph

本文的主要工作如下:

(1) 基于真实的简历数据集构建了人才社会经历知识图谱。

(2) 设计一个分层强化学习智能体,对巨大动作空间进行分类探索,提高知识图谱推理效率,避免大量计算。

(3) 采用一种以起始节点为参照对象的动作剪枝策略,避免了探索-利用的不平衡,并提高推理过程的合理性。

(4) 提出模型能充分利用人才社会经历信息,探索其潜在的需求和兴趣,在工作推荐方面具有新颖性、合理性和可解释性,并有效应对冷启动问题。

2 研究现状

2.1 工作推荐

工作推荐一直是推荐领域研究的热门话题。文

献[9]利用 LinkedIn 中用户的公开资料数据,找出工作和人际交往技能之间的关系。语义关联是通过潜在语义分析(Latent Semantic Analysis, LSA)产生的。利用挖掘的语义得到职位的层次聚类,构建职位推荐系统。端到端的二元感知 APJFNN 模型^[10]利用基于递归的神经网络构建职位要求 and 求职者经验的词级语义表示,设计了 4 种分层注意力机制,以衡量工作要求对语义表征的不同重要性,并对拟合结果提供更好的解释。Meng 等^[11]使用一个内嵌注意力机制的神经网络结构来表征内部和外部的流动工作流动性,同时在训练中考虑个人资料以增强可解释性。Gugnani 等^[6]提出隐藏技能的概念,利用 Doc2Vec 模型得到职位描述中缺失但在相似的职位描述中存在的技能,获得最终的隐式技能集,以匹配从候选人简历中提取的技能与职位描述的显性和隐性技能。

2.2 基于知识图谱的可解释推荐

目前将知识图谱用于推荐的研究逐渐增多,知识图谱的加入主要是用以增强模型的可解释性^[12-13]。当前研究主要以两个方向为主,一是将知识图谱嵌入以丰富现有的信息内容,提升推荐的性能。提取不同类型信息的语义表示,通过异构网络嵌入到知识图谱中,再集成协同过滤,增强了模型的可解释性^[14]。将知识图谱嵌入到记忆网络中,再利用 GRU 和注意力机制得到用户对每个属性的偏好表示,以全面了解用户喜好^[15]。通过自动迭代沿着知识图谱中的链接扩展用户的潜在兴趣,刺激用户对知识实体集偏好的传播。用户历史上点击过的项目所引发的多个“涟漪”因此叠加在一起,形成用户对候选项目的偏好分布,最终推荐物品与用户间的传播路径能够提供可解释性^[16]。另一个方向则是利用知识图谱中的实体和路径信息进行可解释性的推荐。Ai 等^[17]提出一种嵌入异构实体进行推荐的知识库表示学习框架,将用户的各种行为以及与商品属性相关的信息嵌入到知识图谱中,并在此基础上提出一种软匹配算法生成个性化的推荐对象解释。KPRN 利用知识图谱组合实体和关系的语义生成路径表示,借助路径的序列依赖关系进行有效推理,用以推理用户和物品之间的交互关系,同时通过加权池计算连接不同用户和物品之间的路径的优劣,增强模型的可解释性^[18]。

2.3 强化学习

近年来,强化学习引起了研究界的极大兴趣,深度强化学习已经成功应用到很多领域,比如AlphaGo^[19]。已有研究尝试将强化学习运用到推荐系统中,比如广告推荐^[20]、新闻推荐^[21]以及可解释性的因果推荐。Wang等^[22]设计了一个可解释推荐的强化学习框架,用以解释任何推荐模型,并可以根据应用场景灵活地控制解释质量。研究基于自注意力的神经网络实例化了框架中的解释生成器来生成句子级的解释。

也有研究将强化学习与知识图谱结合起来用于知识问答领域,将多跳推理作为一个序列决策问题。Xiong等^[23]提出一种用于学习多跳关系路径的强化学习框架,使用一个基于知识图谱嵌入编码智能体在连续空间中的状态,通过对最有希望的关系进行采样以扩展其路径在知识图谱向量空间中推理。文献[24]使用一个奖励函数训练策略梯度,在奖励梯度中引入不同的标准,有助于更好地控制路径查找。Lin等^[8]采用预先训练的单跳嵌入模型估计未观察到事实的回报,通过强制代理使用随机生成的边缘掩码探索不同的路径集对抗策略上强化学习对假路径的敏感性,利用强化学习训练智能体,在未事先计算好路径的情况下在知识图谱上查找候选答案。

目前的工作推荐模型利用神经网络、注意力机制等^[10-11]提升推荐效果,但是缺乏解释性;而基于知识图谱的推荐可以增强推荐模型的可解释性^[14-18]。将知识图谱和强化学习理论相结合的研究^[8,23-24]可以根据应用场景灵活地控制解释质量,但仍面临着奖励稀疏、动作空间过大、探索与利用难以平衡等挑战。因此,在已有研究^[7]的基础上,本文依据强化学习的理论在知识图谱上训练一个分层强化学习智能体,采用以起始节点为参照对象的动作剪枝策略,为工作推荐的研究提供新的思路。

3 基于人才知识图谱推理的强化学习可解释推荐模型

通过对原始简历数据进行数据清洗,去除异常字符和空值,得到处理后的简历数据集。通过关键字提取出用户职业发展的概念字段,如教育背景、工作经历等,以满足知识图谱构建对元数据的需求。依据概念字段进行实体和属性抽取、实体消歧及实体关系抽取,进而构建人才社会经历知识图谱。将人才社会经历知识图谱中人才的社会经历作为模型的输入,模型输出最后的工作,使用分层推理强化学习模型对数据进行训练,以实现可解释性工作推荐,最后采用MRR、Hit、Precision、Recall等指标评估模型的推荐效果,主要工作流程如图2所示。

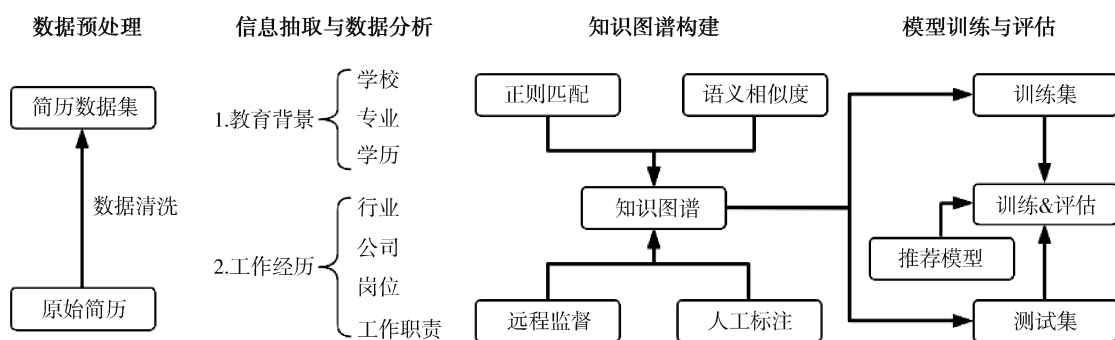


图2 工作流程

Fig.2 Main Workflow

3.1 问题定义

一般地,一个拥有实体集 \mathcal{E} (包含用户实体子集 \mathcal{U})和关系集 \mathcal{R} 的知识图谱 G 定义为 $G = \{(e, r, e') | e, e' \in \mathcal{E}, r \in \mathcal{R}\}$,其中每个三元组 (e, r, e')

表示头实体 e 与尾实体 e' 间存在关系 r 。人才社会经历知识图谱是从海量简历中抽取重要实体信息构成的,它包括用户实体集 \mathcal{U} 和工作实体集 \mathcal{W} ,其中 $\mathcal{U}, \mathcal{W} \subseteq \mathcal{E}, \mathcal{U} \cap \mathcal{W} = \emptyset$,用户实体 $u \in \mathcal{U}$ 与工作实体

$w \in \mathcal{W}$ 之间可以通过关系 $r_{uw} \in \mathcal{R}$ 进行连接,表示用户 u 具有 w 这项工作经历。而基于知识图谱的可解释性推荐就是给定用户实体 u , 找到可能与其存在关系的工作实体 w 作为推荐结果, 并给出实体之间合理的连接路径作为解释。具体地, 将知识图谱上的 k -hop 路径^[7]、1-hop 行走^[7] 以及推荐任务定义如下:

(1) k -hop 路径: 一条从实体 e_0 到实体 e_k 的 k -hop 路径 P_k 定义为由 k 个关系连接的 $k+1$ 个实体, 即 $P_k = \{e_0 \xrightarrow{r_1} e_1 \xrightarrow{r_2} \dots \xrightarrow{r_k} e_k\}$, 其中 $e_{i-1} \xrightarrow{r_i} e_i$ 表示 $(e_{i-1}, r_i, e_i) \in G$ 或 $(e_i, r_i, e_{i-1}) \in G, i \in [1, k]$ 。

(2) 1-hop 行走: 给定 k -hop 路径 $P_k (P_k = \{e_0 \xrightarrow{r_1} e_1 \xrightarrow{r_2} \dots \xrightarrow{r_k} e_k\})$, 则当前节点为 e_k , 那么 1-hop 行走定义为智能体选择了 (r_{k+1}, e_{k+1}) 作为行走路径, 使得当前 k -hop 路径变为 $P_{k+1} (P_{k+1} = \{e_0 \xrightarrow{r_1} e_1 \xrightarrow{r_2} \dots \xrightarrow{r_k} e_k \xrightarrow{r_{k+1}} e_{k+1}\})$, 其中 $(r_{k+1}, e_{k+1}) \in \{(r, e) |$

$(e_k, r, e) \in G, e \notin \{e_0, e_1, \dots, e_k\}\}$ 。

(3) 推荐任务: 给定人才社会经历知识图谱 G , 初始用户节点 $u_0 \in \mathcal{U}$ 和整数 K , 推荐目标是找到与 u_0 存在合理路径 $P_k (2 \leq k \leq K)$ 的工作 w_n 的候选集合 $\{w_n\} \subseteq \mathcal{W}$, 其中 $n \in [1, N]$, N 为候选集容量。

本文提出的完成上述推荐任务的分层强化学习路径推理方法用于在含有丰富异质信息知识图谱上探索 k -hop 路径以寻找可能存在的推荐目标, 如图 3 所示。该方法主要目标是训练一个策略智能体, 使其能够在知识图谱上从起始经历节点开始依照学到的策略进行游走, 寻找潜在的优质推荐目标。当训练完成后, 期望该智能体能够根据用户在知识图谱中所处的环境推理出有效路径并找到合适的工作候选集, 这些推理路径将作为推荐的理由为用户提供推荐解释。

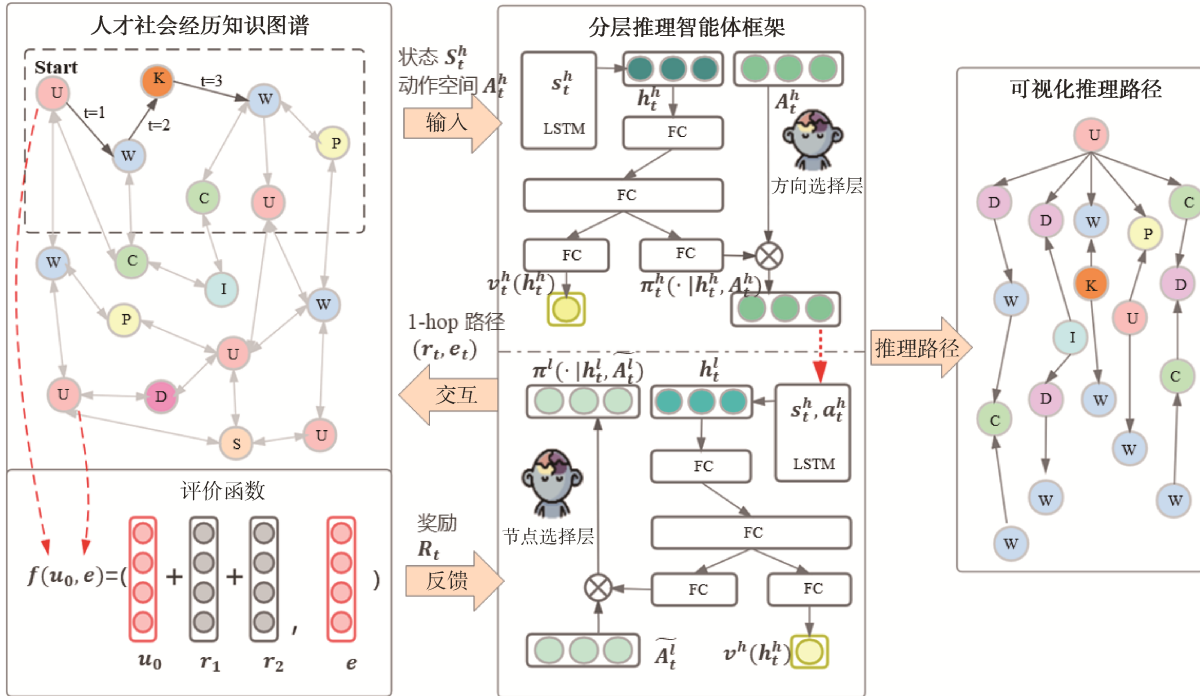


图3 基于人才社会经历知识图谱的分层推理工作推荐模型

Fig.3 Hierarchical Reasoning Work Recommendation Model Based on the Talent Social Experience Knowledge Graph

3.2 人才社会经历知识图谱构建

简历是人才向企业介绍基本信息、教育背景、工

作经历、求职意向等信息的载体, 本身即存在相对结构化的模板, 具备一定的格式。因此在人才社会经

历知识图谱构建中,可以从半结构化的简历信息中通过关键字直接提取用户职业发展的概念字段,如教育背景、工作经历等,以满足知识图谱构建对元数据的需求。

人才社会经历知识图谱的构建主要以简历数据的元数据为依据,对简历数据集中每个用户的工作经历、教育经历进行实体、实体关系抽取,在此基础上构建人才社会经历知识图谱,基于该人才社会经历知识图谱可以提供可解释的推荐路径。知识图谱构建流程主要包括实体抽取、实体消歧、实体关系抽取三部分。

(1) 实体抽取

在原始的简历数据集中,每一份简历都包含求职者的工作经历和教育经历。其中,工作经历包括工作所属行业、工作的公司名、工作岗位、工作职责、工作开始与结束时间;教育经历包括学校、专业、学历、开始与结束时间。对于单个词字段可以直接抽取,而对于长文本字段则通过结巴分词后,使用 TF-IDF 提取长本文中的关键词,将其作为用户完成某个工作所用到的工作技能。

(2) 实体消歧

由于简历数据中的字段文本信息较多,缺乏统一的标准,在用户填写时,相同或相似内容会有不同的文字表达,因此需要对相关的特征字段进行消歧。主要采用关键词匹配、远程监督、文本相似度匹配以及人工标注的方式进行消歧。

①关键词匹配:行业字段虽然类别不多,但是没有统一的标准,且存在重叠的现象。因此采用行业词表^[25],将词表中相关行业的关键词与简历数据集中的行业字段进行关键词匹配,利用关键词进行行业类别归纳和标准化。专业字段存在类似的问题,研究依据国家发布的学科专业目录中二级学科专业目录^[26]对专业进行关键词匹配标注。

②远程监督:对于学校名称存在更名的现象,利用爬虫获取学校的历史名称^[27],然后对学校进行消歧,将2020年中华人民共和国教育部发布的全国高等院校名单^[28]作为标准。以该公司名称作为检索词,利用百度 API 寻找最适合的公司标准名称对公司名进行消歧。

③文本相似度匹配:岗位名称数据类别多且命

名规则不统一,因此采用文本相似度匹配的方法。借助关键词匹配和人工标注从数据集中获取部分准确的结果作为训练集,然后利用 BERT 模型^[29]在该训练集上进行 fine-tune 进而获得领域实体的词向量表示,将词向量表示进行聚类即相似度匹配,最终得到30多个岗位大类200多个岗位中类,示例如表1所示。

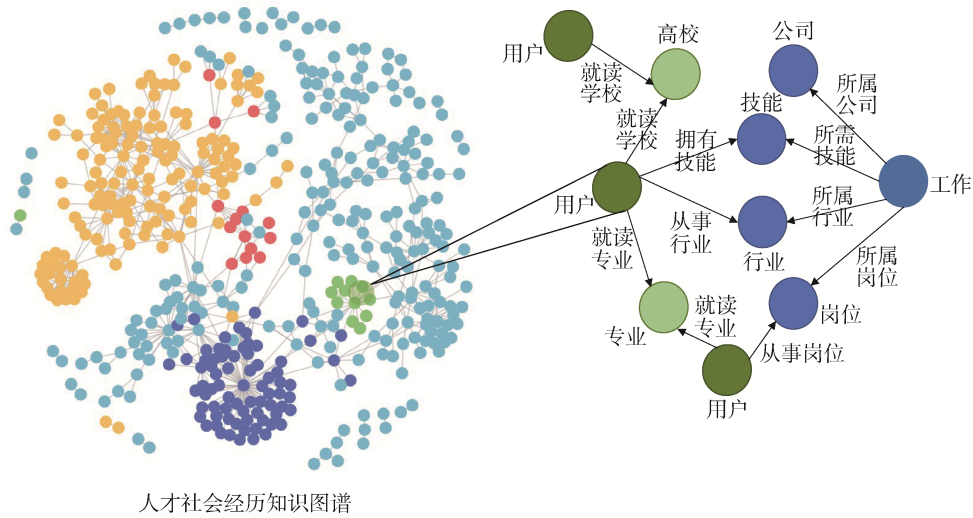
表1 岗位类别示例

大类	中类	原始岗位
计算机技术	测试	软件测试工程师,测试工程师,软件测试,云产品测试开发,集成测试工程师……
	后端开发	后端工程师,web 后端工程师,python 后端开发工程师,高级后端工程师,后端组长,广告后端架构师……
	前端开发	前端工程师助理,js 工程师,Web 前端开发工程师,html 工程师,前端工程师,移动前端负责人……
	信息安全	网络信息安全工程师,信息安全经理,信息安全保障中心网络工程师,安全咨询顾问,密码算法工程师……
	大数据&数据分析	应用分析专员,数据支持专员,数据主管,大数据工程师,数据支撑专员,舆情分析,数据科学家……
	算法	语音算法工程师,推荐算法工程师,图像算法工程师,融合算法工程师,人工智能工程师,数据算法工程师……
	数据库开发	dba 专员,oracle erp 应用工程师,etl+oracle 数据仓库,oracle 高级工程师,数据库设计员,SQL 技术支持……
	……	……

④人工标注:对于无法利用上述手段处理的数据或可能存在模糊的地方,采取人工标注的方式进行数据的标准化处理,数据标注进行了重复标注筛查,以降低数据标注中的人为偏差。比如简历中学校字段会出现不常见的学校缩写,如“黑工程”需要人工标注为“黑龙江工程学院”。

(3) 实体关系抽取

假设共同出现在一段工作经历中的各种实体存在关系,以此进行关系抽取。主要是采用<对象,属性,值域>这种三元组对关系进行抽取,抽取了用户-教育、用户-工作、工作-技能等相关属性关系,这些关系能够较好地反映人才社会经历。获得的知识图谱如图4所示。



人才社会经历知识图谱

图4 人才社会经历知识图谱可视化

Fig.4 Visualization of Talent's Social Experience Knowledge Graph

采用 TransE^[12] 知识图谱表示方法进行嵌入表示用于后续应用,并使用一种可用于学习大型图嵌入表示的分块训练系统 PBG^[30] 进行训练。

3.3 分层推理智能框架

在由大量用户构成的人才社会经历知识图谱中,对于用户类型的节点来说,与其他节点的直接联系相对来说是比较少的,而对于行业、岗位等类型的节点,则连接了海量的工作,整个网络的链接紧密程度是非常不均匀的,而且与一般的商业推荐知识图谱不同的是,工作比商品更难推荐,因为简历数据中没有显式地指出工作是否是同一个,即推荐目标与用户的链接是非常稀疏的,这两个原因使得智能体难以有效探索路径及学习规律。为解决这个问题并完成推荐任务,本文设计了一个包含两层结构的智能体,其中高层次为方向选择层,低层次为节点选择层。简单地说就是将 1-hop 行走分解为选择方向、选择节点两个子过程,当方向选择层完成方向选择后,节点选择层将立即启动。

因此,智能体在知识图谱上进行推荐的大致流程如下:从起始用户节点 e_0 出发,智能体经过选择方向、选择节点两个计算过程,确定 1-hop 行走路径,到达新节点 e_1 ,重复行走过程直到终止状态,终止状态所在的工作节点 e ,则作为推荐对象。不同层次策略的具体定义如下。

(1) 方向选择层

高层次策略即方向选择层的任务是在给定 t -hop 路径 P_t ($0 < t \leq K$) 情景下为智能体选择前进的方向,例如:假设当前路径为 $P_t = \{e_0 \xleftrightarrow{r_1} e_1 \xleftrightarrow{r_2} \dots \xleftrightarrow{r_t} e_t\}$, 则与当前节点 e_t 相连的边可表示为集合 E_h ($E_h = \{(e, r_{t+1}, e_{t+1}) | (e, r_{t+1}, e_{t+1}) \in G\}$), 那么对 E_h 根据 r_{t+1} 的关系类型可分为 R_1, R_2, \dots, R_n ($n \geq 1$) 几个子集,决策层的任务是在 $R_1 \sim R_n$ 中选出 R_{t+1} ($1 \leq t+1 \leq n$) 作为智能体下一步前进的方向。

方向选择层可定义为强化学习策略 μ , 其强化学习有关设定如下:

状态:方向选择层在 t 时刻的状态 $s_t^h \in S^h$ 由上一时刻的 t -hop 路径 P_{t-1} 的序列信息表示,即 $s_t^h = (e_0, r_1, e_1, r_2, e_2, \dots, r_{t-1}, e_{t-1})$, 其中 $r_{i \in [t-1]} \in \mathcal{R}$, $e_{i \in [t-1]} \in \mathcal{E}$, 特别地, $s_0^h = (e_0)$ 。

动作空间:在当前状态 s_t^h 下,对所有与当前节点 e_{t-1} 相连的未遍历的边 $\{(e_{t-1}, r_t, e_t) | (e_{t-1}, r_t, e_t) \in G, e_t \notin \{e_0, e_1, \dots, e_{t-1}\}\}$ 根据 r_t 的关系类型分为子集 R_1, R_2, \dots, R_n ($n \geq 1$), 那么方向选择层的动作空间 $A_t^h = \{R_n | n \in [1, N], N \geq 1\}$ 。

奖励:在工作推荐过程中,不能简单推荐与用户过去的经历相似的工作,因为现实中也存在大量的跨行业、跨专业的工作迁移情况,所以应鼓励智能体

尽可能探索更多的潜在优质推荐目标,以提供新颖的推荐,但同时也需要兼顾推荐的合理性。

因此,当且仅当推荐对象 e_i 属于推荐目标集合 T (工作集合)时,给予智能体一个“软奖励” R_s ,奖励的衡量依据是推荐对象 e_i 与起始节点 e_0 之间在知识图谱结构中的相对相似性,具体的评价函数 $f(e_0, e_i)$ 将在后续介绍。软奖励 R_s 的定义如公式(1)所示。

$$R_s = \begin{cases} \max\left(0, \frac{f(e_0, e_i)}{\max_{e_i \in T} f(e_0, e_i)}\right), & e_i \in T \\ 0, & e_i \notin T \end{cases} \quad (1)$$

此外,为增加推荐的合理性,利用用户的工作经历给智能体增加一个“硬奖励” R_h ,定义如公式(2)所示。

$$R_h = \begin{cases} 1, & (e_0, r_{work}, e_i) \in G \\ 0, & (e_0, r_{work}, e_i) \notin G \end{cases} \quad (2)$$

因此奖励 R_i 定义如公式(3)所示。

$$R_i = R_s + R_h \quad (3)$$

策略:本文设计了符合 Actor-Critic 框架的包含多层线性神经元的策略网络以近似强化学习策略 μ ,实现映射 $\mu: S^h \rightarrow A^h$,同时设计了一个与策略网络共享部分层的估值网络用于优化。为对路径信息进行有效编码,两个网络共享一个 LSTM 网络层,将状态表示为 $h_t^h \in H^h$ 。

策略网络 $\pi^h(\cdot | s_t^h, A_t^h)$ 将当前状态 s_t^h 作为输入,输出动作空间 A_t^h 中每个动作的概率,动作空间外的动作概率为0。估值网络 $v^h(s_t^h)$ 将当前状态 s_t^h 映射为实数作为对当前状态的价值估计,在策略网络优化中发挥作用。具体定义如公式(4)和公式(5)所示。

$$h_t^h = \begin{cases} LSTM^h(0, e_0), & t = 0 \\ LSTM^h(h_{t-1}^h, [r_{t-1}, e_{t-1}]), & t > 0 \end{cases} \quad (4)$$

$$\begin{aligned} x &= \text{dropout}(\sigma(\text{dropout}(\sigma(h_t^h W_1^h))W_2^h)) \\ \pi^h(\cdot | h_t^h, A_t^h) &= \text{softmax}(A_t^h \odot (x W_p^h)) \\ v^h(h_t^h) &= x W_v^h \end{aligned} \quad (5)$$

其中,0为零向量; r_{t-1}, e_{t-1} 由 s_t^h 给定, $r_{t-1}, e_{t-1} \in \mathbb{R}^d$; σ 是非线性的激活函数; $x \in \mathbb{R}^d$ 是从输入状态中学习到的中间特征, d_f 由网络结构决定; \odot 是用于屏蔽动作空间外的非法动作的哈达玛积; A^h 是维度为图 G 中关系类型数量的0-1向量;该层

所有网络的参数定义为 $\theta_\mu = \{LSTM^h, W_1^h, W_2^h, W_p^h, W_v^h\}$ 。

(2) 节点选择层

节点选择层的任务是在给定 t-hop 路径 $P_t (0 < t \leq K)$ 及方向选择层选择了 R_{t+1} 作为前进方向的情景下为智能体选择前进的节点,即在 E_h 的子集 R_{t+1} 中选出一个元素,一条具体的边 $(e_t, r_{t+1}, e_{t+1}) \in R_{t+1}$,那么 (r_{t+1}, e_{t+1}) 即是智能体最终选择的1-hop行走路径。

节点选择层可定义为强化学习策略 ϕ ,其强化学习有关设定如下:

状态:节点选择层在 t 时刻的状态 $s_t^l \in S^l$ 由上一时刻的 t-hop 路径 P_{t-1} 的序列信息及当前时刻方向选择层的决策共同表示,即 $s_t^l = (e_0, r_1, e_1, r_2, e_2, \dots, r_{t-1}, e_{t-1}, r_t)$,其中 $r_{i,i \in [1,t]} \in \mathcal{R}, e_{i,i \in [0,t-1]} \in \mathcal{E}$,特别地, $s_0^l = (e_0, r_1)$ 。

动作空间:在当前状态 s_t^l 及方向选择层选择的动作为 R_t 的情景下,节点选择层的动作空间 $A_t^l = R_t$,其中 $R_t = \{(e_{t-1}, r_t, e_t) | (e_{t-1}, r_t, e_t) \in G, e_t \notin \{e_0, e_1, \dots, e_{t-1}\}\}, r_t$ 为关系集 \mathcal{R} 中一个特定元素。

与方向选择层不同的是,动作空间 A^h 中元素数量大小与图 G 中关系类型数量呈正比从而相对有限,而 A^l 可能非常巨大,因为以同一种关系相连的节点可能会非常多。巨大的动作空间使得探索非常低效,因此需要对动作空间 A^l 的大小进行限制。

一种简单的方式就是随机屏蔽动作空间中部分动作^[31],这样不仅能减小动作空间,还能鼓励进行探索,避免陷入局部最优解。但随机屏蔽过多动作可能导致探索-利用的不平衡,因此还采用了另一种方式约束动作空间。与文献[8]类似,本文采用一种以起始节点为参照对象的动作剪枝策略,选取潜在价值最高的动作,使得动作空间不超过设定最大值。具体来说,使用评价函数 $f((e_{t-1}, r_t, e_t) | e_0)$ 将动作空间 A_t^l 中每个动作 (e_{t-1}, r_t, e_t) 映射为一个实数,依据是 e_0 在图 G 中的空间位置,剪枝后的动作空间 \tilde{A}_t^l 定义如公式(6)所示。

$$\begin{aligned} \tilde{A}_t^l(e_0) &= \{(e_{t-1}, r_t, e_t) | \text{rank}(f((e_{t-1}, r_t, e_t) | e_0)) \leq \alpha, \\ &\quad (e_{t-1}, r_t, e_t) \in A_t^l\} \end{aligned} \quad (6)$$

其中, α 为设定的动作空间最大值, 评价函数 $f((e_{i-1}, r_i, e_i)|e_0)$ 的具体定义将在后续介绍。

奖励: 节点选择层的奖励与方向选择层的奖励相同。

策略: 与方向选择层类似, 本层用包含多层线性神经元的策略网络 $\pi^l(\cdot|s^l, \tilde{A}^l)$ 近似强化学习策略 φ , 实现映射 $\varphi: S^l \rightarrow \tilde{A}^l$, 同时设计了一个与策略网络共享部分层的估值网络 $v^l(s^l)$ 用于优化。具体定义如公式(7)和公式(8)所示。

$$h_t^l = \begin{cases} LSTM^l(0, e_0), t = 0 \\ LSTM^l(h_{t-1}^l, [r_{t-1}, e_{t-1}]), t > 0 \end{cases} \quad (7)$$

$$x = \text{dropout}(\sigma(\text{dropout}(\sigma(h_t^l W_1^l))W_2^l))$$

$$\pi^l(\cdot|h_t^l, \tilde{A}^l) = \text{softmax}(\tilde{A}^l \odot (xW_p^l)) \quad (8)$$

$$v^l(h_t^l) = xW_v^l$$

其中, 与方向选择层不同的是, \tilde{A}^l 是维度为设定的最大动作空间容量的 0-1 向量; 该层所有网络的参数定义为 $\theta_\varphi = \{LSTM^l, W_1^l, W_2^l, W_p^l, W_v^l\}$ 。

(3) 评价函数

动作空间剪枝和奖励给予过程使用的评价函数与文献[8]采用的方法类似。

本文所使用的知识图谱 G 有一个性质, 即给定三元组中头实体类型和关系类型, 那么尾实体的类型也随之确定。因此, 如果给定一个关系序列 $\tilde{r}_k = \{r_1, \dots, r_k\}$ 和头实体 e_0 , 那么任意一条满足 \tilde{r}_k 的 k-hop 路径 $P_k = \{e_0 \xleftrightarrow{r_1} e_1 \xleftrightarrow{r_2} \dots \xleftrightarrow{r_k} e_k\}$ 中所有实体的类型也将确定。由于采用 TransE 方法^[30]对图 G 进行表示, 一个三元组 (h, r, t) 在向量空间中尽可能满足性质: $h + r \approx t$, 因此该性质可用于图 G 中实体关系预测任务。

本文定义多跳评价函数 $f(e_0, e_k|\tilde{r}_k)$ 如公式(9)所示。

$$f(e_0, e_k|\tilde{r}_k) = \left\langle e_0 + \sum_{i=1}^k r_i, e_k \right\rangle \quad (9)$$

其中, $\langle \cdot, \cdot \rangle$ 表示点积操作, $e, r \in \mathbb{R}^d$ 均为实体和关系的嵌入向量, \tilde{r}_k 由环境根据 e_0, e_k 的类型设定。

至此, 用于奖励给予的评价函数 $f(e_0, e_i)$ 即为 $f(e_0, e_i|\tilde{r}_k)$, 用于动作空间剪枝的评价函数 $f((e_{i-1}, r_i, e_i)|e_0)$ 即为 $f(e_0, e_i|\tilde{r}_k)$ 。

(4) 智能体优化

基于上述奖励方式的设定, 智能体为起始用户节点 $u_0 \in \mathcal{U}$ 提供了更优质的推荐意味着获得更多的奖励, 因此目标就是优化智能体的随机型策略最大化期望奖励 $J(\theta_\mu, \theta_\varphi)$, 定义如公式(10)所示。

$$J(\theta_\mu, \theta_\varphi) = E_{\pi^h, \pi^l} \left[\sum_{t=0}^{T-1} \gamma^t R_{t+1} \mid s_0^h = (u_0) \right] \quad (10)$$

与元强化学习优化方式^[32]类似, 将方向选择层策略 μ 和节点选择层策略 φ 分开优化, 此时策略 μ 的动作被策略 φ 当作环境的一部分, 策略 φ 的动作也被策略 μ 当作环境的一部分, 这样当推荐任务增加时, 可以通过共享节点选择层进行优化, 以获得对节点更全面的评价, 各层策略奖励定义如公式(11)和公式(12)所示。

$$J(\theta_\mu) = E_{\pi^h} \left[\sum_{t=0}^{T-1} \gamma^t R_{t+1} \mid s_0^h = (u_0) \right] \quad (11)$$

$$J(\theta_\varphi) = E_{\pi^l} \left[\sum_{t=0}^{T-1} \gamma^t R_{t+1} \mid s_0^l = (u_0, r_0) \right] \quad (12)$$

对每层的优化目标, 具体采用的方法与 A3C 方法^[33]类似, 使用多进程异步采样但是同步更新, 每层策略网络的梯度如公式(13)和公式(14)所示。

$$\nabla_{\theta_\mu} J(\theta_\mu) = E_{\pi^h} \left[\nabla_{\theta_\mu} \log \pi^h(\cdot|s^h, A^h) (R - v^h(s^h)) + \nabla_{\theta_\mu} (R - v^h(s^h))^2 + \beta \nabla_{\theta_\mu} H(\pi^h(\cdot|s^h, A^h)) \right] \quad (13)$$

$$\nabla_{\theta_\varphi} J(\theta_\varphi) = E_{\pi^l} \left[\nabla_{\theta_\varphi} \log \pi^l(\cdot|s^l, \tilde{A}^l) (R - v^l(s^l)) + \nabla_{\theta_\varphi} (R - v^l(s^l))^2 + \beta \nabla_{\theta_\varphi} H(\pi^l(\cdot|s^l, \tilde{A}^l)) \right] \quad (14)$$

其中, R 对每个时刻 t 有 $R_t = r_t + \gamma R_{t+1}$, $t \in \{0, \dots, T-1\}$, T 为终止状态时刻; H 为熵, 是防止策略过于单一的惩罚项; β 为设定参数, 决定熵的权重。

(5) 路径推理策略

使用训练后的分层推理智能体在人才社会经历知识图谱上完成工作推荐任务, 即给定一个用户, 智能体将给出一个推荐工作的候选集 $\{w_n\}$, 同时给出每一个推荐工作的推理路径 $\{p_n\}$ 。一种简单的方式是完全依照分层推理智能体学习到的概率策略进行无约束搜索, 在知识图谱环境中采样 n 次, 得到 n 条路径。

4 实验

4.1 数据集描述

本文实验数据来自e成科技简历脱敏数据集^[34],数据集存储为JSON格式。数据集包括教育经历、工作经历和项目经历三大类近50字段,全面刻画候选人的职业生涯。但由于简历数量庞大,质量参差不齐,需对简历数据进行质量筛选,将缺失值较多或者信息异常的简历数据进行剔除,最终得到39 081份优质简历数据。因为数据比较稀疏,所以按照用户划分,将75%作为训练集,10%为验证集,15%为测试集。

在此基础上构建的知识图谱实体和关系的类型及数量如表2和表3所示。

表2 人才社会经历知识图谱实体简要统计

Table 2 Entity Statistics of Talent Social Experience Knowledge Graph

实体类型	描述	数量
User	用户	39 081
Work	工作	96 990
Corporation	公司/单位名称	61 125
Position	岗位名称	249
Industry	行业名称	38
Keyword	关键词	90 149
School	毕业院校	1 149
Discipline	专业	84

4.2 基准模型

在实验中,将本文提出的模型与一些经典模型进行性能对比。由于候选工作数量较多,某些模型难以对每个候选工作进行计算并召回,因此人为将1个正样本(未来真实存在的工作经历)与20个相似的负样本(同行业的工作经历)混合,仅对该部分模型的排序性能进行测试。

(1) 仅用于排序的模型

逻辑回归(Logistic Regression, LR)^[35]是一种被广泛使用的机器学习算法,通常用于解决二分类问题,在推荐的场景,通过从历史数据中对用户与物品的交互情况进行分类训练(如电商场景,用户是否购买),从而以概率的形式估计出用户对某个物品的偏好,然后将待推荐商品序列按照模型预估的偏好大小进行排序,得到top-N推荐结果。

表3 人才社会经历知识图谱关系简要统计

Table 3 Relationship Statistics of Talent Social Experience Knowledge Graph

关系类型	描述	数量
Employer	Corporation $\xleftrightarrow{\text{Employer}}$ User	97 003
Work_as	User $\xleftrightarrow{\text{Work_as}}$ Position	97 003
Work	User $\xleftrightarrow{\text{Work}}$ Work	97 003
Provide	Corporation $\xleftrightarrow{\text{Provide}}$ Position	97 003
Describe_position	Position $\xleftrightarrow{\text{Describe_position}}$ Key_word	897 510
Field	User $\xleftrightarrow{\text{Field}}$ Industry	97 003
Involve	Position $\xleftrightarrow{\text{Involve}}$ Industry	97 003
Graduation	User $\xleftrightarrow{\text{Graduation}}$ School	46 178
Major	User $\xleftrightarrow{\text{Major}}$ Discipline	46 178
Attr_Position	Work $\xleftrightarrow{\text{Attr_Position}}$ Position	97 003
Attr_Corporation	Work $\xleftrightarrow{\text{Attr_Corporation}}$ Corporation	97 003
Attr_Key_word	Work $\xleftrightarrow{\text{Attr_Key_word}}$ Key_word	897 510
Attr_Industry	Work $\xleftrightarrow{\text{Attr_Industry}}$ Industry	97 003

贝叶斯个性化排序(Bayesian Personalized Ranking, BPR)^[36]是一种经典的推荐算法,它以配对样本方式(Pair-Wise)进行训练,通过降低与偏好矩阵的差距,从而学习到用户和物品的潜在表示向量,然后通过潜在表示向量进行内积运算得到用户与拟推荐物品的匹配度,在实际中被广泛应用于从海量物品中进行top-N推荐。

联合表示学习(Joint Representation Learning, JRL)框架^[37]是一种融合多源信息的深度学习top-N推荐框架,可以融合多种来源与多种数据类型的信息源,比如评论文本、产品图片、产品评分等,学习到用户和物品的表示向量,模型以梯度下降的方法对正样本和负样本的配对损失进行训练。

(2) 用于召回并排序的模型

PGPR^[7]用强化学习方法将知识图谱路径推理用于推荐任务以获得可解释性,该方法用一个A3C网络^[33]近似推理策略,并利用基于图嵌入的打分函数对召回的物品进行排序。

(3) 基线模型的适应性说明

在本文的应用场景中,采用以上基准模型根据用户的简历信息进行工作推荐需要考虑以下问题:

①输入信息的多样性

简历包括用户的教育经历和工作经历,而教育经历包括学校、专业、学历信息,工作经历包括行业、公司、岗位和负责事项。

②推荐工作的稀疏性

不同于传统的电影推荐、商品推荐等,在本实验数据中,可以认为几乎没有同一份工作被两个用户同时做过,这造成了工作的稀疏性。

因此,在应用以上模型时,必须进行一些输入特征或者模型结果上的修改以适应推荐场景,使用浅层神经网络提取特征作为 LR 和 BPR 的输入,对 JRL 框架,根据深度学习特征交互的方式分为表示型(JCL-rep)和交互型(JCL-int)。

4.3 评价指标

对各种模型的推荐效果使用以下指标进行评价: MRR (Mean Reciprocal Rank)、Hit (Hit ratio)、P (Precision)、R (Recall), 根据 top-M 推荐列表的不同,将 P、R 评价指标表示为 P@M、R@M 等。

4.4 实施细节

本文采用的训练和测试方式如下:

对基于知识图谱的推荐模型,为防止预测目标关联节点造成信息泄露,先利用全部样本构建知识图谱,训练知识图谱嵌入向量,然后在训练集中,屏蔽“User”节点与“Work”节点的关系链接,进行模型训练;在测试集中,屏蔽“User”节点与“Work”节点的关系链接以及序列预测目标在知识图谱构建过程中生成的“泄露”该用户信息的链接(如岗位、公司等)。

本文默认的参数设置如下:

(1) 在知识图谱方面,对实体和关系采用 TransE 方法^[12]进行嵌入表示为维度为 250 的向量。在进行推荐任务时,设定最长的推理路径长度为 6。

(2) 在智能体方面,作为状态编码器的 LSTM^[38]网络在上下层策略中均设置为 300 个神经元,层数为 3。策略网络和估值网络在上下层中采用类似的结构,即 $W_1 \in R^{300 \times 512}$, $W_2 \in R^{512 \times 256}$, $W_p \in R^{256 \times M}$, $W_v \in R^{256 \times 1}$, 其中 M 为上下层动作空间的容量,方向选择层为 13,节点选择层为 50。

(3) 在训练过程方面,使用 Adam^[39]优化方法对智能体进行 12 代(Epoch)训练,学习步长为 0.000 1,训练批次为 128。为使智能体学到更多样同时兼具

合理性的策略,在损失函数中加入了策略的熵(Entropy)作为惩罚项,并对惩罚力度进行针对性的设计:将方向选择层在所有时刻的策略熵进行加权平均,权重分别为 1.0, 0.5, 0.05, 0.001, 0.001, 0.001, 节点选择层的权重均为 0.001。

4.5 实验效果

本文模型与各个基准模型的推荐效果对比如表 4 和表 5 所示。

表 4 仅用于排序的模型效果对比

Table 4 Comparison of Effects for Ranking-only Model

模型	MRR@20(%)	Hit@1(%)	Hit@5(%)	Hit@10(%)
LR	28.7	18.0	50.8	73.7
BPR	30.6	20.2	51.7	74.1
JRL-int	34.9	25.3	55.6	77.4
JRL-rep	34.7	25.0	61.8	80.6
Our	81.7	74.8	92.2	97.0

表 5 用于召回并排序的模型效果对比

Table 5 Comparison of Effects for Recall and Ranking Model

模型	P@10(%)	R@10(%)	P@20(%)	R@20(%)	MRR@20(%)	候选集平均大小(个)
PGPR	0.2	1.9	0.3	6.1	0.8	69.6
Our	1.6	16.1	0.9	17.1	11.8	58.1

从表 4 中可以看出, JRL 模型在基准模型中是比较强的, 相比于 LR 和 BPR, 深度学习网络的引入增加了表达能力, 而基于图嵌入的排序方法性能则远远超过基准模型, 在各项指标的评估中都取得了最优的效果。

表 5 显示, 与同样使用强化学习方法进行路径推理的 PGPR 相比, 本文提出的分层强化学习推理策略在各项指标上均有提升, 而且是在召回更小候选集前提下实现的, 这表明本文模型更为高效。

在实验中测试了不同图嵌入维度对推荐效果的影响, 结果如表 6 所示。

表 6 图嵌入维度对模型性能的影响

Table 6 Effect of Graph Embedding Dimension on Model Performance

嵌入维度	P@10(%)	R@10(%)	P@20(%)	R@20(%)	MRR@20(%)	候选集平均大小(个)
200	0.7	7.4	0.4	8.3	5.7	61.8
250	1.6	16.1	0.9	17.1	11.8	59.1
300	0.7	7.2	0.4	7.6	5.5	58.1

在模型参数相同的情况下,图嵌入维度对推荐效果的影响较大,并且效果没有随着维度增加而提升,这说明选择合适的图嵌入维度至关重要。

4.6 案例分析

智能体经过探索、学习后,共发现了 58 种不同长度的有效路径模式(能够发现真实存在的未来工作则认为有效),如表 7 所示,部分模式如图 5 所示。

表 7 不同长度路径模式数量及有效性

Table 7 Number and Effectiveness of Path Patterns with Different Lengths

路径长度	2-hop	3-hop	4-hop	5-hop	6-hop
路径模式数量	3	5	19	6	25
推理路径占比	16.0%	16.2%	63.4%	1.9%	2.5%
有效路径占比	82.3%	4.5%	12.5%	0.2%	0.6%

从模式数量来看,智能体发现了多种路径模式,具有提供多样性推荐的能力。从模式的内容来看,推理的步骤基本合理,且具有一定的可解释性,例如发现了与用户目前同一家公司的同事跳槽到某公司,并成功预测了该用户也会跳槽到该公司。从有

效路径占比来看,较短路径占比较大,表明多跳推理仍然不够有效。

为验证模型的可解释性和直观地感受推荐的有效性,随机抽取几个测试样本进行 top10 推荐,并展示真实存在于用户未来工作经历的成功推荐(Dream Work),以此分析案例,如图 6 所示。

对人才 A 和 D,智能体从毕业院校、所学专业出发,找到了同专业、同学校的其他人才,然后根据他们所从事的行业推荐工作。这表明智能体通过探索学习到了协同过滤的方法,在缺乏工作经历的冷启动情况下也能进行推荐,并给出相应理由。

对人才 B,智能体找到了同专业的人从事的岗位,并根据人才 B 之前工作的行业,推荐符合该岗位该行业的工作。对人才 C,智能体不仅发现有同专业同学校的其他人才在从事测试和开发岗位,还发现人才 C 曾供职的公司也提供了该岗位,基于这些推理路径推荐了该工作。这表明智能体具备多样化的推理能力,能从多个角度给出同一推荐,增加了推荐的说服力和可解释性。

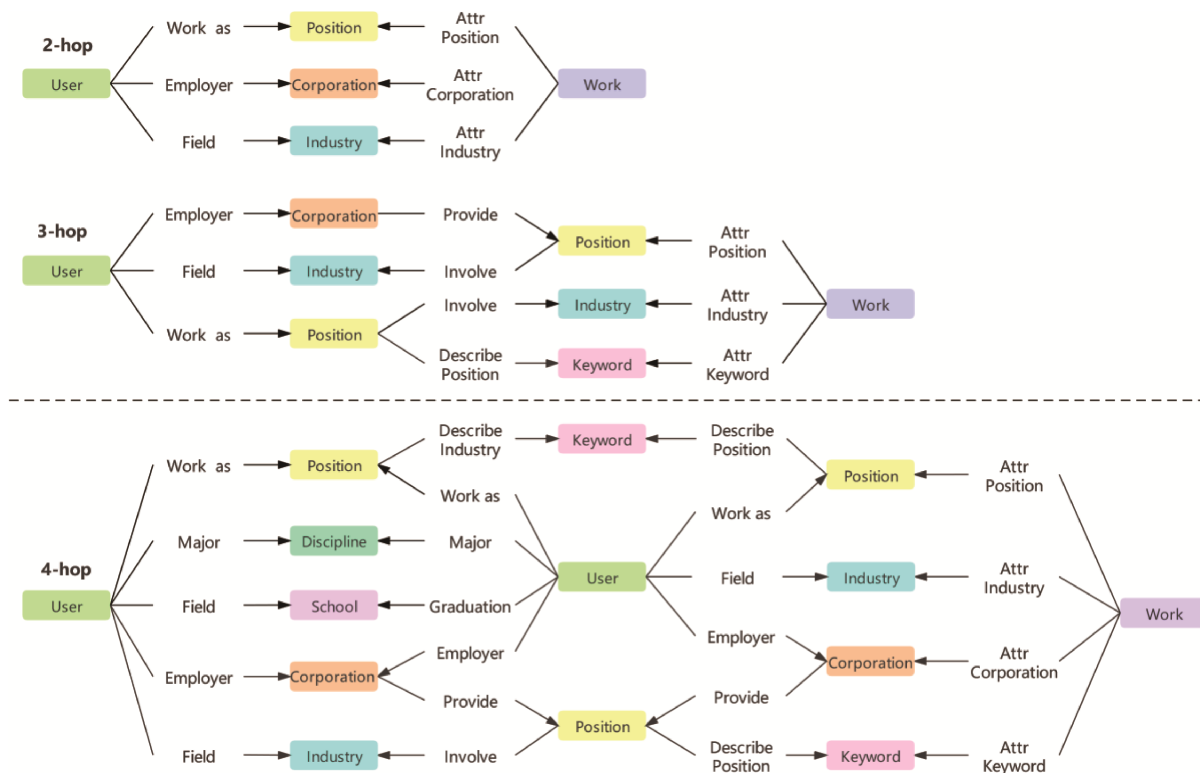


图 5 智能体发现的不同长度的有效路径模式

Fig.5 Effective Path Patterns of Different Lengths Discovered by Agents

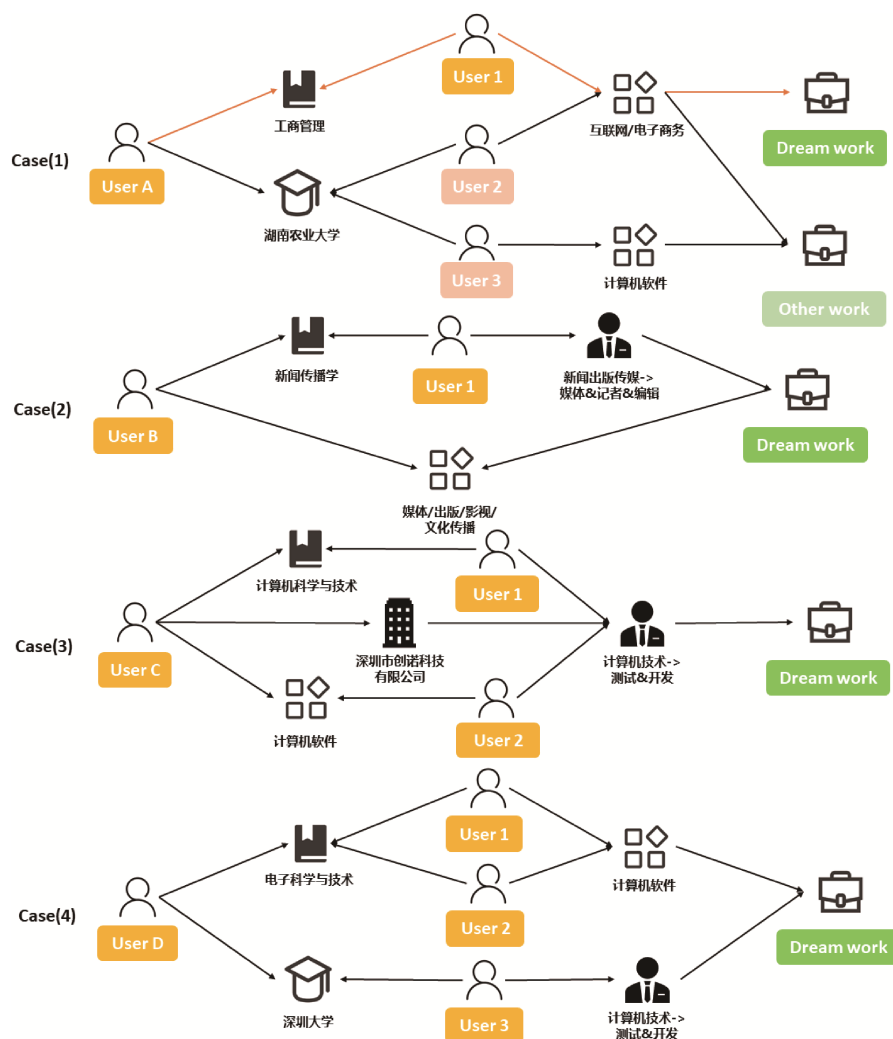


图 6 真实工作推荐案例分析

Fig.6 Case Analysis of Real Job Recommendation

从上述案例中可以发现,智能体能够结合人才的所学专业、所处行业、毕业学校等信息,针对不同人才的社会经历进行多样化、个性化的推荐,并兼顾了新颖性、合理性和可解释性。

5 结 语

本文基于真实的简历数据集构建人才社会经历知识图谱,并基于分层强化学习理论提出了一种新颖的知识图谱分层推理策略,提高了推理效率,为大规模应用提供可能。研究实现了推理路径可视化的可解释工作推荐模型,能够有效结合前人经历进行推荐,同时给出推理路径,有效应对了冷启动和缺乏

新颖性、可解释性问题。研究结果证明了该方法在工作推荐方面的有效性和可解释性。未来研究可以进一步加大数据量,完善人才社会经历知识图谱。本文对于人才社会经历中的时间信息没有充分利用,下一步可以考虑在工作推荐时融入时序特征。未来的推荐不仅仅是工作,还可以进一步告诉用户需要掌握哪些技能等。

参考文献:

- [1] 腾讯网. 2020 年 1-7 月全国就业情况分析全国城镇新增就业 674 万人 [EB/OL]. (2020-09-16). [2020-09-30]. <https://new.qq.com/rain/a/20200916A05NLM00>. (Tencent. Analysis of National Employment from January to July 2020: 6.74 Million New Urban

- Jobs were Created[EB/OL]. (2020-09-16). [2020-09-30]. <https://new.qq.com/rain/a/20200916A05NLM00>.)
- [2] Zhu H, Li X, Zhang P Y, et al. Learning Tree Based Deep Model for Recommender Systems[C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2018:1079-1088.
- [3] Kanakia A, Shen Z H, Eide D, et al. A Scalable Hybrid Research Paper Recommender System for Microsoft Academic[C]//Proceedings of the 2019 World Wide Web Conference. 2019: 2893-2899.
- [4] Kumar B, Sharma N. Approaches, Issues and Challenges in Recommender Systems: A Systematic Review[J]. Indian Journal of Science and Technology, 2016, 9(47):1-12.
- [5] Schein A I, Popescul A, Ungar L H, et al. Pennock Methods and Metrics for Cold Start Recommendations[C]//Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.2002:253-260.
- [6] Gughani A, Misra H. Implicit Skills Extraction Using Document Embedding and Its Use in Job Recommendation[C]//Proceedings of the 34th AAAI Conference on Artificial Intelligence. 2020,34 (8): 13286-13293.
- [7] Xian Y K, Fu Z H. Muthukrishnan S, et al. Reinforcement Knowledge Graph Reasoning for Explainable Recommendation [C]//Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval.2019:285-294.
- [8] Lin X V, Socher R, Xiong C M. Multi-Hop Knowledge Graph Reasoning with Reward Shaping[C]// Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing.2018: 3243-3253.
- [9] Domeniconi G, Moro G, Pagliarani A, et al. Job Recommendation from Semantic Similarity of LinkedIn Users' Skills[C]//Proceedings of the 5th International Conference on Pattern Recognition Applications and Methods.2016:270-277.
- [10] Qin C, Zhu H S, Xu T, et al. Enhancing Person-Job Fit for Talent Recruitment: An Ability-aware Neural Network Approach[C]// Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval.2018: 25-34.
- [11] Meng Q X, Zhu H S, Xiao K L, et al. A Hierarchical Career-Path-Aware Neural Network for Job Mobility Prediction[C]// Proceedings of the 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining.2019: 14-24.
- [12] Bordes A, Usunier N, García-Durán A, et al. Translating Embeddings for Modeling Multi-Relational Data[C]// Proceedings of the 26th International Conference on Neural Information Processing Systems.2013: 2787-2795.
- [13] Nickel M, Tresp V, Kriegl H. A Three-Way Model for Collective Learning on Multi-Relational Data[C]//Proceedings of the 28th International Conference on Machine Learning.2011:809-816.
- [14] Zhang F Z, Yuan J N, Lian D F, et al. Collaborative Knowledge Base Embedding for Recommender Systems[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.2016: 353-362.
- [15] Huang J, Zhao W X, Dou H J, et al. Improving Sequential Recommendation with Knowledge-Enhanced Memory Networks [C]//Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval.2018:505-514.
- [16] Wang H W, Zhang F Z, Wang J L, et al. RippleNet: Propagating User Preferences on the Knowledge Graph for Recommender Systems[C]//Proceedings of the 27th ACM International Conference on Information and Knowledge Management. 2018: 417-426.
- [17] Ai Q Y, Azizi V, Chen X, et al. Learning Heterogeneous Knowledge Base Embeddings for Explainable Recommendation [J]. Algorithms, 2018, 11(9):137.
- [18] Wang X, Wang D X, Xu C R, et al. Explainable Reasoning over Knowledge Graphs for Recommendation[C]//Proceedings of the 33rd AAAI Conference on Artificial Intelligence. 2019: 5329-5336.
- [19] Silver D, Huang A, Maddison C J, et al. Mastering the Game of Go with Deep Neural Networks and Tree Search[J]. Nature, 2016, 529(7587):484-489.
- [20] Theocharous G, Thomas P S, Ghavamzadeh M, et al. Personalized Ad Recommendation Systems for Life-Time Value Optimization with Guarantees[C]//Proceedings of the 24th International Conference on Artificial Intelligence. 2015: 1806-1812.
- [21] Zheng G J, Zhang F Z, Zheng Z H, et al. DRN: A Deep Reinforcement Learning Framework for News Recommendation [C]//Proceedings of the 2018 World Wide Web Conference.2018: 167-176.
- [22] Wang X T, Chen Y R, Yang J, et al. A Reinforcement Learning Framework for Explainable Recommendation[C]//Proceedings of the 2018 IEEE International Conference on Data Mining. 2018: 587-596.
- [23] Xiong W H, Hoang T, Wang W Y. DeepPath: A Reinforcement Learning Method for Knowledge Graph Reasoning[C]// Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. 2017:564-573.
- [24] Das R, Dhuliawala S, Zaheer M, et al. Go for a Walk and Arrive at the Answer: Reasoning Over Paths in Knowledge Bases Using Reinforcement Learning[C]//Proceedings of the 31st Conference on Neural Information Processing Systems.2017.
- [25] 国家统计局. 2017年国民经济行业分类 (GB/T 4754-2017)[EB/OL].(2017-09-29). [2020-09-30]. <http://www.stats.gov.cn/tjsj/tjbz/>

- hyflbz/201710/t20171012_1541679.html. (National Bureau of Statistics. 2017 National Economic Industry Classification (GB/T 4754-2017) [EB/OL]. (2017-09-29). [2020-09-30]. http://www.stats.gov.cn/tjsj/tjbz/hyflbz/201710/t20171012_1541679.html.)
- [26] 中国学位与研究生教育信息网. 学科、专业目录[EB/OL]. [2020-09-30]. <http://www.cdgdc.edu.cn/xwyyjsjyxx/sy/glmd/264462.shtml>. (China Academic Degrees & Education Information. Disciplines and Specialties Directory[EB/OL]. [2020-09-30]. <http://www.cdgdc.edu.cn/xwyyjsjyxx/sy/glmd/264462.shtml>.)
- [27] 中华人民共和国教育部. 1990 年以来高校合并情况(截止到 2006 年 5 月 15 日)[EB/OL].(2006-05-15). [2020-09-30]. http://www.moe.gov.cn/srcsite/A03/moe_634/200605/t20060515_88440.html.(Ministry of Education of the People's Republic of China. Mergers of Colleges and Universities Since 1990 (as of May 15, 2006) [EB/OL]. (2006-05-15). [2020-09-30]. http://www.moe.gov.cn/srcsite/A03/moe_634/200605/t20060515_88440.html.)
- [28] 中华人民共和国教育部. 教育部发展规划司院校设置[EB/OL]. (2020-08-28). [2020-09-30]. http://www.moe.gov.cn/s78/A03/ghs_left/s181/.(Ministry of Education of the People's Republic of China. Colleges and Universities Setting, The Ministry of Education Development Department [EB/OL]. (2020-08-28). [2020-09-30]. http://www.moe.gov.cn/s78/A03/ghs_left/s181/.)
- [29] Devlin J, Chang M W, Lee K, et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding[OL]. arXiv Preprint, arXiv:1810.04805.
- [30] Lerer A, Wu L, Shen J J, et al. PyTorch-BigGraph: A Large-scale Graph Embedding System[C]//Proceedings of the 2nd Conference on Systems and Machine Learning. 2019.
- [31] Yu X, Gu Q Q, Zhou M W, et al. Citation Prediction in Heterogeneous Bibliographic Networks[C]//Proceedings of the 2012 SIAM International Conference on Data Mining. 2012: 1119-1130.
- [32] Frans K, Ho J, Chen, X, et al. Meta Learning Shared Hierarchies [C]//Proceedings of the 6th International Conference on Learning Representations.2018.
- [33] Mnih V, Badia A P, Mirza M, et al. Asynchronous Methods for Deep Reinforcement Learning[C]// Proceedings of the 33rd International Conference on Machine Learning.2016: 1928-1937.
- [34] e 成科技. “e 成科技简历脱敏数据集”[EB/OL].[2020-06-10]. http://hdl.handle.net/20.500.12291/10226_V1 [Version].(Ifchange. “Resume Desensitization Dataset of Ifchange” [EB/OL]. [2020-06-10]. http://hdl.handle.net/20.500.12291/10226_V1 [Version].)
- [35] Wright R E. Logistic Regression[J].Reading and Understanding Multivariate Statistics,1995:217-244.
- [36] Rendle S, Freudenthaler C, Gantner Z, et al. BPR: Bayesian Personalized Ranking from Implicit Feedback[C]//Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence. 2009:452-461.
- [37] Zhang Y F, Ai Q Y, Chen X, et al. Joint Representation Learning for Top-N Recommendation with Heterogeneous Information Sources[C]//Proceedings of the 2017 ACM Conference on Information and Knowledge Management. 2017:1449-1458.
- [38] Hochreiter S, Schmidhuber J. Long Short-Term Memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [39] Kingma D P, Ba J. Adam: A Method for Stochastic Optimization [OL]. arXiv Preprint, arXiv: 1412.6980.

作者贡献声明:

阮小芸:设计研究方案,标注、清洗和分析数据,论文起草与修订;
廖健斌:模型设计,进行实验,论文修订;
李祥:设计研究方案,标注、清洗数据,进行实验,论文修订;
杨阳:设计研究方案,采集、标注和清洗数据;
李岱峰:模型设计,论文最终版本修订。

利益冲突声明:

所有作者声明不存在利益冲突关系。

支撑数据:

支撑数据由作者自存储, E-mail:1090419436@qq.com。

- [1] 阮小芸. Main Model.zip. 实验主模型代码.
- [2] 阮小芸. Data_39081.pkl. 实验数据集.
- [3] 阮小芸. Baseline.zip. 实验基准模型代码.
- [4] 阮小芸. Analysis.zip. 实验数据处理代码.

收稿日期:2020-12-06
收修改稿日期:2021-03-13

Interpretable Recommendation of Reinforcement Learning Based on Talent Knowledge Graph Reasoning

Ruan Xiaoyun Liao Jianbin Li Xiang Yang Yang Li Daifeng

(School of Information Management, Sun Yat-Sen University, Guangzhou 510006, China)

Abstract: [Objective] This paper proposes an interpretable reinforcement learning method for job recommendation based on talent knowledge graph reasoning, which addresses the issues of difficulties in large-scale application, cold start, and lack of novelty. [Methods] First, we constructed a knowledge graph for the social experience of the job applicants based on their resume data. Then, we trained a strategic agent with the knowledge graph and the theory of reinforcement learning. This algorithm, which divided the reasoning process into choosing directions and nodes, could identify potential high-quality recommendation targets from the knowledge graph. [Results] The MRR@20 (81.7%), Hit@1 (74.8%), Hit@5 (92.2%) and Hit@10 (97.0%) of the proposed model were higher than those of the LR, BPR, JRL-int, JRL-rep and PGPR models. [Limitations] The size of the experimental datasets and the task-types needs to be further expanded. [Conclusions] Our model could effectively recommend jobs for applicants based on their previous experience or other successful recommendations. It also provides reasoning paths with the help of knowledge graph.

Keywords: Work Recommendation Knowledge Graph Reasoning Reinforcement Learning Interpretable Recommendation

科学家试图通过传感器数据融合攻克雪地驾驶

全自动驾驶汽车面临的一个主要挑战是在恶劣的天气中行驶。尤其是下雪天,自动驾驶汽车会混淆关键的传感器数据,而自动驾驶汽车需要利用这些数据来测量深度、找到障碍物并保持在黄线正确的一侧行驶。近期,密歇根理工大学的研究人员研究了雪天驾驶场景的解决方案,可以帮助自动驾驶在芝加哥、多伦多等雪地城市的应用。

自动驾驶汽车涵盖了多个级别,从市场上具有盲点警告或制动辅助功能的汽车,到可以切换自动驾驶模式的汽车,再到可以完全自主驾驶的汽车。几大汽车制造商和相关研究人员仍在调整自动驾驶技术的算法。目前市面上的自动驾驶汽车偶尔也会发生事故,要么是由于人工智能的判断出错,要么是人类驾驶员对自动驾驶功能的误用。

人类通过眼睛、平衡感和运动感以及大脑的处理能力帮助人们了解周围的环境。这些看似基本的输入使人们能够在几乎所有场景中驾驶,即使这个场景是全新的场景,因为人类的大脑擅长概括新的体验。而在自动驾驶汽车中,安装在万向节上的两个摄像头使用立体视觉来扫描和感知深度,以模仿人类的视觉,而平衡和运动可以使用惯性测量单元进行测量,但是,计算机只能对它们以前遇到过或被编程识别的场景做出反应。

由于人工大脑还没有出现,因此特定任务中,必须依靠人工智能算法,这意味着自动驾驶汽车必须依赖多个传感器的数据,比如鱼眼相机扩大了视野,其他相机的作用很像人眼,红外线接收器能接受热信号,雷达可以看穿雾和雨,光探测和测距(激光雷达)能穿透黑暗。密歇根理工大学的研究人员致力于通过传感器融合的人工智能过程将传感器的数据整合在一起以改善自动驾驶汽车在恶劣环境下的表现:“每个传感器都有其局限性,同时每个传感器都能辅助另一个传感器。传感器融合技术使用多个不同模式的传感器来理解场景。”通过传感器数据融合可以更好地了解恶劣天气,随着传感器融合等技术进步,未来将能够在积雪的道路上安全行驶。

(编译自: <https://www.sciencedaily.com/releases/2021/05/210527172545.htm>)

(本刊讯)