

基于知识图谱的人岗推荐系统构建*

徐孟奇¹, 熊 熙^{1,2†}, 李斌勇¹, 李中志¹

(1. 成都信息工程大学 网络空间安全学院, 成都 610225; 2. 四川大学 空天科学与工程学院, 成都 610065)

摘要: 为解决传统人岗推荐系统存在的三个常见问题,即数据稀疏性、数据冷启动和数据利用率低,提出了基于知识图谱的人岗推荐系统构建方法。该方法通过改进传统推荐模型,将知识图谱作为辅助边信息融合到推荐系统中进行人岗推荐,有效解决了数据稀疏性和数据冷启动问题;引入知识图谱补全算法提高了数据利用率。提出的方法在人岗推荐上准确率可达92%,比现有人岗推荐方法准确率提高约1%。实验结果表明该方法是可行的,知识图谱的加入可以提升人岗推荐系统的推荐效果。

关键词: 知识图谱; 推荐系统; 人岗推荐; 知识图谱补全

中图分类号: TP18 文献标志码: A 文章编号: 1001-3695(2022)01-034-05

doi: 10.19734/j.issn.1001-3695.2021.06.0228

Construction of person-post matching recommended system based on knowledge graph

Xu Mengqi¹, Xiong Xi^{1,2†}, Li Binyong¹, Li Zhongzhi¹

(1. School of Cybersecurity, Chengdu University of Information Technology, Chengdu 610225, China; 2. School of Aeronautics & Astronautics, Sichuan University, Chengdu 610065, China)

Abstract: To solve the three common problems in the traditional person-post matching recommended system: data sparseness, data cold start and low data utilization rate, this paper proposed knowledge graph-based person-post matching recommended system construction methods. By improving traditional recommendations, it integrated knowledge graph as auxiliary side information to the recommended system to carry out person-post matching recommended system, effectively solved data sparseness and data cold start issues, and introduced knowledge graph completion algorithm to improve data utilization, the proposed method could reach 92% in the recommended accuracy rate, and the accuracy rate was increased by about 1% more than the existing person-post matching recommended methods. The experimental results show that the method is feasible, and the addition of the knowledge graph can improve the recommendation effect of the person-post matching recommended system.

Key words: knowledge graph; recommended system; person-post matching recommended; knowledge graph completion

0 引言

大数据时代,信息过载使人们面临大量商品时需要花更多时间进行挑选。推荐系统^[1] (recommender system) 可以帮助用户从互联网数据中更全面和更精准地筛选内容,比如在购物推荐、书籍推荐、游戏推荐、音乐推荐^[2] 和电影推荐^[3] 等场景均有广泛应用。人岗推荐系统作为推荐系统的主要应用方向之一,能够帮助求职人员快速高效地匹配到求职岗位,同时也能降低招聘企业的筛选工作量,提高招聘效率,因而日渐受到用人单位与社会的广泛关注。主流的推荐系统是基于协同过滤 (collaborative filtering-based recommendation, CF) 的推荐系统^[4]。基于 CF 的推荐主要是以用户或交互数据项的相似度来建模用户偏好,目前基于 CF 的推荐系统仍然被广泛应用,因为它可以有效地捕获用户偏好,并且容易在多个场景中实现,而不需要在基于内容的推荐系统^[5] 中提取特征。虽然基于 CF 的推荐系统多年来广受青睐,但其存在的数据稀疏性和冷启动问题^[6] 不容忽视,为了减少其固有问题的影响,不少研究人员提出将侧边信息整合到 CF 中,如用户/项目属性^[7]、项目评论^[8]、社交网络^[9]、混合推荐系统^[4] 等,有了这些额外整合的

侧边信息,可以有效减少数据的稀疏性和数据冷启动问题。受此启发,本文尝试通过融合知识图谱的方式来进一步提高人岗推荐系统的准确率。

知识图谱以结构化的方式对所有知识作出表示,这种表示方式越来越受学术和工业应用的青睐,简单明了的三元组形式的结构就能组成知识图谱网络,虽然其结构简单,仅由实体—关系—语义描述这三部分组成,但仍能完整地表示出现实世界中的各种知识。其中实体可以是现实世界中的对象和抽象概念,关系表示两个实体之间的关联,实体及其关系的语义描述包含了具有明确含义的类型和属性。为了解决传统推荐系统存在的数据稀疏和冷启动问题,将知识图谱引入推荐系统作为辅助信息已经引起了研究者的关注,并且将知识图谱引入到人岗推荐系统中,可以有效解决简历数据的稀疏性和冷启动问题,使得人岗推荐系统的推荐效果更加精准。

本文在人岗数据库的基础上构建知识图谱,提高了人岗推荐和岗位匹配效果,同时利用知识图谱补齐算法,有效解决了数据稀疏性问题。本文通过融合知识图谱的推荐算法,基于知识图谱进行人岗推荐,大大提高了人岗数据的利用率,同时也有效缓解了人岗推荐无历史参考数据的问题。对 RippleNet 模

收稿日期: 2021-06-06; 修回日期: 2021-08-03 基金项目: 国家自然科学基金资助项目(81901389); 国家教育部“春晖计划”合作科研项目(2019); 中国博士后科学基金资助项目(2019M653400)

作者简介: 徐孟奇(1994-),男,河南新乡人,硕士研究生,主要研究方向为人工智能、推荐系统、知识图谱;熊熙(1983-),男(通信作者),四川成都人,副教授,硕士,主要研究方向为大数据、人工智能、工业互联网(xiongxi@cuit.edu.cn);李斌勇(1982-),男,四川成都人,硕士,主要研究方向为云服务、云计算与大数据安全、网络化协同制造;李中志(1973-),男,四川成都人,副教授,硕士,主要研究方向为计算机应用、信息与信号处理。

型进行改进,增加了知识图谱补全功能,提高人岗数据的利用率,实验证明知识图谱补全后人岗推荐效果有明显提升。

1 相关工作

人岗推荐系统又称职位推荐系统,随着推荐系统研究的快速发展,人岗推荐是推荐系统很重要的应用方向之一。对现有文献进行归纳发现,截止到目前,并没有相关人员将知识图谱算法应用到人岗推荐系统,自知识图谱被提出以来,基于知识图谱的推荐方法多被应用到电影、音乐和书籍推荐系统等方面,对比传统的基于协同过滤等的推荐方法,基于知识图谱的推荐算法确实可以取到意想不到的效果,但是在人岗推荐领域,尚无研究人员将知识图谱结合推荐算法成功应用到人岗推荐上来,并且当前对人岗推荐系统的研究成果并不多见,主要的研究内容如下:2006年何峰^[10]提出将求职要求的文本信息向量化,建立向量空间模型,利用文本分类进行人才推荐系统,该推荐系统能智能分类求职文本信息,并与招聘单位发布的职位智能匹配,进而实现个性化推荐功能;2007年,黄海江等人^[9]设计研发了一款职位推荐系统,在功能上基本实现了职位信息与求职者信息的匹配,但是该推荐系统并没有考虑到求职者的求职意向和招聘单位的招聘意向;郭军军^[11]于2009年提出个性化的职位推荐系统,采用轮盘模型表征用户的求职意向,使用反馈用户空间向量模型对用户和职位建模,最后应用基于内容和协同过滤算法的混合模型进行推荐;2014年路小瑞^[12]提出利用Hadoop平台设计的职位推荐系统,该系统通过用户的点击获取用户偏好并且结合用户的个人信息等判断职位的匹配程度,从而给出推荐;2016年文献^[13]提出了基于社交网络的职位推荐系统,给用户推荐其好友工作过的或是正在工作的公司职位,来解决职位相关度不足和过载的问题。通过对之前的相关文献进行归纳总结后发现,虽然人岗推荐系统的研究已经初具规模,并且这些研究成果在解决数据稀疏性或数据冷启动问题上都有些许改善,但是其推荐效果还是不能令人满意,所以为了提高推荐系统的推荐效果,尝试将知识图谱作为额外的辅助边信息融入到推荐系统上进行实验。

2012年,Google提出知识图谱用于改善搜索体验^[14],提高搜索质量,引起了社会各界纷纷关注。比如清华大学建成了第一个大规模中英文跨语言知识图谱Xlore,上海交通大学构建并发布了中文知识图谱研究平台Zhishi.me,复旦大学GDM实验室推出了中文知识图谱项目等^[15]。受到知识图谱在多种领域上成功应用的启发,研究人员开始尝试利用知识图谱来提高推荐系统的推荐效果,且取得了良好的效果。

现有的基于知识图谱的推荐方法可以分为基于嵌入的方法(embedding-based methods)和基于路径的方法(path-based methods)两类^[16]。基于嵌入的应用知识图谱在推荐系统的方法主要是通过图嵌入的方法对实体和关系进行表征,进而扩充原有物品和用户表征的语义信息,其中包括基于Trans系列的图谱嵌入方法和基于异质信息网络的图嵌入方法。Trans系列的嵌入方法是比较典型的一类对知识图谱中的实体、关系进行表征的方法,主要包括TransE^[0]、TransH^[18]、TransR^[19]和TransD^[20]。知识图谱因其节点和边具有不同的类别,又可以被称做一种异质信息网络图,因此可以使用一些异质信息网络图嵌入的方法对图上的实体和关系进行表征。Wang等人^[21]从异质信息网络图中抽取和用户节点相连的N-hop实体节点,设计了基于RippleNet的方法,该方法利用这些实体节点的嵌入表征更新用户的表征,从而利用用户和物品表征的点积去预测推荐结果。本文也是基于此对RippleNet方法进行改进,应用于人岗推荐系统上。知识图谱在推荐系统上的应用往往能起到很好的辅助推荐效果,所以本文首次尝试在人岗推荐系统上结合知识图谱进行推荐,并取得了不错的效果。

2 人岗推荐系统构建

本文在传统人岗推荐系统上融合知识图谱进行构建,从而提高人岗推荐系统的推荐效果。本文知识图谱总体构建流程如图1所示,原始数据的格式可能是多样的,这些原始数据经过处理后,再用知识抽取的方法表示为三元组形式,然后进行知识图谱补全操作进一步完善所得三元组数据,最后构建知识图谱。其中重要的是对获取的数据进行知识抽取和知识图谱补全操作。

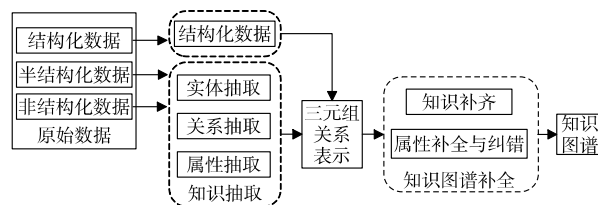


图1 知识图谱流程框架

Fig. 1 Knowledge graph process frame

2.1 图谱构建方案

2.1.1 知识图谱逻辑结构

知识图谱在逻辑上可分为模式层与数据层两个层次,数据层主要是由一系列的事实组成,而知识将以事实为单位进行存储。如果用(实体1,关系,实体2)(实体,属性,属性值)这样的三元组来表达事实,可选择图数据库作为存储介质,本文选用开源的Neo4j图数据库进行数据的存储,对于知识图谱的构建效果,图数据库能更加直观地展示图谱的结构。模式层是构建在数据层之上的,主要是通过本体库来规范数据层的一系列事实表达,通过本体库而形成的知识库不仅层次结构较强,并且冗余程度较小。

2.1.2 知识图谱架构体系

知识图谱构建主要分为自顶向下(top-down)与自底向上(bottom-up)两种构建方式。自顶向下的构建方式适用于知识内容比较明确,关系比较清晰的领域构建知识图谱。自底向上构建方式则是从开放的Open Linked Data中抽取置信度高的知识,或从非结构化文本中抽取知识,完成知识图谱的构建,该方式更适用于常识性的知识,比如人名、机构名等通用知识图谱的构建。然而在实际构建过程中,两种方法并不是孤立进行的,本文在构建人岗推荐系统知识图谱时采用两种方法的结合,首先采用自顶向下的方式构建本体库,然后采用自底向上的方式提取人员的简历信息扩展知识图谱。

2.1.3 数据获取与预处理

知识图谱构建的首要任务就是对数据的处理,本文基于公司人岗数据进行推荐,数据来源于公司招聘收集的个人信息,常见的人岗信息提交格式有PDF文件、JPG图片、Word文档等,对不同的数据进行OCR识别与文本信息提取等操作后,将得到的文本数据进一步进行自然语言处理(natural language processing, NLP)操作,获得标准化人岗数据,从而构建标准化人岗数据库,知识图谱的数据处理流程如图2所示。

对于非结构化数据,比如纯文本数据等,需要进行信息抽取(information extraction)^[22]技术来完成三元组的构建。首先通过对个人简历信息的实体抽取,获取到求职人员的一些基本信息,比如姓名、毕业院校、年龄、期望薪资、就职公司等信息;然后采用半监督算法 Bootstrapping 算法^[23]进行关系抽取,得出求职人员和不同实体之间的关联信息,比如张三—朋友—李四,李四—任职过—公司A,王五一居住于—上海等;同时也进行属性的抽取,比如抽取到张三—年龄—25,李四—学历—硕士,王五一工作经验—3年等信息。

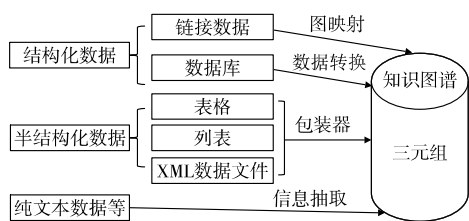


图2 知识图谱数据处理
Fig.2 Knowledge graph data processing

在人岗推荐系统中,将公司获取到的人岗信息经过信息抽取后,得到构建人岗知识图谱的人岗三元组数据,其格式类似于:(张三,年龄,25)(张三,学历,本科)(张三,精通,Python)。但是得到的数据并不全都是完善准确的,有时会出现信息抽取错误的情况,导致抽取到的三元组信息缺失或者错误,这时就需要进行知识图谱补全任务来完善图谱。

2.2 知识图谱补全

由于知识图谱存在不完整的问题,多数现有知识图谱都是稀疏的。由此引出了知识图谱补全(knowledge graph completion, KGC)来向知识图谱添加新的三元组。在初步获取到所需要的人岗三元组数据后,对于缺失的三元组数据,为了达到更好的数据处理效果,需要进行知识图谱补全操作,主要分为头实体缺失补全、关系缺失补全和尾实体缺失补全三部分操作,如图3所示。在补全过程中,需要对图谱中的头尾实体和关系进行知识表示,将这些信息进行嵌入表示后,就可以进行缺失信息的预测。如图4所示,可以将这种预测过程抽象成一个有向图的结构,其中实体表示节点,关系表示边,要实现的知识图谱补全操作其实就是给知识图谱中的不同节点寻找有向边的过程。

在图4中,一共包含 a, b, c, d, e, f 六个节点,补全的目标就是找到 a 和 f 中的对应关系,而且这个关系是 a 与 $\{b, c, d, e\}$ 对应关系中的一个。本文采用最经典的TransE^[17]模型进行知识表示,该模型的核心假设如下:

对于正确的三元组 (h, r, t) 而言,需要满足的是 $h + r = t$,也就是尾实体是头实体通过关系 r 的平移而得到的。通过TransE模型,可以将知识图谱中的关系和实体转换成编码的方式,通过式(1)得到缺失的关系编码表示,最终通过相似度比较的方式来确定实际的关系。对于关系预测问题,目前的方法

多是采用打分的机制,对于一个三元组,给出这个三元组可信的评分。在关系预测问题中,就是给定头节点和尾节点,在所有待选的关系集合中,选出评分最高的作为关系预测的结果。

$$t = h + r \quad (1)$$

通过知识图谱补全技术,可以更好地完善数据内容,使得数据更加健壮,训练效果更好,最终使得推荐的效果得以提高。

head entity prediction: $(?, r, t) \rightarrow (? , \text{熟悉}, \text{PHP})$
relation prediction: $(?, r, t) \rightarrow (\text{陈明}, ?, \text{PHP})$
tail entity prediction: $(?, r, t) \rightarrow (\text{陈明}, \text{熟悉}, ?)$

图3 知识图谱补全
Fig.3 Knowledge graph completion

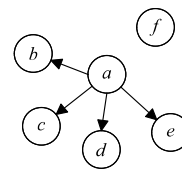


图4 知识图谱节点关系
Fig.4 Knowledge graph node relationship

2.3 人岗推荐算法

人岗推荐算法与常见的传统推荐算法不同的地方在于结合了知识图谱,本文的主要推荐算法基于RippleNet模型,通过改进算法模型将其应用到人岗推荐系统中,最后在人岗推荐上取得了不错的效果。如图5所示,为本文改进后的RippleNet算法框架主要构建架构。其输入为一个招聘需求信息 u 和一个候选的人员简历信息 v ,输出为招聘需求对应人员简历的可能性。其中 Rh 表示关系 R 和头节点 h 的嵌入表示, t 代表尾节点嵌入表示。对输入的招聘需求 u 来说,与求职人员交互的历史信息作为种子存储到图谱网络中,即先将所有和招聘需求有过关联的求职人员作为历史记录形成种子集合构建到图谱网络中,方便提取参照这些历史交互信息,而传播到一跳或者两跳,甚至多跳种子人员集合,就是根据不同的关联信息来划分的历史交互种子集合,类似于直接将招聘需求相关联的人员当做一跳人员集合,然后间接和招聘需求相关联的人员作为两跳关联人员集合。这些种子集合信息和人员简历信息的嵌入表示进行迭代交互,获取招聘需求信息与人员信息的反馈信息,将这些信息联合形成最终的招聘需求表示,最终通过招聘需求的嵌入表示和人员简历信息的嵌入表示共同预测出招聘需求对此简历信息的招聘可能性。

$$\hat{y}_{uv} = \sigma(u^T v) \quad (2)$$

可能性计算公式如式(2)所示,其中 $\sigma(x) = 1 / [1 + \exp(-x)]$ 是sigmoid激活函数。

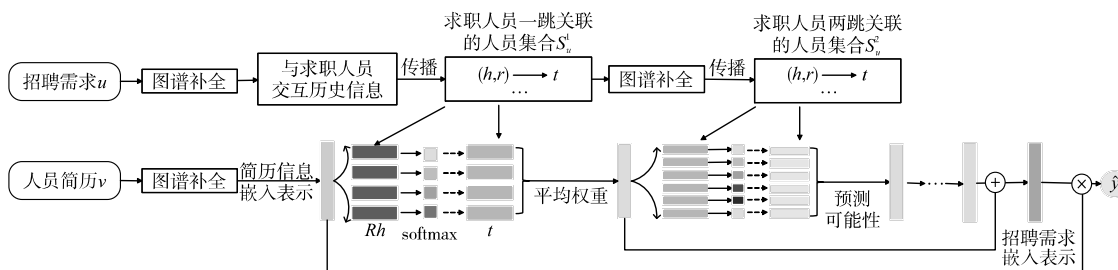


图5 算法框架
Fig.5 Algorithm framework

在进行公司需求和招聘人员简历信息的嵌入表示前,首先将存入知识图谱的不完整信息进行知识图谱补全,把抽取到的三元组信息通过图谱补全技术来完善数据内容,使得推荐算法取得更好的推荐效果。在下面的实验中对对比了图谱补全前后的简历推荐准确率,实验证明这种改进方式可以提高简历推荐效果。

2.4 人岗推荐系统图谱构建效果

基于知识图谱的人岗推荐系统图谱数据构建效果如下:

a) 简历人员信息的图谱构建。将人员的简历信息抽取解析后,抽取三元组信息存放到neo4j图数据库中,如图6所示

(出于隐私保护,图中未显示真实公司名和项目名)。

根据不同需求,建立人员主要的个人信息三元组信息,比如擅长的技术技能、任职过的公司、参与的项目经历、期望薪资水平等,将这些信息构建到图谱中,以便与应聘岗位要求进行关联。

b) 公司招聘需求信息图谱构建。将公司招聘信息、岗位要求抽取解析,以三元组的形式存放到图数据库中,如图7所示。

岗位要求信息主要就是招聘公司对于不同岗位的招聘要求,包括招聘岗位的技术要求、年龄要求、学历要求、技能熟练度要求等。

c) 岗位推荐结果的图谱构建。通过图谱推荐算法计算出招聘公司在某一岗位推荐的人员, 并将推荐结果存放到图数据库, 如图8所示。

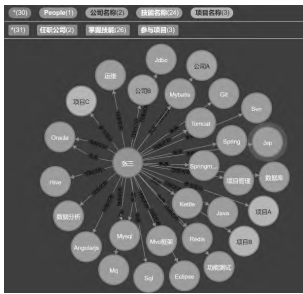


图6 简历人员图谱构建

Fig. 6 Construction of resume personnel knowledge graph

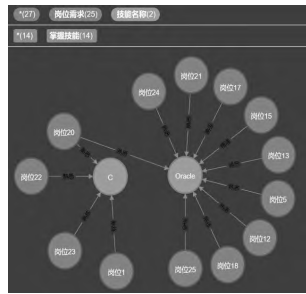


图7 岗位需求信息图谱构建

Fig. 7 Construction of job demand information knowledge graph

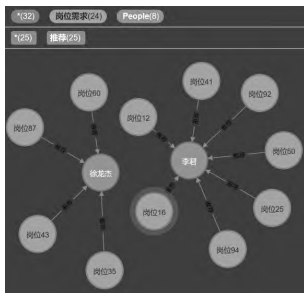


图8 岗位推荐结果

Fig. 8 Job recommended results

岗位推荐结果展示了与公司需求的岗位对应较匹配的人员, 过滤了大量和岗位要求不一致的求职人员, 从而提高了招聘效率。在进行人岗推荐系统构建的过程中, 对比传统人岗推荐, 本文方法的关键技术在于知识图谱的构建和知识图谱补全技术的应用。在人岗推荐算法上, 本文通过改进基于知识图谱的推荐算法, 结合知识图谱补全技术, 使得在人岗推荐上构建的知识图谱数据利用率更高、更完善。通过知识图谱的构建和知识图谱补全算法这两个关键技术的应用, 成功提升了传统人岗推荐系统的推荐效果, 下面的实验部分展示了实验的具体细节。

3 岗位推荐结果展示了公司需求的岗位对应较实验

本章通过在真实的公司求职数据场景中评估基于知识图谱的人岗推荐系统的实现效果。首先简单介绍了实验数据集的选取, 然后解释了实验评估指标的符号及含义, 最后呈现了实验结果。

3.1 实验数据集

本文采取的是真实的公司求职简历数据, 由于影响推荐准确率的因素主要包括数据因素和算法因素, 为了突出本文提出的融合知识图谱及知识图谱补全算法在人岗推荐系统中的有效性, 在进行对比不同算法时, 选取同一数据集, 将基于图谱的人岗推荐算法、图谱补全后的人岗推荐算法分别与传统人岗推荐算法进行对比, 并且训练、评估和测试集的比例均为 6:2:2, 每个实验重复三次, 取平均值进行对比。

3.2 实验评估指标

本次实验结果的评价标准是 AUC (area under curve) 和 ACC (accuracy)。其中 ROC (receiver operating characteristic) 曲线所覆盖的面积称为 AUC, AUC 越大则性能越好。准确率 ACC 如式 (3) 所示。

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (3)$$

其中: TP (true positive) 表示样本的真实类别为正, 最后预测结

果也为正; FP (false positive) 表示样本的真实类别为负, 最后预测得到的结果却为正; FN (false negative) 表示样本的真实类别为正, 最后预测得到的结果却为负; TN (true negative) 表示样本的真实类别为负, 最后预测得到的结果也为负。

3.3 实验结果

本次实验数据来自招聘公司提供的相关简历研究数据, 将这些数据进行数据清洗后, 抽取需要的简历三元组信息, 存放到 neo4j 图数据库中, 然后构建知识图谱, 经过基于知识图谱的推荐算法得出相应岗位推荐的人员, 在三种不同的算法下, 主要对比了有无知识图谱构建算法的推荐准确率, 除了模型训练算法不同, 其余对数据的处理方式和准确率评估方式均保持一致, 将这些个人简历信息对应招聘信息进行评分表示, 比如根据招聘要求, 将技能要求评为 (精通 4 分) (掌握 3 分) (熟练 2 分) (了解 1 分) (其他 0 分)。然后根据个人简历介绍, 对应进行个人标签的评分, 使得每个简历人员都有相应的技能分值。根据简历要求, 比如大于等于 3 分的进行推荐, 这样就可以过滤掉此项要求不符合的求职人员。同理在其他招聘要求项上均可进行此操作, 需要根据具体的招聘需求, 进行不同的评分表示, 比如招聘要求年龄在 35 岁以下, 此项分值为 0 或 1。最后满足每个招聘要求的求职人员, 将被保留下来, 并且保留下来的人员中, 也根据评分有相应的排名, 招聘公司可以根据分值排名取前多少名作为候选人。

如图 9 所示, 展示了传统人岗匹配算法的训练结果 AUC 和 ACC 曲线结果。图 10 是基于知识图谱的人岗推荐系统 AUC 和 ACC 结果, 实验表明经过传统的人岗匹配系统融合知识图谱后, 其推荐效果提高了 0.5% ~ 0.7%。图 11 是经过知识图谱补全算法后的 AUC 和 ACC 结果, 实验表明经过知识图谱补全后, 相比传统人岗匹配方式, 其推荐效果提高了约 1%。

表 1、2 列出了基于传统方式进行人岗推荐和基于图谱方式及图谱补全方式进行人岗推荐的实验 AUC 和 ACC 结果。实验共进行 20 轮训练, 虽然受训练集规模有限的影响, 导致训练提升效果有限, 但仍可证明在人岗推荐系统上结合基于知识图谱的推荐算法进行构建是可行的, 且有不少的提升效果。

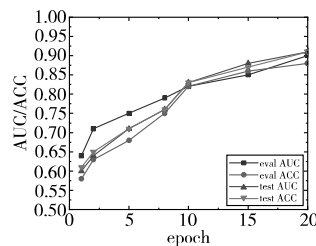


图9 训练 AUC 和 ACC 结果对比(传统方式)

Fig. 9 Comparison of AUC and ACC training results (traditional method)

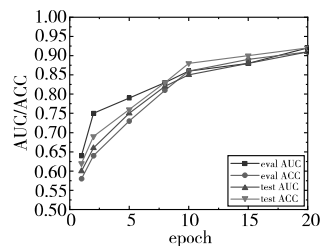


图10 训练 AUC 和 ACC 结果对比(基于图谱)

Fig. 10 Comparison of AUC and ACC training results (based graph)

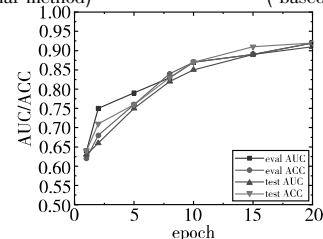


图11 训练 AUC 和 ACC 结果对比(图谱补全后)

Fig. 11 Comparison of AUC and ACC training results (after graph completion)

通过表 1、2 的实验数据可以看出, 在未进行基于知识图谱构建的情况下, 传统推荐算法在人岗推荐上的训练 AUC 值和 ACC 值均低于图谱补全后的人岗推荐算法。基于图谱构建效果的提升, 得益于知识图谱作为辅助信息, 大大降低了人岗推

荐模型训练过程中的数据稀疏和数据冷启动问题,知识图谱的应用使得人岗推荐系统中的个人简历数据与招聘需求信息联系紧密,所以推荐算法能够更加准确地进行两者之间的关联推荐。同时知识图谱补全算法的加入更是充分利用了简历数据,将原本残缺需要抛弃的数据内容补全后加入到图谱网络中,无疑增加了数据的健壮性,使得推荐算法能够获取更多的数据信息,从而更准确地进行模型的训练预测。

表 1 不同方式下训练 AUC 结果

Tab. 1 Training AUC results in different ways

方法	epoch			
	5	10	15	20
传统 RippleNet	0.713	0.834	0.875	0.911
基于图谱	0.754	0.854	0.883	0.909
图谱补全后	0.746	0.853	0.892	0.921

表 2 不同方式下训练 ACC 结果

Tab. 2 Training ACC results in different ways

方法	epoch			
	5	10	15	20
传统 RippleNet	0.711	0.826	0.874	0.906
基于图谱	0.764	0.882	0.903	0.921
图谱补全后	0.763	0.867	0.914	0.924

实验最终结果是在同一数据集下,每种构建方式进行三次实验后的平均结果。通过实验对比,传统的推荐方式构建人岗匹配系统平均准确率约为 91%,本文提出的基于图谱的构建方式在经过图谱补全后平均准确率约为 92%,提升了约 1%,说明人岗推荐系统是可以融合知识图谱来进一步提高人岗推荐效果的。

4 结束语

本文主要介绍了推荐系统、知识图谱和人岗推荐的相关背景知识,提出将知识图谱融合推荐系统进行人岗推荐的方法,并取得了不错的效果。目前在人岗推荐系统上的研究,传统的方法仍然存在很多问题,比如简历数据的稀疏性和简历人员历史数据缺失导致的冷启动问题等,同时由于数据的缺失,导致很多数据不能正常使用。而随着对知识图谱的研究发现,将知识图谱融入人岗推荐系统能够很好地解决传统人岗推荐方法的弊端,故本文尝试构建基于知识图谱的人岗推荐系统,并加入了图谱补全算法,用于提高数据利用率,通过实验得出此方法构建效果良好。

传统人岗推荐系统只是基于推荐方法来改进推荐效果,本文提供了融合知识图谱的思路来改进推荐效果。实验表明基于知识图谱的人岗推荐系统方案是可行的,将知识图谱应用到人岗推荐系统中确实可以取得较好的效果,接下来可以进一步研究如何提高人岗推荐系统的准确率,比如可以通过数据增强、知识图谱补全纠错等方式来优化系统推荐效果,同时表明基于知识图谱的人岗推荐系统的研究是值得探索的有前景的方向。

参考文献:

- [1] Zhang Fuzheng, Yuan N J, Lian Defu, et al. Collaborative knowledge base embedding for recommender systems [C]//Proc of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press 2016: 353-362.
- [2] Zhao Huan, Yao Quanming, Li Jianda, et al. Meta-graph based recommendation fusion over heterogeneous information networks [C]//Proc of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press 2017: 635-644.
- [3] Padhi A K, Mohanty A, Sahoo S. FindMoviez: a movie recommendation system [M]//Intelligent Systems. Berlin: Springer 2021: 49-57.
- [4] Sun Zhu, Guo Qing, Yang Jie, et al. Research commentary on recommendations with side information: a survey and research directions [J]. Electronic Commerce Research and Applications, 2019, 37: 100879.
- [5] Javed U, Shaikat K, Hameed I A, et al. A review of content-based and context-based recommendation systems [J]. International Journal of Emerging Technologies in Learning 2021, 16(3): 274-306.
- [6] Natarajan S, Vairavasundaram S, Natarajan S, et al. Resolving data sparsity and cold start problem in collaborative filtering recommender system using linked open data [J]. Expert Systems with Applications 2020, 149: 113248.
- [7] Paparrizos I, Cambazoglu B B, Gionis A. Machine learned job recommendation [C]//Proc of the 5th ACM Conference on Recommender Systems. New York: ACM Press 2011: 325-328.
- [8] Singh A, Rose C, Visweswariah K, et al. PROSPECT: a system for screening candidates for recruitment [C]//Proc of the 19th ACM International Conference on Information and Knowledge Management. New York: ACM Press 2010: 659-668.
- [9] 黄海江, 杨贵中. 基于本体的学习内容个性化推荐 [J]. 科学技术与工程 2007, 7(14): 3394-3398. (Huang Haijiang, Yang Guanzhong. Ontology-based personalized recommendation in E-learning [J]. Science Technology and Engineering 2007, 7(14): 3394-3398.)
- [10] 何峰. 基于文本分类的人才自动推荐系统 [D]. 苏州: 苏州大学, 2006. (He Feng. An automatic recommended system talents based on text classification [D]. Suzhou: Soochow University 2006.)
- [11] 郭军军. 个性化职位推荐系统研究与实现 [D]. 西安: 西安电子科技大学 2009. (Guo Junjun. Research and realization on personalized post recommender system [D]. Xi'an: Xidian University 2009.)
- [12] 路小瑞. 基于 Hadoop 平台的职位推荐系统的设计与实现 [D]. 上海: 上海交通大学 2014. (Lu Xiaorui. The design and implementation of job recommendation system based on Hadoop [D]. Shanghai: Shanghai Jiao Tong University 2014.)
- [13] 何俊. 基于社交网络的个性化推荐系统的设计与实现 [D]. 贵阳: 贵州大学 2015. (He Jun. Design and implementation of personalized recommendation system based on social network [D]. Guiyang: Guizhou University 2015.)
- [14] Singhal A. Introducing the knowledge graph: things, not strings [M]. [S.l.]: Official Google Blog 2012.
- [15] Cheng Xueqi, Jin Xiaolong, Wang Yuanzhuo, et al. Survey on big data system and analytic technology [J]. Journal of Software 2014, 25(9): 1889-1908.
- [16] Wang Hongwei, Zhang Fuzheng, Wang Jialin, et al. Exploring high-order user preference on the knowledge graph for recommender systems [J]. ACM Trans on Information Systems 2019, 37(3): 1-26.
- [17] Bordes A, Usunier N, Garcia-Duran A, et al. Translating embeddings for modeling multi-relational data [C]//Proc of the 26th International Conference on Neural Information Processing Systems. Red Hook, NY: Curran Associates Inc. 2013: 2787-2795.
- [18] Wang Zhen, Zhang Jianwen, Feng Jianlin, et al. Knowledge graph embedding by translating on hyperplanes [C]//Proc of the 28th AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2014: 1112-1119.
- [19] Lin Yankai, Liu Zhiyuan, Sun Maosong, et al. Learning entity and relation embeddings for knowledge graph completion [C]//Proc of the 29th AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press 2015: 2181-2187.
- [20] Ji Guoliang, He Shizhu, Xu Liheng, et al. Knowledge graph embedding via dynamic mapping matrix [C]//Proc of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. Stroudsburg, PA: Association for Computational Linguistics 2015: 687-696.
- [21] Wang Hongwei, Zhang Fuzheng, Wang Jialin, et al. RippleNet: propagating user preferences on the knowledge graph for recommender systems [C]//Proc of the 27th ACM International Conference on Information and Knowledge Management. 2018: 417-426.
- [22] Zong C, Xia R, Zhang J. Information extraction [M]//Text Data Mining. Berlin: Springer 2021: 227-283.
- [23] Zhang Chunyun. The study of entity relation extraction algorithm [D]. Beijing: Beijing University of Posts and Telecommunications 2015.