[文章编号] 1003－4684(2026)02-0078-07

基于图神经网络的工作名称推理

陈飞飏[1]，钱 奕[2]

(1 香港中文大学(深圳)，广东 深圳 518172；2 湖北交通职业技术学院，湖北 武汉 430073)

[摘 要] 工作名称的准确性对工作推荐系统的效果具有重要影响，能够使系统向用户推荐合适的工作岗位。但是，系统中存在一些较模糊的工作名称，缺乏具体的职业信息。这些模糊的工作名称可能会导致系统的推荐效果变差，甚至引起用户的抱怨。本文提出了一种基于图神经网络的工作名称推理方法(GNN-JT)。它能够根据工作属性信息和工作属性分类表信息将原本模糊的工作名称转化为准确的名称。实验结果表明，GNN-JT 能够有效推断出准确的工作名称，从而提高工作推荐的效果。

[关键词] 图神经网络；深度学习；自然语言处理；推荐系统

[中图分类号] TP183 [文献标识码] A

工作推荐系统(job recommendation system，简称 JRS)[1][2]被广泛应用于许多招聘网站(如“Boss 直聘”“招聘网”等)。JRS 根据工作信息和用户背景，向用户推荐适合的工作机会[3]，帮助用户找到理想的职位。每天有数千万的用户登录招聘网站寻找工作，JRS 的推荐效果直接影响用户求职体验以及招聘网站的营收。作为重要的工作信息，工作名称的准确性对 JRS 的效果有重要影响。准确的工作名称可以帮助 JRS 理解工作性质，从而向具备相关工作背景的用户推荐。但是，工作名称通常由招聘人员手工输入，由于语言表达的灵活性，同一个职位可能会被描述为不同的工作名称。例如，“软件工程师”可能被描述为“软件开发者”“软件程序员”或“软件设计师”等。工作名称的表达灵活性对 JRS 来说是一个挑战，增加了 JRS 机器学习模型训练的难度。目前，工业界主流做法[4]是建立工作名称分类表，并使用该分类表中的标准名称来统一标识不同的工作名称。工作名称分类表包含不同职业的标准工作名称，并使用唯一的 ID 进行标识。JRS 会将招聘人员输入的原始工作名称映射到标准工作名称 ID 上，然后使用 ID 来表示工作名称，进行机器学习建模。这种将原始工作名称映射到工作名称分类表中 ID 的过程被称为工作名称标准化。

工作名称标准化提高了 JRS 机器学习模型的学习效果和效率，使得 JRS 的推荐相关性得到了很大改善，提升了用户的体验。但是，工作名称标准化也带来了新的问题。大量招聘人员在发布新工作时使用一些模糊的工作名称来描述职位。“模糊”是指用户无法通过工作名称的描述了解具体职业信息，如“工程师”“经理”和“专员”等。通过对某招聘网站上工作的随机采样分析，发现大约有 10%～15% 的工作具有模糊的工作名称。这些模糊的工作名称无法进行标准化处理，严重影响了推荐系统的效果，甚至引起用户的抱怨和投诉。例如，“前端工程师”这个工作由于工作名称为“工程师”，很有可能会被推荐给职业背景为“后端工程师”的用户。因此，如何对模糊的工作名称进行标准化是当前 JRS 迫切需要解决的问题。本文提出了一种基于图神经网络的工作名称推理方法(GNN-JT)，旨在解决模糊工作名称的标准化问题。该方法利用图神经网络对目标工作的属性信息(如工作所需技能和工作领域)以及工作属性的分类表信息进行编码，生成目标工作的表征，然后通过多元分类的方法推理出准确的工作名称。相比传统的机器学习方法，GNN-JT 的图神经网络能利用工作和工作属性之间高维联系信息来生成更好的工作表征。实验结果表明，GNN-JT 能有效地将模糊的工作名称转化为准确的名称，从而使得 JRS 的推荐效果得到很大提高。

1 相关工作

1.1 工作推荐系统和工作信息标准化

工作推荐系统(JRS)是属于推荐系统的一

[收稿日期] 2023－10－31

[第一作者] 陈飞飏(2001－)，男，湖北武汉人，香港中文大学(深圳)本科生，研究方向为高性能机器学习。

[通信作者] 钱 奕(1973－)，男，湖北武汉人，湖北交通职业技术学院讲师，研究方向为计算机应用。

种[5]。本质上,推荐系统利用人工智能和机器学习技术为用户提供个性化推荐。它通过分析用户的历史行为、兴趣爱好和个人特征,从大量的信息中筛选出用户可能感兴趣的内容,向其提供精准的推荐建议。目前,深度学习在推荐系统中的应用越来越广泛[15],基于规则或传统机器学习模型逐渐被基于神经网络的深度学习模型所取代。深度学习模型能够从大量的数据中学习到更复杂、更准确的用户兴趣和行为模式。

作为招聘网站的重要功能,JRS 从大量新工作中选择出合适的工作机会,并推荐给相关用户。JRS 的核心是个性化的工作推荐模型。它能根据用户的信息(如教育背景、职业、技能、地点等)和工作的信息(如工作名称、所需技能、地点等)进行机器学习建模。生成的模型能对工作进行个性化排序,从中选择适合的工作提供给不同用户。

工作名称是工作推荐模型中的重要特征。然而,直接使用工作名称的原始文本作为模型特征会面临数据稀疏问题。这是因为招聘人员在发布工作时通常使用很多不同的文本来描述相同的职位,例如,人工输入的工作名称"市场销售""市场销售代表""市场销售专员"实际上都表示同一类工作。工作名称的文本多样性对工作推荐模型提出了挑战。本质上,这是一个高维数据稀疏性问题。直接使用工作名称文本作为特征输入需要大量数据进行模型训练,以避免过拟合问题。

为了应对这个问题,目前工业界 JRS 的主流方法是对工作名称进行标准化处理。其主要思想是将不同的工作名称文本映射到分类表中的唯一标识符 ID 上。标准化处理包含以下三个步骤:

1)根据市场上的职业建立工作名称分类表,为每个职业分配一个标准文本名称和对应的标识符 ID。例如,工作标准名称"软件工程师"对应的标识符 ID"210"。

2)根据建立的工作名称分类表,JRS 对原始输入的工作名称文本进行标准化处理,例如,将原始工作名称"软件开发者"标准化为分类表中的条目"软件工程师(ID:210)"。标准化的过程可以采用基于规则或者机器学习模型的方法进行处理(具体的标准化方法不是本文讨论的重点。有兴趣的读者请参考文献[6-8])。

3)工作推荐模型使用标准化 ID 进行机器学习模型的训练和部署。工作名称的标准化过程实际上是一个特征泛化的过程,能有效解决工作名称的数据稀疏问题。

除了工作名称的标准化,还可以对工作的其他属性进行类似的处理,例如,所需的技能,招聘企业和工作领域等。类似地,需要建立不同的属性分类表。

1.2　图神经网络

许多机器学习任务需要处理和分析图数据。图神经网络(graph neural network,GNN)[9-13]作为一个研究图数据的深度学习方法逐渐受到重视。GNN 的核心思想是将节点和边的特征信息进行传递和聚合。它通过迭代地更新节点的表征(表示向量),使每个节点能够融合周围节点的信息,并将这些信息传递到下一层。具体来说,GNN 的每一层都有两个主要的步骤:信息传递和节点更新。信息传递通过节点之间的连接关系,将节点的特征信息传递给相邻节点。节点更新则利用传递过来的信息和当前节点的特征,更新节点的表示向量。这样,通过多层的迭代更新,GNN 可以逐渐融合和整合全局的图结构信息(例如,图的节点特征和边特征),最终生成图中节点的表征,

图 1 显示了 GNN 的应用示例。目标任务通过节点 $\{A,B,\cdots,F\}$ 和节点之间的边生成图。节点一般是目标任务中的实体,边是实体之间的关系。例如,在社交网络分析中,节点表示人,而节点之间的边表示人之间的关系。各个节点的初始特征在经过 GNN 运算以后,得到新的特征(即嵌入向量)被应用到目标任务中,如社交关系预测。

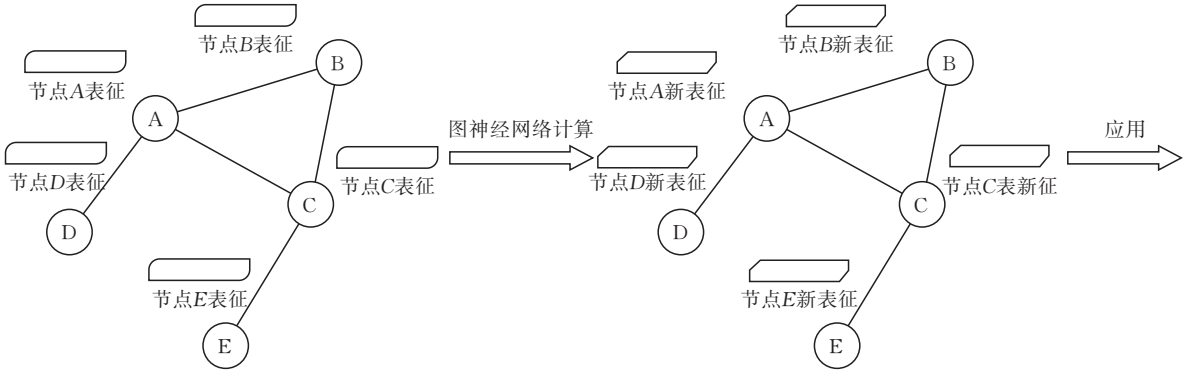

图 1　GNN 应用示例

GNN 的计算过程可以被认为是神经信息传递过程[14]，即图 1 中各个节点之间通过神经网络互相交换和更新向量信息。定义图为 $G=(V,E)$，这里 V 和 E 分别表示节点和边。图中允许自身闭环存在，即 $(v,v)\in E$。定义节点的特征矩阵为 $X\in\mathbb{R}^{m\times|V|}$，其中每一列 $x_v\in\mathbb{R}^m(v\in V)$ 表示一个节点的特征。通过 GNN 生成每个节点的嵌入 n_μ，$\forall\mu\in V$。

GNN 在每一次信息传递过程中，图中每个节点 $\mu(\mu\in V)$ 的隐藏嵌入 $h_\mu^{(k)}$ 会根据图中节点 μ 的所有邻居 $N(\mu)$ 的聚合来进行更新。GNN 的第 k 次信息传递可以定义如下：

$$h_u^{(k-1)}=\mathrm{UPDATE}^{(k)}(h_u^{(k-1)},\mathrm{AGGREGATE}^{(k-1)}\left(\left\{\begin{matrix}h_v^{(k-1)},\\ \forall v\in N(\mu)\end{matrix}\right\}\right))=\mathrm{UPDATE}^{(k)}(h_u^{(k-1)},m_{N(\mu)}^{(k)}) \tag{1}$$

式中，UPDATE 和 AGGREGATE 代表任意可微分的函数。$m_{N(\mu)}$ 表示从节点 μ 的所有邻居 $N(\mu)$ 聚合得到的信息。聚合过程(AGGREGATE)由图 2 所示。更新过程(UPDATE)会使用当前邻居节点的聚合消息和节点上一次的嵌入来进行更新。经过 K 次信息传递后，会得到每个节点的最终的嵌入(也可被理解为神经网络最后一层的输出)。

$$n_\mu=h_u^{(k)},\forall u\in V \tag{2}$$

经过第一次消息传递($K=1$)，图中每个节点的嵌入包含了它一阶(1-hop)相邻节点(即图中一步能到达的相邻节点)的特征；经过第二次消息传递($K=2$)，每个节点的嵌入包含它二阶相邻节点的特征；类似，经过 K 次消息传递以后，每个节点的嵌入包含它 K 阶相邻节点的特征。传递的消息一般包含两方面关于图的信息：1)图的结构信息。经过 K 次消息传递以后，$h_u^{(k)}$ 包含了节点 u 的所有 K 阶相邻节点的度信息，这些节点的度信息能表征图的结构；2)图节点的特征信息。$h_u^{(k)}$ 包含了节点 u 的所有 K 阶相邻节点的特征信息。

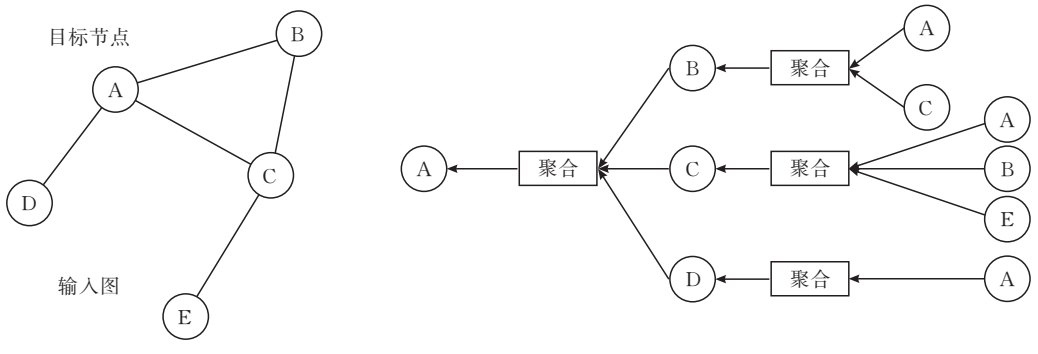

图 2 图神经网络中节点信息聚合过程(节点 A 是目标节点)

2 基于图神经网络的工作名称推理

在这一章将介绍如何通过图神经网络使用工作的属性，属性分类表以及它们之间的关系构建模型，训练得到目标工作的表征。将利用目标工作的表征对模糊工作名称进行推理，得到准确的工作名称推理。

2.1 问题定义

首先需要明确模糊工作名称的推理任务。其相关定义如下。

定义 1：模糊工作名称是指一个招聘启事的工作名称只包含工作的角色信息(例如，“工程师”“经理”“专员”等)，而不包含工作的具体职业信息(例如，“软件开发”“市场营销” 等)。模糊工作名称无法通过自身信息进行标准化处理。

定义 2：确定工作名称是指一个工作名称既包含工作的角色信息也包含工作的具体职业信息。工作名称分类表中每个条目的标准名称都是确定工作名称。

定义 3：工作名称推理任务是指给定一个工作描述 d 和对应的模糊工作名称 v，在 V 大小的工作名称分类表 M 中找到最合适的确定工作名称 s。工作名称推理任务的目标函数定义如下，

$$\underset{s\in M,|M|=V}{\operatorname{argmax}}\,p(s\mid v,d) \tag{3}$$

2.2 基于图神经网络的工作名称推理

根据上述工作名称推理任务的定义，提出了基于图神经网络的工作名称推理方法 GNN-JT。其具体工作流程如下，

针对模糊工作名称推理任务，GNN-JT 首先建立了工作属性分类表(taxonomy)。根据工作的不同属性，GNN-JT 建立了三张工作属性分类表：

1)工作名称分类表：该表包含 869 个标准工作名称代表职业，例如，“软件工程师”“市场营销专员”等。每个标准的工作名称由唯一 ID 标识。一个标准的工作名称由三部分组成：角色，职业方向和职业级别。角色是指工作的一般职称，如“工程师”“专员”和“经理”等；职业方向代表工作的具体方向和内

容,如“软件开发”“市场营销”等;职业级别表示工作所需要的经验水平,如“高级”“中级”“初级”等;

2)工作技能分类表:该表包含 4678 个常用工作技能,如“Java”“软件开发”“机器学习”等。同理,每个技能有标准名称和唯一 ID 标识;

3)工作领域分类表:该表包含了工作市场上常见的 125 个工作领域,如“通讯”“互联网”“市场营销”等。同理,每个工作领域有标准名称和唯一 ID 标识。

值得注意的是,不同分类表之间的条目可以通过某种关系进行相连。例如,工作名称分类表可以与工作所需技能表通过关系“职业特殊技能”相连。例如,工作名称“机器学习工程师,ID:240”可以与工作技能“机器学习,ID:55” 和“深度学习,ID:70”相连,因为“机器学习”和“深度学习”是机器学习工程师的职业特殊技能。表 1 列出了不同分类表条目之间的关系。

表 1　不同分类表条目之间的连接关系

原分类表	目标分类表	条目连接关系	示例
工作名称	工作技能	职业特殊技能	“前端工程师”具有技能“Javascript”
工作名称	工作领域	领域职业	“金融”领域的职业“银行专员”
工作领域	工作技能	领域特殊技能	“金融”领域的特殊技能“财务分析”

GNN-JT 将模糊工作名称的推理问题视为一个多元分类问题。它训练了一个端到端的基于图神经网络的深度多元分类器 C 。C 使用的损失函数是交叉熵损失函数。C 的输入是工作信息(包括,工作属性等)和其模糊工作名称,C 的输出是模糊工作名称在工作名称分类表中每个标准名称上的概率分布。GNN-JT 选择最大概率的标准名称作为模糊工作名称的推理名称。GNN-JT 模型的架构如图 3 所示。

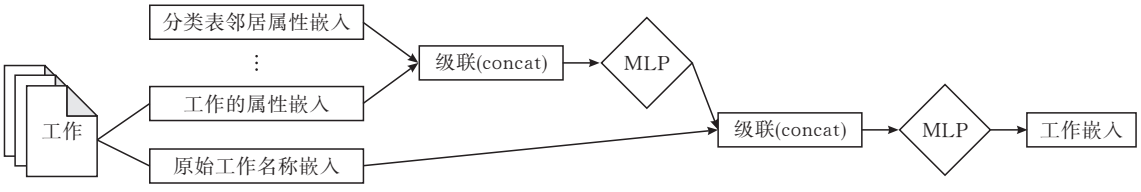

图 3　GNN-JT 模型架构

GNN-JT 模型的计算图 G 包含四类节点:表示工作的节点 V_j ,V_j 的原始工作名称节点 V_t ,V_j 的属性节点集 V_a (V_j 的工作所需技能节点和 V_j 的工作领域节点)和 V_a 在属性分类表中的邻居节点 V_n 。图 G 中的的边 E 是不同节点之间的连接边,即 V_j 与 V_t 和 V_a 连接边 ;V_a 和 V_n 的连接边。后者的连接边来自于不同分类表中条目的关系连接,如表 1 所示。

GNN-JT 模型使用的特征包括:1)输入工作的原始工作名称的词嵌入,其通过 Word2Vec[16] 算法得到; 2)工作属性的嵌入(即工作所需技能和工作领域); 3)工作属性的分类表邻居属性的嵌入。

GNN-JT 通过图结构来计算得到各种不同的嵌入,再对各种嵌入进行级联(concat)和非线性转换(前馈神经网络 MLP),最终得到目标工作的表征。该表征最终通过一个 $softmax$ 函数输出在工作名称分类表中各个条目的概率值。上述嵌入(如工作属性的嵌入和分类表邻居属性嵌入)都是 GNN-JT 通过图神经网络进行训练得到。

本质上,GNN-JT 是一个端到端深度学习模型。它利用图神经网络作为编码器得到输入工作的表征嵌入,再利用该嵌入实现一个多元分类模型。

对于 GNN-JT 模型的训练数据,一共从某招聘网站采样了 1,200,000 个工作作为训练数据。每个工作都有确定的工作名称和相应的工作属性信息。GNN-JT 的正例训练集合 P 是 1,200,000 包含(工作,工作确定名称 ID)训练记录,其负例训练集合 N 的生成则采用随机采样的方式,例如,给定一个正向记录(工作,工作确定名称 A),会在分类表中随机选择另外一个工作确定名称 B 作为负向记录(工作,工作确定名称 B)。将正向训练数据集的记录数目和负向训练数据集的记录数目大致保持在 1 比 1 的比例。

得到正负例训练数据集后,GNN-JT 模型的图嵌入生成算法如下所示。

GNN-JT 图嵌入生成算法

输入:图 $G(V,E)$; 输入特征 $\{x_v, \forall v \in V\}$; 深度 K ; 权重矩阵 $W^k, \forall k \in \{1,\cdots,K\}$; 非线性方程 σ ,聚合方程 $AGGREGATE_k, \forall k \in \{1,\cdots,K\}$;

输出:图节点嵌入表征 z_v , $\forall v \in V$, 目标工作嵌入表征 j 。

步骤 1: 初始化

$h_v^0 \leftarrow x_v, \forall v \in V$

步骤 2:图节点嵌入迭代计算

for $k = 1 \cdots K$ do

for $v \in V$ do

$h_{N(v)}^{k} \leftarrow AGGREGATE_k(\{h_u^{k-1}, \forall u \in N(v)\})$

$h_v^k \leftarrow \sigma(W^k \cdot CONCAT(h_v^{k-1}, h_{N(v)}^k))$

end

$h_v^k \leftarrow \frac{h_v^k}{\|h_v^k\|_2}$

end

步骤 3:得到图节点嵌入输出

$z_v \leftarrow h_v^K, \forall v \in V$

步骤 4:得到目标工作嵌入

$j \leftarrow \sigma(concat(z_{v\text{-}rawtitle}; \sigma(concat(z_{v\text{-}attribute}; z_{v\text{-}taxonomy}))))$

GNN-JT 模型得到目标工作的嵌入表征 j 后,将其输入 $softmax$ 函数,选择最大的概率输出,最终得到的目标工作的确定工作名称。

3 实验及结果分析

在这一章将介绍 GNN-JT 模型的对比实验结果,并对实验结果进行分析。

3.1 评价数据收集

为了更准确地评价 GNN-JT 的模型效果,在某工作招聘网站上随机采集并人工标注了 2000 个具有模糊工作名称的工作作为工作名称推理任务的评价数据集。这 2000 个工作中,每个工作都包含了原始模糊工作名称,人工标注的确定工作名称,标准化的工作属性等。

3.2 算法比较

实现了一个基于深度学习的基本多元分类器来与 GNN-JT 模型进行对比分析。两者均采用相同的正负例训练数据集。该多元分类器具有 869 个类别(即工作名称的分类表条目的数量)。分类器输入一个工作及其对应模糊的工作名称,输出是确定的工作名称。

3.2.1 基本多元分类器 基本多元分类器(baseline)使用了前馈神经网络作为多元分类器的分类模型。该模型使用的特征包括:工作的名称,工作所需技能和工作领域。工作的原始输入名称通过 Word2Vec 得到嵌入,工作所需技能、工作领域等都使用分类表中的 ID 来表示。

3.2.2 GNN-JT 模型 在表 2 中,列出了 GNN-JT 的几种模型参数和特征设置,来分析不同模型参数和特征对模型表现的影响。GNN-JT 模型的一个重要参数是图神经网络的深度 K 。在图神经网络中,使用不同的网络深度 K 做节点的嵌入计算的效果不一样。较小 K 值使 GNN-JT 模型主要聚合计算图中目标节点的邻近节点的特征,而较大 K 值使得 GNN-JT 能聚合计算图中离目标节点更远的节点的特征。但是,K 值取得过大会造成图中各个节点的嵌入趋向一致。GNN-J 模型 T 选取了 $K=1$ 和 $K=2$ 两个深度参数值进行实验。

除了图神经网络的深度 K 值以外,还想分析不同的特征对 GNN-JT 模型的影响。所以,选取了不同特征的组合作为 GNN-JT 模型的输入,例如,只使用工作名称的特征,使用工作名称和工作属性的组合;或者使用工作名称,工作属性和工作属性分类表特征组合。

基本多元分类器和 GNN-JT 模型的各种变体在评价数据集上的实验结果如表 2 所示,例如,GNN-JT 模型的变体"GNN-JT-K2-Title+JobAttribute"是指 GNN-JT 模型选择网络深度 K 值为 2,并选择"工作名称"和"工作属性"作为特征组合。使用准确率(accuracy)作为各个模型效果的评价标准,它表示给定 2000 个具有模糊工作名称的工作,通过模型得到正确的精确工作名称的工作的比例。

表 2 基本模型和 GNN-JT 模型变体的实验结果

模型	准确率
Baseline	0.852
GNN-JT-K1-Full(Title+JobAttribute+TaxonomyGNN)	0.883
GNN-JT-K2-Title	0.874
GNN-JT-K2-Title+JobAttribute	0.892
GNN-JT-K2-Full (Title+JobAtrribute+TaxonomyGNN)	0.918

从表 2 的实验结果可以看出,GNN-JT 模型的变体"GNN-JT-K2-Full",即使用了所有特征且图深度 K 为 2 的模型效果最好,这也表明了模型特征的有效性以及图神经网络的优越性。

除了以上各个模型的离线实验,还对模糊工作名称推理对 JRS 的推荐效果影响进行了在线 A/B 测试。对某招聘工作网站的模糊工作名称进行了标准化处理,并进行了在线 A/B 测试。其实验效果如表 3 所示。这里,在线实验指标"平均页面访问量"指的是在 JRS 中,用户每次平均的工作页面访问数量;"负反馈率"指的是用户在 JRS 推荐的工作列表中直接表示推荐的工作不合适的的比例。负反馈率越高表示用户的抱怨越多。

表 3　A/B 测试中页面访问量和负反馈率的结果

模型	平均工作页面访问量	负反馈率
原 JRS	4.1	0.005
进行模糊工作名称推理的 JRS	4.5	0.004

从表 3 实验结果可以看出，经过模糊工作名称推理的 JRS 与原 JRS 相比，其推荐效果有了明显的提高，具体表现在用户查看工作页面数量增加并且用户的抱怨降低。

图 4 显示了在 A/B 测试中，原 JRS 和进行模糊工作名称推理的 JRS(即图中“改进的 JRS”)对工作点击率的影响。“CRT1”“CRT5”和“CRT10”分别指的是 JSR 推荐的工作列表中，位置 1、位置 5 和位置 10 的工作点击率。从图中可以看出，改进的 JRS 推荐的工作点击率上升，表示推荐的工作相关性变好。

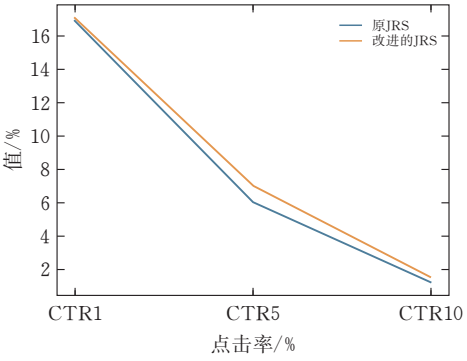

图 4　A/B 测试中工作点击率的结果

从以上实验结果可以看出，GNN-JT 模型通过对模糊工作名称推理，产生的工作推荐效果强于基本推荐系统模型。认为 GNN-JT 模型效果较好的原因主要有以下三点：第一，GNN-JT 模型能更好地利用计算图中所蕴含的高阶联系信号，生成较好的工作表征嵌入；第二，GNN-JT 模型通过图神经网络利用丰富的工作属性分类表信息；第三，GNN-JT 模型能同时训练和利用深度 NLP 模型和图神经网络模型。为了进一步提高 GNN-JT 模型表现，可以收集更多训练数据以及建立更完善的工作属性分类表。

4　结论

工作名称的准确性对工作推荐系统的推荐效果影响很大。解决模糊的工作名称的准确性问题是工作推荐系统需要迫切解决的问题。本文提出了一种基于图神经网络的模糊工作名称推理方法 GNN-JT。该方法利用图神经网络和深度 NLP 模型，结合工作的属性信息和分类表信息生成准确的工作表示，从而进行模糊工作名称的推理。实验证明了 GNN-JT 模型的有效性，并显示模糊工作名称推理对工作推荐系统的推荐效果有显著提高作用。

[参 考 文 献]

[1] AL-OTAIBI S T, YKHLEF M. A survey of job recommender systems[J]. International Journal of the Physical Sciences, 2012, 7(29): 5127-5142.

[2] ZHENG S, HONG W, ZHANG N, et al. Job recommender systems: A survey[C]// Proceedings of the 7th International Conference on Computer Science & Education (ICCSE). IEEE, 2012: 920-924.

[3] YAN R, LE R, SONG Y, et al. Interview choice reveals your preference on the market: To improve job-resume matching through profiling memories[C]// Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2019.

[4] LI S, SHI B, YANG J, et al. Deep job understanding at LinkedIn[C]// Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2020: 2145-2148.

[5] 于蒙，何文涛，周绪川，等. 推荐系统综述[J]. 计算机应用，2022，42(06)：1898-1913.

[6] SHI B, YANG J, GUO F, et al. Salience and market-aware skill extraction for job targeting[C]// Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2020: 2871-2879.

[7] ZHANG M, JENSEN K, SONNIKS S, et al. SkillSpan: Hard and soft skill extraction from English job postings[C]// Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2022: 4962-4984.

[8] KHAOUJA I, KASSOU I, GHOGHO M. A survey on skill identification from online job ads[J]. IEEE Access, 2021, 9: 118134-118153.

[9] HAMILTON W. Graph representation learning[M]. San Rafael: Morgan and Claypool Publishers, 2020.

[10] HAMILTON W, YING Z, LESKOVEC J. Inductive representation learning on large graphs[C]// Advances in Neural Information Processing Systems. 2017, 30.

[11] KIPF T N, WELLING M. Semi-supervised classification with graph convolutional networks[J/OL]. arXiv, 2016. https://doi.org/10.48550/arXiv.1609.02907.

[12] 徐冰冰，岑科廷，黄俊杰，等. 图卷积神经网络综述[J]. 计算机学报，2020，43(05)：755-780.

[13] 吴国栋，查志康，涂立静，等. 图神经网络推荐研究进展[J]. 智能系统学报，2020，15(01)：11.

[14] GILMER J, SCHOENHOLZ S S, RILEY P F, et al. Neural message passing for quantum chemistry[C]// Proceedings of the International Conference on Ma-

chine Learning. PMLR, 2017: 1263-1272.
[15] 黄立威, 江碧涛, 吕守业, 等. 基于深度学习的推荐系统研究综述[J]. 计算机学报, 2018, 41(07): 1619-1647.
[16] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space [J/OL]. arXiv, 2013. https://doi.org/10.48550/arXiv.1301.3781.

A Job Title Inference Approach (GNN-JT) Based on Graph Neural Network

CHEN Feibiao[1], QIAN Yi[2]
(1 *The Chinese University of Hong Kong, Shenzhen, Shenzhen* 518172, *China*;
2 *Hubei Communications Technical College, Wuhan* 430073, *China*)

Abstract: The precision of job title has great impact on the results of job recommendation system. A precise job title enables the system to recommend relevant jobs to users. However, some job titles are vague in the system, which do not contain specific occupation information. These vague job titles will lead to the suboptimal recommendation results, even cause user complaints. The paper proposed a job title inference approach (GNN JT) based on graph neural network. It can infer a precise job title from the original vague one by leveraging job attribute information and job attribution taxonomies. The experiments show that GNN JT can infer precise job title effectively and then improve job recommendation relevance.

Keywords: graph neutral network; deep learning; natural language processing; recommendation system

[责任编校: 闫 品]

(上接第 23 页)

Research on Fast Equalization Strategy for Series Lithium-ion Battery Packs Based on Fuzzy Logic Control

WU Tiezhou[1], ZHU Lei[1], ZHANG Ming[2], WANG Yueyang[2], LI Yi'an[1]
(1 *Hubei Key Laboratory of Solar Energy Efficient Utilization and Energy Storage Operation Control, Wuhan* 430068, *China*;
2 *Hangyu Weichuang Technology (Beijing) Co., Ltd., Beijing* 100097, *China*)

Abstract: To effectively enhance the balancing speed of series-connected lithium-ion battery packs, this paper proposes a two-stage balancing topology. Internally, the battery pack employs a Cuk circuit combined with a dual layer selector switch balancing topology, while inter pack balancing utilizes a Buck Boost circuit. By adopting SOC as the balancing variable and designing a Fuzzy Logic Control algorithm, dynamic adjustment of the balancing current is achieved, thereby significantly reducing the required balancing time. A model was established and simulated using a 9-cell series-connected lithium-ion battery as an example. Results demonstrate that under identical idle and charge/discharge conditions, the proposed method reduces idle balancing time by approximately 13%, charge balancing time by about 9%, and discharge balancing time by roughly 11% compared to the maximum value balancing method, validating the feasibility of this balancing scheme.

Keywords: cuk circuit; buck boost circuit; SOC; fuzzy logic control

[责任编校: 赵 瞳]