

## boss: 探索用于在线职位推荐的图形数据理解的大型语言模型



SmartMindAI

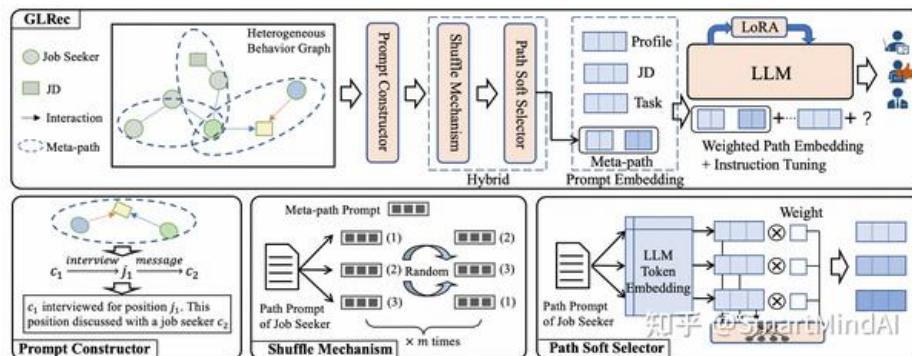
专注搜索、广告、推荐、大模型和人工智能最新技术，欢迎关注我

已关注

3 人赞同了该文章

### Introduction

在线招聘推荐旨在基于求职者的偏好和能力，推荐相关的工作机会。为了解决互动数据稀疏的问题，我们提出了一种元路径+提示构造器，将图的交互信息编码到自然语言提示中，通过大语言模型+（LLM）直接生成职位推荐。在本文中，我们首次将微调的大语言模型+作为职位推荐器，通过LLM的语义丰富性和大量知识提高匹配准确性。同事我们提出了一种元路径提示构造器，利用LLM推荐器理解行为图，并设计相应的路径增强模块以缓解提示偏差。



### Methodology

#### Problem Formulation

考虑候选集合  $C = \{c_1, c_2, \dots, c_{n_1}\}$  和工作集合  $J = \{j_1, j_2, \dots, j_{n_2}\}$ ，每个候选人和工作都有简历和工作要求的文本文档，并与招聘平台+互动记录相关联。互动表示为  $A_{c_i}$  和  $A_{j_k}$ ，我们的目标是预测职位与候选人的匹配度。

#### Generative Large Language Models

生成型大语言模型例如GPT-3和GPT-4能生成连贯的上下文相关文本。微调策略如提示调优和指令调优能够使模型适应特定领域。指令调优通过设计指令数据来限制输出，例如，指令输入为“喜不喜欢这个物品？”，指令输出为“喜欢”。最终损失公式如下：

$$\mathcal{L}_f = \max_{\theta} \sum_{(x,y) \in \mathcal{T}} \sum_{t=1}^{|y|} \log(\mathcal{P}_{\theta}(y_t | x, y_{<t}))$$

#### Task-specific Instruction

在本文中，我们设计了两个任务来测试LLM推荐的表现。假设有一个求职者，我们将其命名为candidate，其Candidate Profile Prompt和推荐JD Prompt如下所示：

**JD Prompt:** 职位名称: 全栈工程师<sup>+</sup>, 教育要求: 本科学历, 工作经历: 1-3年, 技能要求: HTML/JAVA/Spring Boot/SQL.

对于point-wise任务, 我们让LLM推荐器预测求职者对推荐职位的满意度, 具体指令如下:

**point-wise指令:** 你是一个推荐器, 决定求职者是否会对推荐的职位感到满意。请回答“是”或“否”。

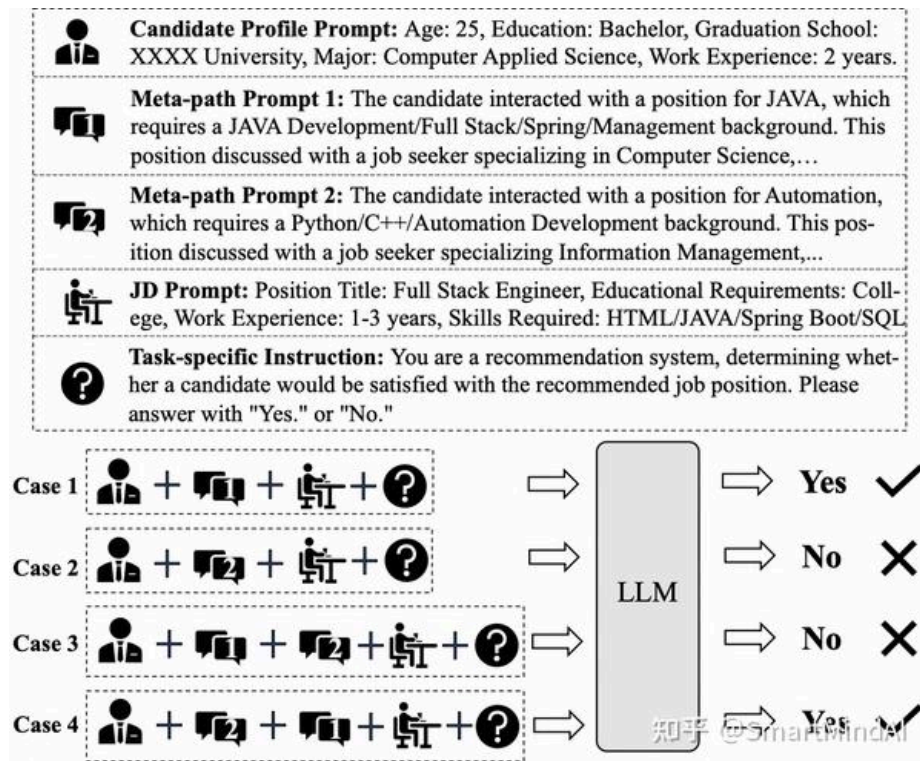
对于pair-wise任务, 我们让LLM推荐器验证求职者对推荐职位的偏好, 给定两个职位的JD Prompt A"和 B", 具体指令如下:

**pair-wise指令:** 你是一个推荐器, 决定哪个职位更适合求职者。请回答 "[A]" 或 "[B]" 。

通过上述设计的提示和指令, LLM能够适应领域推荐的任务。为了确保训练的稳定性, 我们将JD提示添加到真实数据的末尾, 以增加预测长度。

### Behavior Meta-path Prompt Generation

为了使大语言模型理解图数据中的交互关系, 我们提出了一种基于元路径的提示构造方法, 以获取能够表示局部子图的提示输入。异质图由节点集合和边集合组成, 与节点类型映射函数和链接类型映射函数相关联。元路径定义为节点之间的复合关系路径, 描述了节点之间的复合关系。在在线招聘背景下, 行为图是一个异质图, 其中不同的节点类型包括(求职者, 职位描述), 不同的边类型包括消息、面试、匹配等。由于每种类型的边具有独特的定义语义, 自然地考虑将图数据格式的元路径转换为大语言模型可以接受的自然语言描述。节点信息填充到元路径提示模板中, 以生成最终提示数据, 其中实际案例可以参考图。



为了避免元路径过于相似导致冗余, 我们定义了一个简单的相似度量, 确保最终选择的路径之间的相似度不超过超参数 $\gamma$ 。

### Path Debiasing and Soft Selection

不同于传统的网络嵌入, 基于序列的元路径提示会对LLM理解候选人的行为子图带来两个挑战: 路径权重的影响和路径提示顺序的位置偏差。路径权重的影响指不同的元路径会对模型决策呈现不同的权重; 路径提示顺序的位置偏差指路径提示顺序的改变带来了不稳定的答案。为应对这些挑战, 我们设计了增强模块, 包含shuffle机制、路径软选择器和混合机制。shuffle机制通过随机shuffle

和路径软选择器，期望能够解决路径权重的影响和路径提示顺序的位置偏差带来的挑战。

LLM Instruction Tuning and Recommendation

对于指令调优，我们使用屏蔽提示损失位置的监督微调方法。为了解决直接微调模型带来的计算问题，我们提出了一种使用LoRA的轻量级微调方法，该方法有助于减少GPU内存占用。最终的学习目标为：

$$\mathcal{L}_f = \max_{\Theta_L} \sum_{(x,y) \in \mathcal{T}} \sum_{t=1}^{|y|} \log(P_{\Theta+\Theta_L}(y_t | e_x, y_{<t}))$$

其中  $\Theta_L$  表示LoRA参数，仅在训练过程中调整。至于推荐过程，我们设计的答案解析方式较为简单，捕获模型输出对应于地面真实值位置的softmax概率。

Experiments

Datasets

我们在中国大型在线招聘平台上收集了两个数据集RecrX和RecrY。数据集包含匹配行为（Match）和交互行为（Interaction），每个候选人和职位都关联有描述性文本。整体统计数据显示在表1中，职位推荐是一个交互稀疏的场景，训练集和测试集的比例为5:1，所有敏感信息均已去除。

| Dataset | # Candidates | # Jobs | # Match | # Interaction |
|---------|--------------|--------|---------|---------------|
| RecrX   | 12,440       | 19,318 | 23,879  | 54,147        |
| RecrY   | 18,260       | 26,576 | 47,725  | 119,329       |

Quantitative Comparison.

我们在两个数据集上进行了定量性能实验。正如方法部分所述，我们采用point-wise和pairwise设置进行评估。我们还探讨了 OOD 情况对不同模型的影响。以下是随机、OOD\_position 和 OOD\_JD 的实验划分设置：

**随机:** 我们基于每个用户的交互记录，随机将训练集和测试集分开。**OOD\_position:** JD的“职位”特征在训练集和测试集之间没有交集。**OOD\_JD:** 在JD项目之间，训练集和测试集之间没有交集。

| Task       | Point-wise   |              |              |              |              |              | Pair-wise    |              |
|------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Split      | Random       |              | OOD_position |              | OOD_JD       |              | Random       |              |
| Dataset    | RX           | RY           | RX           | RY           | RX           | RY           | RX           | RY           |
| RobertaRec | 0.710        | 0.734        | 0.503        | 0.528        | 0.506        | 0.536        | 0.727        | 0.740        |
| HGT        | 0.744        | 0.756        | 0.572        | 0.595        | 0.576        | 0.593        | 0.747        | 0.751        |
| DPGNN      | 0.727        | 0.743        | 0.596        | 0.603        | 0.588        | 0.617        | 0.744        | 0.756        |
| TALLrec    | 0.842*       | 0.829*       | 0.770*       | 0.788*       | 0.766*       | 0.798*       | 0.849*       | 0.825*       |
| GLRec      | <b>0.891</b> | <b>0.876</b> | <b>0.810</b> | <b>0.843</b> | <b>0.814</b> | <b>0.852</b> | <b>0.905</b> | <b>0.883</b> |
| Improve ↑  | 18.4%        | 14.1%        | 25.2%        | 28.3%        | 26.4%        | 27.8%        | 15.5%        | 13.2%        |

实验结果表明，GLRec模型在所有基线模型中表现最佳，尤其是在OOD任务上。尽管其性能在随机设置下略有下降，这实际上反映了利用知识关联进行模型泛化的必要性。此外，基于图的HGT和DPGNN在职业推荐中优于传统的双塔模型，表明学习关系的重要性。大多数模型在pair-wise任务上表现优于point-wise任务。

Qualitative Comparison.

为了提供更直观的可视化，表中展示了由模型产生的某些定性比较结果。

|  |  |   |  |     |     |     |     |
|--|--|---|--|-----|-----|-----|-----|
| Bachelor's degree, Business Administration, 9 years of work experience, skills...    | Project Assistant, Skill requirements: Project Engineering Management          | Bachelor's degree, International Economics and Trade, 3 years of work experience, skills... | Project Assistant, Qualification: Bachelor's degree, 3 years or more, Skill requirements: Project Engineering Management         | Yes | Yes | Yes | Yes |
| Bachelor's degree, Computer Applications Tech, 2 years of work experience, skills... | JAVA, Skill requirements: JAVA/Spring/Team Management Experience               | Associate's degree, Internet of Things Technology, 4 years of work experience, skills...    | Full Stack Engineer, Qualification: Associate's degree, 1-3 years of work experience, Skill requirements: JAVA/Spring/HTML       | Yes | No  | Yes | Yes |
| Bachelor's degree, Finance, 10 years of work experience, skills...                   | Functional Testing, Skill requirements: Software Testing/Requirement Alignment | Bachelor's degree, Financial Engineering, 2 years of work experience, skills...             | Test Engineer, Qualification: Bachelor's degree, 3 years of work experience, Skill requirements: Functional Testing/Unit Testing | Yes | No  | No  | Yes |

具体来说，前两行是直接的，多个模型可以准确预测。在第三行中，仅使用用户资料不足以进行预测。值得注意的是，用户互动过的 Java 位置（节点 1）与目标职位在技能要求上的匹配非常好。因此，只有 TALLrec 和 GLRec 生成了正确的预测。最后一行强调了高阶交互路径，在 LLM 推荐中的重要性。尽管候选人金融专业的背景与目标职位之间存在感知上的不匹配，但在测试工程师和金融科技领域的互动提供了细微的线索。对于这种复杂的情况，依赖于过去行为的 TALLrec 模型出错，而只有 GLRec 模型预测正确。

论文《Exploring Large Language Model for Graph Data Understanding in Online Job Recommendations》

发布于 2024-10-22 14:25 · IP 属地北京

推荐系统 LLM 图神经网络 (GNN)

赞同 3 添加评论 分享 喜欢 收藏 申请转载 ...



理性发言，友善互动



还没有评论，发表第一个评论吧

推荐阅读

【AAAI 2024】解锁深度表格学习 (Deep Tabular...

近日，阿里云人工智能平台PAI与浙江大学吴健、应豪超老师团队合作论文《Arithmetic Feature Interaction is Necessary for Deep Tabular Learning》正式在国际人工智能顶会AAAI-2024上发... 阿里灵杰

指针分析/Point-to Analysis/Reference Analysis

### 目录 @toc 指针分析指针分析紧接数据流分析，是静态分析中的一个难点。遗憾的是能把相关内容讲得通俗易懂的资料非常少，中文的更基本没有。所以我斗胆在这里以例子的形式讲一下静态... 信息门下皮... 发表于Resea...

NLP技术高手—  
技术流程图  
——意AI增效家

把一段话/文档/书转为大模型的数据集！适用微调+知识库！...  
——意AI增效家

91%

用数据讲故事 (上) | 如  
图表类型？视觉认知的  
看得见的城... 发表于看得