

## 【A15】基于知识图谱的大学生就业能力评价和职位推荐系统【锐捷网络】

### 1.命题方向

智能计算

### 2.题目类别

计算类

### 3.题目名称

基于知识图谱的大学生就业能力评价和职位推荐系统

### 4.背景说明

#### 【整体背景】

高校毕业生是国家宝贵的人才资源，数据显示，2023 届高校毕业生达 1158 万人，毕业生对自己就业能力的认知和精准提供就业信息，有利于实现就业选择与用人需求的双赢，达到精准有效，促进高校毕业生高质量充分就业。

大学生就业质量也是评价学校培养质量的重要组成部分，也是社会各界关注的焦点。目前，网络招聘是获取就业信息的重要途径，然而，当前网络招聘市场中的职位信息和求职者数量庞大，使得求职者面临信息过载问题，同时也给企业招聘带来效率低下问题。

因此，如何能够帮助求职者快速找到理想职位，并且让其知晓对于意向岗位自身知识和技能上的缺陷，找到短板，有针对性提升自我能力，同时降低企业和求职者的匹配成本，提高求职和招聘的效率，具有很好的实用价值和市场潜力。

#### 【公司背景】

锐捷网络。

#### 【业务背景】

系统需实现岗位的智能推荐和胜任度评价，具有以下几方面的功能：

（1）**知识图谱构建：**职位推荐系统需要构建一个包含相关职位、技能要求、行业信息等多个领域知识的知识图谱。

（2）**聚合和分析用户信息：**与知识库相结合，对用户的个人简历、求职意向、工作经验等信息进行聚合和分析处理。根据用户提供的信息学习用户的兴趣和特征，并利用这些特征在知识图谱上匹配职位需求和其他相关信息。

（3）**建立推荐模型：**根据用户信息设计推荐算法和预测模型。通常来说，推荐算法有协同过滤、关键因素模型、深度学习模型等，具体选择应该根据实际情况来确定。

（4）**推荐结果展示：**将推荐结果呈现给用户，并支持用户进行选择和反馈。需要注意的是推荐结果的呈现方式也应该根据不同用户群体的需求和喜好来定制化。

一是信息展示差异化：提供筛选功能。

济南 公司行业 职位类型(1) 求职类型 工作经验 薪资待遇(1) 学历要求 公司规模 融资阶段

二是交互方式个性化：提供列表概览（关键词）和详细信息两种展示。

java开发工程师 [济南·历下区·高新会展中心]

15-30K·14薪 3-5年 本科 王先生 项目经理

后端工程师 | Java | MyBatis | MySQL | Java开发经验

 华为

计算机软件 不需要融资 10000人以上

加班补助, 餐补, 补充医疗保险, 五险一金, 零食下午茶, ...



## 5.项目说明

### 【问题说明】

实现一套大学生就业能力评价和智能岗位推荐系统，根据提供的岗位信息样例库，设计一套含智能算法的软件系统方案：基于知识图谱的岗位信息，利用职位和用户信息，结合推荐算法和相关技术，为用户提供符合其需求和兴趣的职位推荐结果和能力评价结果。

**能力评价：**用户**上传个人简历**，并明确自己**期望的职位**，系统自动判断用户与期望职位间的**契合度，差异性**，给出**提升建议**，让其知晓对于意向岗位自身知识和技能上的缺陷，找到短板，有针对性提升自我能力。



**岗位推荐：**用户上传个人简历，系统自动分析简历内容，生成推荐职位，用户可以给出**推荐是否有效的反馈**。如果不满意，可修订简历部分内容，重新进行推荐。

#### 选择原因，减少不合适职位推荐

##### 不太感兴趣

“济南”城市	“山东嘉友互联软件...”
不是实习职位	实习月数不合适

##### 效率反馈

职位停招/招满	面试过/入职过
---------	---------

##### 其他反馈



**招聘推荐：**企业招聘人员输入或上传职位要求，系统自动分析匹配求职人员简历，筛选出符合期望的人员列表。

### 【用户期望】

对开发的产品方案期望如下：

- (1) 算法优化合理，求职与招聘推荐结果与用户期望一致性高；
- (2) 胜任度能力评价结果合理，给出的提升建议符合用户短板；
- (3) 保护简历中个人数据的安全，不侵犯用户隐私；

**(4) 扩展功能：可以对注册用户的数据和操作行为进行统计分析，给出热门职位、热门技能、热门专业等的一些趋势图等。**

为了直观展示智能匹配系统的相关数据，本文通过使用开源第三方库 ECharts 来实现数据可视化。该类库使用纯 Javascript 编写，对任何版本的浏览器的兼容性都比较好。本文运用 barChart 和 lineChart 组件实现了条形图、饼状图、散点图、折线图以及气泡图等，直观地展示了求职者偏好、企业偏好、用户活跃度、热门职位类别、求职者学学历年龄比、专业跨度等。可视化界面如图 5-10 所示。



图 5-10 人职匹配数据可视化界面

## 6.任务要求

### 【开发说明】

本项目要求完成上述所有功能的软件系统，实现方法不限。

### 【技术要求与指标】

(1) **推荐有效性达到 80% 以上**（用户调查：电子或计算机类相关专业毕业生简历与岗位样例库进行匹配）；

(2) 系统数据库中个人隐私信息（姓名、手机、邮箱、通讯地址等至少 4 项）**进行加密，只能被授权人员解密**；

加密

**(3) 架构设计上可并发支持 1000 人以上同时在线使用，推荐响应时间在 5s 以内；使用消息队列系统来处理异步任务，如 RabbitMQ**

(4) 技术不限，开发工具不限，可采用开源技术。

#### 【提交材料】

- (1) 项目概要介绍；
- (2) 项目简介 PPT；
- (3) 项目详细方案；
- (4) 项目演示视频；
- (5) 企业要求提交的材料：
  - ①项目详细方案中详述推荐算法；
  - ②软件安装包；
  - ③安装部署手册；
- (6) 团队自愿提交的其他补充材料。

#### 【任务清单】

- (1) 完成需求分析与程序设计；
- (2) 编码开发与功能实现软件 Demo；
- (3) 测试验证功能与性能等主要指标；
- (4) 洞察应用场景与市场机会。

#### 【开发工具与数据接口】

无限制，自行选择

#### 7.其他

无

#### 8.参考信息

命题企业提供测试用的岗位信息（含专科、本科、研究生）样例库（csv 文件，可用 Excel 打开），下载地址：

链接：<https://pan.baidu.com/s/1BsrKlVRbWe6i9DTUc0cGtg?pwd=vmeh>

提取码：vmeh

(BOSS 直聘)

#### 9.评分要点

赛题评分要点见附件一：A 类企业命题初赛统一评分标准。

1	问题	A15中的数据集是否已经给全，理论上应该还有一个简历数据集
	答复	简历没有相关数据集，可以使用学生自有数据集。
2	问题	【问题说明】中“用户上传个人简历”，这个“上传”如何定义，是只需要用户输入信息吗？\n【技术要求与指标】（1）推荐有效性达到80% 以上。这个80%如何定义，题目只给了岗位信息，没有给任何的简历信息、或者推荐信息。
	答复	企业会根据简历样本做匹配正确性验证。

## 命题介绍 – 命题内容

实现一套**大学生就业能力评价和智能岗位推荐系统**，根据提供的岗位信息样例库，设计一套含智能算法的软件系统方案：基于知识图谱的岗位信息，利用职位和用户信息，结合推荐算法和相关技术，为用户提供符合其需求和兴趣的职位推荐结果和能力评价结果。

### 能力评价

- 学生上传简历
- 提供期望职位
- 系统给出契合度评分、差异点、提升建议

### 岗位推荐

- 学生上传简历
- 自动分析简历
- 自动生成推荐职位
- 用户反馈有效性
- 可换一批

### 招聘推荐

- 招聘者提供职位要求
- 自动分析职位
- 自动生成推荐求职人员列表
- 招聘者反馈有效性
- 可换一批

## 命题介绍 – 答题要求

1

推荐有效性达到80% 以上（用户调查：电子或计算机类相关专业毕业生简历与岗位样例库进行匹配）

2

系统数据库中个人隐私信息（姓名、手机、邮箱、通讯地址等至少4项）进行加密，只能被授权人员解密

3

架构设计上可并发支持1000人以上同时在线使用，推荐响应时间在5s以内



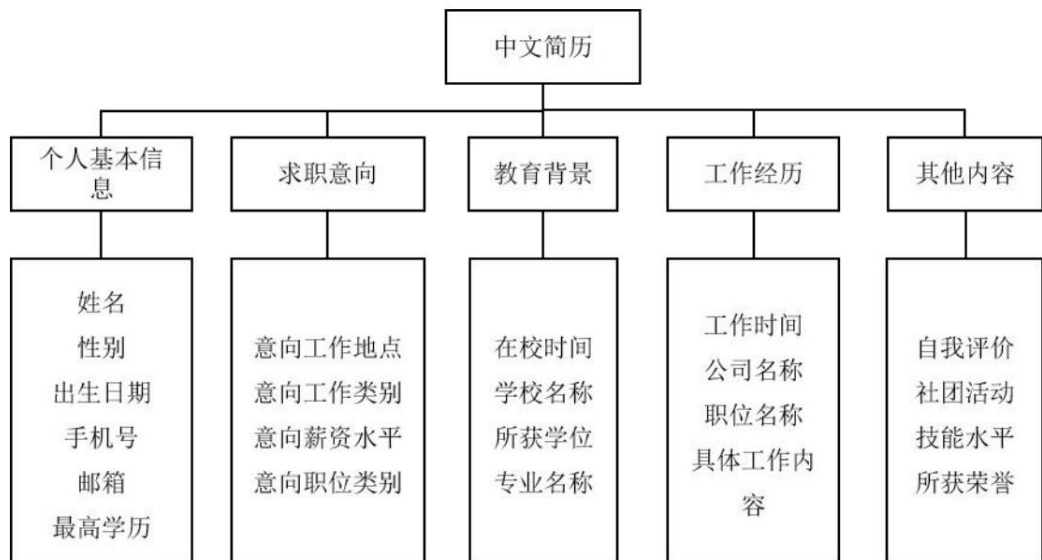


图 3-1 中文简历一般形式

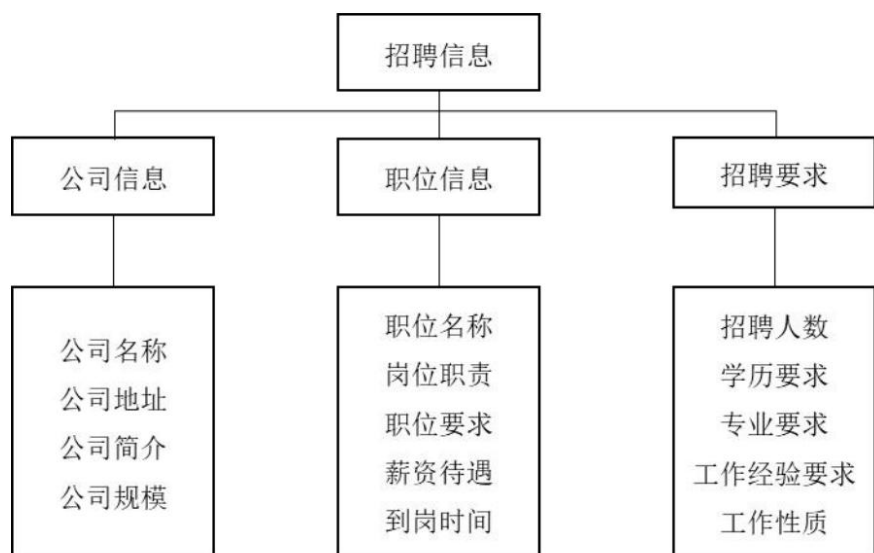


图 3-2 招聘信息一般形式

3.5.1 实验数据集

目前在招聘信息与简历方面没有经典权威的数据集，因此本文所使用的数据集主要来自商业合作的招聘信息和简历数据。本文将针对这些简历信息和招聘信息的特点完成智能匹配工作，在进行结构化文本特征满意度计算时，从求职者简历的角度出发度量企业职位信息中薪资、工作地点、职位性质、职位类别等字段的满意度，从企业职位信息的角度出发度量求职者简历中学历、专业、工作经验、专业技能、期望职位类别的满意度。此外，在计算非结构化文本的相似度时，需要对简历中工作经验以及职位信息中的职位描述等字段进行提取。智能匹配算法所需要数据字段如表 3-4 所示。

表 3-4 智能匹配算法数据字段

表名	字段	字段说明	表名	字段	字段说明
Resume	resumeID	简历 id	Job	jobID	职位 id
	education	最高学历		r_education	学历要求
	major	专业		r_major	专业要求
	exp_year	工作经验年限		r_exp	工作经验要求
	skill	专业技能		r_skill	技能要求
	h_salary	期望月薪		salary	薪资
	h_place	期望工作地点		place	工作地点
	h_type	期望职位类别		type	职位类别
	h_property	期望工作性质		property	工作性质
	experience	工作经验		description	职位描述

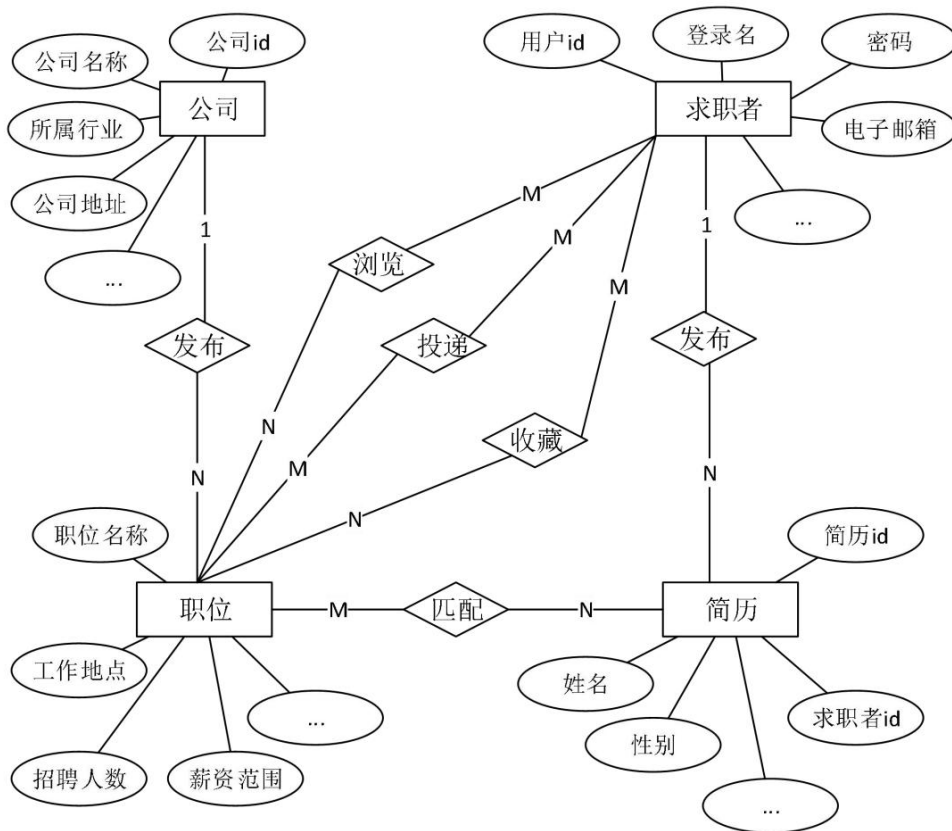


图 4-13 系统数据库实体关系图

即使没有简历数据集, 仅仅有岗位知识图谱, 也可以设计基于岗位知识图谱的推荐算法。这种算法可以根据岗位的要求和特征, 结合候选人的个人信息 (如果有), 从而进行岗位推荐。

以下是一些可能的方法:

**岗位相似度匹配:** 利用岗位知识图谱中的岗位要求、技能需求、工作内容等信息, 计算不同岗位之间的相似度。当候选人搜索或表达对某一岗位的兴趣时, 系统可以推荐与该岗位相似度较高的其他岗位。(协同过滤)

**内容过滤推荐:** 根据岗位知识图谱中的关键词和特征, 结合候选人的搜索历史或输入的关键词, 利用文本相似度或匹配算法推荐与其相关的岗位。

**基于规则的推荐:** 建立一些规则, 根据岗位知识图谱中的特征, 例如技能要求、工作经验等, 筛选出与候选人匹配度较高的岗位。

虽然缺少简历数据集会限制个性化程度, 但是仍然可以通过分析岗位知识图谱中的信息, 提供与候选人兴趣和能力相匹配的岗位推荐。然而, 简历数据集能够更全面地了解候选人的背景和技能, 从而提高推荐的准确性和个性化程度。

任职要求: 1. 负责仪器仪表的现场安装调试 **专业能力**、对客户的技术培训 **专业能力**、解答客户问题, 提供相关技术支持工作; 2. 大专以上学历, 机械 **专业知识**、机电 **专业知识**、环保 **专业知识** 等相关专业, 可接受应届生, 能适应出差; 3. 动手能力 **个人素养** 强, 具有较强的服务意识 **个人素养**, 工作积极热情 **个人素养**, 勤奋敬业 **个人素养**, 做事细心认真 **个人素养**; 薪酬福利: 有竞争力的薪资

• 在职培训教育机会, 完善的员工培训制度及职业生涯规划。• 试用期即办理六险一金 (养老保险、失业保险、工伤保险、生育保险、医疗保险、住房公积金、商业意外险)。• 福利完善: 免费工作午餐、过节费、生日礼金、结婚礼金、生育津贴、通讯补贴、免费体检、年终奖、父母生日礼物、特殊关怀慰问金、子女儿童节礼品、子女夏令营、子女国学亲子班等各种福利以及每年一至两次员工旅游等活动。• 假期充足: 法定假日、带薪年假、婚假、丧假、产假、陪产假、工伤假、带薪病假。• 周末双休, 每周五天工作制, 每天7.5小时

#### 标注结果

实体标注  
专业能力 2  
专业知识 3  
个人素养 5

请输入关系标签名称

实体标签

专业能力

专业知识

个人素养

关系标签



职位 id 所在公司 xxxx  
职位 id 所在城市 xxxx  
职位 id 职位名称 xxxx  
职位 id 职位类别 xxxx  
职位 id 是否实习 xxxx  
职位 id 薪资 xxxx  
职位 id 学历要求 xxxx

职位 id 技能要求 xxxx  
职位 id 知识要求 xxxx  
职位 id 素养要求 xxxx

**能词**。技能词可以看成是特定专业领域内的命名实体<sup>[2]</sup>或者术语<sup>[3]</sup>。因此，网络招聘文本技能词的抽取任务可以借鉴命名实体识别或者术语抽取的方法。目前，虽然命名实体识别的相关工作很多，但重点都是在识别正式文本中的人名，地名和机构名，而术语抽取的相关工作主要针对特定领域内的术语识别，缺乏通用性和可移植性。其次，大多数命名实体识别或者术语抽取研究均采用人工规则或传统的机器学习方法，基于一个学科或一个领域进行抽取，需要人工设计规则进行特征提取。近年来，许多相关研究也将命名实体识别或术语抽取任务转化为序列标注问题，而最新技术主要基于深度神经网络进行端到端训练以捕获上下文信息。这类方法可以将输入字符转换为输出标签，而无需显式的人工设计特征提取。但是，基于深度学习的序列标注方法仅专注于域内监督学习，这需要大量带标注的数据。对于网络招聘数据而言，人工标注数据既费时又昂贵，只能依靠领域专家手工标注少量语句。因此，很难获得足够的带标注的语句来训练深度神经网络。如何使用少量标注数据来快速、准确地抽取技能词是非常具有挑战的。由此可见，中文网络招聘文本中的技能词抽取研究仍然是一项颇为艰巨的，也是非常具有价值的任务。

的适当水平的能力来表征每个大数据工作族。在国内，詹川<sup>[42]</sup>根据已有的关于电子商

## 第一章 引言

务行业的专业术语构建出该专业的**技能词词典**，对采集的 66925 条电子商务岗位的招聘信息进行**关键词抽取**，并对抽取出的**高频技能关键词进行分类**，分析出电子商务行业中不同类型岗位的**通用技能需求和特定岗位技能需求**。俞琰<sup>[43]</sup>等人从前途无忧招聘网站中抓取了 10000 条计算机领域和 30000 条非计算机领域的招聘信息，利用依存

而目前大部分的推荐系统采用协同过滤的推荐算法，存在冷启动和数据稀疏问题，比如对于新发布的职位，系统中没有用户的反馈信息，所以也无法将这个职位推荐出去，而用户没有历史行为记录，无法通过用户历史行为找到相似用户群体，也就很难进行推荐。