

# 针对医院多元素质人才招聘的双向混合推荐模型构建研究

刘俊, 徐丹

(南京脑科医院, 江苏 南京 210009)

**摘要:** 为了提高医院对于多元素质人才招聘与医学生就业的双向满意度, 研究构建面向医院招聘的双向混合推荐模型。首先提出基于内容过滤的双向推荐方法, 建立向量空间模型。其次针对数据稀疏性问题, 将预测评分引入评分矩阵。最后通过引入时间因子, 改进相似度计算。结果表明求职满意度提升了9.9%。而得到职位推荐准确率也提高了10.7%。医院在使用研究模型后的满意度提高了8.9%, 并且获得的人才推荐准确度也提高了7.4%。此次研究为其他招聘平台的改进和优化提供了参考方向, 也推动了医疗招聘平台的更新和完善。

**关键词:** 医院; 招聘; 多元素质人才; 协调过滤; 双重聚类

**中图分类号:** TP18 **文献标识码:** A **文章编号:** 1003-7241(2025)08-0178-06

## Research on Bi-directional Hybrid Recommendation Model Construction for Hospital Multi-quality Talent Recruitment

LIU Jun, XU Dan

(Nanjing Brain Hospital, Nanjing 210009, China)

**Abstract:** To improve the two-way satisfaction of recruitment of multi-quality talents and employment of medical students in hospitals, a two-way mixed recommendation model for hospital recruitment is constructed. Firstly, a bidirectional recommendation method based on content filtering is proposed, and a vector space model is established. Secondly, to solve the problem of data sparsity, the predictive score is introduced into the scoring matrix. Finally, the similarity calculation is improved by introducing time factor. The results show a 9.9% increase in job satisfaction. The accuracy of getting job recommendations also increased by 10.7%. Hospitals experiences an 8.9% increase in satisfaction with the study model and a 7.4% increase in the accuracy of their talent recommendations. This study provides a reference direction for the improvement and optimization of other recruitment platforms, and also promotes the update and improvement of medical recruitment platforms.

**Keywords:** hospital; recruitment; multi-quality talents; coordinated filtering; double clustering

### 0 引言

随着技术的发展与成熟, 在互联网招聘平台上, 企业的招聘人员可以在较短时间内发布招聘岗位需求并应聘合适的求职者。同样求职者也可以利用行业、职位、地点等关键词搜索技术来筛选出招聘信息, 从而找到一份满意的工作。在高等医学教育中, 医学生的就业是一个非常重要的环节。近年来, 我国医学院校毕业生的就业与录用问题受到了社会各界的广泛关注。随着医学教育的不断发展, 以及医药就业制度的革新, 医疗机构的人员规模也在逐渐壮大<sup>[1]</sup>。在医疗事业员工的岗位接近饱和的就业环境下, 医学生的就业率持续走低, 同时医院的招聘质量也未得到提升。由于医疗机构的特殊性, 医学生的就业与招聘是受多因素共同影响的复杂过程。医学生出现低就业率, 以及医院对多元素质人才招聘不理想的原

因主要在于学生与医院对于双方认识与评估的不足<sup>[2]</sup>。面对求职与招聘两头难的困境, 则需要分析医学生与医疗单位各自的能力优势与需求间的匹配程度, 采取个性化推荐方案。医院招聘工作规范性要求高、招聘工作量大、涉及面广。医院人才的竞争, 归根结底是软实力的竞争, 也是硬实力的竞争。在日益精细化的招聘需求下, 结合医院的实际情况, 构建出符合当前环境下高质量的招聘简历系统, 高效率地完成各项招聘任务, 是医院医疗人才争夺战中的首发战场。当前, 国外关于职位推荐系统的研究主要集中在简历信息提取方面<sup>[3]</sup>。Kevin等提出了一种基于候选人能力简历的推荐方法, 该方法是基于内容的过滤和相似性指数系统的混合模型<sup>[4]</sup>。但这种推荐模型通常不能同时考虑求职者的个人意愿及招聘方的客观需求, 缺少对双方共同利益的综合考虑。为此, 研究创新性地提出了基于内容过滤的双向推荐, 以及双重聚类协调过滤的混合推荐模型。研究旨在提高医院招聘与人才选岗的推荐准确率, 量化用户双方属性值, 提高医学生

\*基金项目: 江苏省中医药科技发展计划项目(YB2020023)

收稿日期: 2023-12-25

和医院之间的双向推荐满意度。

## 1 面向医院多元素质人才招聘的双向混合推荐模型设计研究

### 1.1 基于内容过滤的双向推荐方法

医学生就业与招聘关系的建立必须强调其供需关系的契合度,只有双方的需求得到满足,能力与供给符合要求,才能引导人才流向合适的市场。传统的基于协作过滤的推荐算法在推荐系统中的应用已越来越难以满足人们的要求。针对目前求职招聘平台中存在的问题和“岗位-人才”双向选择的特征,研究将推荐算法引入到招聘网站中,构建基于双向推荐的求职招聘模型。双向推荐既能综合考虑双方需求偏好,其在用户信息的获取上也有诸多优势<sup>[5]</sup>。考虑到招聘中双方的需求对等,这就要求将双方的偏好均引入到双向推荐模型中,该框架如图1所示。



图1 双向推荐模型

面向人才招聘的双向推荐模型主要分为如图1所示的三大模块,分别为招聘模块、求职模块、推荐模块。招聘模块主要为企业进行职位与简历等信息的管理处理。求职者可通过求职模块创建个人信息,以待被推荐。推荐模块用于分析招聘双方的属性信息,量化用户属性,计算相似度等。并基于相似度值,向求职者与用人单位推荐职位、人才,实现双向推荐。职位与简历信息均属于文本信息,其含有大量基本数据。建立双向推荐模型首先要解决的问题就是如何获得用户的数据,并对其进行分析。与英文文本相比,中文文本需要额外的中文分词过程。常见的分词工具包括Ansj、jieba分词等<sup>[6]</sup>。其中,Ansj中文分词是一种Java开源分词工具,使用效果较佳,因此研究将其作为推荐模型的分词工具。为了减少系统的计算成本,还需对文本数据进行预处理。该处理过程如图2所示。

在对中文进行分词处理,去除停用词之后,就可得到简历和工作岗位有关的属性词。而各向量的分量则表示不同的属性条件,一般分为求职向量和职位向量。两个

向量的各自分量一一对应,研究采用用户的性别、专业、学历、报酬、居住地,以及工作城市作为主要分量。将企业招聘的职位向量和毕业生的求职向量分别用表示 $E$ 、 $P$ ,其各自的分类用表示 $e_i, p_i$ 。研究以向量相似性计算为依据,实现双向推荐。其中,相似度的计算结果对匹配具有决定性的作用。首先是对性别的量化,当向某医院推荐多元素质人才时,将职位的性别属性值设为1,符合性别要求则为1,否则为0。当向求职者推荐职位时,与上述量化标准相同。学历量化则分为两种情况,第一种为当向医院推荐人才时,求职者学历为博士、硕士、本科、专科的量化结果依次为1、0.8、0.6、0.3。第二种为当向求职者推荐职位时,若该求职者学历低于职位要求学历,则学历量化值为0。若其学历与职位相符或高于职位要求时,则按式(1)进行量化。

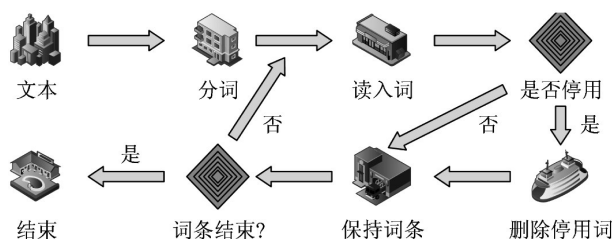


图2 文本预处理步骤

$$v = \frac{w_e}{w_p} \quad (1)$$

式中, $w_e$ 表示按第一种情况量化的职位学历。 $w_p$ 代表按照第一种情况量化的求职者学历。将被推荐用户的专业属性作为标准向量1,若求职与招聘的专业相符则为1,否则为0。对于报酬的量化,在条件相同的条件下,收入不超过医院开出的工资的求职者则优先被推荐给医院。假设医院所给报酬为 $a$ ,存在 $n$ 个求职者,其期望报酬为 $a_1, a_2, \dots, a_n$ 。其中, $M = \max\{a_1, a_2, \dots, a_n\}$ ,  $m = \min\{a_1, a_2, \dots, a_n\}$ 。在向医院推荐求职者时,职位报酬作为标准向量分量,即为1。在推荐求职者时,期望报酬的量化计算表达如式(2)为

$$e = \frac{a-x}{M-m} + \frac{M-m}{M-a}, x \in (a_1, a_2, \dots, a_n) \quad (2)$$

式中,若 $e < 0$ ,则取值为0。

在向求职者推荐岗位时,医院给出的报酬越是高于期望报酬,则被优先推荐的概率越大。同样假设求职者的期望报酬为 $c$ ,存在 $n$ 个职位,所给的报酬分别为 $d_1, d_2, \dots, d_n$ 。其中, $M = \max\{d_1, d_2, \dots, d_n\}$ ,  $m = \min\{d_1, d_2, \dots, d_n\}$ 。当向求职者推荐职位时,职位报酬量化的过程如式(3)为

$$p = \frac{x-a}{M-m} + \frac{a-m}{M-m}, x \in (d_1, d_2, \dots, d_n) \quad (3)$$

在向医院推荐求职者时,同等条件下,对食宿要求低的求职者被优先推荐的概率越大。因此建立映射关系如

下,提供食宿、提供住宿、提供工作餐、食宿自理的量化值分别为0.1、0.5、0.7、1.0。而对于求职者来说,提供住食、住宿、工作餐、食宿自理的量化值分别为1.0、0.7、0.5、0.1。此外,研究还从邮编的角度对地址进行分类。研究将以量化的结果为基础,进行后续的相似性计算及人才招聘领域的双向推荐。式(4)为相似度计算过程为

$$sim\left(\vec{u}, \vec{v}\right)=\frac{\vec{u} \cdot \vec{v}}{\left\|\vec{u}\right\| \left\|\vec{v}\right\|} \quad (4)$$

### 1.2 融合内容过滤与双重聚类的医院多元素质人才招聘混合双向模型构建

在基于内容过滤的推荐中,用户属性内容属于静态信息,只能作为初筛条件。而为了实现高满意度的双向推荐,则还需对用户的操作行为进行分析,并挖掘更多的潜在偏好。在内容不易解析的领域,采用协同过滤推荐是较好的选择<sup>[7]</sup>。基于用户的协同过滤推荐算法是基于一种假设,若用户对某项目的评分比较相似,则其对其他项目的评分也会较为相似。算法根据目标用户的最近邻域对某个项目的评分逼近目标用户对该项目的评分。该算法步骤如下,首先从项目的得分出发,对用户进行建模。之后,将该模型表示为用户-项目评分矩阵。该矩阵如式(5)为

$$R(i, j)=\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1j} \\ r_{21} & r_{22} & \cdots & r_{2j} \\ \cdots & \cdots & \cdots & \cdots \\ r_{i1} & r_{i2} & \cdots & r_{ij} \end{bmatrix} \quad (5)$$

式中,  $r_{ij}$  表示用户  $i$  对项目  $j$  的评分,该值与用户对项目的偏好度相关。在协调过滤推荐中,相似用户的确定需由相速度来确定。然而,余弦相似度没有考虑到用户评论习惯对相似性计算的影响<sup>[8]</sup>。为此,研究引入用户平均分,改进了余弦相似度计算方法,其表达如式(6)为

$$sim(u, v)=\frac{\sum_{i \in I_{uv}}\left(r_{u, i}-\bar{r}_u\right)\left(r_{v, i}-\bar{r}_v\right)}{\sqrt{\sum_{i \in I_{uv}}\left(r_{u, i}-\bar{r}_u\right)^2} \sqrt{\sum_{i \in I_{uv}}\left(r_{v, i}-\bar{r}_v\right)^2}} \quad (6)$$

在确定目标用户的邻域集后,也就是按照相似用户对项目的评价来预测没有触及的项目,该计算公式如式(7)为

$$p(u, i)=\bar{r}_u+\frac{\sum_{v \in N} sim(u, v)\left|r_{v, i}-\bar{r}_v\right|}{\sum_{v \in N} sim(u, v)} \quad (7)$$

式中,  $N$  代表目标用户  $u$  的邻居集。在建立用户偏好模型之前,必须采集具有时间特征的用户行为数据。在招聘领域中,时间因子的影响表现为用户兴趣的漂移<sup>[9]</sup>。与简

历或职位信息生命周期这两个方面。在特定应用领域,信息都存在一定生命周期,招聘一般都有时间节点<sup>[10]</sup>。研究引入时间因子后,得到的推荐模型如式(8)为

$$f(t)=e^{-\lambda t} \quad (8)$$

在协同过滤推荐中,数据稀疏性也是一个亟待解决的难题。为了解决该问题,需基于相关规则对矩阵数据进行填充<sup>[11]</sup>。研究针对医院招聘领域的特点,分别对职位与建立进行双重聚类分析,该步骤如图3所示。

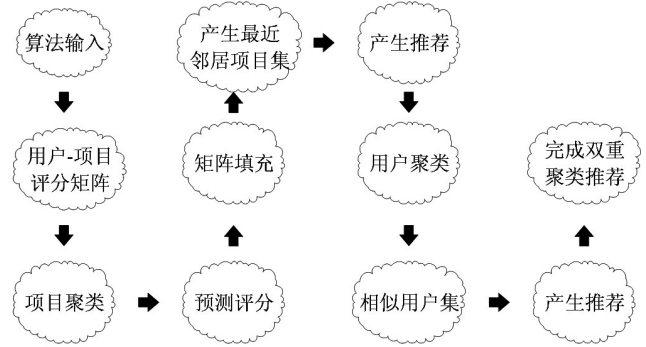


图3 双重聚类推荐流程

由于K-means算法在数据挖掘领域中具有便捷、高效的优点,因此研究采用该算法对医院招聘职位与求职者进行聚类<sup>[12]</sup>。首先,选取初始聚类中心,计算评分矩阵中各数据与其的距离。选择与各个数据最接近的聚类集合,然后将其合并到同一个类别中。而在此时,一个新的聚类中心即产生。重复计算聚类中心,直至各个聚类中心的平均值保持一致,或变化范围一定截止。考虑到时间因子对用户兴趣漂移的影响,研究改进了相似度的计算方法,改进后的方法如式(9)为

$$sim'(u, v)=\frac{\sum_{i \in I_{uv}}\left(r_{u, i}-\bar{r}_u\right)\left(r_{v, i}-\bar{r}_v\right) f(t)}{\sqrt{\sum_{i \in I_{uv}}\left(r_{u, i}-\bar{r}_u\right)^2} \sqrt{\sum_{i \in I_{uv}}\left(r_{v, i}-\bar{r}_v\right)^2}} \quad (9)$$

之后,再对目标用户未接触过的项目进行评分预测,改进后的计算表达如式(10)为

$$p(u, i)=\bar{r}_u+\frac{\sum_{v \in N} sim'(u, v)\left|r_{v, i}-\bar{r}_v\right|}{\sum_{v \in N} sim'(u, v)} \quad (10)$$

将项目评分预测值回填入用户-项目评分矩阵,即可得到数据稀疏性的改善。最终的推荐效果由基于内容过滤与协调过滤推荐共同决定,因此可获得混合双向推荐模型,如式(11)为

$$\begin{cases} a+b=1 \\ T=a \times R_c+b \times sim \end{cases} \quad (11)$$

式中,  $a, b$  代表基于内容过滤和协调过滤的推荐所占权重;  $R_c$  表示基于内容过滤的推荐匹配度;  $sim$  表示协同过滤推荐中相似度。

## 2 双向混合推荐模型的实验结果分析

研究运用问卷调查法,以当前的招聘网站为实验对象,来统计并分析大学生在求职网站上的推荐满意度。实验共实施4次问卷调查,满意度评分总分为10分。在首份调查表中,尚未使用满意反馈功能。在进行第二次问卷调查的时候,将系数 $a, b$ 均设为0.5。在第三、四次问卷调查中, $a$ 值分别取值为0.6、0.7。之后,继续增大系数 $a, b$ 值。图4为问卷收发情况及用户满意度。

由图4可知,第二次问卷调查比首次调查得到的用户满意度有了很大提高。医院的用户满意度与求职者的满意度均提高了3.3。在第三次问卷调查中,当 $a$ 取值为0.6时,医院与求职者的满意度均达到峰值,分别为8.64、8.65。因此将 $a$ 值设为0.6, $b$ 值设为0.4。实验通过对研究的基于内容过滤的双向推荐算法与传统基于内容过滤的推荐算法进行 $F_1$ 分值、平均绝对误差的对比,来验证其推荐效果,结果如图5所示。

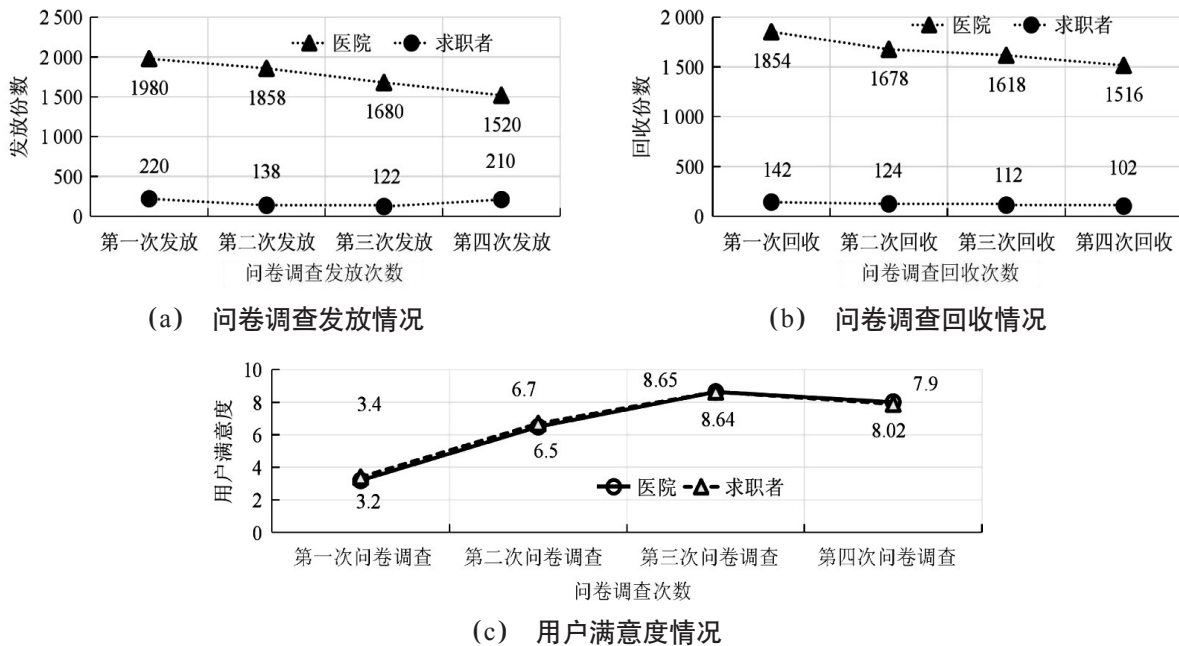


图4 问卷收发情况及用户满意度

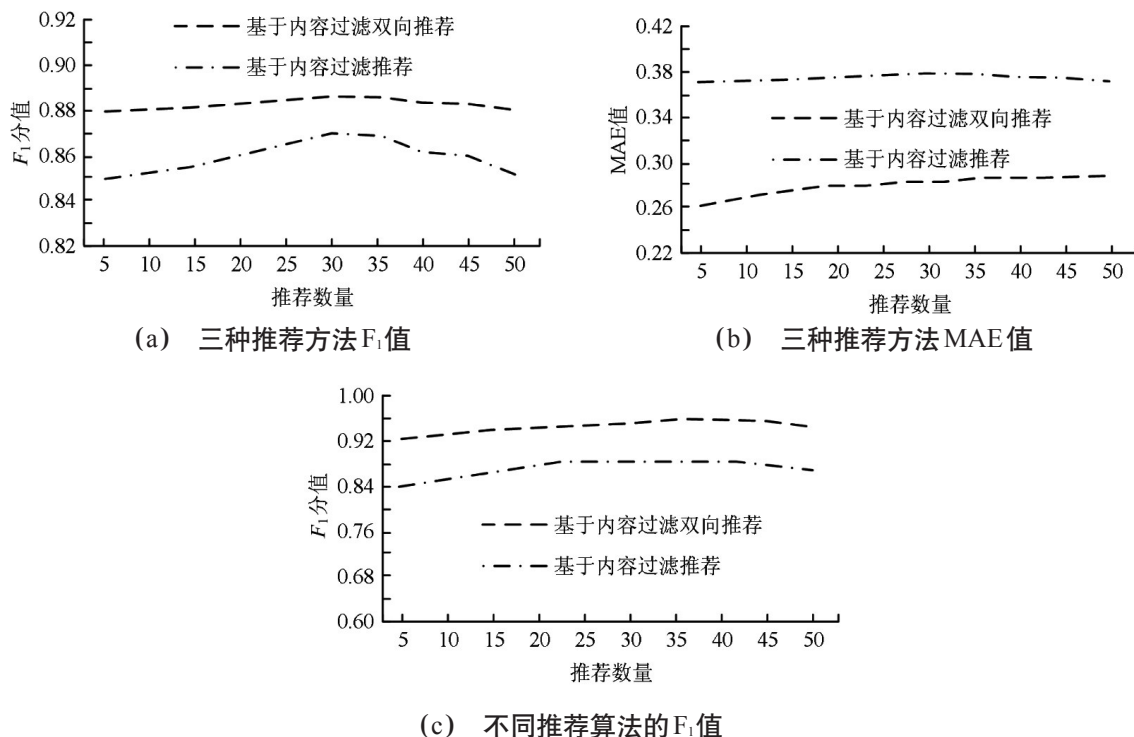
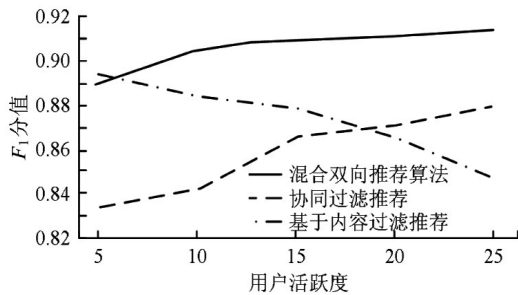
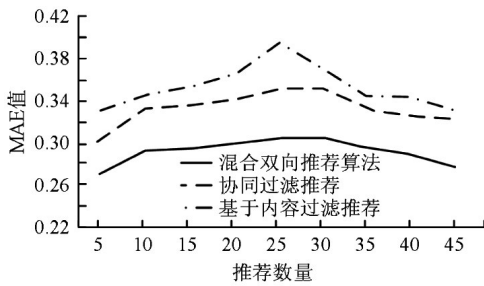


图5 两种推荐方法的 $F_1$ 值、MAE值对比情况

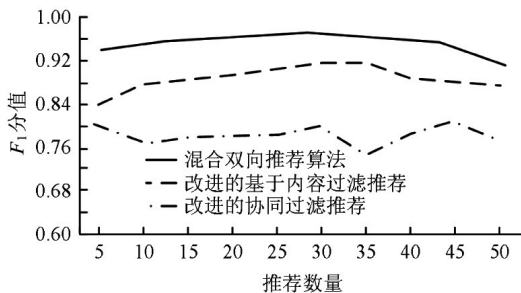
由图5(a)可知,研究算法的 $F_1$ 值的范围在0.87~0.89之间,而传统基于内容过滤的算法 $F_1$ 值位于0.85~0.87的范围内。研究算法较传统算法的 $F_1$ 值略高。由图5(b)可知,研究改进后的算法平均绝对误差(mean absolute error, MAE)值的范围在0.26~0.29之间,而传统算法的MAE均值约为0.37。这表明研究改进后的算法推荐误差更低,更容易实现满足双方要求的高质量推荐,验证了其在医院招聘领域中的优势。由图5(c)可知,当有新用户加入时,研究算法的 $F_1$ 值其平均值约为0.94,而传统算法的 $F_1$ 值均值约为0.87。在求职者新加入后,获得了更好的推荐效果。为了进一步验证研究的混合双向推荐算法的性能,研究将其与基于内容过滤的推荐算法和经典协同过滤算法进行对比,结果如图6所示。



(a) 三种推荐方法 $F_1$ 值



(b) 三种推荐方法MAE值

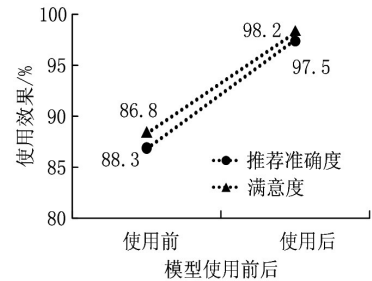


(c) 不同推荐算法的 $F_1$ 值

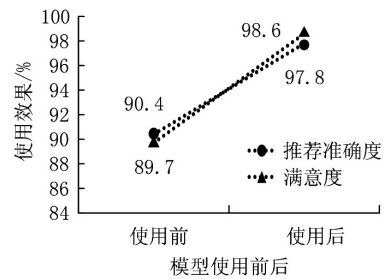
图6 不同算法的推荐性能

由图6(a)可知,全部推荐算法均随用户活跃程度的提高而发生了显著的变化。而基于内容过滤的推荐算法的 $F_1$ 值呈下降趋势。研究的双向推荐方法的 $F_1$ 值则基本保持稳定上涨趋势,并且较其他两种算法均要高。由图6(b)可知,研究设计的混合算法的MAE值相较于其他两

种推荐算法均要低,其MAE值均值为0.27。由图6(c)可知,在加入新用户后,混合双向推荐算法的 $F_1$ 值仍然高于其他两种改进算法,其 $F_1$ 值的均值为0.94。向之前同一批问卷调查对象征集使用双向混合推荐模型后的满意度值,得到如图7所示结果。



(a) 求职者



(b) 医院

图7 研究模型使用前后效果对比

由图7(a)可知,求职者在使用研究模型前后的满意度得到了较大提升,求职满意度提升了9.9%。而得到职位推荐准确率也提高了10.7%。由图7(b)可知,医院在使用研究模型后的满意度提高了8.9%,并且获得的人才推荐准确率也提高了7.4%。综上可知,研究模型针对医院多元素质人才招聘的双向混合推荐效果有了显著提升。

### 3 结束语

为了解决医院对于多元素质人才的招聘与医学生就业双向难的问题,研究将内容过滤与协同过滤方法相结合,建立一种双向混合的招聘推荐模型。在此基础上,利用问卷调查和实证研究的方法,对混合推荐中的各项权重系数进行了确定。最终,由统计得到用户使用研究模型前后的应用感受。实验表明,研究改进后的基于内容过滤的双向推荐算法的 $F_1$ 值的范围在0.87~0.89之间,其MAE值的范围在0.26~0.29之间,改进后的算法推荐误差更低,更容易实现满足双方要求的高质量推荐。研究设计的混合算法的MAE值相较于其他两种推荐算法均要低,其MAE值均值为0.27。在加入新用户后,混合双向推荐算法的 $F_1$ 值仍然高于其他两种改进算法,其 $F_1$ 值的均值为0.94。求职满意度提升了9.9%。而得到职位推荐准

确率也提高了10.7%。医院在使用研究模型后的满意度提高了8.9%,并且获得的人才推荐准确度也提高了7.4%。综上可知,研究在混合了改进基于内容过滤的双向推荐算法与协同过滤算法后,其在招聘领域中的推荐效果得到了较大的提升。

#### 参考文献:

- [1] 李懿,田玉玲.远程医疗信息系统中的三因素匿名认证协议[J].计算机工程与应用,2023,59(10):280-287.
- [2] 李红兰,巫蓉.基于决策分类算法的本科毕业生就业形势分析系统设计[J].现代电子技术,2021,44(10):73-76.
- [3] 赵勇进.基于人工智能技术的智媒体就业推荐平台构建研究[J].自动化技术与应用,2024,43(1):84-87.
- [4] KÉVIN CÉDRIC GUYARD, DERIAZ M. A scalable recommendation system approach for a companies-seniors matching[J]. International Journal of Semantic Computing, 2023, 17(2): 275-291.
- [5] 王世杰,高鑫,许舒翔.一种基于深度学习的推荐系统协同双向约束算法分析[J].微型电脑应用,2021,37(2):163-165.
- [6] 师娅杰.基于文本挖掘的客户重复诉求智能分析应用[J].新型工业化,2021,11(12):120-122.
- [7] 周晶,刘丹,李慧超,等.考虑用户兴趣的个性化协同过滤推荐方法[J].微型电脑应用,2022,3(8):74-78.
- [8] 钱梦薇,过弋.融合偏置深度学习的距离分解Top-N推荐算法[J].计算机科学,2021,48(9):103-109.
- [9] 代晓菊,张杰,董斌,等.基于预训练模型的语义相似性计

算在智能工单推荐中的应用[J].电信科学,2023,39(Z1):216-220.

[10] 张晓莹,陆慧娟.基于多层特征表征与级联模型的医疗简历筛选[J].小型微型计算机系统,2022,43(3):655-660.

[11] 武森,董雅贤,魏桂英,等.面向稀疏数据的协同过滤用户相似度计算研究[J].计算机科学与探索,2022,16(5):1043-1052.

[12] 何建强,张玉萍,滕志军.基于K-means和改进KNN算法的风电功率短期预测系统[J].计算机测量与控制,2022,30(5):156-162.

作者简介:刘俊(1984—),男,本科,助理研究员,研究方向:人力资源。

通信作者:徐丹(1985—),女,硕士,助理研究员,研究方向:人力资源。

(上接第164页)

了三种技术的AUC值,由此证明所研究技术的准确性更高,实现了研究目标的可行性分析。

#### 参考文献:

- [1] 马俊杰.绿色建筑理念下的建筑自动化技术应用设计[J].建材世界,2023,44(2):116-119.
- [2] 陈士成.自动化技术在电力系统中的应用[J].集成电路应用,2022,39(11):262-263.
- [3] 张丽娟,保富.基于改进SVM的电力用户异常用电行为检测方法研究[J].电测与仪表,2022,59(12):163-168.
- [4] 周啸辉,余磊,何茜,等.基于改进Res Net-18的红外图像人体行为识别方法研究[J].激光与红外,2021,51(9):1178-1184.
- [5] 王莹,项雯,张群,等.泛在电力物联网下基于DPSO-Kmeans的客户用电行为分析[J].哈尔滨理工大学学报,2022,27(2):106-113.
- [6] 张洁,夏飞,袁博,等.基于特征优选策略的居民用电行为聚类方法[J].电力系统自动化,2022,46(6):153-159.
- [7] 邓春宇,吴克河,张玉天.基于近邻相似度图聚类的用户用电行为分析[J].计算机工程与设计,2021,42(7):1933-1939.
- [8] 徐瑶,李栓,韩英华.基于CNN-GS-SVM的用户异常用电行为检测[J].控制工程,2021,28(10):1989-1997.
- [9] 刘钊瑞,高云鹏,郭建波,等.基于深度自编码器高斯混合模型的窃电行为检测[J].电力系统保护与控制,2022,50(18):92-102.

[10] 吕笃良,刘梦爽,桓露,等.基于重加权策略平衡损失与LSTM的窃电行为检测研究[J].智慧电力,2022,50(4):15-20,58.

[11] 武超飞,孙冲,刘厦,等.基于改进FCM聚类的窃电行为检测[J].电力科学与技术学报,2021,36(6):164-170.

[12] 万磊,陈成,黄文杰,等.基于BRB和LSTM网络的电力大数据用电异常检测方法[J].电力建设,2021,42(8):38-45.

作者简介:杨宇坤(1990—)男,本科,工程师,研究方向:用电检查及客户侧电力保障等。

通信作者:刘倩如(1990—),女,硕士,工程师,研究方向:优质服务管理和营销稽查管理等。