

基于深度学习算法的高校人才个性化智能推荐系统

陈 慧

(西安航空职业技术学院 党委组织部, 陕西 西安 710089)

摘要: 针对当前高校人才推荐系统无法处理海量稀疏性高校人才信息的缺陷, 提出了基于深度学习算法的高校人才个性化智能推荐系统, 从信息获取层获取高校人才个人简历、求职信、推荐信等信息后发送至数据处理层, 通过深度学习算法的卷积神经网络算法进行处理, 卷积神经网络的输入层接收海量信息, 选取文本向量化方法处理后发送至卷积层, 卷积层利用多组卷积核实现高校人才信息卷积特征的获取, 池化层对卷积层所获取特征实施降维并发送至全连接层, 全连接层选取 Relu 函数作为激活函数, 计算后发送至输出层, 输出层利用矩阵分解方法输出推荐结果, 数据处理层完成推荐后利用推荐服务层的个性化人机界面输出推荐结果。系统测试结果表明, 该系统具有较高的推荐成功率, 且覆盖率与实时性高。

关键词: 深度学习; 高校人才; 个性化智能推荐; 卷积神经网络; 文本向量化

中图分类号: TN911

文献标识码: A

文章编号: 1674-6236(2021)21-0038-05

DOI: 10.14022/j.issn1674-6236.2021.21.008

Personalized intelligent recommendation system for college talents based on deep learning algorithm

CHEN Hui

(Organization Department of the Party Committee, Xi'an Aeronautical Polytechnic Institute, Xi'an 710089, China)

Abstract: In view of the shortcomings of the previous college talents recommendation system that cannot deal with massive sparse college talent information, the personalized intelligent recommendation system for college talents who studied deep learning algorithms, the information acquisition layer obtains the college talents' personal resume, cover letter, recommendation letter and other information and sends it to the data. The processing layer is processed by the deep learning algorithm's convolutional neural network algorithm. The input layer of the convolutional neural network receives massive information, selects the text vectorization method, and sends it to the convolutional layer. The convolutional layer uses multiple sets of convolution kernels to achieve college talents. Information convolution feature acquisition, the pooling layer implements dimension reduction on the features obtained by the convolution layer and sends it to the fully connected layer. The fully connected layer selects the Relu function as the activation function and sends it to the output layer. The output layer uses the matrix decomposition method to output recommendations. As a result, the data processing layer uses the personalized human-machine interface of the recommendation service layer to output the recommendation results after completing the recommendation. System test results show that the system has a high

收稿日期: 2020-05-07 稿件编号: 202005036

基金项目: 西安航空职业技术学院 2017 年度科研计划项目 (17XHGX-018); 2018 年西安航空职业技术学院教育教学改革研究项目 (18XHJG-026)

作者简介: 陈 慧 (1983—), 女, 陕西富平人, 硕士, 馆员。研究方向: 经济管理。

- 38 -

recommendation success rate and high coverage and real-time performance.

Keywords: deep learning ; college talents ; personalization intelligence recommendation ; convolution neural network ; text vectorization

毕业生就业问题以及用人单位招聘人才问题已成为近年来急需解决的社会问题,高校人才推荐系统可有效提升高校人才信息共享自动化以及智能化,令人才接收单位以及人才推荐单位之间实现良好沟通,优化高校人才资源的配置^[1-2]。随着人类步入信息化以及智能化,网络信息量呈爆炸式增长,人类生活均离不开网络,从众多信息中选取需要的信息有助于提升工作效率节省时间。推荐系统是高效的过滤工具,采用推荐系统可有效解决信息量过载的问题^[3-5]。

卷积神经网络是一种性能优异的深度学习算法,为了解决用人单位用人难以及高校人才就业难的问题,提出了深度学习算法的高校人才个性化智能推荐系统。

1 基于深度学习算法的高校人才个性化智能推荐系统介绍

1.1 系统总体架构

深度学习算法的高校人才个性化智能推荐系统包括信息获取层、数据处理层、推荐服务层,高校人才个性化智能推荐系统总体结构如图1所示。

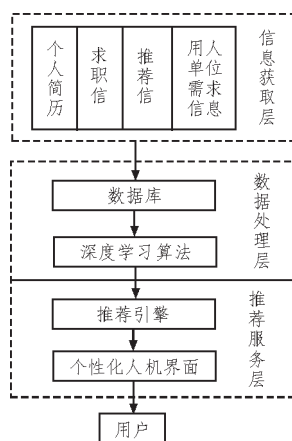


图1 系统总体结构

系统通过信息获取层采集高校人才信息,所采集信息包括个人简历、求职信以及推荐信等各种与就业相关的文档,个人简历中信息包括高校人才姓名、年龄、所属院校、所学专业、获奖情况、各种证书等具体信息;所采集信息发送至数据处理层通过深

度学习算法中的卷积神经网络算法实现高校人才个性化智能推荐^[6],并发送至推荐服务层;推荐服务层针对不同招聘单位需求,将数据处理层获取的推荐结果通过个性化人机界面实现高校人才智能推荐。

系统通过服务器实现客户端以及数据库之间数据交互,系统服务器利用 Web Service 技术以及 MyBatis 框架技术实现,利用 http 通信技术实现服务器与客户端之间的沟通^[7],服务器接收客户端用户请求后返回服务器端,通过深度学习算法实现高校人才个性化智能推荐。

选取 MySQL 数据库为系统提供全部数据支持,利用分散式存储方式提升数据库访问效率,选取 MyBatis 框架技术实现服务器与 MySQL 数据库的连接。MySQL 数据库可满足海量高校人才以及用人单位信息的完整性、一致性以及安全性^[8]。

1.2 系统功能结构

依据高校人才个性化智能推荐需求建立系统的功能结构如图2所示。从图2可以看出,深度学习算法的高校人才个性化智能推荐系统包括高校人才管理模块、招聘方管理模块以及管理员模块。

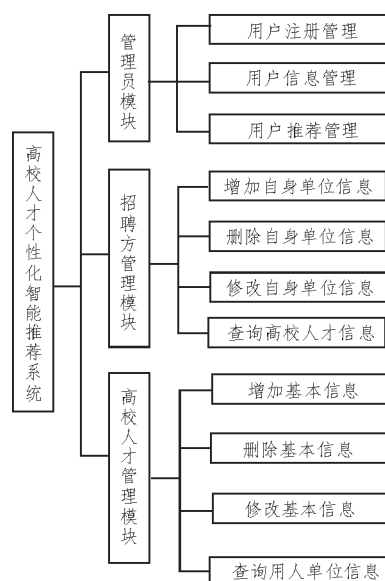


图2 系统功能结构图

高校人才管理模块负责录入以及修改高校人才基本信息,具有增加、删除、修改高校人才基本信息的功能,并可依据自身需求查询全部用人单位信

息^[9]。高校人才用户登录系统后,系统依据高校人才需求为高校人才推荐用人单位信息。

招聘方管理模块负责录入以及修改用人单位基本信息,具有增加、删除、修改自身单位基本信息的功能^[10],并可依据自身需求查询全部高校人才信息。系统依据用人单位需求为招聘方推荐高校人才信息。

管理员模块负责管理高校人才以及用人单位的注册信息和账号信息。

1.3 卷积神经网络

选取深度学习算法中卷积神经网络算法实现高校人才个性化智能推荐^[11],该算法包括输入层、卷积层、池化层、全连接层以及输出层五部分。

1.3.1 输入层

选取文本向量化方法作为卷积神经网络输入层处理方式,将输入海量信息通过单词嵌入方式映射至 n 维向量空间,将全部的输入信息合并^[12]。设输入的海量高校人才数据格式为 $(user_i, c_{i0} \oplus c_{i1} \oplus \dots \oplus c_{in}, \frac{1}{n} \sum_{j=1}^n o_{ij})$,其中 \oplus 表示高校人才信息通过卷积符连接,可得高校人才用户信息 $(user_i, g_i, \bar{o}_i)$ 。其中, g_i 表示高校人才的全部信息, \bar{o}_i 表示用人单位全部信息,利用文本向量化技术将全部信息合并后利用向量 vec_i 表示,即:

$$vec_i = Doc2VecC(g_i) \quad (1)$$

通过式(1)将函数 $Doc2VecC$ 返回至 n 维向量,将用户全部信息表示为文本嵌入的简单平均值,令所获取向量体现训练学习期间全信息的语义内容,用户样本信息通过文本向量化技术转化为:

$$A = (user_i, vec_i, \bar{o}_i) \quad (2)$$

1.3.2 卷积层

在表示矩阵上将所输入的经过文本向量化处理的信息实施卷积,信息卷积与图像卷积形式并不相同^[13],信息可通过一维空间实施卷积,用 $c^j \in R^1$ 表示文本卷积后特征,卷积公式如式(3)所示:

$$c_i^j = f(W_c^j * g_{(i:(i+WS-1))} + l_c^j) \quad (3)$$

式中, $*$ 表示卷积操作, $l_c^j \in R$ 表示偏置向量, f 为卷积层的Relu激活函数,该激活函数可有效避免梯度消失情况^[14]。

特征向量 $c^j \in R^{i-WS+1}$,利用卷积核 W_c^j 的卷积公式如下:

$$c^j = [c_1^j, c_2^j, \dots, c_i^j, \dots, c_{i-WS+1}^j] \quad (4)$$

通过以上过程利用多组卷积核实现高校人才相关信息卷积特征获取。

1.3.3 池化层

利用池化层对卷积层获取的特征实施降维,通过降维保留卷积神经网络中相对重要的特征,避免卷积神经网络过拟合情况。用 $K_t = \{k_1, k_2, \dots, k_{z-s+1}\}$ 与 q_t 表示第 t 个卷积层获取特征图以及池化的结果,选取 K_t 作为池化层中的最大值,可得卷积神经网络池化层处理公式如下:

$$Q_t = \max(K_t) = \max\{k_1, k_2, \dots, k_{z-s+1}\} \quad (5)$$

1.3.4 全连接层

将池化层处理完成的输入结果输入全连接层中,设卷积神经网络全连接层中含有神经元数量为 m ,选取Relu函数作为全连接层激活函数,所获取的固定向量 v_i 即为高校人才信息隐含特征,全连接层计算公式如下:

$$v_i = \text{Relu}(w_i Q_t + l_i) \quad (6)$$

式中, $v_i \in R^m$, Q_t 与 w_i 分别为卷积神经网络池化层输出以及全连接层权重, l_i 表示偏置系数。

1.3.5 输出层

将矩阵分解方法应用于卷积神经网络输出层,矩阵分解方法通过矩阵相乘方法将全连接层内高维并具有稀疏性质信息的矩阵分解为低维度特征矩阵^[15-16],所获取低维度特征矩阵即为最终高校人才智能推荐结果。矩阵分解方法的卷积神经网络输出层公式如下:

$$e_j = CNN(Z, X_j) \quad (7)$$

式中, e_j 表示高校人才推荐信息向量, X_j 与 Z 分别表示用户信息输入以及卷积神经网络权重集合。

2 系统测试结果

选取CPU为Windows XP,硬盘为4 G、内存为8 G的计算机作为文中系统的服务器;选取MySQL8.0作为文中系统数据库;选取Eclipse软件作为文中系统的开发平台;选取Java语言作为文中系统开发语言搭建该系统。统计文中系统于2019年7月10日-2019年7月19日所收集高校人才实例26 806个,其中包括机械、建筑、信息等众多专业的本科以及本科以上的高校人才,统计10天高校人才和用人单位推荐成功率,统计结果如表1所示。从表1测试结果可以看

出,文中系统具有较高的推荐成功率。

表1 高校人才和用人单位推荐结果

日期	高校人才成功率(%)	用人单位推荐成功率(%)
7月10日	76.6	56.1
7月11日	71.9	70.1
7月12日	71.6	61.6
7月13日	72.1	72.7
7月14日	83.5	62.9
7月15日	81.9	56.4
7月16日	73.1	62.1
7月17日	71.2	71.1
7月18日	77.3	58.9
7月19日	75.5	70.3

选取均方根误差作为推荐系统评价指标,均方根误差常应用于推荐系统,均方根误差可有效对比不同系统间的推荐精度,均方根误差值越小表明系统推荐性能越高。选取协同过滤系统以及玻尔兹曼机系统作为对比系统,对比结果如图3所示。从图3测试结果可以看出,采用文中系统推荐高校人才在不同信息量大小均方根误差均较低,文中系统的均方根误差明显低于过滤系统以及玻尔兹曼机系统,验证了文中系统具有较高的推荐精度。

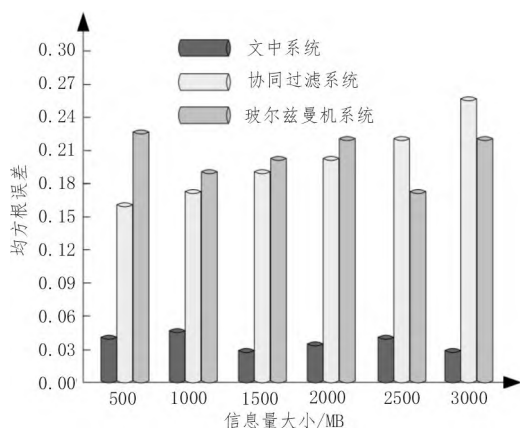


图3 均方根误差对比

统计不同系统推荐过程中的覆盖率,覆盖率直接影响系统推荐性能,对比结果如图4所示。

从图4测试结果可以看出,文中系统具有较高的覆盖率,协同过滤系统覆盖率次之,玻尔兹曼机系统覆盖率最低。测试结果表明,文中系统覆盖率较高,可有效提升系统推荐性能。

统计文中系统、协同过滤系统以及玻尔兹曼机系统作为对比系统的高校人才推荐响应时间,结果如表2所示。从表2可知,文中系统有较高的响应速

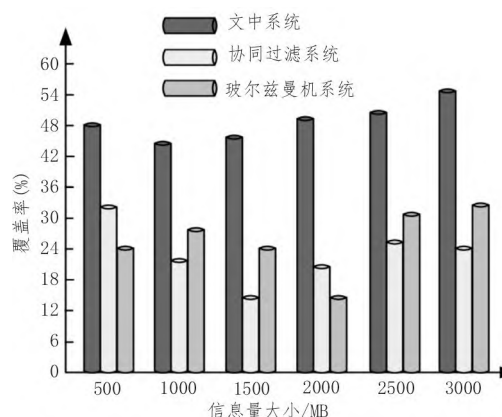


图4 覆盖率对比

度,响应速度均明显高于协同过滤系统和玻尔兹曼机系统。

表2 高校人才推荐响应时间对比

系统名称	登录时间/ms	推荐时间/ms
文中系统	47.9	76.2
协同过滤系统	262.1	447.6
玻尔兹曼机系统	289	425.3

文中系统服务对象为海量用户,系统安全性极为重要,统计文中系统在不同攻击频率下的推荐性能,统计结果如表3所示。由表3测试结果可知,文中系统在不同攻击频率下仍可保持较高的推荐成功率及实时性,文中系统可有效保证数据安全性以及完整性,为高校人才个性化智能推荐提供技术基础,为用人单位以及高校人才供需平衡提供支持。

表3 不同攻击频率下系统运行情况

系统名称	推荐成功率(%)	推荐时间/ms
文中系统	73.92	58.4
协同过滤系统	63.96	454.6
玻尔兹曼机系统	61.91	537.8

3 结束语

由输入层、卷积层、池化层、全连接层以及输出层组成的卷积神经网络构建高校人才智能推荐系统,该系统可有效改善以往人才推荐系统的缺陷,使高校人才以及用人单位解决人才资源合理调配问题,提升人才利用率,增加人才推荐性能,为用人单位、服务企业以及高校人才提供高效、精准的服务,为社会发展创造效益。

参考文献:

[1] 岳昌君,周丽萍.中国高校毕业生就业趋势分析:

- 2003-2017年[J].北京大学教育评论,2017,15(4):87-106,187.
- [2] 武毅英,王志军.质量评价体系探析——基于教育部直属高校2013年就业质量年度报告的数据[J].江苏高教,2015(1):100-104.
- [3] 张学敏,柴晓旭.我国高校毕业生就业率与高校教育质量评价研究[J].东北师大学报(哲学社会科学版),2019(3):131-141.
- [4] 王妍,唐杰.基于深度学习的论文个性化推荐算法[J].中文信息学报,2018,32(4):114-119.
- [5] 李策,张栋,杜少毅,等.一种迁移学习和可变形卷积深度学习的蝴蝶检测算法[J].自动化学报,2019,45(9):1772-1782.
- [6] 高峰,曲建岭,袁涛,等.基于改进差分域特征和深度学习优化的航空发动机剩余寿命预测算法[J].电子测量与仪器学报,2019,33(3):21-28.
- [7] 苑威威,彭敦陆,吴少洪,等.自注意力机制支持下的混合推荐算法[J].小型微型计算机系统,2019,40(7):1437-1441.
- [8] 王瑞琴,吴宗大,蒋云良,等.一种基于两阶段深度学习的集成推荐模型[J].计算机研究与发展,2019,56(8):1661-1669.
- [9] 俸世洲,周尚波.基于深度自编码网络的高校招生咨询算法研究[J].计算机应用,2017,37(11):3323-3329.
- [10] 李宇琦,陈维政,闫宏飞,等.基于网络表示学习的个性化商品推荐[J].计算机学报,2019,42(8):1767-1778.
- [11] 朱建楠,梁玉琦,顾复,等.基于深度学习的机械智能制造知识问答系统设计[J].计算机集成制造系统,2019,25(5):1161-1168.
- [12] 洪雁飞,魏本征,刘川,等.基于深度学习的椎间孔狭窄自动多分级研究[J].智能系统学报,2019,14(4):708-715.
- [13] 廖彬,张陶,于炯,等.QPR-NN:一种结合二次多项式回归与神经网络的推荐算法[J].西安交通大学学报,2019,53(9):79-87,136.
- [14] 王威,张彤,王新.用于图像超分辨率重构的深度学习方法综述[J].小型微型计算机系统,2019,40(9):1891-1896.
- [15] 姚凯,涂平,陈宇新,等.基于多源大数据的个性化推荐系统效果研究[J].管理科学,2018,31(5):3-15.
- [16] 张志鹏,张尧,任永功.基于覆盖约简的个性化协同过滤推荐方法[J].模式识别与人工智能,2019,32(7):607-614.

(上接第37页)

- [7] 葛晓华.翻译学员双向翻译能力比较的实证研究[J].外语教学,2018,39(4):80-85.
- [8] 包乌格德勒,赵小兵.基于RNN和CNN的蒙汉神经机器翻译研究[J].中文信息学报,2018,32(8):60-67.
- [9] 吴丽华.医学文本机辅翻译质量与效率实证研究[J].中国科技翻译,2018,31(1):33-36,52.
- [10] 张丽林,李茂西,肖文艳,等.机器翻译自动评价中领域知识复述抽取研究[J].北京大学学报(自然科学版),2017,53(2):230-238.
- [11] 赵雪琴,徐晗宇,陈莹.视译过程中认知负荷与译文质量相关性研究——以汉语逻辑连词英译为例[J].外语研究,2019,36(4):12-15.
- [12] 阮翀,施文娴,李岩昊,等.基于多译文的中文转述语料库建设及转述评价方案[J].中文信息学报,2018,32(12):67-75.
- [13] 普次仁,侯佳林,刘月,等.深度学习算法在藏文情感分析中的应用研究[J].计算机科学与探索,2017,11(7):1122-1130.
- [14] 沈卉卉,李宏伟.基于动量方法的受限玻尔兹曼机的一种有效算法[J].电子学报,2019,47(1):176-182.
- [15] 吕淑宝,王明月,翟祥,等.一种深度学习的信息文本分类算法[J].哈尔滨理工大学学报,2017,22(2):105-111.
- [16] 李奉栖.基于神经网络的在线机器翻译系统英汉互译质量对比研究[J].上海翻译,2021(4):46-52.