

队伍成员：

曾永家右 黄泰北 赵辉

龙兰心 罗雪蕾

基于大数据的人岗匹配 系统“职达” 求职者画像算法

职达信息技术有限公司

达芬奇队

指导教师：徐进 尹帮旭

目录

一、用户画像简介..... 1

二、求职者画像算法..... 1

 2.1 求职者画像简述..... 1

 2.2 求职者画像数据建模..... 1

 2.3 求职者用户行为分析..... 2

 2.4 求职者画像层次..... 2

 2.5 求职者画像的生成..... 2

 2.5.1 求职者基本信息..... 2

 2.5.2 求职者行为偏好..... 3

 2.5.3 用户的活跃程度..... 3

三、参考文献..... 3



一、用户画像简介

用户画像又称用户角色,作为一种勾画目标用户、联系用户诉求与设计方向的有效工具,用户画像在各领域得到了广泛的应用。我们在实际操作的过程中往往会以最为浅显和贴近生活的话语将用户的属性、行为与期待联结起来。作为实际用户的虚拟代表,用户画像所形成的用户角色并不是脱离产品和市场之外所构建出来的,形成的用户角色需要有代表性,能代表产品的主要受众和目标群体。

用户画像的 PERSONA 七要素

P 代表基本性(Primary): 指该用户角色是否基于对真实用户的情景访谈;

E 代表同理性(Empathy): 指用户角色中包含姓名、照片和产品相关的描述,该用户角色是否引同理心;

R 代表真实性(Realistic): 指对那些每天与顾客打交道的人来说,用户角色是否看起来像真实人物;

S 代表独特性(Singular): 每个用户是否是独特的,彼此很少有相似性;

O 代表目标性(Objectives): 该用户角色是否包含与产品相关的高层次目标,是否包含关键词来描述该目标;

N 代表数量性(Number): 用户角色的数量是否足够少,以便设计团队能记住每个用户角色的姓名,以及其中的一个主要用户角色;

A 代表应用性(Applicable): 设计团队是否能使用用户角色作为一种实用工具进行设计决策。

二、求职者画像算法

2.1 求职者画像简述

这里,我们以一例来说明。张三、男、XX 大学毕业、本科学历、计算机专业、熟悉 Hadoop 体系架构、精通 java 编程语言、3 年工作经验。

这样一串描述即为用户画像的典型示例。如果用一句话来描述,即: 用户信息标签化。由于我们要做的是求职者画像,所以,侧重于能体现出求职者工作能力的标签。

对于求职者而言,必须要完善自己的各种信息,才有更大的可能找到满意的工作。同时,考虑到求职者在找工作时,都是有一定的方向的,与自己的专业或爱好相关,。所以,我们采用数据建模和用户行为分析相结合的方法进行求职者画像。

2.2 求职者画像数据建模

首先,对于求职者而言,基本信息是不可或缺的,包括: 基本信息: 姓名、年龄、性别、毕业院校、学历、专业、技能、工作经验。

其次、就是求职者求职意向，想要在那座城市工作生活，是否有意向岗位。

最后，就是对企业的一些附加要求，想去国企，想去大公司等。

2.3 求职者用户行为分析

统计求职者一个月内浏览岗位纪录，分析其意向城市、意向岗位。

根据求职者浏览频率和时间，预测求职者目前就业状态。

2.4 求职者画像层次

对求职者进行画像的目的有三个：

(1)将求职者的信息及求职意向标签化，方便对其进行个性化的岗位推荐；

(2)通过对求职者在网站浏览行为的分析，发掘用户的偏好，进行辅助的岗位推荐；

(3)通过对用户活跃程度的分析，可以预测求职者是准备跳槽、仍在求职或已找到意向岗位，以此来进行辅助的岗位推荐

所以，我们将求职者画像分为三个部分：

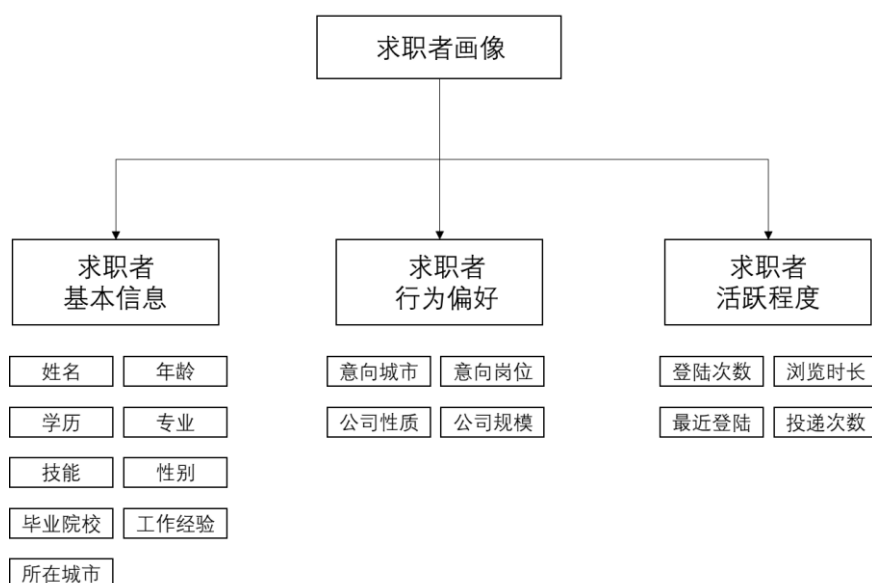


图 2.1

2.5 求职者画像的生成

2.5.1 求职者基本信息

此部分为用户自主填写

2.5.2 求职者行为偏好

此部分信息用户可以选择是否填写。如果用户未填写，我们将通过对用户的历史浏览记录进行统计分析等方法，生成这些标签。

具体的，选取用户最近一个月的浏览记录，从中分别统计出各意向城市、意向岗位、公司性质、公司规模出现的次数，已出现次数最多的作为标签。

2.5.3 用户的活跃程度

登陆次数、登陆时长、最近一次的登陆、投递次数都可以通过对后台数据库的查询得知。

我们采用神经网络的方法对用户目前的就业状态进行预测，首先邀请一批用户填写他们目前真实的就业状态，然后以此为训练集，训练一张神经网络（输入层 4 个神经元，隐含层 9 个神经元，输出层一个神经元）。训练完成后，输入登陆次数、登陆时长、最近一次的登陆、投递次数，即可对用户当前就业状态做出预测。（神经网络部分详见个性化推荐算法）

三、参考文献

[1]李冰,王悦,刘永祥.大数据环境下基于 K-means 的用户画像与智能推荐的应用[J].现代计算机(专业版),2016(24):11-15.

[2]刘海,卢慧,阮金花,田丙强,胡守忠.基于“用户画像”挖掘的精准营销细分模型研究[J].丝绸,2015,52(12):37-42+47.

[3]席岩,张乃光,王磊,张智军,刘海涛.基于大数据的用户画像方法研究综述[J].广播电视信息,2017(10):37-41.

[4]孟巍,吴雪霞,李静,王婧,杜颖,梁雅洁,王艳玫,徐美玲,林晓兰.基于大数据技术的电力用户画像[J].电信科学,2017,33(S1):15-20.

[5]胡媛,毛宁.基于用户画像的数字图书馆知识社区用户模型构建[J].图书馆理论与实践,2017(04):82-85+97.