阿里巴巴大数据智能云上编程大赛

—— 智联招聘人岗智能匹配

1. 目标及数据集描述

详见<https://tianchi.aliyun.com/competition/entrance/231728/information>

2. 特征构造

|  |  |  |
| --- | --- | --- |
| 特征 | | 实现思路 |
| 用户基本特征 | 现居城市 |  |
| 期望工作城市个数 |  |
| 居住地在第几个期望工作城市 | 居住地与用户第一期望工作地点相符返回3，第二返回2；第三返回1；没有返回0 |
| 期望行业个数 |  |
| 第1期望行业 | 行业有2层结构，有两个特征，下同 |
| 第2期望行业 |  |
| 第3期望行业 |  |
| 第4期望行业 |  |
| 期望行业最多父节点 |  |
| 期望行业第二多父节点 |  |
| …… |  |
| 期望职位个数 |  |
| 第1期望职位 | 职位有3层结构，有3个特征，下同 |
| 第2期望职位 |  |
| 第3期望职位 |  |
| 第4期望职位 |  |
| 期望职位最多父节点 |  |
| 期望职位第二多父节点 |  |
| …… |  |
| 期望职位最多子节点 |  |
| 期望职位第二多子节点 |  |
| …… |  |
| 最近工作行业 | 行业有2层结构，有两个特征 |
| 最近工作职类 | 职位有3层结构，有3个特征 |
| 最近工作行业和第一期望的匹配 | 父和子两个特征 |
| 最近工作行业和第二期望的匹配 | 父和子两个特征 |
| …… |  |
| 最近工作行业父节点排总期望行业的第几 |  |
| 最近工作职位和第一期望的匹配 | 父、子、孙三个特征 |
| 最近工作职位和第二期望的匹配 | 父、子、孙三个特征 |
| …… |  |
| 最近工作职位父节点排总期望职位的第几 |  |
| 最近工作职位子节点排总期望职位的第几 |  |
| 期望薪水最高限 |  |
| 期望薪水最低限 |  |
| 最近薪水最高限 |  |
| 最近薪水最低限 |  |
| 期望薪水最高限-最近薪水最高限 |  |
| ……（另外三种相减） |  |
| 学历 | 初中对应10，高中/中专/中技对应15，大专20；本科对应30；硕士/博士/mba/emba 40；其他和空白也为20（学历下同） |
| 年龄 |  |
| 开始工作时间 |  |
| 职位特征 | 公司 | 数据均为空 |
| 城市 |  |
| 职位子类 | 3层结构，有3个特征，下同 |
| 需求人数 |  |
| 最高月薪 |  |
| 最低月薪 |  |
| 开始日期 |  |
| 结束日期 |  |
| 开始和结束时间差 |  |
| 是否要求出差 |  |
| 最小工作年限 |  |
| 最大工作年限 |  |
| 最低学历 |  |
| 最高学历 |  |
| 用户-岗位特征 | 用户当前居住地（live\_city）与岗位城市（city）是否一致 |  |
| 岗位城市（city）与期望工作城市1（desire\_city）一致，返回3；与期望工作城市2一致，返回2；与期望工作城市3一致，返回1，都不一致返回0 |  |
| 职位子类（jd\_sub\_type）与最近工作职类（cur\_jd\_type） | 有三层，是否都匹配 |
| 职位子类（jd\_sub\_type）与第一期望职类（desire\_jd\_type） | 有三层，是否都匹配 |
| 职位子类（jd\_sub\_type）与第二期望职类（desire\_jd\_type） | 有三层，是否都匹配 |
| …… | …… |
| 职位子类（jd\_sub\_type）父节点排总期望职位的第几 |  |
| 职位子类（jd\_sub\_type）子节点排总期望职位的第几 |  |
| 职位标题 |  |
| 岗位最高月薪（max\_salary）vs 用户期望薪水（desire\_jd\_salary\_id） | 有高低2种 |
| 岗位最高月薪（max\_salary）\*最近薪水（cur\_salary） | 有高低2种 |
| 岗位最低月薪（min\_salary）\*用户期望薪水（desire\_jd\_salary\_id） | 有高低2种 |
| 岗位最低月薪（min\_salary）\*最近薪水（cur\_salary） | 有高低2种 |
| 最低学历（min\_edu\_level）-用户学历 |  |
| 最高学历（max\_edu\_level）-用户学历 |  |
| 用户工作年限与最低年限要求 |  |
| 用户工作年限与最高年限要求 |  |