

第二届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

演讲者：咸鱼打酱油

OTTO

Contents 目录

- 01 团队介绍
- 02 赛题分析
- 03 特征工程
- 04 模型构建
- 05 Demo实例
- 06 总结

团队介绍

01

团队介绍

青禹小生

- 2018融360决赛二等奖
- 天池全球城市计算AI挑战赛季军
- 2019拍拍贷第四届魔镜杯大赛 Rank4

OTTO.
Expect Excellence.

Obey Truth Trust Oath

Grand Rookie

- 2019JData亚军
- 2018甜橙金融亚军
- 2019KDD Rank10

咸鱼打酱油

- 第一届云上编程大赛 Rank8

赛题分析

02

赛题解析

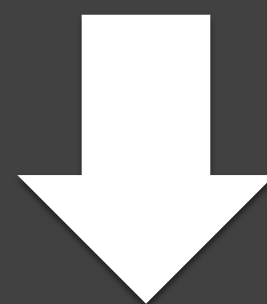
训练排序模型，对求职者的职位候选集进行排序，尽可能使得双端都满意的职位优先推荐。



投递简历



评估简历



求职者是否会投递

用人单位是否满意



特征工程

03

基础特征群

- user的基础信息，如文凭，工作年限，城市，最近薪水，期望薪水
- 最近行业和期望行业是否一致
- user期望的行业count，职类count，经验count
- user投递的jd、jd_title和jd_sub_type的count



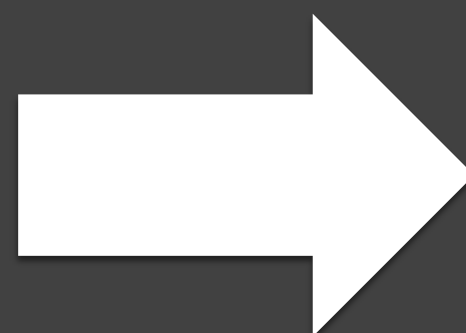
- jd的基础信息，如城市，月薪，需求人数，学历，是否出差，工作年限以及jd有效时长
- 同一jd_title和jd_sub_type下对应的：
 - jd的count，
 - jd的月薪统计 (min, max, mean, std)
 - jd的需求人数统计 (min, max, mean, std)
 - 上述特征继续细化到城市维度
- jd被投递的人数总count，同一jdtitle被投递的总count以及同一jd_sub_type被投递的总count，并细化到城市维度

- jd月薪和user薪水相除
- user和jd需求人数，月薪，jd有效时长的交叉统计
- jd和user的期望月薪，年龄的交叉统计

竞争力特征群

求职者角度：

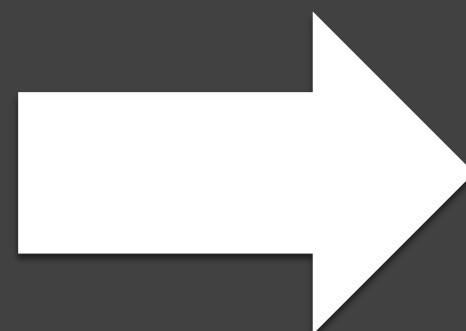
岗位薪资、招聘需求人数



- 当前岗位薪资与同行业整体岗位薪资（均值）对比
- 当前岗位招聘需求人数与同行业整体岗位（均值）对比

用人单位角度

求职者年龄、工作年限、学历、工作经历



- 当前求职者年龄、工作年限、学历与同岗位其他求职者对比
- 当前求职者工作经历和jd描述的匹配与其他求职者对比

文本特征群

字符串相似度

房地产/建筑/建材/工程, 加工制造(原料加工/模具), 贸易/进出口

房地产/建筑/建材/工程

desire_jd_industry_id

jd_sub_type

Levenshtein距离

- 对user的投递岗位和实际对应的jd_no计算字符串相似度

文本相似度

停车|现场|凤凰|预算编制|建设|实习|专家|公园|预算软件|勘察|合同|知识|商务|单位|...
调色员|彩妆|护肤|布料|光源|客户|调色
互联网参考模型osi七层|建设|英语|项目主管|签证|谈判|合同|crm系统|pl/sql|...
行政|酒店|后勤|温泉|文员|文秘|旅游
上色|产品开发|插画|纸张|老师|制作|3dmax|美术|视觉设计|视觉|vi|道具|设计师...



tf-idf



Word2Vec



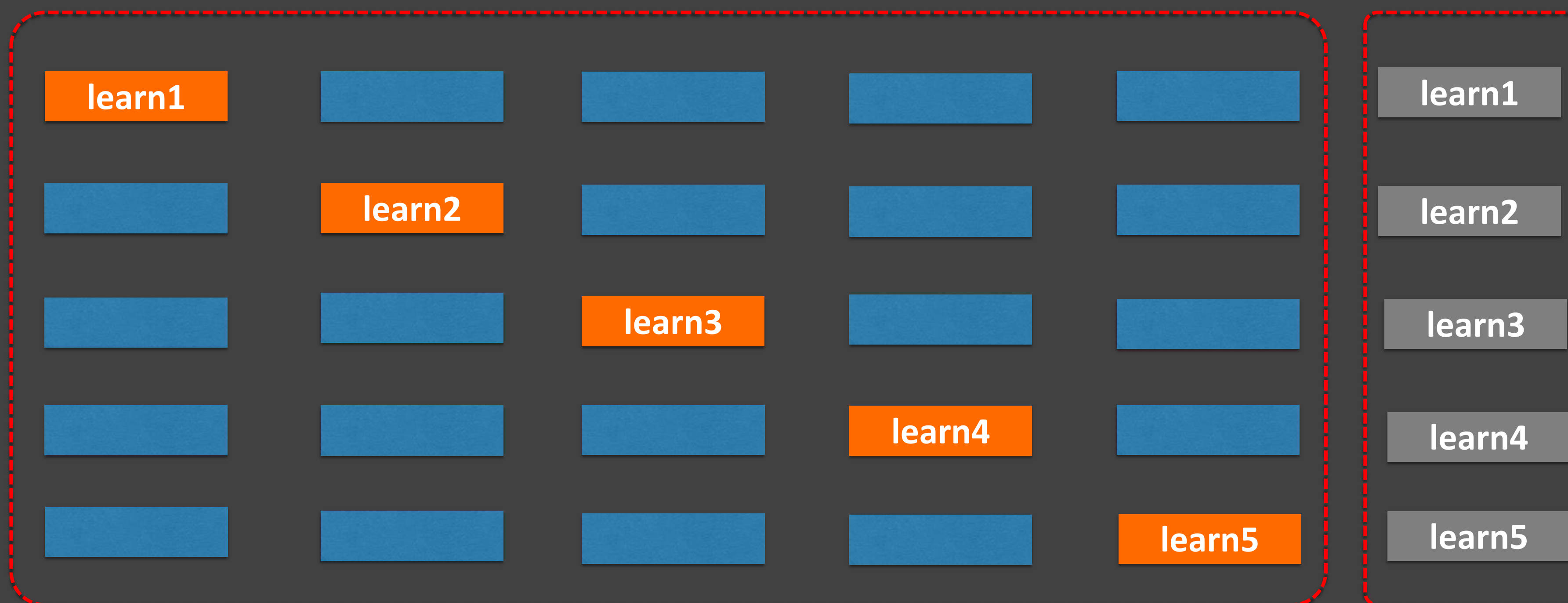
Doc2Vec

- 对user的经验和jd的工作描述分词后做tfidf, Doc2Vec特征
- 对user的经验和jd的工作描述计算文章相似度
- 聚合user投递的jd_title取关键词后做Word2Vec, 并对得出的向量继续聚合统计min, max, mean, std

CTR特征群

Train

Test



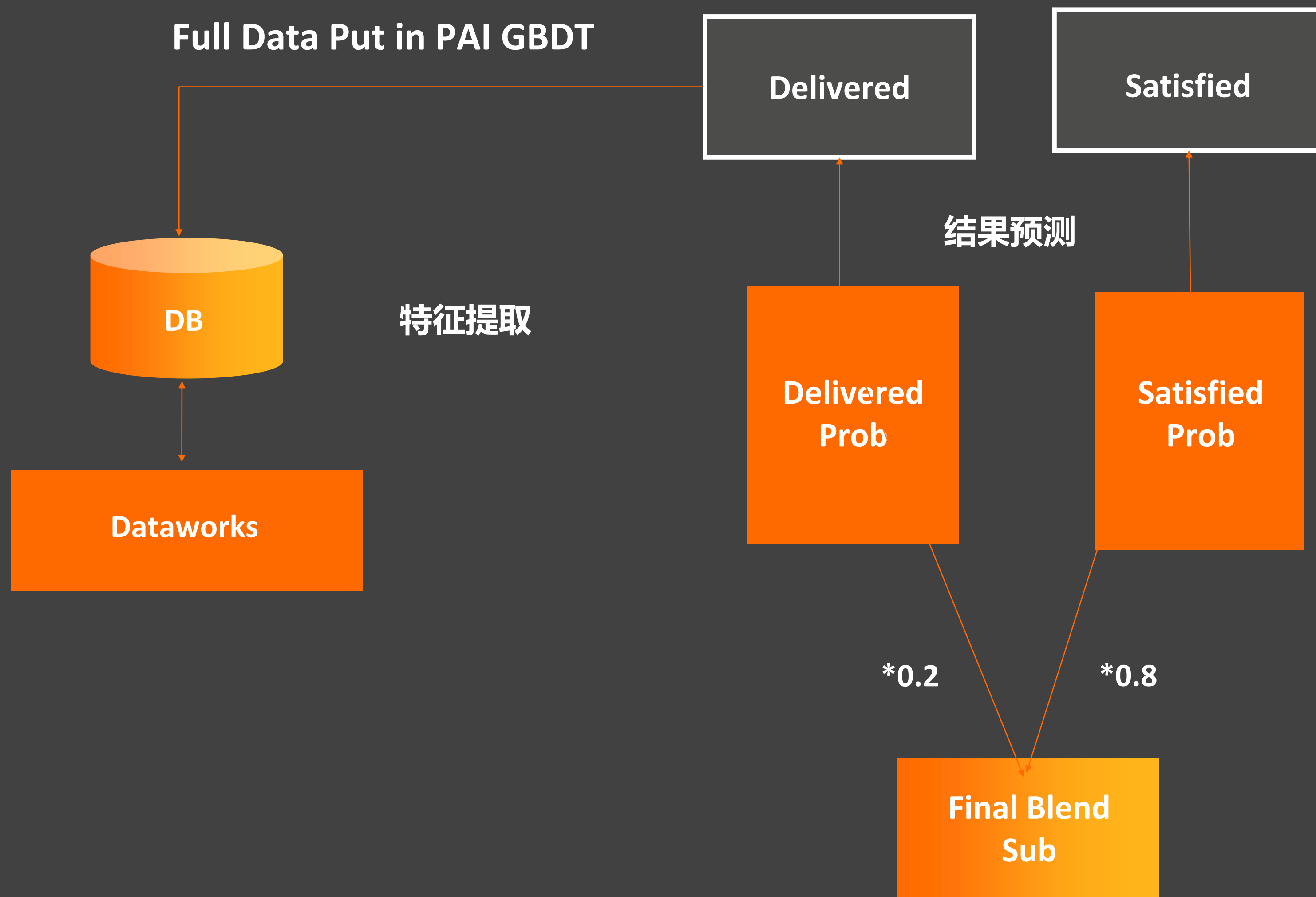
五折交叉构造CTR特征:

- 对jd_no、jd_no+user+学历等分别做satisfied和delivered的CTR
- 对jd_title、jd_sub_type做CTR，以弥补未出现在训练集中的jd而缺失的信息
- 用其他4折预估另外1折样本的CTR，以防止过拟合

模型构建

04

模型构建



Demo实例


05

Demo实例



智联招聘
zhaopin.com

用户首页 企业首页



个人版 > 用户063d436b45ca6719ffc0829a0d6e6c23

OTTO
Expect Excellence.

* 城市: 854

* 职位名称: 会计

搜索

个人投递情况

已投递	数量	岗位ID	岗位名称	月薪	工作年限	要求学历	是否要求出差	截止日期	匹配率
会计	43	73d44	会计	4-6k	1年以内	大专	不要求	2019-12-08	28.7%
会计助理	2	01ffc	会计	4-6k	1年以内	大专	不要求	2019-11-10	26.0%
会计经理	4	60298	会计	6-9k	1年以内	大专	不要求	2019-12-14	24.1%
其他	1	f6878	会计	3-5k	1年以内	大专	不要求	2019-04-11	5.05%
成本会计	1	c61c7	会计	4-6k	3-5年	大专	不要求	2019-04-11	4.03%

用户首页



- ① 用户个人历史投递数据
- ② 用户实时搜索岗位明细及与自身简历匹配度

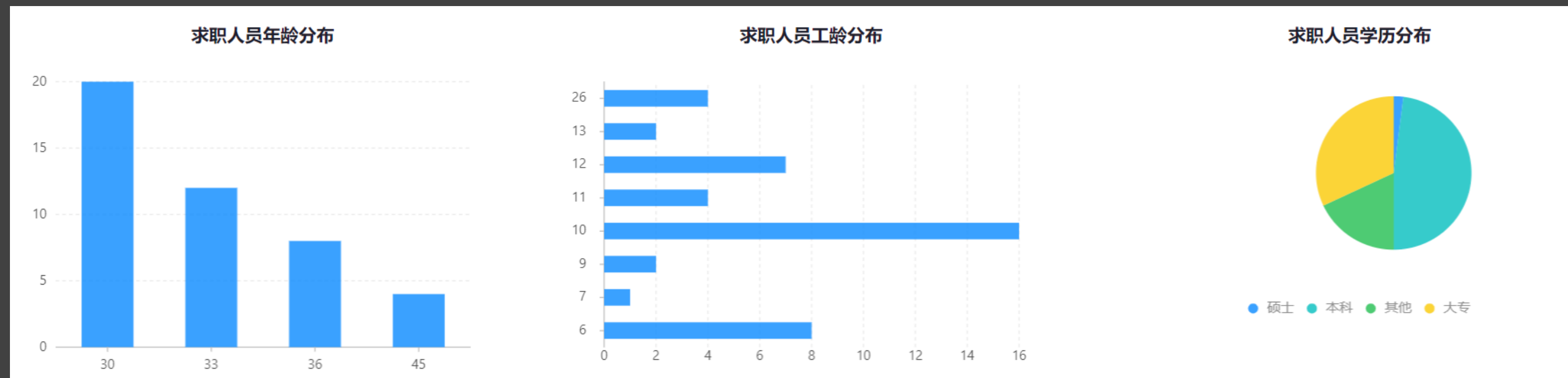


- ① 了解需求岗位的当前投递情况，评估自身竞争力
- ② 投递岗位用户的从业经验热词

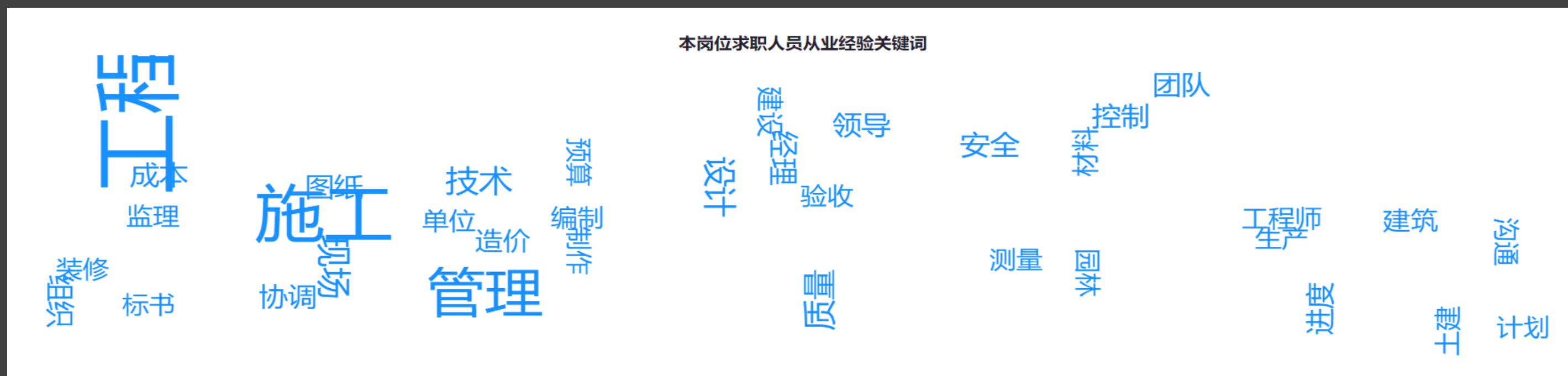
企业首页



- ① 实时获取简历相关度高的优质用户
- ② 当前投递岗位用户的基本特征及与岗位的匹配度



掌握投递岗位用户的用户画像



投递岗位用户的从业经验热词

广告位



顶部轮播广告banner

底部企业宣传视频



总结

06

总结

- 本次比赛团队各成员付出了很大心血，在此想对未能到现场的队友说声辛苦了
- 感谢主办方的辛苦付出，特别是在开发Demo时阿里员工的无私帮助
- 路漫漫其修远兮，勿止于求索，不忘初心，方得始终。

THANKS !