计算机系统应用 ISSN 1003-3254, CODEN CSAOBN
Computer Systems & Applications,2019,28(6):45−52 [doi: 10.15888/j.cnki.csa.006956]

E-mail: csa@iscas.ac.cn
http://www.c-s-a.org.cn
Tel: +86-10-62661041

用户画像在电网设备供应商管理中的应用①

李瑞祥, 黄文涛, 郭欣沅, 张子炎

(国家电网有限公司 国网辽宁省电力有限公司物资分公司, 沈阳 110000)
通讯作者: 郭欣沅, E-mail: 1510282564@qq.com

摘　要: 国家电网在国民经济发展中扮演着重要角色, 选择高质量的供应商合作是其提供优质服务的前提. 许多专家学者在供应商的管理方面做了许多贡献, 但仍存在解决问题单一, 主观因素过强的问题. 针对项目中供应商管理的多应用和研究现状提出了基于用户画像的供应商管理方案, 该方案主要通过用户画像对供应商信用评分、供应商产品选择和供应商长期合作可行性等方面进行辅助决策. 文章主要对画像的标签体系和内容进行了详细阐述, 重点分析了基于自然语言处理的模型标签生成方法. 并结合供应商产品选择案例简述了用户画像在辅助决策上的应用方法, 证明了在画像的基础上的方法可以提高供应商选择的精确度, 简化了业务流程.

关键词: 用户画像; 大数据; 辅助决策; 自然语言处理

引用格式: 李瑞祥,黄文涛,郭欣沅,张子炎.用户画像在电网设备供应商管理中的应用.计算机系统应用,2019,28(6):45–52. http://www.c-s-a.org.cn/1003-3254/6956.html

Application of User Portrait in Grid Equipment Supplier Management

LI Rui-Xiang, HUANG Wen-Tao, GUO Xin-Yuan, ZHANG Zi-Yan

(Material Branch, State Grid Liaoning Electric Power Co. Ltd., STATE GRID Corporation of China, Shenyang 110000, China)

Abstract: The State Grid plays an important role in the development of national economy. Choosing high-quality supplier cooperation is a prerequisite for providing quality services. Many experts and scholars have made many contributions to the management of suppliers, but there are still problems that solve the problem unitarily and subjective factors are too strong. According to the multi-application and research status of supplier management in the project, a user portrait-based supplier management scheme is proposed, which mainly assists the supplier’s credit score, supplier product selection, and supplier's long-term cooperation feasibility through user portraits. The article mainly elaborates on the labeling system and content of the portrait, and focuses on the analysis of the model label generation method based on natural language processing. Combined with the supplier product selection case, the application method of user portrait in assistant decision-making is briefly described. It proves that the method based on the portrait can improve the accuracy of supplier selection and simplify the business process.

Key words: user portrait; big data; decision support; natural language processing

随着国民经济的不断发展电网基础设施建设的规模不断扩大, 国家电网对设备的需求量不断增加, 因此采购的设备量也在稳步增长. 同时, 在信息化不断发展的今天, 电网行业内的大数据特征也越来越明显. 电力系统作为一个庞大的生产系统, 每时每刻都在产生着大量的数据. 单就电力企业内部与电网建设、电网运行和电网检修相关的运行数据以及电网设备供应商的信用数据、区域经济发展数据等外部数据来说就已经

① 收稿时间: 2018-12-17; 修改时间: 2019-01-10; 采用时间: 2019-01-22; csa 在线出版时间: 2019-05-25

很可观, 这些数据中蕴含着较多的商业信息, 如供应商的生产能力、履约能力等. 在大数据环境下, 运用数据分析方法对设备供应商进行全方位量化提取有价值的信息是十分有必要的. 本文阐述的主要内容来自于"东北国网供应商管理系统"中供应商辅助决策模块. 该模块主要用来为国网建立供应商企业信用评级、供应商设备选择推荐和供应商可长期合作可行性分析等多个功能, 系统结构简图如图 1 所示.

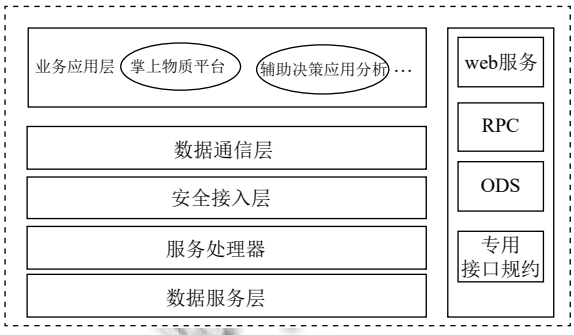

图 1 系统结构简图

针对供应商管理方法的研究已经取得了不错的成果. 以研究过程中典型的三种方案为例来说, 张元新、宋婷等在专家打分的基础上提出了基于 AHP-模糊综合管理[1]的办法对电网物资供应商进行管理, 该方法结合专家打分和 AHP 层次分析法对供应商进行管理, 但是 AHP 层次分析法中各个因素的权重大小需要专家指定, 带有较强的主观性, 模型的鲁棒性较差, 并且对模型的泛化能力不强, 需要针对各个企业进行建模. 随着机器学习的兴起, 樊鹏[2]提出了基于优化的 xgboost-LMT 模型, 该模型通过使用机器学习的自动化学习方法, 有效的缓解了 AHP 方法中主观性过强的问题. 同时席一凡、王超等将模糊神经网络[3]应用于供应商管理中, 与 xgboost-LMT 相比模型的预测精度得到了提升, 但是以上两种机器学习的方法都是针对一个供应商的一个方面进行考察, 难以全方面真实的反应供应商的整体情况. 想要准确的对供应商进行刻画管理, 就需要充分考虑供应商的各方面. 基于以上的思考和调查, 提出了使用用户画像的方法对供应商进行管理的方案. 本文所述的辅助决策模块使用用户画像主要是基于以下考虑: 数据库中, 相关供应商的部分信息已经存在, 但比较散乱无法从数据库中直接获取有用的信息. 但将供应商的数据标签化后, 对供应商就有了一个直观的认识; 分析模块的多个需求发现, 使用用户画像十分有效. 供应商的企业形象是根据其商业行为不断变化的, 使用用户画像可以及时自动更新供应商的标签内容, 进而保持供应商整体形象的动态更新, 动态满足上层需求的调用. 基于以上思考构建了用户画像管理模块, 并在行业专家[4–11]和大数据工程师的合作下构建了供应商的画像标签体系.

1 供货商画像标签体系构建

本文的标签体系分为三级标签, 其中一级和二级标签属于抽象标签, 没有使用意义, 只有统计意义, 在构建过程中只对第三级标签进行填充. 在选择标签时, 我们首先咨询企业管理专家和供应链管理专家在实际工作中常用的考察指标. 然后借鉴贺绍鹏[4]、杨志和[5]、徐晋[9]等学者在标签选择时的经验. 最终推演得到, 要全面的考察一个供应商, 需要从供应商的产品和服务入手, 并且需要考虑企业的基本信息、信用记录和财务信息. 在最终确定三级标签内容时, 由相关行业专家来再次提炼标签尽可能减少标签间映射信息的交叉, 简化整个标签体系. 例如选择反应企业财务状况的标签时, 咨询了企业财务总监, 由专家结合我们实际需求进行标签的再次的提炼简化, 使得最终的标签体系具有更多的科学性和客观性.

最终我们为供应商建立了图 2 所示 (其中标号 U 表示该标签的代号) 的三级标签体系.

1.1 标签体系分析

1.1.1 信息 (U1)

一级标签信息 (U1) 下属有三个二级标签, 企业基本信息 (U11), 信用状况 (U12) 和财务状况 (U13), 图 3 所示.

企业基本信息 (U11) 下属企业简介 (U111)、企业地址 (U112)、企业官网 (U113) 和企业规模 (U114) 四部分. 信用状况 (U12) 下属信用中国 (U121) 至生产能力评估 (U125) 五个部分, 其中 U121 是指信用中国官方对一个企业的信用评估; U122 表示建立合作关系后, 供应商履行合约的能力; U123 指供应商企业参与的诉讼案件, 主要关注于诉讼案件的案由和最终的诉讼结果; U124 指供应商企业具有的国家相关部门颁发的资格; U125 是行业专家通过对企业的技术实力、生产装备和试验设备的考察给出的评估. 财务状况 (U13) 下属总资产周转率 (U131), 资产负债率 (U132) 和流动资金 (U133) 三个部分.

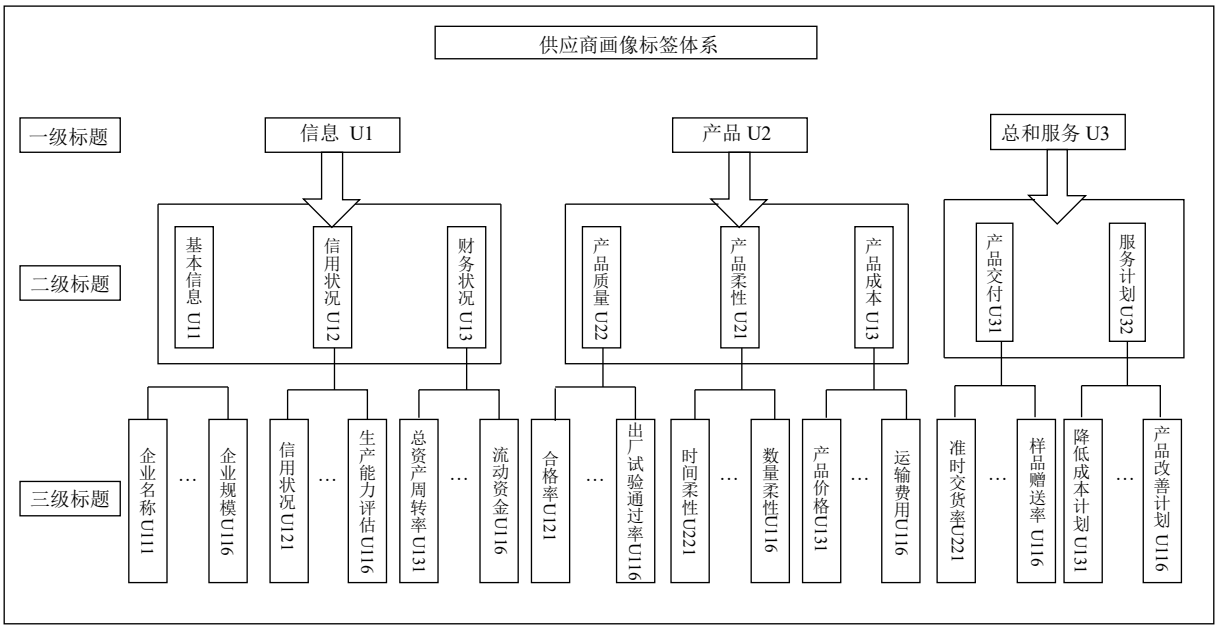

图 2 供应商画像标签体系

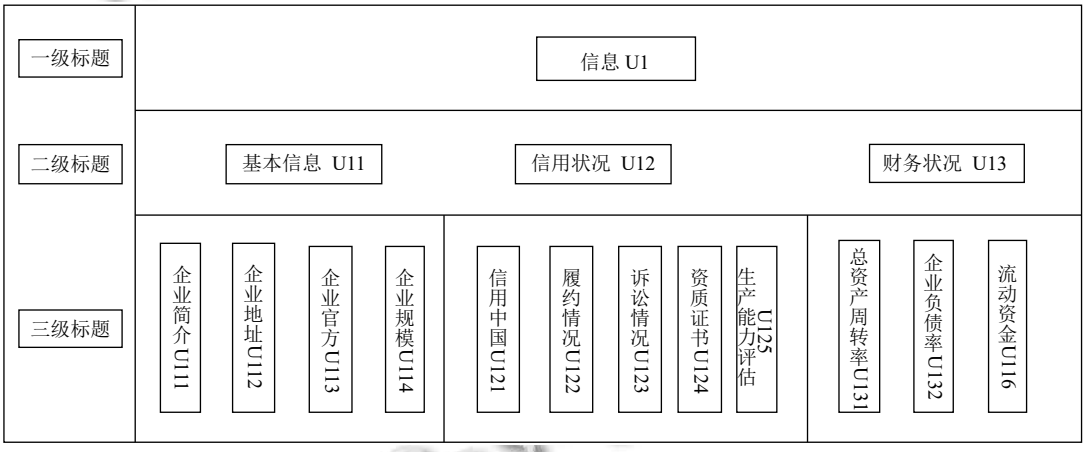

图 3 信息标签体系

1.1.2 产品 (U2)

一级标签产品 (U2) 下面有三个二级标签, 产品质量 (U21), 产品柔性 (U22) 和产品成本 (U23), 如图 4 所示.

产品质量 (U21) 下属合格率 (U211) 至出厂试验通过率 (U214) 四个较为直观的标签. 产品柔性 (U22) 是指供应商在应对外部环境改变时有效的处理能力, 一定程度上反映了企业的承受能力. U221 指供应商从接到产品需求单到正常送货到达时间可以调整的幅度; U222 指企业在一定时期生产新产品的能力, 反映了供应商的产品研发能力; U223 指企业一定时间可以承受的产品订购数量的变化能力. 产品成本 (U23) 下属产品价格 (U231)、获得成本 (U232) 和运输费用 (U233) 三个标签, U232 指企业在一次采购活动上整个链条的总成本, 也就是供应商将产品送达企业整个过程的费用.

1.1.3 综合服务 (U3)

一级标签综合服务 (U3) 下面有两个二级标签, 图 5 所示. 产品交付 (U31), 服务计划 (U32).

产品交付 (U31) 下属准时交货率 (U311) 至样本赠送率 (U315) 四个标签. 其中 U312 表示供应商能否及时响应客户订单, 国网部分对订单的响应时间越来越敏感, 因此订单的响应能力很多程度上反应了供应商

的交货能力. U314 反应了可以节省的资金额度同时也可以测试产品供应的稳定性.

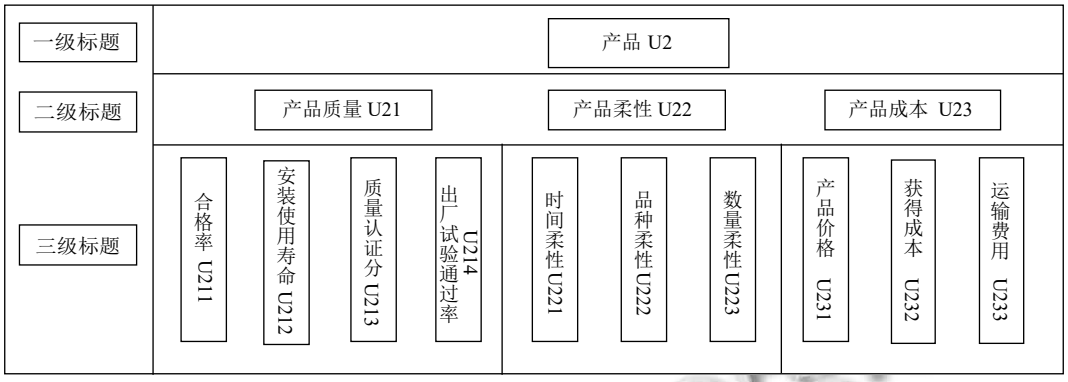

图 4 产品标签体系

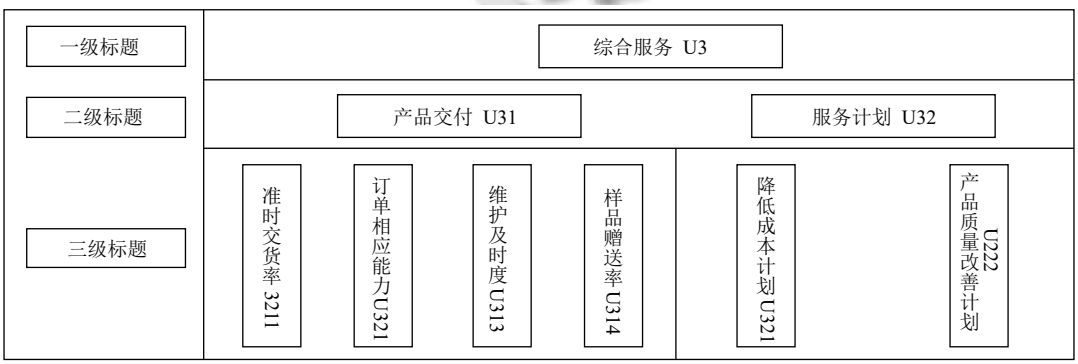

图 5 综合服务标签体系

1.2 标签内容构建

用户画像构建标签分为事实标签和模型标签. 事实标签可以从数据库直接获取或者通过简单的统计得到. 这类标签的构建比较容易但需咨询行业专家给出具体的衡量指标比如 U122, U124, U211, U212, U22, U23, U31 和 U32. 模型标签的构建是标签体系的核心需要机器学习和自然语言处理的知识, 如 U111, U112, U113, U114, U121, U122, U123 和 U13.

1.2.1 事实标签构建

事实标签是指可以解释得到的标签, 具有可量化性. 选取代表性的几个标签来说明事实标签的构建过程.

(1) 合格率 (U211)

在一定时间 T 内, 国网企业采购了 M 件某电力设备, 其中合格的产品数量为 H. 则这一产品的合格率 R 如式 (1) 所示:

$$R=\frac{H}{M}\times 100\% \tag{1}$$

合格率是衡量产品质量的一个重要指标.

(2) 时间柔性 (U221)

供应商相比正常到货时间可以调整的幅度为 A, 合同送货时间为 B, 则时间柔性 R 的计算如下:

$$R=\frac{A}{B} \tag{2}$$

(3) 品种柔性 (U222)

时间为 T, 新产品的种类数量为 N_{new}, 产品品种总数为 N, 则品种柔性的计算公式为:

$$R=\frac{N_{\text{new}}}{T\times N} \tag{3}$$

(4) 数量柔性 (U223)

在一定时间 T 内, 企业可以生产产品的最大数量为 $N_{\max}$, 最小数量为 $N_{\min}$, 这段时间产品的平均需求量为$\overline{N}$. 则计算公式如下:

$$R=\frac{N_{\max}-N_{\min}}{T\times N} \tag{4}$$

(5) 获得成本 (U232)

在一定的时间 T 内单位产品的获得成本是 P, 供应商的产品的成交价格是 C, 采购量是 N, 其所花费的订货费用是 F, 其中的订货费用和企业与供应商合作的方式有关系, 如果采用供应商管理库存的方式, 那么企业的运输与库存就分给供应商, 订货费用是从供应商那里出厂的价格, 那么高额的运输和库存管理将由企业来承担.

$$P=\frac{C\times N+F}{N} \tag{5}$$

(6) 订单响应能力 (U312)

相关行业专家给出了刻画订单响应能力的评价指标描述, 如表 1 所示.

表 1 订单响应能力评分

等级	描述
优	有详细的订单响应措施, 并且该措施得到良好执行
良	有详细的订单响应措施, 并且该措施执行情况一般
中	有详细的订单响应措施, 并未执行
差	无详细的订单相应措施

(7) 资质证书 (U124)

资质证书评价表如表 2 所示.

表 2 资质证书情况评分

等级	描述
优	营业执照, 税务登记证, 组织机构代码证等证件齐全
良	营业执照, 税务登记证, 组织机构代码证等证件齐全, 特殊资质缺少
中	营业执照, 税务登记证, 组织机构代码证等证件齐全, 无特殊资质
差	营业执照, 税务登记证, 组织机构代码证及特殊资质均缺少

分析以上举例发现, 事实标签的数据主要来自数据库, 并且对相关描述的指标构建也来自相关领域的专家. 但是在通过指标构建相应评分时, 为了防止加入过多的主观因素对后面标签的使用产生影响, 使用 Sigmod 函数作为相关标签的自动打分函数. 首先将函数变量初始化为–1, 然后根据与描述中各项的对应情况进行累加. 例如, 在订单响应能力评分中无详细的订单响应措施加 0, 有订单响应能力加 1, 得到良好执行的再加 1, 执行情况若一般则加 0.5, 有但是未执行的加 0. 通过上述方式进行描述情况的表示, 然后送入 Sigmod 函数中得到打分结果. 将函数变量初始化为–1 是为了将打分结果扩展到 0~1 之间. 这样在使用标签时就不用再次进行数据的归一化操作. 通过使用打分函数可以有效的避免专家直接打分中的主观因素, 使得标签内容更加客观可信.

1.2.2 模型标签的构建

模型标签的构建主要考虑两个方面, 标签内容来源和算法的选择. U111, U112, U113, U114, U123, U124 来自网站企查查, 该网站数据整合自官方数据库真实全面. U122 一部分信息来自北极星电力新闻网的电力供应商专栏, 另一部分来自数据库中过去合作的信息. U121 来自网站信用中国的评分. U13 标签是非必要生成标签, 如果供应商是上市企业则通过发布的年报得到, 非上市公司则不予生成.

(1) 部分关键算法

① 中文分词

中文不同于其它语言, 中文表达的基本单位是词而不是字, 所以要想理解中文首先将句子划分为词. 分词算法分为规则分词和统计分词两种方法. 当前比较成熟分词工具一般都是结合两个方案的混合分词技术. 一般是先基于词典的方式进行分词, 然后再用统计的分词方法进行辅助. 这样在保证词典准确律的基础上, 对词典中的未登录词也有较好的识别.

本文使用了当下效果比较好的分词工具——jieba 作为分词器, 作为文本处理的第一步工作.

② 关键词提取算法

关键词是代表文章重要内容的一组词. 在文本自动摘要、关键词提取等起重要的作用. 本文使用了 LDA 算法[12]作为主要的算法模型. LDA 算法拟合出词-文档-主题的分布, 算法假设文档中主题的先验分布和主题词的先验分布都服从狄利克雷分布. 通过对训练文本的统计, 就可以得到每篇文档中主题的多项式分布和每个主题中的词的多项式分布, 然后通过贝叶斯学派的方法, 通过先验的狄利克雷分布和观测数据得到的多项式分布来推断文档中主题的后验分布和主题中词的后验分布.

算法模型如图 6, 其中最大的虚线框 D 表示训练语料的文档集合, K 表示主题的集合. θ_d 表示文本 D 中的主题分布中抽样得到的主题, 这个分布服从参数为 $\bar{\partial}$ 的狄利克雷分布 (DIR), 即

$$\theta_d \sim \frac{\Gamma(\partial_0)}{\Gamma(\partial_1)\cdots\Gamma(\partial_k)}\prod_{k=1}^{K}\mu_k^{\partial_k-1} \tag{6}$$

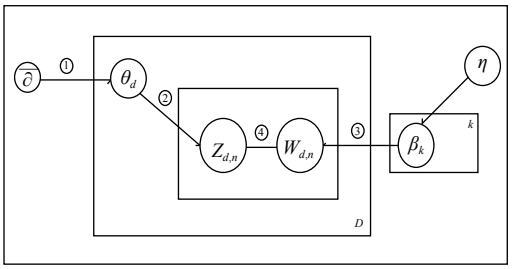

图 6　LDA 算法模型

η表示每个主题分布对应的参数, β_k表示用第K个主题来生成文字. $Z_{d,n}$表示从主题分布中产生主题, 服从多项式分布即

$$Z_{d,n} \sim \frac{n!}{n_1!\cdots n_d!} p_1^{n1}\cdots p_k^{n_k} \sum_{i=1}^{d} n_i = n \tag{7}$$

$W_{d,n}$表示从确定的主题d中产生文字, 同样服从多项式分布.

综上所述, 可以将 LDA 的算法流程整理得到:

算法 1. LDA 算法

for all topics $k\in[1,K]$ do
　sample mixture component $\beta_k \sim Dir(\eta)$ · ①
end for
for all documents $d\in[1,D]$ do
　sample mixture proportion $\theta_d \sim Dir(\bar{\partial})$····· ②
　for all words $n\in[1,N]$ do
　　sample topic index $Z_{d,n} \sim Mult(\theta_d)$·····③
　　sample term for word$W_{d,n} \sim Mult(\beta_{Z_{d,m}})$···· ④⑤
　end for
end for

LDA 算法属于统计模型, 使用之前需要进行预训练得到概率分布的参数. 求解模型的参数一般使用 Gibbs 采样或者 EM 算法来求解. 本文所述的 LDA 算法主要用在电网行业的文本中, 所以使用来自于北极星电力新闻网的网页组成的语料库作为训练语料进行模型训练.

③ 命名实体识别

文中在生成某些标签时, 需要关注供应商名称、机构名称或者事件发生的时间等, 这些名词在语言中被称为命名实体. 本文采用基于条件随机场的命名实体识别算法实现命名实体的识别.

条件随机场是一种在给定观察的标记序列下, 计算整个标记序列的联合概率的方法. 如$X=(X_1, X_2,\cdots,X_n)$和$Y=(Y_1,Y_2,\cdots,Y_n)$是联合随机变量, 若随机变量 Y 构成一个无向图 $G=(V,E)$表示的马尔科夫模型, 则其条件概率分布$P(Y|X)$ 称为条件随机场:

$P(Y_v|X,Y_w,w\neq v)=P(Y_v|X,Y_w,w\sim v)$, 其中$w\sim v$表示图$G=(V,E)$中与结点v右边连接的所有节点, $w\neq v$表示结点v 以外的所有节点. 其图结构如图 7 所示.

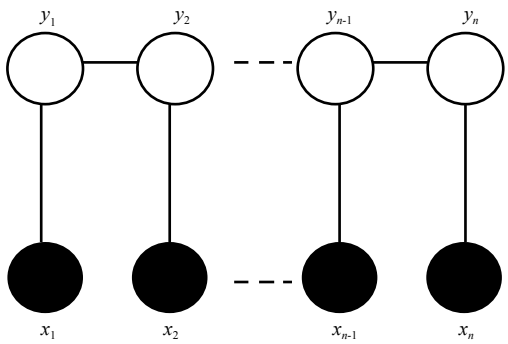

图 7　马尔科夫图模型

假设现在以企业名称识别为例. 做如下标记, 表 3:

表 3　命名实体识别结构构建

标注	含义
B	当前词为企业名称命名实体的首部
M	当前词为企业名称命名实体的内部
E	当前词为企业名称命名实体的尾部
S	当前词单独构成企业名称命名实体
O	当前词不是地理命名实体或组成部分

在训练样本中每个字的标签都在已知的标签集合中选择 (“B”, “M”, “E”, “S”, “O”),

x 是字序列, y 是字对应的标签序列. 训练条件随机场模型的过程就是将已经标注好的训练样本输入初始模型中, 迭代求解特征函数和对应特征函数权重的过程, 训练的目标函数为:

$$P(y|x)=\frac{1}{Z(x)}\exp\left(\sum_j\sum_i w_j f_j(y_{i-1},y_i,x,i)\right) \tag{8}$$

其中, $Z(x)=\sum_y \exp(\sum_j\sum_i w_j f_j(y_{i-1},y_i,x,i))$, w_j表示第j个特征函数的权重, f_j表示特征函数. 本文训练模型的数据集是开放的人民日报语料.

(2) 应用分析

与供应商企业相关的文本处理相比于传统的文本处理更加困难. 因为相关文本大多是短文本, 而传统的文本处理方法会导致文本语义特征稀疏和语义敏感等问题. 所以对企业相关的文本预处理时使用了词性标注以及拼音序列的表征.

考虑到标签数量较多且标签之间有重复使用算法

的现象, 所以选取几个典型的标签来举例. 这里上海某电器集团为来说明.

① 企业简介

本标签主要是对爬取的企业简介文本做关键词提取分析. 这部分相对于其他模型标签构建方法简单, 直接对文本进行分词处理, 分词时要对常见的企业词重点关注比如“上市”、“融资”等. 分好词的文本直接输入的训练好的 LDA 模型中然后输出相应的关键词. 原文和关键词对比见表 4.

表 4 原文与 LDA 处理结果对比

部分原文	使用 LDA 后得到的标签内容 (10 个)
xx 集团股份有限公司是中国装备制造业最大的企业集团之一, 具有设备总成套、工程总承包和提供现代装备综合服务的优势…正在成为一个主业突出、优势明显, 可持续发展的现代化、国际化大型装备集团.	装备制造业, 装备服务, …, 变压设备, 国际化

② 诉讼情况

分析企业的诉讼情况需要关注案件发生的企业双方, 缘由和最终的判决结果. 但是有关诉讼的文本比较短, 且关键性的词语和命名实体比较密集. 所以本质上需要对文本的主要的内容进行语义分析. 获取一条诉讼文本后, 首先进行句法分析得到句法分析树, 根据句法分析树和基于条件随机场的命名实体算法识别出原告和被告的关系和名称.

表 5 诉讼标签提取情况举例

部分原文	标签内容 (原告、被告、案由、结果)
原告 xxx 有限公司与被告 xx 有限公司买卖合同纠纷一案, … . 裁定如下: 准予原告 xx 有限公司撤回对被告 xx 有限公司的起诉.	xxx 有限公司, xx 有限公司 , 买卖合同纠纷, 撤回起诉

诉讼情况的得分的计算方案为:

$$\text{score}=\frac{\sum_{j=1}^{t} w_j c_j}{\sum_{i=1}^{k} w_i c_i} \tag{9}$$

其中, w_i表示i类纠纷的权重, c_i表示i类纠纷的计数, 如果裁定结果为撤诉则不参与计数. 分子的表示与实际需求最相关的t类诉讼案件, 比如当关注于供应商的产品时, 则主要选择与产品相关的诉讼案件作为分子.

2 画像效果评估

本文评估用户画像效果的方法是计算准确率、和是否有时效性机制, 这也是用户画像评估中最常用的方法.

2.1 准确率

准确率指被打上正确标签的用户比例. 准确率是用户画像最核心的指标, 计算公式是:

$$\text{准确率}=\frac{|U_{tag=true}|}{|U_{tag}|} \tag{10}$$

其中, $|U_{tag}|$表示被打上标签的用户数, $|U_{tag=true}|$表示有标签用户被打对标签的用户数.

具体的评估方法为: 随机抽取 15 家合作过的供应商企业, 行业专家首先对供应商进行标注, 并把经过两轮审核后得到的标注结果当作准确的样本. 然后再有新一批专家和自动化模型通过进行标注, 并根据准确样本计算两者标注的准确率, 为了提高评估结果的准确性, 进行 3 组相同的标注过程. 3 组的对比情况如表 6 所示.

表 6 模型准确率测试结果 (%)

组号	专家标注	模型标注
1	93.1	92.6
2	94.8	95.9
3	91.5	91.3

3 画像效果评估

假设国网现在想选择一家变压器供应商购进一批变压器, 首先给出一系列期望的变压器参数, 比如使用寿命, 价格, 安装时间等. 然后将这些参数组合成目标模板. 选择多家供应商的相关标签计算与目标模板的相似度. 根据相似度的分值, 对供应商进行排名. 排名越靠前表示推荐力度越高.

具体实验过程为: 从历史最优采购记录中选取了 20 种设备. 每种设备选取了同时期的 39 家供应商作为干扰项, 加上最优供应商一共 40 家. 然后对每家设备供应商使用 GloVe 算法提取特征, 此其中 GloVe 算法百万数量级的词典和上亿数据集上可以进行快速训练. 提取特征后进行与目标模板进行相似度计算得到一个结果. 同时使用常见的 AHP 和 Xgboot_LMT 算法进行分析得到的最终精确度比较见表 7.

表 7 模型应用准确率 (%)

算法	AHP	Xgboost_LMT	本文
准确率	87. 4	86.6	91.3

4 总结与展望

本文以“辅助决策模块”为实际应用背景. 通过使用用户画像的方案对供应商的数据进行了有效的组织. 在行业专家和大数据工程师的共同参与下, 使用自然语言处理和机器学习的方法, 构建了自动更新的供应商画像标签体系, 通过评估该画像体系取得了比较高的得分. 通过使用用户画像技术简化了开发流程, 提高了系统的工作质量.

但是系统在标签构建的内容上比较繁琐, 并且在构建算法的调优上还有所不足. 后期需要逐步探索更加便捷的标签内容, 并且随着数据量的增加需要对相关算法进行重新训练提高标签内容提取的准确率.

参考文献

1 张元新, 宋婷, 何灵, 等. 基于 AHP-模糊综合评估方法的电网物资供应商评估模型构建. 管理科学与工程, 2017, 6(4): 147–153. [doi: 10.12677/MSE.2017.64018]

2 樊鹏. 基于优化的 XGboost-LMT 模型的供应商信用评价研究[硕士学位论文]. 广州: 广东工业大学, 2016.

3 席一凡, 王超, 聂兴信. 基于模糊神经网络的供应链绩效评价方法研究. 情报杂志, 2007, 26(9): 77–79. [doi: 10.3969/j.issn.1002-1965.2007.09.025]

4 贺绍鹏, 李屹, 邹兰青. 大数据环境下供应商评价设计与分析. 物流技术, 2018, 37(2): 96–100. [doi: 10.3969/j.issn.1007-1059.2018.02.013]

5 杨志和. 基于大数据技术的供应商情报管理系统. 上海电机学院学报, 2018, 21(2): 21–27. [doi: 10.3969/j.issn.2095-0020.2018.02.004]

6 刘海鸥, 孙晶晶, 陈晶, 等. 用户画像模型及其在图书馆领域中的应用. 图书馆理论与实践, 2018, (10): 92–97–23.

7 雷海燕. 基于 Hadoop 的供应商评价系统的研究与设计[硕士学位论文]. 西安: 西安科技大学, 2015.

8 徐晋, 綦振法. 基于神经网络专家系统的供应商信用等级分析. 情报科学, 2004, 22(2): 2.

9 贾安超, 周刚. 基于粗糙集和 BP 神经网络的供应商选择研究. 物流技术, 2012, 31(12): 229–232, 267. [doi: 10.3969/j.issn.1005-152X.2012.12.079]

10 刘昌法. 基于模糊神经网络的供应商选择技术研究. 航空制造技术, 2014, (1-2): 137–139.

11 Li HP, Liu J, Zhang SW. Hierarchical latent dirichlet allocation models for realistic action recognition. Proceedings of 2011 IEEE International Conference on Acoustics, Speech and Signal Processing. Prague, Czech Republic. 2011. 1297–1300. [doi: 10.1109/ICASSP.2011.5946649]

12 Wu T, Shunk D, Blackhurst J, *et al*. AIDEA: A methodology for supplier evaluation and selection in a supplier-based manufacturing environment. International Journal of Manufacturing Technology and Management, 2007, 11(2): 174–192. [doi: 10.1504/IJMTM.2007.013190]