

打败人类围棋手，会看病会开车，它还有什么不能做的呢？

# 政协大数据分析民意报告火了 未来AI要做参政议政“办公神器”

又到一年一度的“两会时间”。就像挑战人类棋手AlphaGo一样,大数据和人工智能与“两会”又能碰撞出什么火花?大数据分析能否将真实的舆情与民意挖掘出来,辅助委员代表参政议政?

本报记者 蔡宇丹

## 这份报告为什么火了

今年省两会期间,山东省政协首次利用大数据精准分析舆情热点,形成民意分析报告提供给政协委员做参考的这条新闻被刷屏。

“谁都没想到,这个分析报告在全国引发这么大关注。”这个项目的合作方,浪潮云服务集团数据服务事业部总经理徐宏伟对齐鲁晚报记者说,这个分析报告对委员确定提案选题提供了一定参考,其实它的市场空间和想象空间非常大。

齐鲁晚报记者了解到,这个分析报告采用了240万条数据,来源于省政协“民声连线”活动的网友评论。浪潮天元大数据平台对这240万条信息产生的数据的热度进行了挖掘。比如,山东网民最关注养老、楼市、环保问题,而最关心山东两会的外地城市是北京;关注省两会的男性是女性的5倍。

徐宏伟表示,这个报告还不像“百度迁徙”那样,是真正的大数据产品,但它在社会上引发的关注,足以说明公众对以大数据分析为技术驱动,来辅助委员代表参政议政的期待。

那些点击量高的社会事件能代表真实的舆情吗?互联网上的民声民意多了,论坛、社区、新闻……徐宏伟说,这就像巨大的岩浆的运动,基于互联网的数据搜索得出的舆情,就像我们看到的火山,舆情可能更倾向于已经公开的这部分意见,或者愿意表达、能够表达的这部分人的意见。实际上,地下剧烈的岩浆的变化还需要社会调查,还需要结合大数据来完成。

## 精准画像到社区

“大数据的真正价值在于多维度的交叉运用。”徐宏伟说,就这个分析报告来说,如果能对12345的数据进行全面深度挖掘,和互联网上的消费数据、城市交通、食品安全、教育、娱乐消费、医疗健康等各方面数据进行关联,再和政府公开数据进行交叉关联,得出的结果就非常精准了。政府手里掌握每一个城市居民的年龄、性别、收入、身体健康、教育状况。这种组织数据是结构化数据、网状数据,比较均匀,上述三方面数据进行交叉关联,就能对不同年龄不同阶层的人群进行精准画像。这种人群画像可以精准到济南市某小区老年活动场所出现的问题,这就能对政府部门落实代表委员的提案议案提供真正有操作性的指导,这样的大数据舆情分析就能产品化。

而互联网数据是爆发性的,不规则的,越热越显现。针对网上某一热点进行热度分析的热搜、排行榜等,它们不是大数据产品,而是新闻或广告。

徐宏伟说,作为国内政务云



图为山东省政协官方网站“民声连线”页面。这个民意分析报告就来源于省政协“民声连线”活动的240万条网友评论等数据。

第一把交椅,浪潮积攒了大量政府数据。现在首先要通过大数据技术,建立一套甄别机制进行聚类分析。比如针对城市交通拥堵,退休人员会呼吁发展城市公交,但年轻人可能更欢迎摩拜单车。这并不是说,谁在互联网上喊得响,就要解决谁的问题。

最典型案例是2010年奥巴马与罗姆尼的总统选战,奥巴马竞选团队制作了数据模型,将收集上来的选民信息输入这个模型,生成了一系列关于选民态度和偏好的报告。技术人员据此开发出一套软件,通过这个系统得出的选民支持率与实际误差不到2.5%。

徐宏伟说,浪潮现在和济南市政府联合打造数创公社,把济南市政府公开数据和政务资源建成一个统一的大数据中心,进行大数据分析,挖掘,建模,解决智慧城市关注的城市治理每一个领域的民生舆情。

## 像AlphaGo一样聪明

今年全国两会期间,浪潮准备同央视合作搞一个民生经济预测栏目,数据来源于浪潮在建的100个数据所中宏观经济数据所的数据,拟请一些国家级智库对数据模型进行指导。

“这就跟AlphaGo一样,正是因为跟棋手下棋,AlphaGo才被训练出来。你越不跟它下,AlphaGo越笨,人类越跟它下,它越厉害。”徐宏伟说,在这个数据模型中,专家贡献智力,专家的智慧被这个系统吸收,这就是一个大数据系统完善的过程。在国务院发展研究中心等专家的指导下,它会变得越来越聪明,预测结果会更准确。

大数据分析如何辅助两会委员代表参政议政?徐宏伟描述的未来场景是,一位IT界委员拟准备一个关于强化国内IT基础研发的提案,但由于工作繁忙难做进一步社会调研。那么,点击进入大数据分析系统后,输入提案主体,

系统会自动推荐社会各阶层关于这方面的舆情,委员再针对这些推送内容进行编选,从专业角度提出解决思路。

同时,这个系统会汇集国内一些公司的研发状况,作为素材推荐给委员。最终,这个提案会反馈到具有深度学习能力的大数据系统。这个系统结合社会治理的关键指标,比如城市森林覆盖率等建立AI模型,这个数据模型会吸收委员代表的意见建议进一步完善。

整个系统会将更多委员代表的提案或议案放在一起进行关联分析、挖掘,对提案和议案质量进行判断,之后推送给政府相关部门。

而委员代表提出的问题,解决方案,哪些是最应该首要解决的呢?可以把智慧城市采集的数据纳进来,比如把雾霾监测数据,交通拥堵数据融入这个系统里,对委员相关提案进行进一步交叉、关联分析,从而评判这个提案的社会价值,或者城市出现的问题点究竟在哪。这个系统通过不断训练、学习,会变得越来越聪明,向政府部门提出的建议会越来越精准。

徐宏伟表示,这是理想状态下的一个未来场景。这个AI模型需要不断进化,才能达到上述水平。不过,人工智能已进入加速创新和成长阶段,这是一个指数级增长的过程,今年是1%的进度,明年就是4%,后年就是32%,这也是人工智能产品吸引众多投资的原因所在。企业完全可以将其开发成产品,由政府购买,也可以通过PPP方式引入社会资本进行应用开发。

但对浪潮而言,当前重点要解决的还是数据问题,建立一个数据资源平台,用类似区块链的技术把数据连接起来。徐宏伟说,浪潮的天元数据网整合了各种算法放在上面,但重点还是聚集数据,握有深度学习算法引擎的人很多,但他们都缺数据。

## C 财富边上

## 这家澳洲葡萄酒庄 拒绝中国经销商的三个理由

□张娜

“We regret to inform you that we are not in a position to cover your need for theSurvivor Vine Shiraz.”

看着最新收到的邮件,苏宇心猛地一缩。虽然早有预感,但真的发现沟通协商了2个多月后,还是没法买到这支酒,他还是感到遗憾郁闷。

在省城一家大型进口葡萄酒贸易公司做采购经理的苏宇,一项重要的工作就是寻找高性价比的酒款。在一次澳大利亚葡萄酒品鉴会上,苏宇发现了一支来自南澳洲巴罗萨谷的老藤设拉子红葡萄酒让人回味无穷。

关键是这支优质葡萄酒的价格相当亲民,酒庄更是历史悠久。在中澳间关税下调,澳洲葡萄酒日益受到中国消费者认可下,如此捡漏式的佳酿极有可能大卖。

苏宇的判断得到了老板的认可,于是,与酒庄谈判,争取买断该酒款成了他近期最重要的工作。

但在与这家澳洲家族酒庄接触后,苏宇屡屡碰钉子。

首先是酒的名称。这家酒庄的设拉子葡萄藤已经有97年历史,苏宇按照中国人的常规思维,询问能不能把这支酒叫做百年老藤珍藏。“就差3年而已,说出来也有感觉。”但是酒庄主人一听就急了,在越洋电话里声音提高八度强调:

“按照澳洲和巴罗萨葡萄酒协会的规定,只有树龄真的超过100年的老藤,才能叫Centurion Vine百年老藤,就是差一年,都只能叫Survivor Vine(幸存老藤)。要是虚标年份被查出来,我在巴罗萨谷和整个澳大利亚都没法立足了!”

好吧,名字不改,多买些酒总可以吧。苏宇转而进入第二轮谈判:份额。

作为一家历经四代人的老牌酒庄,其销售网络早已遍布全球。苏宇看中

的这款老藤设拉子,年均产量2万多瓶,按照酒庄的全球份额分配,整个亚洲也就3000支份额。“我们多加20%的钱,多买一些酒可以吗?”

面对主动送上门的钱,视频电话那头朴实憨厚如老农的庄主(事实上,他就管自己叫农民)反而不安起来:“哦,那的确是不错的价格,可我不喜欢贪婪,我只想保持简单。我们的酒是面向全世界消费者的,我不想因为今年中国人喜欢它,就全卖给中国市场。明年或许韩国人喜欢了,我再全卖给韩国,那样我在其他地方的朋友就喝不到了。我不想让他们失望,所以,加价也不行!”

自夸不要,加价不行,优化一下酒的工艺总可以吧。这款老藤设拉子的瓶盖坚持手工蜡封环节。如今瓶盖都是锡箔,热缩工艺。手工蜡封必须精挑细选优质橡木塞,由工人在滚烫的蜡盆中,一板一板挑出热蜡,封住瓶口,不仅成本高,而且效率低。在澳洲酒庄传来的视频里,蜡封葡萄酒至少要2分钟才能完成封装,而灌装线的产品,2分钟能封装100瓶。

但这个要求依然遭到拒绝,“老藤葡萄酒是酒庄的顶级产品,代表着我们家族的传统和技艺,你们觉得它是酒,在我们看来,这款酒是一件艺术品,是我们的精神寄托。所以,我爷爷开始用最传统的蜡封,致敬这款酒。我不想改变,我也不会让我的儿子改变。”

从第一次与酒庄发邮件开始,苏宇在历时两个月、十几轮的沟通中,没有让酒庄主松口半分,相比圆熟、精明的法国酒商,豪爽、随性的美国酒商,精细但善于变通的智利酒商,这个澳洲酒庄主面对采购商简直是又轴又硬。

不过,在种种碰撞与摩擦里,苏宇也不得不承认,“澳洲酒能在中国越卖越火,还真不像中国当年只靠廉价低质。人家确实有坚守和定力。”