



环境保护部举办大数据与环境管理转型专题培训引追捧

# 「高冷」大数据开始「接地气」

◆本报记者徐丽莉

为落实党中央和国务院对生态环境监测大数据的要求,助推环境管理转型,近期,环境保护部行政管理与人事司主办了大数据与环境管理转型专题培训班。

“这次培训的可贵之处在于,这是环境保护部第一次以先进信息化理念为主要培训内容的专题培训班。”环境保护部信息中心主任程春明表示。

培训班的规模之大也前所未有,基本覆盖了环境保护部、省环境保护厅(局)、市级环保局三级环保系统。现场参加培训的有环境保护部机关各司局

级、处级以上干部,在京各部直属单位主要负责人。全国各省(自治区、直辖市)环保厅(局)工作人员、部机关其他工作人员则通过远程视频会议系统设置的45个分会场,以及内网共171个视频会议终端同步观看了视频直播。据统计,培训人数多达4850人。

主办方邀请的主讲嘉宾也都是重量级专家,中国科学院虚拟经济与数据科学研究中心常务副主任石勇,曾从事中国人民银行个人信用评分研究、实时信用卡审批系统研发等;阿里巴巴副总裁涂子沛是著名的信息管理专



## 大数据应用要避免哪些误区?

大数据冠以“大”字,很容易让人误以为大数据是要穷尽全部的数据,方能破解难题。

### ●大数据不只是“大”和“全”

石勇坦言道,很多人确实认为大数据时代就是要研究数据整体,数据量越大越好,某些著作也不乏这样的误导。他解释说,大数据条件下,既要全数据,也需要样本数据,不能有了全体,就抛弃样本。

“没有经过机器学习的模型和算法就不能发挥预测作用。”石勇说,任何大数据都不能违背机器学习的过程,具有预测作用的模型都是机器学习的结果。并且开发一项应用可能同时有很多算法和模型,机器学习成为挑选最优算法和模型关键。

而这一切都依赖于样本,模型和算法需要通过样本进行学习,从而具有可推广型和预测性。如在1000个全样本中挑取

300个进行机器学习,将结果推广到另外700个样本中,验证模型和算法。同时,大数据中抽取的样本普适性更强。

郑宇强调说,不能盲目迷信大数据。数据并不是种类越多、数量越大,模型效果就越好。他认为,不同数据代表的含义不一样,不能把所有的数据放到“黑盒子”里进行计算,纳入模型计算的每一类数据都要经过验证,对数据进行有机融合才能发挥1+1>2的效果,否则既浪费计算资源,也影响模型和算法的性能。

郑宇提出了3条验证数据相关性的途径,一要靠前人的经验和知识进行判断,二要对数据相关性做可视化分析,三要用实践证明数据与目标的相关性。

### ●大数据不只是发现相关性

大数据时代最重要的一点是,可以



## 大数据怎么助推信息经济?

后互联网时代,数据经济为代表的信息经济已经初露端倪。涂子沛认为谁拥有数据,谁就将拥有金山,谁使用好数据,谁就会脱颖而出。

### ●大数据提升效率

涂子沛认为信息经济的首要表现是大数据引发的效率革命。

滴滴快车正是大数据时代下信息经济的典型应用。滴滴快车通过整合司机和公众的供求信息,为双方提供更加对称的信息,达到了一石三鸟的效果,道路行车效率提高,司机收入增加,公众需求得到满足。据介绍,用了滴滴快的等打车软件,司机的平均收入增加了约三分之一。

石勇则提到实时信用卡审批系统,5秒钟就可以办一张信用卡,怎么实现的呢?申请人的信息输入个人信用系统后,系统可以在成千上万的信息中搜索跟申请人相似的信息,通过类比,系统对

申请者的贷款申请做出同意或拒绝的决策。如果申请人的贷款申请通过了,系统还会给出建议信贷额度。

石勇表示,现代社会高效率运作只有大数据可以实现,这在以往完全是无法想象的。

### ●大数据催生众创众筹

大数据不仅带来了效率革命,而且带来了新的资源和蓝海。移动互联网带来的公众大数据,是大数据的新蓝海。

涂子沛举例道,华尔街日报要把129年的纸质报纸电子化,但是由于报纸年代久远,扫描的准确率非常低。如果用人工打字,日夜兼程也需要将近48年的时间。恰巧,当时校验码的研发者发现,校验码一天有几十亿次的应用。于是,他们将报纸扫描后,分割成一个一个的校验码,运用公众的力量完成报纸的电子化工作。原理是,如果3-4人输入



## 大数据应用需要哪些先决条件?

大数据是现代经济的助推器,但是大数据的实现还需要整个社会做好准备。

### ●数据整合

石勇认为,大数据系统和模型开发过程中,数据的筛选和整理是大数据非常重要和关键的环节,但过程是非常繁复的浩大工程,比如中国人民银行个人信用评分研究中,要梳理480个参数、八亿五千万个数据,数据处理工作就长达一年多;中国金融期货交易所项目,为从历史交易数据中挖掘出对结算会员风险预测,数据的收集和整理也花费了半年多时间。

石勇将数据挖掘分为6个步骤,其中第二、第三步都是有关数据收集和整合工作。首先要对数据进行筛选,严格要求进入模型的数据,避免就是垃圾数据,数据模型中最忌讳“零”,可能导致模型清空;其次,要对筛选的数据进行统计计算、回归分析等预处理工作。

“大数据建设的前提就是要对数据整理和整合,数据不整合就没有力量。”石勇表示,同样,环保部门数据的整理和整合也是环保大数据的第一步。

### ●数据开放

与数据整合同样重要的是数据开放,数据只有开放,有价值才能成为共享

家,著作《大数据》、《数据之巅》成为我国研究大数据的标杆;微软亚洲研究院主管研究员郑宇是多家高校的客座教授,提出了“城市计算”理论并获得国内外多个奖项。此外,清华大学软件学院副院长王建民、百度总编辑魏承、IBM中国研究院研发总监尹文君也都是大数据领域的资深专家。

专家们的精彩讲座令听众大开眼界,不少上午因业务未能参加的人,闻听反应热烈,下午纷纷挤时间赶来参加。许多听众表示,大数据培训不仅开阔了思路,也化解了平时工作的迷茫。

将碎片化的、看似毫无关联、只反映某个方面表面现象的信息进行关联分析,从而发现彼此之间的联系,为决策提供参考。

但石勇认为发现数据之间的相关性还不够,还要通过相关性研究数据之间的因果性和必然性,实现大数据的预测功能,为决策服务,这才是大数据魅力所在。

郑宇也认为,大数据的魅力就在于通过不确定加不确定,形成确定性的结果,从而形成决策指导实践。

微软配合环保部门开发的智慧城市大气主要污染物时空分布大数据模型——U-Air,正是将各种不确定的数据融合在一起,产生了确定性的结果。U-Air可以预测任何时段任何地点1km×1km的空气质量,准确率已经超越传统模型。

同样的内容,就显示内容识别正确,之后再通过软件将正确识别的验证码整合起来,仅用24个月就完成了所有报纸的电子化。华尔街日报正是通过众包的方式,借助大众的力量,低成本、高效率地完成了报纸电子化。

具体到环保领域,涂子沛认为未来政府可以充当“接单员”的角色,借助市场和公众的力量参与环境保护,开展环境管理工作。贵州就已经开始应用随手拍,市民对污染信息拍照后可以直接上传到政府部门,政府可以根据公众的举报进行处理,这样有利于形成全民参与环保的局面。

此外,大数据时代还可以调动公众、企业、环保组织等社会资源,共同挖掘数据,为环境管理贡献力量。原来仅靠环保部门一家完成的事情,现在可以由整个社会一起分担,真正实现环境管理的众创众筹。

的大数据。上世纪80年代立法后,美国的企业开始开放数据,90年代企业的排放量就下降了50%。

涂子沛认为数据开放其实是一种管理社会的手段,开放数据是政府与社会实现共治的纽带。政府可以利用数据实现管理,规范数据开放范围,并且可以将数据变为价值,有偿供应一些数据。

数据资源的开放也为环境管理实现众筹提供了条件,社会有很多可以帮助环保部门共治的资源。通过开放数据,公民、环保组织、高校、研究所等机构便可以开展更多的数据挖掘工作,为环境管理提供多种解决方案。

## 看大数据如何应用于环境管理

具有实时性,并且不同人不同时间对噪音的感知不同,所以对噪音的监管非常困难。

计算区域噪音指数需要解决3个问题,即纽约什么地方有噪音、什么样的声音算噪音以及这个地方为什么有噪音。如果用传统添加硬件的方法进行解决,整个纽约市至少需要一万多个噪音传感器,显然这并不可行,并且噪音传感器只能反映分贝数,无法与公众感受结合起来。

郑宇的大数据团队则注意到与噪音相关的周边数据。纽约有一部311电话,用于市民投诉,这是公众感受的重要数据源。于是,郑宇团队对311的数据进行了整合,对噪音投诉的时间、地点、类别、次数、性质进行了整理,通过与道路结构、兴趣点、人的社交媒体数据进行融合,计算出每个区域每段时间噪音指数,反映了公众对噪音的整体感受情况。

纽约城市噪音模型可以回放24小时任何时间任何地点的噪音污染情况,并直接显示噪音污染区域的排名情况。



### 省时省力 精准测算

和历史数据,第二类数据是空气质量相关性数据,包括气象数据、交通数据、人流数据、道路结构数据、兴趣点。通过对这些数据建立模型,可以查看任意点之前6小时的空气质量状况,并预测6-12小时、12-24小时的空气质量状况。

U-Air已在61个城市进行验证,平均准确度比传统方法高出7个百分点。北京的准确度可以达到75%,广州和深圳可以达到80%。

从精度来说,目前U-Air可以实现1km×1km的细粒度的空气质量预测,未来,空气细粒度将越来越细。

从计算速度来说,只要花费几秒钟,U-Air就可以实现全市空气质量的预测。

由此,从准确度、精度、速度来说,基于大数据研发的城市细粒度空气质量模型远远超越传统空气质量模型。

### ■计算尾气排放量

过去一个小时,北京市路面上总油

### ■计算区域噪音指数

纽约长期被噪音污染困扰,但噪音



## 2014年鄱阳湖遥感监测数据发布

湖面最大和最小时面积相差三倍

本报综合报道 近日,江西省遥感信息中心发布2014年鄱阳湖遥感监测简报,简报显示鄱阳湖最大的水面面积是最小水面面积的近3倍。据了解,下一步江西省遥感信息中心将构建空地一体的遥感监测网络,保护好鄱阳湖“一湖清水”。

“2014年,鄱阳湖水面积最大出现在7月28日,为3224平方公里;水面面积最小出现在2月2日,为840平方公里,两者之间相差近3倍。”江西遥感信息中心副主任方豫分析说,当年,鄱阳湖水面积大于2500平方公里的时间段为5月25日至10月8日,历时约136天;水面面积低于1000平方公里的时间段主要为1月1日至3月30日,历时90天。

与2014年相比,监测数据显示,2014年鄱阳湖的水面面积比2013年略大一些。2013年,鄱阳湖湖面面积最大的时候出现在7月4

日,为2830平方公里。

据了解,在对鄱阳湖实施遥感监测的内容中,有一项重点内容是生物量,可为鄱阳湖候鸟保护提供科学依据。

监测数据表明,2014年,鄱阳湖湿地草洲生物量变化曲线呈波浪形。其中,上半年变化趋势为先减后增,下半年则表现为先减后平缓增加趋势。夏季的生物量变化表现尤为明显。“与2013年相比,2014年鄱阳湖的生物量有所降低。”方豫说,2014年,鄱阳湖草洲生物量的最低值和最高值分别发生在3月上旬和9月下旬,最低不到800克/平方米,最大超过了1800克/平方米。

方豫透露,下一步,江西省遥感信息中心计划在鄱阳湖湖区放置浮标,设立无线传感器,对鄱阳湖实施异地调控、远程监控,构建江西省空地一体的遥感监测网络。

## 拓展监测指标 开展遥感监测 大连建天地一体化监测体系

本报综合报道 据辽宁大连市环保局透露,大连市环境监测中心将不断拓展环境监测领域,拓展大气、水、土壤环境质量标准和污染物排放标准中的全部监测指标,同时结合遥感监测,建设天地一体化监测体系。

据介绍,在现有空气质量预报预警系统的基础上,大连环境监测中心将建立灵活性更强、可持续发展的应用平台,对系统内部进行适时扩充,升级完善空气预报预警系统,提高预报精度,提高重污染应急措施的实用效果,更新完善污染源排放清单、开展多项污染物的相关研究工作。

在水质监测方面,环境监测中心站将继续规范饮用水源监测网络,逐步将监测断面调整为与各级

保护区界限基本一致,对海域环境功能区划分进行必要修改。环境监测中心将开展开展藻、藻、鱼类、微生物等的急、慢性毒性实验,开展毒理学监测,将环境生态学监测手段应用于日常环境监测工作,特别是环境污染事故应急响应监测。

配合地面监测,大连环境监测中心同时开展了环境遥感技术应用,辅以无人机等手段,大连将建成大气、地表水、生态常态化遥感业务应用系统,形成环境遥感影像数据的自动接收、处理、分析、制图等能力。天地一体化监测体系不仅有效弥补了地面监测在空间覆盖和时间连续性上的不足,也有利于深入地了解大连环境污染成因和规律,及时发现环境问题,充分发挥环境监测预警作用。

## 拉网排查污染隐患 一企一档动态监测 齐河拧紧汛期水环境“安全阀”

本报记者董若义 通讯员刘芹芹 齐河报道 汛期临近,为杜绝汛期水环境污染隐患,山东省齐河县环保局日前在全县范围内组织开展了汛期水环境安全专项整治行动,全力拧紧汛期水环境“安全阀”。

截至目前,齐河县环保局已对重点河流断面和污染源巡回检查30余次,为40余家重点企业建立了环保健康档案。为从源头上控制水源污染,齐河县环保局制订完善了水污染事故应急预案和应急处置监测方案,抽调10余名骨干力量成立汛期环境应急监测小分队和环境监察巡

查小组,对全县出入境河流断面、重点企业、污水处理厂、集中式饮用水水源等进行拉网式排查,实行无缝隙监管,全天候监测。

同时,齐河县环保局还建立了重点污染源动态监测档案,实行“一企一档”,对市控以上重点企业进行不定期人工监测和在线监测数据比对,确保企业外排水全面稳定达标排放。同时,齐河县环保局畅通12369环保举报热线电话,实行24小时轮流值班制度,确保一旦接到污染投诉,执法人员能够第一时间赶赴现场进行应急处置,保障水环境安全。

## 衢州“爱环保”APP下载量达5万

动态显示30项环境内容,可投诉污染事件

本报综合报道 浙江省衢州市环境信息中心牵头开发的手机APP——“爱环保”,下载量已达5万余次。公众只要拿出手机一点,全市所有空气站PM<sub>2.5</sub>、饮用水源质量等实时数据便可动态显示。

为鼓励全民参与环保,接受公众监督,目前,衢州市环保局依托智慧环保项目,实时公开发布包括全市环境空气质量指数、饮用水源质量、出境水质量、交接断面水质情况等8大方面30项内容。

衢州市“爱环保”手机APP不仅能让市民实时查询全市的环境质量和重点企业污染源自动监测数据,还能学习了解环保相关知识,及时掌握环保热点事件,投诉反映环境污染问题。

目前,衢州市已有1200人通过APP参加环保知识竞赛活动,200余名公众通过APP报名参加环保公益活动,通过留言、上传录音等方式上传投诉、举报50余起。

智慧环保  
整体解决方案专家

中科院图天下科技有限公司特约刊登