



扫二维码
看科学周末



总第 6381 期
今日 12 版

2015年9月18日 星期五
乙未年八月初六

官方微博
新浪: http://weibo.com/kexuebao
腾讯: http://t.qq.com/kexueshibao-2008

国内统一刊号: CN11-0084
邮发代号: 1-82

大数据计算你和学霸的距离

■本报记者 胡珉琦 实习生 许雪松

在成都最冷的 20 天里还能坚持早起吃早餐,总是在晚上 10 点到 11 点之间洗澡;在教学楼打水近 80 次……正值开学季,电子科技大学教育大数据研究所的数据显示,普通的你,和学霸之间,恐怕就差了这些“微不足道”的行为。

这个颇有意思的发现,来源于覆盖了电子科大两万万余名本科生的大数据系统——“学生画像”, 将其每名学生的几乎所有的在校活动轨迹与成绩之间建立了关联。利用这些数据,不仅能预测出学生的学习状况,研究者还希望借此引导他们更好地规划各自的学业和就业方向。

学霸的生活轨迹

传统教育认为,学生有规律的生活,是学生提高成绩的重要保证。

“普遍情况下,良好的行为习惯与学习成绩是呈正相关的,这基本是得到公认的。”21 世纪教育研究院副院长熊丙奇表示,这些行为习惯,有的与成绩变动直接相关,例如学生按时上下课,常去图书馆等;还有一些与成绩的变动是间接相关的,比如养成早起的习惯,经常洗衣服,有规律的打水等。“虽然不直接作用于学生的学习,但是,好的生活习惯,反映的是学生积极的状态,说明学生自我管理的能力较强。那么,这些学生用在学习上时间也相对有保证,也就势必会对学习成绩产生影响。”

不过,必须承认,这个被普遍接受的结论很难被定量描述。“如果我们能定量地证明已有的依据,并提出科学的可参照的建议,这项研究就是有价值的。”电子科技大学教育大数据研究所副所长连德富这样解释这项研究的初衷。

如今,研究团队已经花了一年多的时间记录学生的校内行为,包括吃饭、购物、打水、进出图书馆、借阅图书、宿舍门禁、洗澡、使用洗衣机、乘坐公交等。结果发现,学霸有着与一般学生完全不同的学习生活轨迹。

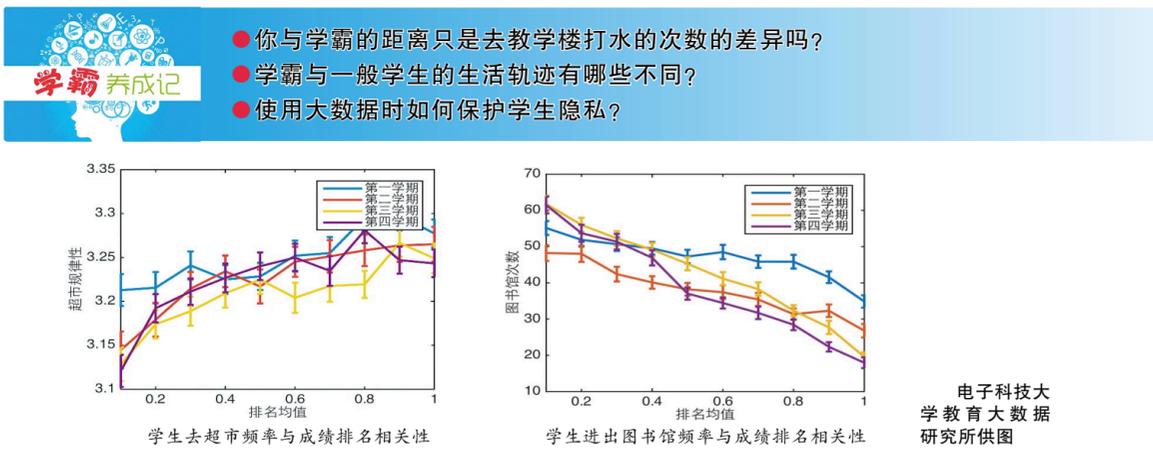
以某专业排名第 3 的学生为例,她几乎每天固定在 8 点、12 点、14 点三个时间点出门,留在宿舍的总时长低于专业平均水平。而该专业成绩排名第 61 的小石每天进出宿舍的时间很随机,而且通常每次外出的时长不超过 2 小时,“宅指数”明显高于专业平均水平。

此外,成绩最好的学生吃早餐次数在 110 次,成绩最差的学生吃早餐次数仅为 60 次。9 点前出现在食堂吃早餐的同学,成绩相对更好。成绩较好的学生集中在晚上 10 到 11 点之间洗澡,而成绩较差的学生,洗澡时间无明显规律。

除作息规律以外,进出图书馆次数也是重要指标。第一学期,成绩最好的学生进入图书馆的次数为 55 次,成绩最差的学生进入图书馆的次数为 35 次;到了第四学期,成绩最好的学生进入图书馆的次数为 61 次,成绩最差的学生进入图书馆的次数为 18 次。

同样的情况还存在于教学楼。学生去教学楼饮水机打水次数越多,就说明学生长期在教学楼里活动。第三学期的数据显示,成绩最好的学生在教学楼打水近 80 次,成绩最差的学生在教学楼打水不到 10 次。

尤其让连德富印象深刻的是,不同成绩的学生在借阅图书的种类上也有明显的区别。他发现,《蜘蛛公墓》《变态心理学》等带有悬疑色



彩的图书,借阅者的成绩普遍不理想。

过去,要想清楚地知道不同成绩水平的学生群体与他们的行为特征之间——的对应关系是非常困难的,但有了大数据的帮助,一切变得简单起来。

“学生画像”还能做什么

仅仅根据学生行为习惯的数据统计,就可以制定出学霸路线吗?

要想精确刻画一个人需要用无数的数据,但连德富认为,校园就是一个拥有丰富数据的很小的社会系统,只要这些数据能与目标联系起来,就有一定的指示作用。

事实上,“学生画像”的首要功能,就是算出每名学生的学习、生活状态,并设计出一系列辅助他们更好规划各自学业的功能模块。目前已经实现的挂科预警。

据了解,研究团队设计了一个针对挂科率的公式,即过去的学习基础 + 一段时期内的努力程度。学习基础是根据已考科目成绩、已考与将考科目之间的关联性计算得出的,而努力程度则主要依据教学楼打水频率、进出图书馆的时间与次数等。

如果有学生正处于挂科率高风险的边缘,系统就会自动向负责该名学生的辅导员发送预警信息。而在过去,只有当学生已经出现挂科的情况,辅导员才能得知,即便如此,事后也很难分析学生挂科的具体原因。

“目前,教育大数据暂时的定位还是辅助传统教育,管理、引导学生。”连德富表示。除了关注学生成绩,他相信,现在的校园里依然充满了像当初的自己那样对未来迷茫的人。“我们都曾迷茫过,不知道自己喜欢什么,可以做什么。”

如何利用“学生画像”帮助学生找到适合自己的路,是研究团队正在努力的方向。目前,数据库不仅有现有学生的行为轨迹,还有已经毕业的学长们的行为轨迹。连德富告诉《中国科学报》记者,最终选择考研、出国或者创业的学生,在生活、学习方式上是存在一些差异的。

比如,打算出国的学生在选修课程、借阅图书时都会偏向语言方面的内容,而偏好创业

的学生则与一项很有意思的数据产生关联。由于电子科大在成都郊区,学生进市区需要乘坐一趟班车,有的学生去市区的频率明显要高于其他同学,这也意味着他们的社交行为可能更为丰富。

“学生画像”可以将现有学生的行为轨迹与已经毕业的学长们的行为轨迹进行对比,如果在选课、借阅图书、参加社团活动等方面的轨迹与某类去向的毕业生比较相似,学校就可以提供相关方面的建议、指导。

对此,21 世纪教育研究院副院长熊丙奇也表示,传统的大学教育对学生的管理是比较松散的,因此,通过对学生的生活、学习的数据分析,掌握学生的目标动向,有助于学校对学生合理的引导,无论是学习还是参与社会活动,都能够有的放矢地帮助学生作进一步规划。

中科院心理所研究员尹文刚则将关注点瞄准了当下大学生的心理健康问题。

“学生从高压的中学时期,进入完全需要自我管理的大学生活,一时间无法适应,容易出现心理问题。一旦受挫,通常选择回避的态度,甚至会出现抑郁的情况,更严重的可能危及生命。”尹文刚直言,近年来,大学生频繁曝出跳楼、伤害同学事件,都与心理健康密切相关。

他认为,通过教育大数据,可以及时掌握学生的行为习惯特征,一旦发生明显异常,比如长期独处、很少参与公共生活等情况,学校就可以适当关注学生的心理健康问题,采取相应对策。

研究团队正在挖掘“学生画像”在关注学生心理健康方面的作用。他们可以根据学生的行为习惯来量化“孤独”。

性格孤僻、有强烈的孤独感,往往被认为是抑郁易感人群具有的一些共性特征。他们发现,学校最孤独的一群人出现心理问题的概率比普通同学高一个数量级。因此,研究团队正在设计算法,依据“30 天内,两个素不相识的人,有两次或两次以上前后脚打水、打饭、进公寓、进出图书馆或坐公交车车经历的概率,不超过十二万分之一”的结论,可算出每名学生的“在校朋友圈”,以此量化学生孤独的程度。

在尹文刚看来,预知大学生的行为习惯所

隐含的心理问题,特别是关注学生的变化,对于开展学生工作是非常有帮助的。

从这些角度看,大数据在教育领域的应用是有一定价值的。

如何保护学生隐私

打水刷卡、进图书馆刷卡、进宿舍房间刷卡……在校园逐步实施一卡通的时代,学生大数据已经变得易得而且可控。学校一方面依靠大数据加强对学生的管理,一方面也要照顾到学生的隐私,尊重学生的行为习惯。因此,如何合理使用大数据,同时又保护学生隐私,就成了一个重要的问题。

“如果让学生感觉生活在学校的监控下,那么即使学校的出发点好的,希望能够对学生的行为及心理进行合理引导,这种举措无疑会让学校反感。”熊丙奇特别提示学校在运用大数据的时候,要考虑到学生的隐私及习惯。

对此,连德富表示,“学生画像”在保护数据隐私方面是非常谨慎的。

“设计系统功能时,很重要的一方面是数据 PK。”连德富介绍说,希望每名学生在除了能在系统上看到自己的“画像”外,还能看到自己与同专业同学比较后的相对优势和劣势。“但是,比较的根本目的是找到学习的模板,而不是与具体的某个个体比高下。”

因此,在申请 PK 时,系统是存在强权限管理的。发起对比的学生必须经过对方的允许,才能看到对方的画像。但是连德富透露,由于该功能涉及的数据隐私比较敏感,目前并没有向全体学生开放。

而己有的面向辅导员的端口,所涉及的数据经过严格的加密处理,系统不保存学生的真实姓名及学号,只有出现重要预警信息,才会自动给相应的辅导员进行短信推送。这意味着,任何一个技术人员都看不到学生个人的信息,而辅导员最多能够掌握自己负责学生可能出现的重大问题。

连德富坚持认为,教育大数据不会刻意强调每个个体的情况,而是反映学生整体的生活、学习状况,以及时预测预警学生的异常状况,从而为学校的决策提供数据支撑。

看点

中央国家机关公车改革全面完成

9 月 16 日,国家发展改革委副主任连维良表示,今年年底前,将基本完成对地方公车改革方案的审批。

目前,中央国家机关 140 个参改单位已经全面完成车改,取消车辆 3868 辆,而且都已经规范处置,收入已经上交国库。涉及的参改人员接近 5 万人,压缩的车辆达到 62%,安置司勤人员 2000 多人。

对于目前正在推进的地方车改,连维良介绍,已经上报到国家层面的车改方案有 20 个,已经批复 16 个,有信心在年底前完成对地方公车改革方案的审批。

杭州获 2022 年亚运会举办权

9 月 16 日,在土库曼斯坦首都阿什哈巴德举行的第 34 届亚奥理事会代表大会上,亚奥理事会主席艾哈迈德亲王宣布,杭州获得 2022 年亚运会举办权。这样,杭州成为继北京、广州之后,第 3 个举办亚运会的中国城市。

杭州作为长三角地区的经济发达城市,已经具有较为完善的城市基础设施条件。一个多月前,亚奥理事会终身名誉副主席魏纪中曾在考察杭州后向记者表示,杭州利用现有的和正在建设的体育设施,无须再进行更多的城市建设就完全可以承办 2022 年亚运会。

但由于杭州拿到亚运举办权之前的申办过程非常短暂,因此在今后的筹办过程中,可以做好城市形象在国内、国际的传播工作。

24 条城际铁路将密布京津冀

9 月 17 日发布的《京津冀城际铁路网规划修编环境影响报告书》透露,京津冀城际铁路网将以“京津唐、京保石、京唐秦”三大通道为主线,以京、津、石三大城市为核心,形成“四纵四横一环”为骨架的城际铁路网络,覆盖区域中心城市、重要城镇和主要产业集聚区。

其中,“四纵”包括京石邯城际、京霸衡城际、津承沧城际、环渤海城际。“四横”包括京津唐城际、京唐城际、津保城际、石沧黄城际。“一环”为环北京城际。初期目标是到 2020 年先期完成 8 条,2050 年全部建设完成。

根据规划目标,中部核心区主要城市间要形成 0.5 至 1 小时交通圈,中心城市之间形成 1 到 2 小时交通圈,中心城区与周边城镇形成 0.5 至 1 小时通勤圈。

2015 年中国薪酬增长趋缓

中智上海 9 月 16 日公布的一份专业调查显示,2015 年中国的总体薪酬增长为 8.0%,相比 2014 年 8.6% 有所减少,总体薪酬趋势上节奏放缓。

从不同城市薪酬幅度来看,二线城市(总体薪酬 8.8%)高于一线城市 1%;从总体薪酬水平实际差距来看,呈现薪酬差距缩小趋势。

该报告比较了北京、上海、广州、深圳四地薪酬幅度,上海的金融行业薪酬增长幅度领跑其他三座城市,达 30%,但在互联网领域,上海(9.6%)明显低于北京(10%至 13%)。

天津港爆炸事故善后 推十项综合性政策

9 月 15 日,“8·12”事故善后综合服务领导小组会同有关部门,对近期社会各界关心的热点问题逐一解答。

回应的问题主要包括商业用房个体经营户受损财产如何补偿、对因事故导致经营特别困难或濒临解体企业及其员工如何帮扶、对事故中居民非营运机动车辆损失如何评估和对因事故造成生产经营或固定资产受损的企业有何支持政策等。

截至目前,天津滨海新区已经先后推出了住院、伤残人员综合救助政策、支持企业加快发展政策、居民室内财产评估定损政策、严重受损住宅收购、修缮、退房政策等十项综合性政策,基本覆盖了受事故影响的各个方面。

民政部宣布将废止单身证明

据民政部新闻办官方微博 9 月 16 日消息,为落实简政放权、方便群众办事创业,民政部近期与教育部、住建部、银监会等部门沟通并达成一致:除对涉台和哈萨克斯坦等 9 个国家和地区的公证事项外,群众将无须提供婚姻登记记录证明。

婚姻登记记录证明是证明个人已经在民政部门登记过结婚,而无婚姻登记记录证明则是证明个人未办理过婚姻登记,即通常所说的“单身证明”。

申领低保、购房贷款等诸多事项可能都需要证明单身或已婚,今后如果还需要证明该怎么办?北京市民政局相关负责人表示,规范此类证明,目的在于减轻群众的办事负担,具体如何操作将于近期发布。

栏目主持: 廖中平

视野



“硬”科技“亮”风景

——中关村智能硬件创客节掠影

图/文 本报记者 赵广立

9 月 15 日~16 日,中关村智能硬件创客节在北京中关村创业大街上演。智能家居、智能制造、可穿戴设备等一系列架构在智能硬件基础上的物联网概念产品被搬上展台与观众见面。

全球知名半导体企业飞思卡尔携“Internet of Tomorrow 巡展——驶向智能未来”大篷车也在中关村智能硬件创客节亮相。先进的汽车驾驶员辅助系统(ADAS)、智能充电器、体感手环等 104 个物联网产品和演示搭乘大篷车,吸引了众多参观者驻足。

除物联网产品展示外,主办方还安排了智能硬件培训讲座以及创客路演活动。比如飞思卡尔就通过大篷车上的车载培训室向“创客”和年轻企业家开展基于安全可靠物联网的培训,并携手中国矿业大学等 5 所大学合作成立了创客空间实验室。

栏目主持: 罗萨

值班主任: 魏刚 李芸
主 编: 郭勉愈 朱子峡
编 辑: 王剑 温新红
校 对: 王心怡 么辰