

无处不在的算法

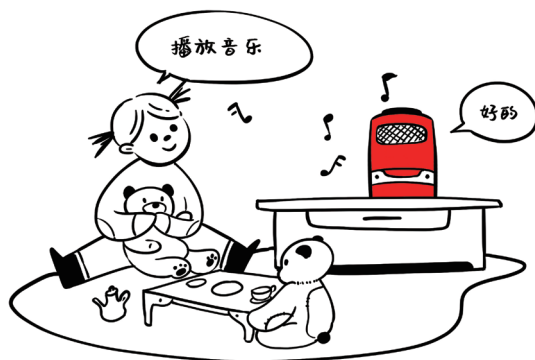
在人工智能的外壳下，算量（数据）、算法、算力构成了三大核心要素。而在“三兄弟”中，算法又凭借着强大的可操作性与实用性、广泛的应用范围、优越的预测能力，成为当仁不让的“大哥”。

接下来，让我们走近算法，了解它是如何大显神通的。

万物互联时代，计算无处不在

要了解算法，先要从开启了万物互联时代的物联网说起。物联网是指在互联网基础上延伸扩展的网络，它能够将各种信息传感设备与网络相结合，密切关注需要监控、连接、互动的物体，采集其声、光、热、电、力学、化学、生物、位置等信息或数据。通过各类可能的网络接入，实现物与物、物与人的泛在联接，最终连接起万事万物，这便是物联网在万物互联时代的设想和愿景。

物联网在数字世界和物理世界之间架起了一座桥梁。根据场景的不同，物联网可以分成工业物联、农业物联、城市物联、家居物联等等。物联网能够让我们的社会环境变得更智能和可测量，比如智能音箱可以让我们轻轻松松播放音乐、设置闹钟或查询信息；家庭安全系统可以让我们随时察觉家里是否有异样发生；智能温控器可以在我们到家前提前调整好房间温度……



由物联网连接起的万物互联时代里，计算无处不在。以上种种看似寻常的生活场景，实际涉及到了大量的数据运算和处理。物联网设备每运行一分钟，都会产生大量的数据，这些数据都蕴含着一定的价值，值得收集、存储和分析。将上述数据“运输”到人工智能系统后，系统可以利用这些数据进行深入分析、预测，让物联网越变越强。

算法的闪耀登场

在万物互联时代，人与物、物与物之间的交互，人们日常生

活工作的种种场景都离不开“计算”二字。在这一背景下，算法闪亮登场，加速推进着社会的发展。

简单来说，算法是为了解决某个问题而采取的有限长度的具体计算方法和处理步骤。从计算机程序设计的角度看，算法由一系列求解问题的清晰指令构成，能够针对一定规范的输入，经过连续的计算过程后，在有限时间内获得所要求的输出。通常来说，算法的产出物有两种，第一种是算法产出的结果（分群、分类、预测值等），第二种是算法产出的规则。

下面我们通过点外卖的例子，直观感受一下算法在后台是如何发挥作用的。

今天，你因为加班太累而不想做晚饭，于是打开外卖 APP 给自己点了份披萨套餐。外卖 APP 后台接到你的新订单后，需要决定把订单分配给哪个骑手。

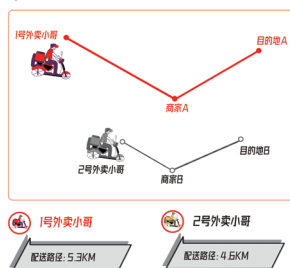


首先，后台会对配送范围内所有骑手的送餐情况进行分析，包括骑手当前的位置和手头已有的订单数量；基于骑手情况预估出接新订单后所需要的配送时间，对现有订单是否会产生超时影响等等。

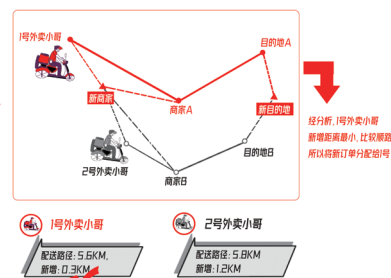
其次，后台会进一步计算出时间充裕的骑手当前的送餐距离和送餐路线，预估他们如果接你的订单，新的送餐路线和新增送餐距离。

最后，后台把订单分配给时间充裕且最为顺路的骑手。

1、2号外卖小哥的骑行路线如下：



如若接新订单, 1、2号外卖小哥的骑行路线如下：



在上面的例子中，订单分配的结果实际上就是算法应用的结果。后台根据骑手当前位置、手头已有订单数量等数据，计算当前送餐距离和送餐路线，预估新的送餐距离和送餐路线，最终把订单分配给时间充裕且最为顺路的骑手，实际上就是通过一系列明确的计算步骤来进行判别和预测，这也是算法的本质。

正如我们可以用不同的方法和步骤解答一个数学问题，我们也可以用不同的算法解决同一问题。值得注意的是，就像同一道

数学题的不同解法计算量不同一样，针对同一问题的不同算法在运行时也存在性能方面的差异。

一个好的算法设计，除了能够解决问题之外，还应该计算工作量越少越好（运行效率最高）、所需内存空间越小越好（耗费资源最少）。更重要的是，算法可以从相同任务的不断重复中“汲取营养”，变得更加明确、简单又有效。

此外，不同算法之间还可以“强强联手”，进行组合优化，产生的新算法通常可以处理单一算法难以解决的问题。比如，在人机围棋大战中一战成名的阿尔法围棋（AlphaGo），其综合使用了线性模型、深度学习、强化学习、蒙特卡洛搜索等算法，这些算法已经存在并发展了数十年，但在组合优化后成功超越了人类的围棋水平，将原先预计短期内不可能完成的任务变成了可能。



被算法包围的世界

购物、短视频、资讯等平台背后的智能推荐算法不断分析计

算着人们的购物偏好、浏览习惯，然后为大家推荐可能喜欢的商品、短视频、文章；价格算法影响着人们购买商品时的价位；导航算法、网约车平台算法和自动驾驶算法影响着人们的出行；外卖平台算法在决定人们收到外卖时效的同时，也控制着骑手们的劳动……

如今，算法正日益渗透到社会生活的方方面面，甚至可以说，我们早已生活在算法的世界里，享受着算法为我们带来的高效和便捷。

除了商业服务领域，人工智能算法在公共服务、政务服务等领域也被广泛应用，用以解决判别、预测、分类、解析、处置和干预等问题，帮助人们不断提高生产和工作效率。新冠肺炎疫情防控期间启用的行程码和健康码，可以视为是算法在社会治理方面的重要实践。



从经济角度看，算法推动了新经济模式的发展，特别是诸如共享经济这样的新经济模式，也改变着传统经济模式。在不远的

未来，我们也许会迎来算法经济的时代，即可以将生产经验、逻辑和规则总结提炼后“固化”在代码上，使生产经营活动无需人工干预，即可自动执行的经济模式。算法这个好帮手将大幅提升经济社会运行效率，为大众生活的方方面面提供便利。

算法的产业化

算法与产业发展走向深度融合，是信息社会发展的必然趋势之一。然而，移动互联网时代产生的海量信息，以及物联网发展过程中催生的数据大爆炸，导致当下信息严重过载且碎片化。

对于个体而言，如何从大量信息中找到自己感兴趣的信息，常常困难重重。尽管搜索引擎可以满足人们的部分需求，但仅适用于需求明确的场景。现实中，普通用户常常无法准确描述自己的信息搜索需求，甚至对“怎样精准提出问题”都不太了解。对于行业来说，海量数据的处理和运算是进行科学决策、合理预判的重要前提，而依赖人力以及简单的运算规则显然无法满足行业发展带来的大量信息处理需求。在“计算无处不在”的情况下，我们更需要借助算法的力量来重新认知、定义世界，以解决个体及行业遇到的种种瓶颈问题。

在这样的背景下，算法产业化发展凸现出三大趋势特征。

首先，算法与行业、领域的结合更加紧密。无论是产业经济还是社会治理领域，算法在“解决好某一领域的某个问题”方面已经崭露头角，显示出极强的实力。越来越多的认知与决策都

需要在算法的辅助甚至主导下完成，而通过算法的“神力”来判断舆情态势、社会风险，也正在成为各类机构和管理部门的常用手段。

以法律界为例，现已能够使用“风险评估工具”算法来确定罪犯刑期，这种算法参考了数十年的量刑判例，并能结合十几个参数来评估被告在一定时期内重新犯罪的可能性，使量刑更为合理。



其次，围绕算法进行数据挖掘有了更多可能的尝试。以新词发现算法为例，其不仅可以实时处理社交媒体产生的大规模数据，可以通过全局特征提取、使用标注模型等方式发现新词，还可以通过更高效算法的候选词提取、命名实体过滤、新词特征选择、特征计算与候选排序四个步骤发现更多新词，大大提升社交媒体数据挖掘的广度及深度。

最后，算法正在逐渐“众智化”。所谓“众智化”，顾名思义，就是更依赖于大众的智慧来进行判断和决策。以骚扰电话的处理为例，如果有十几个人在某一段特定时间内将同一个电话号码标注为“骚扰电话”，那么敏锐的算法就可以通过“群众的智慧”来判定它是骚扰电话。“众智”意味着算法能够借助更多人的力量，在更快、更准反映众人智慧的同时，不断为自己注入新鲜的血液和活力。

未来，算法产业化最值得期待的地方在于：算法可用于解决特定领域的特定问题，并在实践过程中助推更多优质算法脱颖而出。

目前，在社会生活、城市管理、工业检测等方面，依然有许多亟待解决的问题。以灾难预测为例，尽管气象部门能够预报暴雨的发生，但却无法做出更进一步的精准预测。实际上，灾难预测还有更多种可用算法，比如借助“机警”的算法判断社交网络上零星出现的异常，如判断识别微信、微博发布的灾难信息的真伪，并做出是否要扩大传播的决定。



此外，未来的算法可能会像 APP 一样，社会或行业发展中出现任何需求，都可以像下载 APP 一样下载相应的算法，接入数据就可以立刻“跑起来”。尽管这更多地涉及 to B、to G 方向，而非 to C 方向，但在算法产业化发展进程中，算法必然会更进一步地嵌入到我们生活中，从根本上改变我们的工作、学习和生活方式。

由于大规模市场的存在，国内在算法应用上有着远超其他国家的优势。在不远的将来，中国的算法产业创业者可以在全球领先的算法技术基础上结合中国实际，充分拓展算法的用武之地，并借助中国在大数据、大计算平台、大应用场景方面的优势，将算法产业化做得更加全面和深入，让算法充分发挥作用，提升全社会的效率。

诈骗层出不穷，算法筑起反诈“防火墙”

和朋友聚会的间隙，你有没有拿起手机，发现一条内容为“尊敬的XXX您好，恭喜您中奖了！请点击以下链接领取奖品！”的短信？

昏昏欲睡的清晨，你有没有被急促的电话铃声惊醒，看一眼手机屏幕却发现是来自几内亚的奇怪号码？



为避免大家因接听一通电话或点击一则短信中的链接便白白

损失巨额钱财，我们机智负责的“智慧公安”闪亮登场！

近年来，各地公安部门纷纷引入智慧公安，帮助公安部门“出任务”。智慧公安最亮眼的功绩之一，就是开发出越来越智能的防范电信诈骗系统（即反诈系统），从源头建立起无辜群众和电信诈骗犯之间的屏障。其实，打造训练有素的反诈系统需要三步。

第一步，运用特征算法，为大量呼出号码打标签。

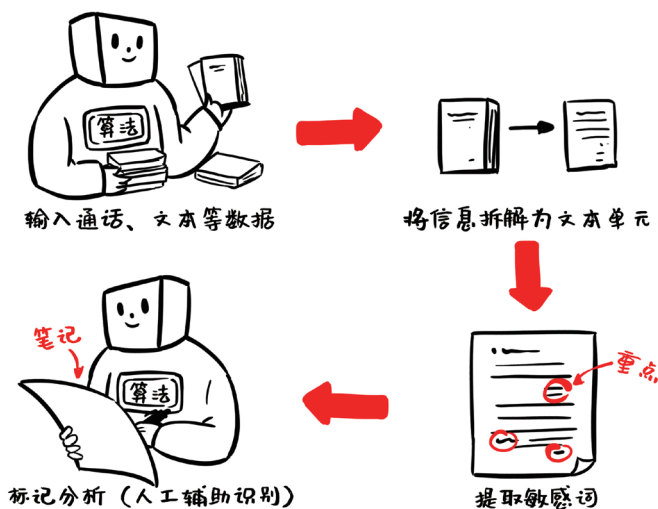
反诈系统通过多频率单号短时呼叫、超长通话、呼叫连续的企业通信行为模式来判断呼出号码所属企业是否有诈骗风险。打个比方，如果某一号码在 5 分钟内呼出了 300 通电话，或者某一通电话长达 300 分钟，甚至呼出每通电话间隔不超过 2 分钟，那么，反诈系统就会将其标记为“有诈骗嫌疑的号码”。



第二步，运用语义分析系统，训练反诈系统的诈骗敏感度。

如果只根据反诈系统对通话数据的智能判断，便笼统地为号

码贴上“好”或“坏”的标签，也存在误判的可能。因此，反诈系统还会运用语义分析工具对大量通话文本进行语义分析，将通话文本拆解为由关键字、关键词构成的文本单元，找出其中的敏感关键字，标记此类关键字，再由人工识别来修正系统的判断结果，以提高对于违规号码的检测效率和准确性。



最后一步，精准标记违规号码。

如果某一企业的呼出通话中涉及到人名、地名以及“身份证号”“卡号”“公检法”“礼品”等敏感词，触发了敏锐的反诈系统，呼出号码所在企业则有可能被停机、停号，还可能被上报给运营商，被添加到呼出黑名单中，这也是我们在接到陌生来电时会被提示“该号码被多人标记为诈骗骚扰”的缘由。

反诈系统通过在呼出阶段提醒接听方此通电话存在嫌疑，为接听方筑起心理防线，提醒其留意潜在风险。



总而言之，智慧公安通过“运用算法将大量数据标签化→引入多种系统提升算法准确性→将危险信号及时传递给人民群众”三个步骤，完成保护人民群众生命财产安全这一艰巨任务。

数据显示，仅 2021 年 1 ~ 4 月短短几个月，某省共破获电信网络诈骗案件 1.18 万起，采取刑事强制措施 8771 人，挽回经济损失 8.8 亿元，足以见得智慧公安的功力。



此外，近年来某省警方紧扣大数据、算法等信息技术发展智慧公安系统，开发建设全省警情大数据应用服务平台。平台按秒汇聚全省 72 个 110 接警区的接处警数据，实现对全省警情数据的实时监测、多维展示和深度应用。警情大数据平台建成后，每个接警区产生的新 110 接处警信息，都会实时汇聚至全省警情数据库，确保省市县三级公安机关都能在第一时间获取相关数据信息，达到“实时监测治安，态势一网掌控，深度研判警情，资源全警应用”的目的。

有了算法加持的智慧公安帮忙，警察叔叔再也不用大海捞针式地寻找诈骗分子，公安部门运转效率大大提升。相信在不远的未来，越来越多的公安部门将会借助算法装备上“三头六臂”，筑起保障人民群众生命财产安全的有效屏障。