

009 | 数据科学家高阶能力之分析产品

2017-10-23 洪亮劼

AI技术内参

[进入课程 >](#)



讲述：初明明

时长 10:29 大小 4.80M



人工智能工程师和数据科学家的主要工作是什么？很多人认为，他们的主要工作是利用复杂的机器学习模型和算法来解决产品中的难题。这样的认识既“对”也“不对”。“对”的地方是说，机器学习模型和算法的确是人工智能技术在产品上落地的核心步骤。“不对”的地方是说，这种认识往往片面地总结了人工智能从业人员的工作范畴。

实际上，要想真正地提出一个好的人工智能解决方案，分析产品的能力是必需的。从较高的层次来讲，就是分析一个产品目前遇到的难题是什么，为什么需要用人工智能技术去解决，哪些是可以用人工智能技术解决的，哪些不能。

今天，我就来分享一下，站在人工智能工程师和数据科学家的角度，我们如何理解并提升分析产品的能力，学会了解产品的需求。

产品需求的庖丁解牛

一个数据驱动的产品往往是一个复杂的复合体。这里面当然有很多数据、人工智能的元素，也有不少其他元素，比如设计、人机交互、商业规则、心理学等等。那么，如何在这么一个综合复杂的体系中找到人工智能技术的合适位置，以及技术究竟要扮演什么样的角色，其实是一个数据驱动型产品能否成功的核心问题。

想要对这个问题有一个比较全面的认识，我们首先需要回答这么一个问题，那就是人工智能技术到底能够给产品带来什么？

很多朋友可能觉得这个问题不言自明。人工智能技术难道不是解决产品的核心算法难题吗？

这种看法其实不够全面。**人工智能技术给产品带来的其实不仅是一些核心的模型和算法，更重要的是，带给产品一项根本性的能力：数据驱动的决策过程。**

什么叫作“数据驱动的决策过程”呢？我们还是要从人工智能技术的特点说起。

人工智能技术的特点有两个方面：第一，数据驱动。第二，在不确定的因素下智能决策。

数据驱动

我们先来谈谈第一个方面，数据驱动。这里其实是两个部分，“数据”和“驱动”。

一个产品要想利用人工智能，第一步，也是非常重要的一个步骤，就是要有“数据”的概念。什么是“数据”的概念？就是一个产品需要有数据收集的意识，并有数据收集的机制。然后，一个有“数据”的产品需要慢慢建立数据管道，并开始建立数据的检测系统。这些都是人工智能介入的先决条件和基础设施。

没有数据，没有流畅的数据链条，是无法构建一个健康的人工智能决策环境的。这一点说起来容易，要真正做到其实需要很扎实的技术基础。很多团队、很多产品最终无法利用人工智能技术的方方面面，一个关键原因就是数据链条上出了各种问题。

有了数据以后，第二个环节就是“驱动”。也就是说，只有数据是不行的，还必须有一个意识，主动利用这些数据来驱动产品的发展，驱动产品方方面面的进化。这个步骤其实不仅是针对产品的决策人员，比如产品经理、项目负责人，也是针对这个产品的所有参与人员的。

参与产品的各方面人员，包括工程方面的、设计方面的、市场方面的，大家有没有意识，在有了数据链条之后，通过数据检测、数据分析不断加深对产品的认识，提出更好的想法。当产品遇到各种问题时，大家有没有一个意识，那就是先到数据中去找答案，先去看数据是不是出了什么问题，去理解数据中显示出来的内容。

如果说数据驱动的第一部分是有关“硬件”的，是数据链条的技术以及实现，那么第二部分就是有关“软件”的，是项目人员的意识和责任。

智能决策

当一个产品有了数据驱动的基础以后，下一步，就需要建立“智能决策”的理念。

“智能决策”是什么意思？

要想明白“智能决策”的意义，我们首先要来想一想“非智能决策”是什么样的。

很多传统的产品或者不是数据驱动的科技产品都是非智能决策的产物。非智能决策主要是指，不依靠数据，或者依靠很少量的数据，由产品经理或产品负责人人工地进行决策。注意，非智能决策并不一定无法带来好的产品。实际上，在历史的很长时间内，各行各业都是依靠非智能决策在进行演化。

非智能决策的一大特点是决策的主观性。通俗地讲，就是决策者依靠自我的认识，主观“拍板”决策关于产品的方方面面。因为没有有一个系统的方法论，或者说是没有一整套机制给决策者相应的信息，来帮助决策者完成决策，非智能决策所带来的产品结果往往有很大的偶然性。

这种偶然性来自于决策者本人的各种能力，来自于执行者的能力。由于这样的偶然性和主观性，非智能决策的另外一个特点，或者说是结果，就是不可复制性。这是因为决策的方法和方式都不能动态地随着时间、随着数据的变化而变化。

非智能决策在什么时候会变得比较困难呢？数据量太大的时候，需要做选择的可能性太多的时候，需要做的选择本身复杂度变高的时候。这些特征也正是新时期下互联网产品或者人工智能产品的特点。因此，在将来非智能决策会越来越多地让位给智能决策。

简单来说，智能决策就是产品的决策者依据产品的特点，把一些复杂的、需要依靠大量数据、选择面太广的决策交给人工智能模型和算法，并且建立起一整套体系，利用人工智能手

段依靠数据来对整个产品进行快速迭代。

如果给这种产品决策找一些典型的类比，就像现代搜索引擎技术，代替了传统的图书馆管理员的角色；现代的电子商务网站为用户推荐各类商品，代替了传统商店里的导购；智能自动驾驶汽车，代替了人类的手工驾驶。

智能决策不仅仅是某一项任务的智能化，更重要的是一种理念的提升。也就是说，一旦产品的决策中出现了有需要大量数据、有复杂选项的时候，作为产品的决策者就需要马上意识到，这部分决策任务应该逐渐从人转移到智能模型和算法上，依靠数据驱动流程来加快迭代。这一点是智能决策的关键。

我们可以接着之前的电子商务网站的例子来说明智能决策的理念。最开始的时候，也许这个网站只有一个简单的界面，可以根据用户的一些历史信息来推荐商店的商品。这个时候，智能决策的部分还仅限于推荐模块这一部分。紧接着，越来越多的用户开始使用这个网站，于是任务就变得越来越复杂，也越来越困难。

比如，如何设计下一版的网站界面？设计师、前端工程师、用户体验工程师、甚至产品经理都会有自己的看法。这个任务本身就很困难，怎么能让上百万的用户满意？怎么能体现出不同用户的不同选择偏好？怎么能体现出这个产品自身的美学价值？你看，这就是一个需要基于数据的复杂的决策任务。

很多团队能够意识到推荐模块需要智能决策，却意识不到“下一版的界面”问题可能也需要智能决策。其实，一旦一个问题变成智能决策问题，我们反而有章可循。

比如这个界面问题，所有人的意见、想法或者创意，依据一定的规则，可以用一些人工智能模型和算法来表达。然后通过现代的 A/B 在线测试手段，可以针对不同的人群展示不同的界面。随后通过数据链条来对测试进行监控和分析。

这时候，决策反而变得简单。因为我们不需要为数百万用户拿一个主意，而需要做的是为智能决策提供足够多的创意，然后由智能模型和算法以及实验流程来选择用户喜欢的界面。最后，下一版的网站，不只有一个界面，而是有几十甚至是几百种不同的界面，为百万千万的用户服务。

小结

我们之前提到了数据驱动，提到了智能决策。那么回到我们今天最开始的主题，作为人工智能工程师或者数据科学家的一个高阶技能，就是能够培养这样一种理念，对产品进行持续分析，检测产品是否遵循了数据驱动的理念，挖掘产品有哪些需求可以进行智能决策。

一旦有了这个分析产品的能力以后，我们可以发现人工智能技术将成为产品进化的驱动器和核心机制，而不仅仅是锦上添花的一种噱头。这就是对产品的一种完全革新式的思维。

同时，我这里需要指出的是，今天在这里提到的数据驱动也好、智能决策也好，你可以认为这些都不是什么新思想。但是，如何把这些思想真正地应用到产品实践中却是一件非常困难的事情。

另外一点需要注意的是，这些理念本身也不是教条，也是一个与时俱进的过程。并不是所有的产品在所有的阶段都需要做数据驱动，都需要做智能决策。我们之前提到了，有很多没有真正数据驱动的产品也依然获得了成功。所以，对产品的分析能力，其实需要你在产品的迭代过程中逐步提升。

今天我为你讲了数据科学家和人工智能工程师如何提升产品分析能力。一起来回顾下要点：第一，产品分析的能力其实就是对产品需求的一个分解，而分解之后的产品迭代很大程度上依赖于数据驱动和智能决策。第二，我详细地阐述了什么是数据驱动，什么是智能决策，究竟怎样可以为产品带来这两项核心能力。

最后，给你留一个思考题，什么样的产品不太适合数据驱动和智能决策呢？

欢迎你给我留言，和我一起讨论。

AI 技术内参

你的360度人工智能信息助理

洪亮劼

Etsy 数据科学主管
前雅虎研究院资深科学家



新版升级：点击「👤 请朋友读」，10位好友免费读，邀请订阅更有**现金**奖励。

© 版权归极客邦科技所有，未经许可不得传播售卖。页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 008 | 曾经辉煌的雅虎研究院

下一篇 010 | 数据科学家高阶能力之评估产品

精选留言 (11)

写留言



Xuan

2017-10-24

👍 2

一般做了a/b后，不是会选择最优的那个方案来全量展示吗？

展开



Netfox

2017-10-23

👍 2

电信行业各种计费，鉴权，承载系统应该不需要的！

展开





黄德平
2018-11-30



想了几个，比如畜牧业，蔬菜种植业，这些远离互联网，变化慢的行业。或许基因编辑技术可以用来做农产品的AB测试。。。



东东
2019-05-05



这样的产品很少吧，ab测试的应该不少
展开 ∨



晏
2018-07-18



数据驱动:主动利用这些数据来驱动产品的发展，驱动产品方方面面的进化。
展开 ∨



晏
2018-07-18



一旦产品的决策中出现了有需要大量数据、有复杂选项的时候，作为产品的决策者就需要马上意识到，这部分决策任务应该逐渐从人移到智能模型和算法上，依靠数据驱动流程来加快迭代。这一点是智能决策的关键。
展开 ∨



晏
2018-07-18



数据驱动的决策过程:数据，数据收集意识，建立数据管道;驱动，在数据链条上，项目人员的意识，如何用数据驱动的决策。



张弛
2018-06-22



没有足够的用户数据的产品
展开 ∨



yy
2018-06-09



老师，我懂你，哈哈

展开 ▾



lzh

2017-12-03



垄断行业

展开 ▾



vincent

2017-10-23



没有这样的产品

展开 ▾