

AI赋能

AI重新定义产品经理

连诗路◎著



AI赋能 AI重新定义产品经理

连诗路◎著

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

随着 AI 在越来越多的行业被应用，AI 赋能的价值逐步体现出来。本书从 AI 的本质出发，介绍 AI 技术过往的发展历程和最新的理论成果，然后讲解如何站在移动互联网和大数据的基础上，系统地学习、应用 AI 技术。本书旨在向读者提供学习 AI 技术的资料、路径，以及打磨 AI 产品的观点、思路。

此外，本书通过介绍笔者接触、打磨 AI 产品的实际经历，指出 AI 赋能过程中需要避免的“坑”。期待我们在 AI 时代共同发展自己，在未来遇到更好的 AI 产品、更好的自己。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有，侵权必究。

图书在版编目（CIP）数据

AI 赋能: AI 重新定义产品经理 / 连诗路著. — 北京: 电子工业出版社, 2019.6
ISBN 978-7-121-35044-3

. ①A... . ①连... III. ①企业管理—产品管理 IV. ①F273.2

中国版本图书馆 CIP 数据核字(2018)第 212562 号

策划编辑：田志远

责任编辑：牛 勇

印 刷：

装 订：

出版发行：电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本：720×1000 1/16 印张：21.25 字数：392 千字 彩插：2

版 次：2019 年 6 月第 1 版

印 次：2019 年 6 月第 1 次印刷

定 价：69.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：(010) 88254888, 88258888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：010-51260888-819, faq@phei.com.cn

推荐序

自 20 世纪 80 年代以来，互联网改变了世界；自 2010 年起，移动互联网改变了世界；未来，人工智能（AI）将改变世界。

工业机器人、服务机器人、特种机器人的前端交互和后台决策的智慧能级正在全面跃升，无人机、无人车、无人超市、无人商店、无人物流等技术正在更为广泛的领域和行业得到试水和应用，移动化、屏幕化、可视化、音频化的发展趋势正在为产品注入更多的智能化要素，万物互联的智能感知、整合一体的数据平台、智慧数据的算力与算法等都为“计算机大脑”提供了智能和精细化思考的能力……我们正迎来 AI 赋能万物的新时代。

主动追寻 AI 赋能万物的创新密码，着力自我思维、自我知识的更新，努力实现 AI 赋能、万物计算、万物安全，是产品经理未来的必经之路。本书从 AI 赋能具体应用场景出发，讲述了 AI 技术、AI 产品的打造方法、AI 赋能产品的具体场景，为从事 AI 赋能的人提供了一本基础书籍，非常适合 AI 领域的产品人、运营人、跨界者、创业者阅读。

在 AI 赋能万物的新时代，产品、运营、技术等人员正面临深刻重塑。在摒除旧技术、运用 AI 新技术的转型过程中，需要具有 AI 赋能的新思维，本书对 AI 赋能具体产品的技术、思维、方法进行了成体系的讲述。大数据之“大”，在于数据的联通，这迫切需要打破部门间的“数据烟囱”，拆除产业间的“数据孤岛”；数据+算法再辅以算力定将赋能一款产品新的未来。产品经理应该发掘 AI 的潜力，充分利用数据分析方法，运用数据挖掘技术为产品赋能。

AI 带来的 AI 赋能时代正在向我们走来，它不仅将给经济社会的发展插上腾飞的翅膀，也将带给 AI 产品经理新的发展机遇。面对 AI 技术，许多大型企业都在摩拳擦掌，期待着在 AI 赋能上赢得先机。但是，这个先机不仅是巨头们的专利，更是新一代 AI 产品经理的事业。如今，年轻一代产品经理对创业抱有很高的热情，而在 AI 技术之中有着无限的创业可能。

AI 离我们的生活并不遥远。AI 技术不是只能用在“高大上”的实验室、研究所或智能工厂里，它也能用在人们的衣、食、住、行之中。AI 中潜藏着巨大的机会。本书从 AI 技术出发，以打造 AI 产品的过程为例，讲解了 AI 在医疗影像、智能小车、智能安防、金融科技、消费电子、工业制造、云计算等领域的具体赋能方法。没有 AI 赋能，这些技术创新就不可能实现，而这些产品的成功离不开 AI 产品经理的智慧。

湖畔大学第一期学员、“互联网大篷车”行动发起人

混沌大学社长、芬尼克兹 CEO

宗毅

前言

不论我们做哪个行业，也不论我们有无相关基础，时下和未来，我们定将看到越来越多的 AI（人工智能）技术赋能具体产业、具体场景的产品出来。站在算法和大数据的肩膀上，AI 必将迎来爆发式增长。

AI 是带有生物的某些特征、具有自我迭代能力的技术，时下和未来一段时间是 AI 工程化的阶段。产品经理需要面对的日常工作变化快、时效强、场面大，产品经理在 AI 时代如果不加紧学习，就会产生焦虑，甚至成为 AI 时代的“无用阶层”。

笔者多年来一直从事 AI 技术赋能具体产品的实践工作，因此整理并总结了 AI 赋能多行业、多场景、多产品的真实案例，从 AI 技术本质的角度解析 AI 的理论知识，从 AI 赋能产品的角度解析打造 AI 产品的创新方法，期待能给想要学习 AI 技术，掌握打造 AI 产品的方法、具体流程的读者一份 AI 赋能的“大礼”。

此外，书中也根据 AI 赋能的实际情况对 AI 的未来发展进行了预测，希望能给求职者、创业者、投资者提供一些参考。

本书讲的是广义上的 AI 赋能，以便能够基于更全面的视野、更系统地搭建 AI 赋能的知识框架。

科学界对 AI 多是从应用领域进行分类的，而本书将创造性地从 AI 的发展趋势、技术特征等层面对 AI 进行分类。

狭义的 AI 是指以计算机为载体，用一组程序或者指令把所有预测情况表达出来，并且通过判断相应的条件，给出合适选择。狭义的 AI 绝大部分是被动的，因为所有算法都在可预测范围之内，如果超出预测范围就直接中断执行。狭义的 AI 具有机械

特性、高效率性和高准确性，然而这种智能缺乏灵活性和自主性。

广义的 AI 是指以计算机或其他物体(比如具备生物特性的传感器) 作为载体，能够自主根据所处环境自发编写程序或指令，并且能够产生一个合适的算法，进而自主执行。广义的 AI 包含狭义的 AI。广义的 AI 具有可控的自主特性，如自主学习、自主编程等。广义的 AI 还具备生物的某些特性，比如生物最基本的新陈代谢特性。

本书认为 AI 算法是构成 AI 的核心要素之一，实现 AI 的方法主要分为两大类：工程学方法和模拟法。工程学方法仅从逻辑层面去设计并编程以实现 AI，不会考虑是否与人或动物所用的方法相同或者相似；而模拟法更多地通过模拟人或动物机体所用的方法来设计算法并最终实现 AI。

机器学习是指过模拟人类的学习行为，让机器也能够像人类一样具备学习能力并不断获取新的知识和技能。机器学习的分类比较复杂，一般从两个方面来对机器学习进行分类。第一个方面是根据学习方式分为监督学习、无监督学习、半监督学习、强化学习和迁移学习。第二个方面是根据算法相似性分为线性分类算法、贝叶斯方法、决策树、关联规则学习、聚类算法、集成学习算法、降维算法和神经网络等。

机器学习被科技界普遍认为是实现 AI 的根本途径，其中神经网络是机器学习多个分支中的重点，而由神经网络延伸出来的深度学习更被认为是机器学习整个领域的重中之重。不少科技界的专业研究人员认为：在未来，深度学习会淘汰其他所有 AI 算法。

遗传算法是模拟人类或生物的遗传进化机制的计算模型，主要的规则来源于达尔文的生物进化论：适者生存和优胜劣汰。遗传算法的具体实现方法是，从初代群体里选出环境适应性表现良好的个体，利用遗传算子对这些适应性良好的个体进行组合交叉和变异，如此产生第二代群体，再从第二代群体中选出环境适应性良好的个体进行组合交叉和变异以形成第三代群体，如此不断进化，直至产生末代种群，即我们问题的近似最优解。遗传算法通常应用于路径搜索问题，如迷宫寻路问题、8 字码问题等，这导致遗传算法并没有像机器学习那样取得令人瞩目的关注度。不过值得注意的是，遗传算法正在和其他 AI 算法互相渗透与结合，尤其是在机器学习的神经网络方面。

通过以上对算法的概述可以发现 AI 发展的基本规律：简易的 AI 能用工程学方法

完全实现，这个阶段可以定义为狭义的 AI 时期；随着 AI 复杂程度的加深，工程学方法已经无法实现或者要付出的代价（如人力、脑力、精力、时间、经济成本等）已经超出了可承受范围，此时迫切需要新的方法来实现 AI，此时就进入了从狭义的 AI 到广义的 AI 的过渡期。笔者认为 AI 正处于这个过渡期，这个过渡期会产生很多的算法，然后通过不断的实践和实验角逐出最佳算法。

AI 的发展趋势应是先具备自主能力，再具备生物特性。

AI 的自主能力包含自主学习能力和自主编程能力。就如同工业革命把人类从繁重的工业劳动中解脱出来一样，AI 具备的自主能力会把技术工程师从繁杂的逻辑思考中解脱出来。这个阶段的 AI 不仅应用于企业生产和商业应用，还会真正走进千家万户。任何一个人在任何时候、任何地点都会和 AI 产生千丝万缕的联系。自主能力仍然属于算法的范畴，AI 算法层出不穷，经过积累和不断发展，自主能力算法被研究出来是可预期的。

AI 具备生物特性的主要表现是具备了独立思考能力，这个阶段的 AI 是真正意义上的独立个体或者说是某一方面的绝对专家。

生物科技的两大基石技术是基因重组和细胞融合，许多新的技术和新的应用都是在这两个基石上发展起来的。截至 2018 年，科学家已经通过生物纳米技术研发出了生物传感器和生物晶片，生物传感器可以完成生物信号和电信号之间的转化，生物晶片可以利用核酸探针完成对基因序列的检测。然而，这还不够，让 AI 具备生物特性还有很长的路要走，期待未来生物科技和计算机技术能够快速融合，尽快创造出能够承载 AI 的新型生物载体。

在 AI 时代，不管你学不学习，AI 会一直学习；不管你进不进步，AI 会一直进步。所以，产品经理在做产品、求职、投资、创业时应该具有永不枯竭的求知欲望。

在 AI 时代，万事万物每时每刻都在变化，AI 无时无刻不在重新对产品经理提出新的要求，唯有求知欲能让我们跟上这种变化。

轻松注册成为博文视点社区用户（www.broadview.com.cn），扫码直达本书页面。

◆ **提交勘误：**您对书中内容的修改意见可在“提交勘误”处提交，若被采纳，将获赠博文视点社区积分（在您购买电子书时，积分可用来抵扣相应金额）。

◆ **交流互动：**在页面下方“读者评论”处留下您的疑问或观点，与我们和其他读者一同学习交流。

页面入口：<http://www.broadview.com.cn/35044>





目 录

Contents

第 章 新手认识

| | | |
|----------|----------------------------|----|
| -----1.1 | AI 的历史和算法 | 1 |
| 1.1.1 | AI 的历史 | 1 |
| 1.1.2 | AI 的三大学派 | 3 |
| 1.1.3 | AI 的三次大发展 | 5 |
| -----1.2 | AI 领域的奠基者 | 6 |
| -----1.3 | AI 的发展现状 | 12 |
| 1.3.1 | AI 的发展概况 | 12 |
| 1.3.2 | AI 相关行业的发展现状 | 15 |
| 1.3.3 | AI 相关公司的发展现状 | 18 |
| 1.3.4 | AI 相关行业的人才现状 | 21 |
| -----1.4 | AI 产品经理的现状 | 26 |
| 1.4.1 | 岗位需求与人员素质不匹配 | 26 |
| 1.4.2 | AI 产品经理来源广泛 | 27 |
| 1.4.3 | 成为 AI 产品经理可能遇到的主要问题及解决方案 | 29 |
| -----1.5 | AI 产品经理认识 AI 的 3 个阶段、9 个过程 | 31 |
| -----1.6 | AI 产品经理应该了解的算法 | 32 |
| 1.6.1 | 传统算法 | 32 |

| | | |
|-----------------------------|-------------------------|-----------|
| 1.6.2 | AI 算法 | 33 |
| 1.6.3 | 传统算法和 AI 算法的区别、联系 | 36 |
| 1.7 | AI=数据+算法+算力 | 38 |
| 第 2 章 从零开始成为 AI 产品经理 | | 42 |
| 2.1 | AI 带来的改变与挑战 | 42 |
| 2.1.1 | AI 带来的改变 | 42 |
| 2.1.2 | AI 带来的挑战 | 45 |
| 2.2 | AI 产品经理与传统产品经理的区别 | 46 |
| 2.2.1 | 面向的用户不同 | 46 |
| 2.2.2 | 工作内容不同 | 48 |
| 2.2.3 | 技术方案的维度不同 | 49 |
| 2.3 | 打造 AI 产品的流程 | 50 |
| 2.3.1 | 打造传统软件的流程 | 50 |
| 2.3.2 | 打造硬件产品的流程 | 54 |
| 2.3.3 | 打造 AI 产品的流程 | 58 |
| 2.3.4 | 打造 AI 产品的实际案例 | 59 |
| 2.3.5 | 打造 AI 产品的创新因素 | 59 |
| 2.4 | AI 产品经理应该具备的思维 | 60 |
| 2.4.1 | 多样化思维 | 60 |
| 2.4.2 | 差异化思维 | 61 |
| 2.4.3 | 穿透思维 | 61 |
| 2.4.4 | 大道思维 | 62 |
| 2.5 | 成为 AI 产品经理的方法 | 63 |
| 2.5.1 | 非技术背景的人员如何转型成为 AI 产品经理 | 63 |
| 2.5.2 | 有技术背景的人员如何转型成为 AI 产品经理 | 68 |
| 2.5.3 | 有技术背景的人员转型成为 AI 产品经理的案例 | 69 |
| 2.6 | AI 产品经理的创新方法 | 70 |
| 2.6.1 | 用户协同创新 | 71 |
| 2.6.2 | 影响用户协同创新的因素 | 72 |

| | | |
|--|------------------------|------------|
| 2.6.3 | AI 产品的创新方法 | 76 |
| 2.7 | 成为 AI 产品经理的核心路径 | 79 |
| 第 3 章 非技术出身的 AI 产品经理应该了解的 AI 技术 | | 84 |
| 3.1 | 常见的 AI 技术 | 84 |
| 3.1.1 | 自然语言生成 | 85 |
| 3.1.2 | 语音识别 | 92 |
| 3.1.3 | 虚拟代理 | 97 |
| 3.1.4 | 机器学习平台 | 101 |
| 3.1.5 | GPU 和其他设备 | 116 |
| 3.1.6 | 决策管理技术 | 118 |
| 3.1.7 | 深度学习技术 | 121 |
| 3.1.8 | 生物特征识别技术 | 133 |
| 3.1.9 | 机器人技术 | 146 |
| 3.1.10 | 文本分析和自然语言处理 | 151 |
| 3.2 | 从企业视角看 AI 技术 | 157 |
| 3.3 | AI 产品经理应该知道的 AI 技术 | 160 |
| 3.3.1 | AI 产品经理对 AI 技术的理解 | 162 |
| 3.3.2 | 热门的 AI 算法 | 163 |
| 3.3.3 | 深度学习大热背后的力量 | 165 |
| 3.3.4 | AI 产品经理应该了解的深度学习算法及其应用 | 167 |
| 3.3.5 | AI 技术的落地 | 170 |
| 3.4 | AI 技术小白的学习资料 | 170 |
| 3.5 | AI 技术小白的学习方法 | 173 |
| 第 4 章 AI 时代产品需求的特点和落地 | | 184 |
| 4.1 | AI 时代产品需求的新特点 | 184 |
| 4.2 | AI 时代寻找产品需求的方法 | 188 |
| 4.2.1 | 重新定义寻找产品需求的方法 | 188 |
| 4.2.2 | 外部方法 | 192 |

| | | |
|--------------------------|-----------------------|------------|
| 4.2.3 | 内部方法 | 193 |
| 4.2.4 | AI 时代的产品创新源自产品经理的自我增强 | 194 |
| 4.2.5 | 小结：成为 AI 产品经理的心路历程 | 197 |
| 4.3 | AI 时代产品需求落地的方法 | 197 |
| 4.3.1 | AI 产品的落地 | 197 |
| 4.3.2 | 裸眼 3D 智能试衣镜 | 202 |
| 4.3.3 | AI 衣柜 | 203 |
| 第 5 章 AI 重新定义敏捷开发 | | 207 |
| 5.1 | AI 时代更需要敏捷开发 | 207 |
| 5.1.1 | 什么是敏捷开发 | 207 |
| 5.1.2 | 产品总监带领团队执行敏捷开发的方法、步骤 | 214 |
| 5.1.3 | 认识 AI 时代的敏捷开发 | 217 |
| 5.1.4 | 机器学习 MVP 的成本 | 222 |
| 5.2 | 创新的敏捷开发管理 | 225 |
| 第 6 章 AI 思维 | | 226 |
| 6.1 | AI 产品经理需要 AI 思维 | 226 |
| 6.1.1 | AI 产品经理需要 AI 思维矩阵 | 226 |
| 6.1.2 | AI 产品经理如何构建 AI 思维 | 227 |
| 6.2 | AI 思维的拓展 | 229 |
| 6.2.1 | AI 思维的作用和重要性 | 229 |
| 6.2.2 | AI 生态思维 | 229 |
| 6.2.3 | AI 工程思维 | 230 |
| 6.3 | AI 产品经理要懂得心理学知识 | 230 |
| 6.4 | AI 产品经理的心力 | 231 |
| 6.4.1 | 认知力：产品方向的致知力 | 231 |
| 6.4.2 | 方法论力：事上练 | 232 |
| 6.4.3 | 迭代力：初心不动、随机而动 | 233 |
| 6.4.4 | 验证力：知行合一 | 235 |

| | |
|-------------------------|------------|
| 第 7 章 AI 产品美学 | 236 |
| 7.1 AI 产品美学的科学性 | 236 |
| 7.1.1 一根：定义落地 | 236 |
| 7.1.2 二脉：行为落地 | 237 |
| 7.1.3 三通：认知落地 | 237 |
| 7.1.4 四达：交互落地 | 237 |
| 7.2 将美学融入 AI 产品 | 238 |
| 7.3 AI+场景美学 | 239 |
| 7.3.1 懂用户之美 | 239 |
| 7.3.2 懂痛点之美 | 239 |
| 7.3.3 爆款产品之美 | 240 |
| 7.4 AI 产品的美学目标 | 240 |
| 7.5 AI 产品美学综述 | 240 |
| 7.6 如何培养产品美学观 | 241 |
| 第 8 章 AI 赋能产品的实例 | 242 |
| 8.1 AI+金融保险 | 242 |
| 8.2 AI+医疗健康 | 245 |
| 8.3 AI+工具软件 | 249 |
| 8.4 AI+商业服务 | 253 |
| 8.4.1 数字化转型下的 AI | 253 |
| 8.4.2 AI 一定会颠覆所有商业模式 | 254 |
| 8.4.3 AI 在商业中的具体应用 | 255 |
| 8.4.4 预测未来最好的方法就是创造未来 | 255 |
| 8.5 AI+新零售 | 258 |
| 8.5.1 零售行业更需要个性化 | 258 |
| 8.5.2 新零售化解供需矛盾 | 258 |
| 8.5.3 AI 赋能新零售的案例 | 260 |

| | | |
|--------|-----------------------|-----|
| 8.6 | AI+机器人 | 262 |
| 8.7 | AI+安防 | 267 |
| 8.7.1 | 安防市场的发展特点 | 267 |
| 8.7.2 | 安防企业进行规划建设的典型实例——佳都科技 | 268 |
| 8.7.3 | AI 赋能安防的实战化应用 | 270 |
| 8.7.4 | AI 赋能安防的四大方向 | 273 |
| 8.8 | AI+农业 | 274 |
| 8.8.1 | 农业正在迅速实现数字化 | 275 |
| 8.8.2 | 农业 AI 技术迅猛发展 | 275 |
| 8.8.3 | AI 技术在农业领域应用的前景 | 276 |
| 8.8.4 | AI 技术不仅仅是为了精确农业 | 277 |
| 8.8.5 | AI 对农民的影响 | 278 |
| 8.8.6 | AI 赋能农业的具体应用 | 278 |
| 8.9 | AI+交通 | 282 |
| 8.10 | AI+文化产业 | 286 |
| 8.11 | AI+教育 | 289 |
| 8.12 | AI+物流 | 293 |
| 8.13 | AI+工业 | 298 |
| 8.13.1 | AI 早已进入工业领域 | 298 |
| 8.13.2 | 从 AI 到工业智能 | 299 |
| 8.13.3 | AI 赋能工业的具体应用 | 300 |
| 8.14 | AI+家居 | 302 |
| 8.15 | AI+法律服务 | 304 |
| 8.15.1 | 政策方向 | 304 |
| 8.15.2 | 技术变革 | 304 |
| 8.15.3 | 行业态度 | 304 |
| 8.15.4 | 市场规模 | 305 |
| 8.15.5 | 案例分析 | 305 |
| 8.16 | AI+社交 | 307 |

| | | |
|-----------------------|---------------|------------|
| 8.16.1 | AI 如何赋能社交 | 307 |
| 8.16.2 | AI 赋能社交对生活的改善 | 307 |
| 8.16.3 | 社交 AI 的现实应用 | 308 |
| 8.16.4 | AI 赋能社交的案例 | 308 |
| 8.17 | AI+人力资源管理 | 309 |
| 8.18 | AI+房地产 | 310 |
| 8.19 | AI+旅游 | 314 |
| 8.20 | AI+信息通信 | 316 |
| 8.21 | AI+城市公共服务 | 318 |
| 后记：AI 重新定义产品经理 | | 322 |

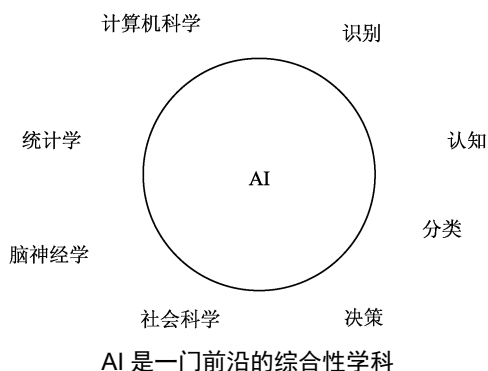
第1章 AI 新手认识 AI

1.1 AI 的历史和算法

1.1.1 AI 的历史

这里的 AI (Artificial Intelligence, 人工智能) 是指广义范畴上的。什么是 AI? 问不同的 AI 专家会得到不同的答案, 总结下来可以说: AI 是一门融合了计算机科学、统计学、脑神经学和社会科学的前沿综合性学科。它的目标是希望计算机拥有和人一样的智力, 可以替代人类完成识别、认知、分类和决策等多种工作。

广义范畴上的 AI 的知识内容如下图所示。



我们首先来回顾一下 AI 的历史。

1950 年, 图灵发表论文提出著名的图灵测试, 预言了创造智能机器的可能性, 这成为 AI 历史上第一个严肃提案。

1956 年, 达特茅斯会议聚集了最早的一批 AI 研究者, 确定了 AI 的名称与任务。

达特茅斯会议成为 AI 学科诞生的标志。

20 世纪 60 年代涌现了大批的 AI 研究方向，包括搜索式推理、自然语言处理、机器视觉等。不少研究者对“机器可在十年左右达到人类的智能水平”这一观点持乐观态度。与此同时，英美政府每年也在研究经费上给予支持。

1974 年到 1980 年，AI 的发展遇到瓶颈，迎来了第一次低谷。数据的缺失和算力的不足使得计算的复杂性难以被解决，AI 的发展陷入莫拉维克悖论。莫拉维克悖论是由 AI 学者和机器人学者所发现的一个和常识相左的现象：和传统假设不同，人类所独有的高阶智慧能力，如推理，只需要非常少的计算能力，但是无意识的技能和直觉却需要极大的运算能力。这个理念由汉斯·莫拉维克、布鲁克斯、马文·明斯基等人于 20 世纪 80 年代所阐释。莫拉维克说过：“要让电脑如成人般下棋是相对容易的，但是要让电脑有一岁儿童般的感知和行动能力却是相当困难甚至不可能的。”

由于研究缺乏进展，期望落空，而且还受到了业内的很多人士的批评，多国政府停止了对 AI 研究的资金支持。

1980 年，一款名为 XCON 的专家系统因为每年可以为企业节省数千万美元而受到热捧，进而带动了大公司在 AI 上的投入，为其提供支持的产业也随之而生，日本、英国、美国等政府重新斥巨资投入 AI，拨款数额高达亿级，AI 又经历了短暂的繁荣期。

1987 年，由于 XCON 等最初大获成功的专家系统的实用性仅仅局限于某些特定场景，而且难以升级，维护费用也居高不下，第二波 AI 浪潮迅速由热转冷。由于对 AI 的高预期再次落空，政府开始大幅削减费用支出，认为 AI 并非为下一个浪潮。

1997 年，深蓝计算机战胜国际象棋冠军，这成为 AI 史上的里程碑事件。

2000 年，互联网发展进入拐点，数据量开始逐步增长。

2012 年，深度学习通过 ImageNet 大规模视觉识别挑战赛（ILSVRC）而名声大噪，进而被广泛应用。

2016 年，AlphaGo 战胜围棋选手李世石，再次将 AI 推向高点，资本加码、政策扶持的轮回再次开启。

AI 在近年“火”起来是有原因的，核心算法的改进、数据量的增长和应用模式的成熟为 AI 的发展提供了天时、地利、人和的环境，AI 开始从技术上突破了在产业上

大规模应用的瓶颈，语音识别、图像识别等 AI 技术在同一时间取得了较大的突破。三个要素的成熟具体体现在以下几个方面。

核心算法的提升。这体现在以深度学习为代表，基于统计模型的算法的整体突破。深度神经网络（DNN）和递归神经网络（RNN）是可以很好提升图像识别和语音识别能力的核心算法。

大数据的发展。AI 在很多方面是基于统计模型的，统计模型里非常重要的一点就是要有足够多的、符合真实世界客观事实的数据样本。以前在实验室环境下很难收集足够多的数据，现在，数据在数量、覆盖性和全面性方面都获得了大幅提升，人们可以获得的数据量更大，数据更加多维，同时覆盖更全面。此外，处理海量数据的能力非常重要，需要超级计算机、海量存储等工程技术体系的支撑。

应用模式的建立。现在与以往最大的不同在于：研究技术、工程、产品、应用、最终用户的整个产业链条已经被打通，形成了大数据、移动互联网、云计算和智能化的整个循环。

1.1.2 AI 的三大学派

AI 在发展过程中产生了很多的学派，符号主义、连接主义和行为主义是其中最主要的三家，这些学派相辅相成，共同推进了 AI 的发展。

符号主义学派，又称为逻辑主义学派、心理学派或计算机学派，其原理主要为物理符号系统（即符号操作系统）假设和有限合理性原理。

连接主义学派，又称为仿生学派或生理学派，其主要原理为神经网络及神经网络间的连接机制与学习算法。

行为主义学派，又称为进化主义学派或控制论学派，其原理为控制论及“感知——动作”模式控制系统。

三个学派对 AI 发展有着不同的看法。

符号主义学派认为 AI 源于数理逻辑。数理逻辑从 19 世纪末开始迅速发展，到 20 世纪 30 年代开始用于描述智能行为。计算机出现以后，符号主义学派又在计算机上实现了逻辑演绎系统，代表性的成果为启发式程序——LT 逻辑理论家证明了 38 条数学定理，这表明了人类可以应用计算机研究人的思维构成，模拟人类的智能

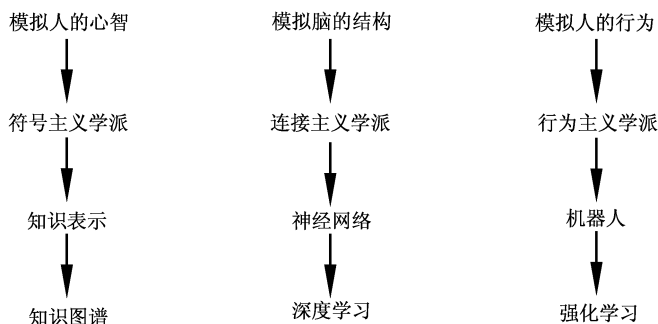
活动。符号主义学派在 1956 年首先采用“AI”这个术语,后来又发展了启发式算法、专家系统、知识工程理论与技术,并在 20 世纪 80 年代取得很大发展。符号主义学派曾长期一枝独秀,为 AI 的发展做出了重要贡献,特别是专家系统的成功开发与应用,对 AI 走向工程应用、实现理论联系实际具有特别重要的意义。在 AI 的其他学派出现之后,符号主义学派仍然是 AI 的主流派别。这个学派的代表人物有纽厄尔、西蒙和尼尔逊等。

连接主义学派认为 AI 源于仿生学,特别是对人脑模型的研究。它的代表性成果是 1943 年由生理学家麦卡洛克和数理逻辑学家皮茨创立的脑模型,即 MP 模型。MP 模型开创了用电子装置模仿人脑结构和功能的新途径。它从神经元开始进而研究神经网络模型和脑模型,开辟了 AI 的又一发展道路。20 世纪 60 年代至 20 世纪 70 年代,连接主义学派对以感知机(Perceptron)为代表的脑模型的研究成为热潮,但由于受到当时的理论模型、生物原型和技术条件的限制,对脑模型的研究在 20 世纪 80 年代初期进入低谷,直到霍普菲尔德教授在 1982 年和 1984 年发表两篇重要论文,提出用硬件模拟神经网络以后,连接主义学派才重新抬头。1986 年,鲁梅尔哈特等人提出多层网络中的反向传播(Back Propagation, 后文简称 BP)算法。此后,连接主义学派势头大振,从模型到算法,从理论分析到工程实现,各项技术的发展为神经网络计算机走向市场打下了基础。当下,人们对人工神经网络(ANN)的研究热情仍然较高,但研究成果没有预想的好。

行为主义学派认为 AI 源于控制论。控制论思想早在 20 世纪 40 年代至 20 世纪 50 年代就成为时代思潮的重要组成部分,影响了早期的 AI 工作者。维纳和麦卡洛克等人提出的控制论和自组织系统以及钱学森等人提出的工程控制论和生物控制论影响了许多领域。控制论把神经系统的工作原理与信息理论、控制理论、逻辑以及计算机联系起来。早期的研究工作重点是模拟人在控制过程中的智能行为和作用,如对自寻优、自适应、自镇定、自组织和自学习等控制论系统的研究,并进行“控制论动物”的研制。在 20 世纪 60 年代至 20 世纪 70 年代,上述这些控制论系统的研究取得一定进展,播下了智能控制和智能机器人的种子,并在 20 世纪 80 年代诞生了智能控制和智能机器人系统。行为主义学派在 20 世纪末才以 AI 新学派的面孔出现,并引起许多人的兴趣。这一学派的代表作首推布鲁克斯的六足行走机器人,它被看作新一代的“控制论动物”,是一个基于“感知—动作”模式模拟昆虫行为的控制系统。

总结：符号主义学派认为 AI 源于数理逻辑；连接主义学派认为 AI 源于仿生学，特别是对人脑模型的研究；行为主义学派认为 AI 源于控制论。

AI 三大学派的路线图如下。



AI 三大学派的路线图

1.1.3 AI 的三次大发展

AI 的第一次高潮始于 20 世纪 50 年代。在算法方面，感知机模型被提出用于模拟人的神经元反应过程，并能够使用梯度下降法从训练样本中自动学习，完成分类任务。此外，由于计算机技术的发展，利用计算机实现逻辑推理的一些尝试取得成功，理论与实践的发展带来神经网络的第一次浪潮。然而，感知机模型的缺陷在之后被发现，即它本质上只能处理线性分类问题，就连最简单的异或问题都无法正确分类。许多应用难题并没有随着时间推移而被解决，神经网络的研究也陷入停滞。

AI 的第二次高潮始于 20 世纪 80 年代。反向传播算法被提出，并用于多层神经网络的参数计算，以解决非线性分类和学习问题。此外，针对特定领域的专家系统也在商业上获得成功应用，AI 迎来了新一轮高潮。然而，人工神经网络的设计一直缺少相应的、严格的数学理论的支持，之后反向传播算法更被指出存在梯度消失问题，因此无法对前层进行有效的学习。专家系统也暴露出应用领域狭窄、知识获取困难等问题。AI 的研究进入第二次低谷。

AI 的第三次高潮始于 2010 年。深度学习的出现引起了广泛的关注，多层神经网络学习过程中的梯度消失问题被有效地抑制，网络的深层结构能够自动提取并表征复杂的特征，避免传统方法通过人工提取特征的问题。深度学习被应用到语音识别以及

图像识别中，取得了非常好的效果。AI 在大数据时代进入了第三次发展高潮。

AI 的三次大发展的小结。

第一阶段（20 世纪 50 年代至 20 世纪 80 年代）。这一阶段 AI 刚诞生，基于可编程数字计算机已经出现，符号主义学派快速发展，但由于很多事物不能形式化表达，建立的模型存在一定的局限性。此外，随着计算任务的复杂性不断加大，AI 发展一度遇到瓶颈。

突破点：解代数应用题、证明几何定理、学英语。

第二阶段（20 世纪 80 年代至 20 世纪 90 年代末）。在这一阶段，专家系统快速发展，数学模型的发展取得重大突破，但由于专家系统在知识获取、推理能力等方面存在不足，以及开发成本高等原因，AI 的发展又一次进入低谷。

突破点：出现能够进行人机对话、翻译、图像识别的机器。

第三阶段（21 世纪初至今）。随着大数据的发展、算法的革新、计算能力的提升，AI 在很多应用领域取得了突破性进展，迎来了又一个繁荣时期。产生了 IBM 的深蓝战胜国际象棋世界冠军卡斯帕罗夫，谷歌的 AlphaGo 战胜围棋世界冠军李世石等一系列轰动事件。

突破点：机器学习。

1.2 AI 领域的奠基者

在 AI 璀璨的星空中，不仅有浩如烟海的 AI 知识，更有无数为 AI 知识做出自己独特贡献的奠基者。笔者在这本书中将列出为 AI 相关领域做出重要贡献的学者和专家，从这些 AI 相关专家的身上学习 AI 知识，能帮助我们更好地理解、运用 AI。（注：以下人物排列不分先后顺序。）

1. 艾萨克·阿西莫夫：突出贡献是提出“机器人学三定律”

艾萨克·阿西莫夫（1920 年 1 月 2 日—1992 年 4 月 6 日），美国著名科幻小说家、科普作家、文学评论家，美国科幻小说黄金时代的代表人物之一。

阿西莫夫一生著述近 500 本，题材涉及自然科学、社会科学和文学艺术等许多领域，他与儒勒·凡尔纳、赫伯特·乔治·威尔斯被并称为科幻史上的三巨头，还与罗伯特·海因莱因、亚瑟·克拉克并列列为科幻小说的三巨头，同时也是著名

的门萨学会会员，并在后来担任副会长。其作品中的“基地”系列、“银河帝国三部曲”和“机器人”系列被誉为“科幻圣经”。他曾获代表科幻界最高荣誉的雨果奖和星云终身成就大师奖。小行星 5020、《阿西莫夫科幻小说》杂志和两项阿西莫夫奖都是以他的名字命名。他提出的“机器人学三定律”被称为“现代机器人学的基石”。

2. 艾伦·图灵：突出贡献是图灵试验和图灵机

艾伦·图灵（1912 年 6 月 23 日—1954 年 6 月 7 日），英国数学家、逻辑学家，被称为计算机科学之父、AI 之父。

1931 年，图灵进入剑桥大学国王学院，毕业后到美国普林斯顿大学攻读博士学位，第二次世界大战爆发后回到剑桥，后曾协助盟军破解德国的 Enigma 密码系统，帮助盟军取得了二战的胜利。

2013 年 12 月 24 日，在英国司法部部长克里斯·格雷灵的要求下，英国女王向图灵颁发了皇家赦免。英国司法部部长宣布：“图灵的晚年生活因为其同性取向（同性恋）而被迫蒙上了一层阴影，我们认为当时的判决是不公的，这种歧视现象现在已经遭到了废除。为此，女王决定为这位伟人送上赦免，以此向其致敬。”

图灵对于 AI 的发展有诸多贡献，不仅提出了一种用于判定机器是否具有智能的试验方法，即图灵试验（至今每年都有相应的比赛），他提出的图灵机模型还为现代计算机的逻辑工作方式奠定了基础。

3. 约翰·麦卡锡：突出贡献是组织并发起达特茅斯会议，提出 AI 的概念

1927 年 9 月 4 日，麦卡锡生于美国波士顿的一个共产党员家庭。他因在 AI 领域的贡献而在 1971 年获得图灵奖。实际上，正是他和明斯基在 1956 年的达特茅斯会议上提出了 AI 这个概念。

麦卡锡对 AI 的兴趣始于他当研究生的时候。1948 年 9 月，他参加了一个“脑行为机制”的专题讨论会，会上，冯·诺伊曼发表了一篇关于自复制自动机的论文，提出了可以复制自身的机器的设想，这激起了麦卡锡极大的兴趣和好奇心，自此，他开始尝试在计算机上模拟人的智能。1949 年，他向冯·诺伊曼谈了自己的想法，后者表示赞成和支持，鼓励他搞下去。在达特茅斯会议前后，麦卡锡的主要研究方向是计算机下棋。下棋程序的关键之一是如何减少计算机需要考虑的棋步。麦卡锡经过艰苦探

索,终于发明了著名的 α - β 搜索法,使搜索能有效进行。在 β 搜索法中,麦卡锡将结点的产生与求评价函数值(也称返上值或倒推值)两者巧妙地结合起来,从而使某些子树结点根本不必产生与搜索。之所以称为 β 搜索法,是因为将处于取最大值的结点的返上值或候选返上值(PBV)称为该结点的 α 值,而将处于取最小值的结点的返上值或候选返上值称为该结点的 β 值。这样,在求得某结点的 β 值时,就可与其前辈结点的 β 值相比较,若 $\beta < \beta_{\text{parent}}$,则可终止该结点以下的搜索,即从该结点处加以修剪,这叫 β 修剪;而在求得某结点的 α 值时,就可与其前辈结点的 α 值相比较,若 $\alpha > \alpha_{\text{parent}}$,则可终止该结点以下的搜索,即从该结点处加以修剪,这叫 α 修剪。为了说明 α - β 修剪,我们举一个最简单的例子。设在取火柴棍的游戏中,A、B 两人轮流从一定数量的火柴中取 1 根或 2 根,不得多取,也不能不取,取走最后一根火柴者胜。用 $A(n)$ 、 $B(n)$ 表示轮到 A 或 B 时有 n 根火柴的状态,当 $n=5$ 时轮到 A 取,A 有两种可能,一种是取 2 根火柴进入 $B(3)$,另一种是取 1 根火柴进入 $B(4)$ 。显然,进入 $B(3)$ 后,不管 B 取几根,A 必胜,故 A 必走这一步,余下的分支不必再搜索了。 α - β 搜索法至今仍是解决 AI 问题中一种常用的高效方法。

至于达特茅斯会议,东道主麦卡锡是主要发起人,另外 3 个发起人是当时在哈佛大学的明斯基、在 IBM 公司的罗杰斯特、信息论的创始人香农。麦卡锡发起这个会议的目标非常宏伟,是想通过 10 来个人 2 个月的共同努力设计出一台具有真正智能的机器。会议的经费是洛克菲勒基金会资助的,包括每位代表 1200 美元的参会费用加上外地代表的往返交通费用。会议的原始目标虽然由于不切实际而不可能实现,但由于麦卡锡在下棋程序尤其是 α - β 搜索法上所取得的成功,以及卡内基·梅隆大学的西蒙和纽厄尔(这两人是 1975 年的图灵奖获得者)带来了能证明数学名著《数学原理》第二章的 52 个定理中的 38 个定理的启发式程序——逻辑理论家(LT),明斯基带来了名为 Snare 的学习机的雏形(主要学习如何通过迷宫),使得会议参加者仍能充满信心地宣布 AI 这一崭新学科的诞生。

4. 马文·明斯基:突出贡献是提出关于思维的理论,提出 AI 概念

明斯基是框架理论的创立者,他和麦卡锡在 1956 年发起达特茅斯会议并提出 AI 这个概念。明斯基获得了 1969 年的图灵奖,是第一位获此殊荣的 AI 学者。其后,麦卡锡(1971 年)、西蒙和纽厄尔(1975 年)、费根鲍姆和劳伊·雷迪(1994 年)5 名 AI 学者先后获奖,AI 学者获得图灵奖的人数占获奖总人数的近 1/6,可见 AI 学科的

影响之深远。明斯基的代表作包括《情感机器》《心智社会》等。明斯基的主要贡献包括如下几点。

提出关于思维的基本理论。1951 年，他提出了关于思维如何萌发并形成的一些基本理论，并建造了一台学习机，取名为 Snare。Snare 是世界上第一个神经网络模拟器，其目的是学习如何穿过迷宫，其组成包括 40 个代理（Agent，国内资料也把它译为“主体”“智能体”）和一个对成功给予奖励的系统。基于 Agent 的计算和分布式智能是当前 AI 研究中的一个热点，明斯基也许是最早提出 Agent 概念的学者之一。Snare 虽然还比较粗糙且不够灵活，但毕竟是 AI 研究中早期的尝试之一。在 Snare 的基础上，明斯基综合利用他多学科的知识，解决了使机器基于过去行为的知识预测其当前行为的结果这一问题，并以“神经网络和脑模型问题”（Neural Nets And The Brain Model Problem）为题完成了他的博士论文，明斯基于 1954 年取得博士学位。

发起并组织 AI 学会。明斯基学成以后，选择留校工作 3 年，其间他与麦卡锡、香农等人一起发起并组织了成为 AI 起点的达特茅斯会议。1958 年，明斯基从哈佛转至 MIT，同时麦卡锡也由达特茅斯来到 MIT 与他会合，他们在这里共同创建了世界上第一个 AI 实验室。

在 AI 方面的贡献。明斯基在 AI 方面的贡献是多方面的。1975 年他首创框架理论（Frame Theory）。框架理论的核心是以框架这种形式来表示知识。框架的顶层是固定的，表示固定的概念、对象或事件。下层由若干个槽（Slot）组成，其中可填入具体值，以描述具体事物特征。每个槽有若干个侧面（Facet）对槽做附加说明，如槽的取值范围、求值方法等。这样，框架就可以包含各种各样的信息，例如描述事物的信息、如何使用框架的信息、对下一步发生的事件有什么期望、期望如果没有发生该怎么办等。利用多个有一定关联的框架组成框架系统，就可以完整而确切地把知识表示出来。

开发出最早的机器人。明斯基还把 AI 技术和机器人技术结合起来，开发出了世界上最早的能够模拟人活动的机器人 Robot C，使机器人技术跃上了一个新的台阶。明斯基的另一个大举措是创建了著名的“思维机公司”（Thinking Machines Inc.），以开发具有智能的计算机。在 20 世纪 80 年代中期，思维机公司开始推出著名的“连接机”（Connection Machine）系列 CM-1、CM-2 和 CM-5，把大量简单的“存储—处理单元”连接成一个多维结构，在宏观上构成大容量的智能存储器，再通过常规计算机

实现控制、I/O 和用户接口功能，处理智能信息。CM-1 由 4 个象限组成，每个象限包含多达 16384 个 1 位处理器，全部处理器则被分为 4096 组，组间形成 12 维超立方体结构，其集成峰值速度达到每秒 600 亿次。CM-5 的结点数更多，功能更强。

“虚拟现实”的倡导者。明斯基也是“虚拟现实”(Virtual Reality, 后文简称 VR)的倡导者，虽然 VR 这个名词与概念是 20 世纪 90 年代才出现与明朗起来的。早在 20 世纪 60 年代，明斯基就自己造了一个名词 Telepresence，直译应为“遥远的存在”或“远距离介入”。明斯基所谓的 Telepresence 是这样一种设备或环境：它允许人体验某种事件，而不需要真正介入这种事件，比如感觉自己在驾驶飞机、在战场上参加战斗、在水里游泳，等等，实际上这些事都没有发生。明斯基提出过利用微型摄像机、运动传感器等设备来实现 Telepresence 的一些方案。

明斯基作为 AI 的倡导者之一，坚信人的思维过程可以用机器去模拟，机器也可以有智能。他的一句流传颇广的话是：大脑无非是肉做的机器而已 (The brain happens to be a meat machine)。他的这种观点遭到了另一位图灵奖获得者威尔克斯针锋相对地抨击，孰是孰非有待进一步的科学研究和实践去判断。

5. 克劳德·香农：突出贡献是提出信息论和熵的概念

香农是美国数学家，信息论的创始人。香农理论的重要特征是熵 (Entropy) 的概念。熵 (Entropy) 指的是体系混乱的程度，它在控制论、概率论、数论、天体物理、生命科学等领域都有重要应用，其在不同的学科中引申出更为具体的定义，是各领域十分重要的参量。熵由鲁道夫·克劳修斯提出，并率先应用在热力学中。后来，香农第一次将熵的概念引入信息论。他证明熵与信息内容的不确定程度有等价关系。我们可以把熵理解为分子运动的混乱度。信息熵也有类似的意义，例如，在处理中文信息时，汉字的静态平均信息熵比较大，中文是 9.65 比特，英文是 4.03 比特。这表明中文的复杂程度高于英文，反映了中文词义丰富、行文简练，但处理难度也更大。信息的熵大，意味着不确定性也大。因此我们应该深入研究，以寻求处理中文信息的深层突破，不能盲目认为汉字是世界上最优美的文字，从而引申出汉字最容易处理的错误结论。

6. 沃尔特·皮茨：突出贡献是提出算法逻辑及数据关联关系，尝试构建一个大脑的模型

皮茨推动了数字计算、机器学习中的神经网络方法、联结主义学派的发展，这三

点为 AI 三大学派的理论基础。

皮茨建立了控制论和 AI 的基础，他将精神医学的方向从弗洛伊德式的分析转到对思想机械论的理解。他展示了大脑是如何进行计算的，提出了心理活动是信息处理活动的观点。

同时，他还展示了机器可以进行计算，这为现代计算机的架构提供了灵感。因为他的工作，在历史上一段时间内，神经科学、精神医学、计算机科学、数理逻辑以及 AI 是统一的，这实现了莱布尼茨最先提出的理念：信息是人类、机器、数字以及精神的“通用货币”。表面看起来迥然不同的各种成分——大块的金属、脑中的灰质、纸上的墨痕，在深层次上都是可相互交换的。

7. 李飞飞：突出贡献是机器视觉

李飞飞，1976 年出生于北京，幼年和少年时代在四川度过，16 岁随父母移居美国新泽西州，现为美国斯坦福大学教授，曾任斯坦福大学 AI 实验室与视觉实验室负责人、谷歌云 AI 和机器学习首席科学家。2015 年 12 月 1 日，李飞飞入选 2015 年“全球百大思想者”；2018 年 3 月，获“影响世界华人大奖”。

李飞飞主要研究方向为机器学习、计算机视觉、认知计算神经学。

8. 吴恩达：突出贡献是深度学习

吴恩达，华裔美国人，斯坦福大学计算机科学系和电子工程系副教授，AI 实验室主任。吴恩达是 AI 和机器学习领域国际上最权威的学者之一，他也是在线教育平台 Coursera 的联合创始人。

2014 年 5 月 16 日，吴恩达加入百度，担任百度公司首席科学家，负责百度研究院的领导工作，尤其是百度大脑计划（已经离职）。

2017 年 10 月，吴恩达出任 Woebot 公司董事长，该公司拥有一款同名聊天机器人。

吴恩达早期的工作包括斯坦福自动控制直升机项目，他的团队开发了世界上领先的自动控制直升机。

吴恩达同时也是机器学习、机器人技术和相关领域的 100 多篇论文的作者或合作者，他在计算机视觉的一些工作被一系列的出版物和评论文章所重点引用。

作为在线教育平台 Coursera 的联合创始人，吴恩达在 2008 年发起了 Stanford Engineering Everywhere（SEE）项目，这个项目把斯坦福的许多课程放到网上，供学

员免费学习。他也亲自讲授了一些课程,如“机器学习”课程包含了 he 录制的视频讲座和斯坦福 CS299 课程的学生材料。

吴恩达的理想是让世界上每个人都能够接受高质量的、免费的教育。Coursera 和世界一些顶尖大学合作,一起提供高质量的免费在线课程。

9. 小结

AI 时代的来临无疑让科学走向极致,让技术承载未来。AI 再也不仅仅是实验室里冷冰冰的研究,而是逐步成为科技公司最核心的竞争力,并逐渐影响着人类的生活。我们在享受 AI 带来的便捷和高效的同时,要特别感谢 AI 领域的国内外学者。

1.3 AI 的发展现状

1.3.1 AI 的发展概况

当下 AI 上升为我国的国家战略,AI 在各个行业和领域中应用、普及的速度也超乎想象,智能机器人逐渐切入我们生活的方方面面,智能时代正在开启。

在这一部分,笔者聚焦 AI 领域,解析相关内容,希望为读者提供一个全景式的 AI 发展图景。

总体来说,笔者认为当下的概况是多方焦虑。

1. AI 成为国家竞争的新焦点

2017 年 3 月, AI 首次被写入政府工作报告。近几年,国家先后出台了很多与 AI 相关的政策,从产业升级到深度教育,可以说,国家已经越来越重视 AI,并且希望 AI 能帮助传统行业升级改造。不夸张地说, AI 正在成为国际竞争的新焦点。

2018 年 6 月,中国科学院文献情报中心和科睿唯安联合发布了《G20 国家科技竞争格局之辩》系列报告,这份报告指出:在 G20 国家中,中国在科研和技术创新力方面表现突出,在 AI 领域的科技实力仅次于美国,且增速明显。

大环境的变化也反映在图书出版领域,由腾讯团队和工业和信息化部高端智库联合创作的《人工智能:国家人工智能战略行动抓手》(中国人民大学出版社出版)指出: AI 如同互联网一样,正在揭开时代变迁的新篇章,在这一社会变迁过程中,中国将会成为 AI 应用最大的市场。而在国际竞争中,中美两国技术竞赛的

格局初步显现，两国都有自己的优势和劣势。国家层面的支持为中国在 AI 领域的发展打下了良好的基础。

《AI 真的来了》(江苏凤凰文艺出版社出版)为媒体人杨澜的跨界之作。2016 年，杨澜带领团队历时一年，采访了 30 多个全球重量级科研机构的 80 多位行业专家，制作出《探寻 AI》纪录片。在剪辑纪录片的过程中，杨澜回味和沉浸其中，用媒体人的人文视角、独特的亲身体验以及细腻的情感，通过文字记录了一个文科生的 AI 探寻之旅。杨澜在书中写道：“AI 改变世界，谁来改变 AI？我们创造了 AI，并在它的身上看到了我们自己的希望、想象和恐惧，以及我们与这个世界相处的另一种可能性，但更让我们发现了人类智能的种种奇妙之处。”《AI 真的来了》书写了高科技背后的浪漫与理性，也赢得了专业人士的认可。百度创始人、董事长兼首席执行官李彦宏评价：智能革命是一场对人类生存和发展方式的根本性变革。当 AI 正在不知疲倦地学习人类时，我们对于 AI 的深刻影响是否有足够的认知和重视？杨澜的观察和思考恰当其时。

《纽约时报》记者、普利策奖得主约翰·马尔科夫的《人工智能简史》(浙江人民出版社出版)讲述了 AI 的前世今生，它聚焦了 AI 在 60 年间的惊天裂变，从横空出世到进入低谷，再到今日的黄金时代。《人工智能简史》(约翰·马尔科夫著)一书责任编辑安焱在接受《中国新闻出版广电报》记者采访时表示：“虽然《人工智能简史》这本书对于国家战略没有明确的建议，但是书中体现了科学家对科技的发展起着十分关键的作用，所以大力培养和发展 AI 界的科技人才，是我们掌握人类未来的关键一环。”

2. 重视普及与梳理

当前，很多读者对 AI 概念及其发展脉络十分感兴趣，因此，很多关于 AI 的图书重视概念的普及以及相关历史知识的梳理。

在人民邮电出版社的任芮池看来，这类图书应站在普及概念的角度，讲述的故事应力求通俗易懂：“AI 几经浮沉，走过了第一个 60 年，在这个时间点向大众普及 AI 应该是开放的，因为 AI 的未来充满了不确定性和不可预知性。”

任芮池表示：“《人工智能简史》(尼克著，人民邮电出版社出版)站在了一个中立角度，从多个方面介绍了 AI 的发展历史，重点介绍了 AI 领域杰出的科学家，以及

他们创造非凡成果的有趣故事。更为重要的是，作者的叙述浅显易懂，使得这本书变成了一本青少年、文科生都能看懂的 AI 简史。”

不同于向民众普及 AI，《人工智能：国家人工智能战略行动抓手》立足产业，阐述国家在 AI 领域的政策布局。中国人民大学出版社大众图书出版中心主任曹沁颖认为，这本书的作者包括产业界人士、AI 科学家、法律专家、政策制定者等众多专家，因此具有很强的政策指导性。“这本书的技术篇揭示了 AI 这项技术神秘的面纱，帮助读者了解这项技术的前世今生；产业篇带领大家一览 AI 发展的全貌；战略篇展现了世界大国如何在战略领域对 AI 进行全方位的布局；法律篇、伦理篇和治理篇则直面 AI 给个人和社会带来的挑战，研究如何降低 AI 风险，拥抱 AI 带来的变革；最后，未来篇带领读者‘大开脑洞’，畅想 AI 带给人类社会的‘天方夜谭’。”

同时，她认为这本书虽然立足产业，但是叙述角度新颖。比如它对 AI 的概念界定非常巧妙，从概念界定出发引出了 AI 的基本技术框架和发展历史，让读者不是单纯记住一个概念，而是了解了 AI 技术的多个层面和维度。这种娓娓道来的方式，让读者更容易理解 AI 这个概念。

3. 人与机器人共舞

未来，AI 的前景究竟如何是我们关心和讨论的主要问题之一。仍有很多人持“AI 威胁论”的看法，要想人与 AI 和谐共生，我们应该做些什么呢？在人民邮电出版社的《人工智能简史》（约翰·马尔科夫著）中，作者以“将良知注入机器人内心的初步思考”为题，做了独特的阐述。

约翰·马尔科夫的《人工智能简史》提出：大规模取代人类劳动力的智能机器的出现，无疑将会引发一部分人的焦虑。约翰·马尔科夫认为，AI 并不是只有替代人类这一条路径，我们完全可以用 AI 来强化人类，未来一定是“人机共生”的世界，是“人与机器人共舞”的时代。

《人工智能：国家人工智能战略行动抓手》则认为，AI 最大的挑战针对的是监管者：创新意味着不断突破已有边界，或者在原本没有边界的地方探索。如何面对那些突破边界的实践？如何在没有边界的地方划定边界？如何在保护公民和鼓励创新之间平衡？太多抉择需要监管者做出，对他们而言这将是巨大的挑战。

我国 AI 相关的政策如下表所示。

我国 AI 相关的政策

| 时 间 | 发布机关 | 相关政策文件 | 政策内容 |
|-------------|------|------------------------------------|--|
| 2015 年 5 月 | 国务院 | 《中国制造 2025》 | 提出“以推进智能制造为主攻方向” |
| 2015 年 7 月 | 国务院 | 《国务院关于积极推进“互联网+”行动指导意见》 | 将人工智能作为“互联网+”的十一个重点布局领域之一 |
| 2016 年 3 月 | 国务院 | 《国民经济和社会发展第十三个五年规划纲要（草案）》 | 提出要“重点突破新兴领域人工智能技术” |
| 2016 年 5 月 | 发改委 | 《“互联网+”人工智能三年行动实施方案》 | 培育发展人工智能新兴产业，推进重点领域智能产品创新，提升终端产品智能化水平 |
| 2016 年 7 月 | 国务院 | 《政府工作报告》 | 人工智能首次被写入《政府工作报告》 |
| 2017 年 7 月 | 国务院 | 《新一代人工智能发展规划》 | 提出了“三步走”的战略目标，宣布举全国之力在 2030 年抢占人工智能全球制高点 |
| 2017 年 12 月 | 工信部 | 《促进新一代人工智能产业发展三年行动计划（2018—2020 年）》 | 促进人工智能产品发展，提升制造业智能化水平，推动人工智能和实体经济深度融合 |

1.3.2 AI 相关行业的发展现状

AI 应用涉及专用型应用和通用型应用两个方面，这也是机器学习、模式识别和人机交互这三项 AI 技术的落地形式。其中，专用型应用涵盖了国内 AI 企业的大多数应用，如英飞拓的人脸识别项目，其他专业型应用包括语音识别以及服务型机器人等方面；而通用型应用则侧重于金融、医疗、智能家居等领域的通用解决方案。目前（2018 年至 2019 年），国内 AI 的应用正处于由专业型应用向通用型应用过渡的发展阶段。

1. 计算机视觉

在国内计算机视觉领域，图像识别和人脸识别是主要的研究方向。

图像识别是计算机对图像进行处理、分析和理解，以识别各种不同模式的目标和对象的技术，识别过程包括图像预处理、图像分割、特征提取和判断匹配。

人脸识别是基于人的脸部特征信息进行身份识别的一种生物识别技术。人脸识别技术是用摄像机或摄像头采集含有人脸的图像或视频流，并自动在图像中检测和跟踪人脸，进而对检测到的人脸进行处理的一系列相关技术，通常也叫作人像识别、面部识别。

由于动态检测与识别技术的发展特性,静态图像识别与人脸识别的研究暂时处于领先位置,参与的企业既有腾讯、蚂蚁金服、百度和搜狗这样基于社交、搜索大数据进行整合的互联网公司,也有微软亚洲研究院、英特尔中国研究院、三星中国技术研究院这类传统硬件与技术服务商。同时,旷视科技等新兴技术公司也在各自专业技术和识别准确率上取得了不错的突破。

图像识别领域的代表企业有英飞拓、百度、搜狗、三星中国技术研究院、微软亚洲研究院、英特尔中国研究院等。

人脸识别领域的代表企业有英飞拓、旷视科技、腾讯优图、蚂蚁金服、FaceID、汉王科技、三星中国技术研究院、微软亚洲研究院、中科奥森、深圳科葩、Linkface、SenseTime 等。

动态视觉检测领域的代表企业有格灵深瞳、东方网力、Video++等。

2. 语音识别

语音识别的关键在于基于大量样本数据进行识别处理,因此,国内大多数语音识别技术提供商都在平台化的方向上发力,希望通过不同平台以及软硬件方面的数据和技术积累不断提高识别准确率。在通用识别率上,各企业的业绩基本都维持在 95%左右,真正的差异化在于对垂直领域的定制化开发。类似百度、科大讯飞这样的上市公司凭借深厚的技术、数据积累站在市场前列,并且通过软硬件服务的开发不断进化着自身的服务能力。此外,在科大讯飞之后发布“语音识别公有云”的云知声在各项通用语音服务市场也占据着不小的市场空间。

3. 智能机器人

由于工业发展和智能化生活的需要,对智能机器人的需求越来越大。国内智能机器人行业的研发主要集中于家庭机器人、工业机器人、企业服务和智能助手三个方面。

根据智研咨询的数据,2017 年全球物联网设备(笔者认为智能机器人是包含智能物联网设备的,即 IoT 是智能机器人的一种)的数量达到 84 亿台,同比增长 31%,预计到 2020 年,物联网设备的数量将达到 204 亿台。随着物联网蓝海市场的开启,服务类机器人的发展有望迎来机遇。其中,工业及企业服务类机器人研发企业依托政策支持和市场需要处于较为蓬勃的发展阶段,代表性企业包括依托中科院沈阳自动化研究所的

新松机器人、聚焦智能医疗领域的博实股份，以及大疆、优爱宝机器人、Slamtec 等专注工业生产和企业服务的智能机器人公司。

4. 智能家居

与家庭机器人不同，智能家居和物联网企业的主要着力点在于智能设备和智能中控两个方面。

在这其中，以海尔和美的为代表的传统家电企业依托自身渠道、技术和配套产品优势建立起了实体化的智能家居产品生态，而以阿里巴巴、腾讯、京东、小米等互联网企业为代表的公司则通过各自平台内的数据和终端资源提供不同的软硬件服务。

值得关注的是，科沃斯、BroadLink、感居物联、风向标科技、欧瑞博、物联传感和华为等技术解决方案供应商在通用硬件和技术、系统级解决方案上已成为诸多智能家居和物联网企业的合作伙伴。

综合来看，智能家居和物联网企业由于市场分类、技术种类和数据积累的不同，各自提供着差异化的解决方案。在既定市场中，没有绝对意义上的排斥竞争，各企业之间的合作融合度较高。

5. 智能医疗

国内智能医疗领域的研究主要集中于医用机器人、医疗解决方案和生命科学领域。

由于起步较晚和受技术门槛的限制，国内医用机器人的研发水平和普及率相较于国际一线水平仍存在一定的差距，从事企业主要集中在手术机器人和康复机器人两大领域，这其中以新松机器人、博实股份、妙手机器人、璟和技创等企业为代表。

在医疗解决方案方面，主要玩家是以腾讯、阿里巴巴、百度和科大讯飞为代表的国家新一代 AI 开放创新平台（备注：依托百度公司建设自动驾驶国家新一代人工智能开放创新平台；依托阿里云公司建设城市大脑国家新一代人工智能开放创新平台；依托腾讯公司建设医疗影像国家新一代人工智能开放创新平台；依托科大讯飞公司建设智能语音国家新一代人工智能开放创新平台）。上述公司通过和政府、医疗机构合作，为脑科学、疾病防治、医疗信息数据等领域提供智能解决方案。

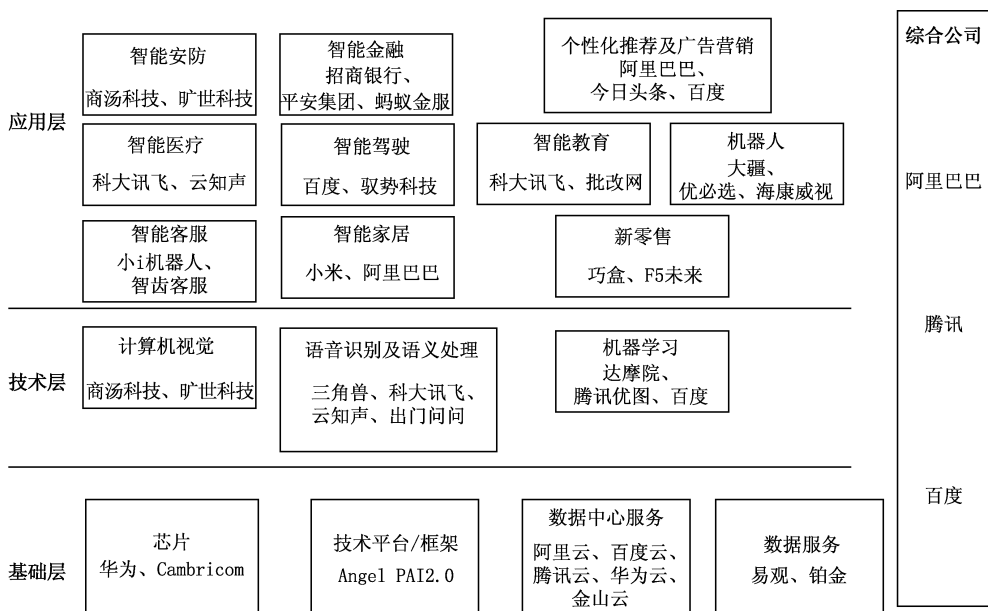
而在生命科学领域，研究的着眼点在以基因和细胞检测为代表的前沿研究领域，

代表企业有华大基因、碳云智能、贝瑞和康等。

在笔者看来，国内 AI 产业链的基础技术链条已经构建成熟，AI 技术和应用集中在人脸和图像识别、语音助手、智能生活等专用领域的场景化解决方案上。就趋势来看，未来国内 AI 领域的差异化竞争和突破将主要集中在 AI 相关技术的突破和应用场景升级两个层面。

1.3.3 AI 相关公司的发展现状

各 AI 公司的现状如下图所示。



各 AI 公司的现状

在国内 AI 产业链中，虽然 80% 的企业属于 B 轮或 B 轮以前的初创企业，但值得注意的是，仍有一批关键势力依托自身技术、资金与数据门槛构成了国内 AI 行业的第一梯队。这其中既包括以 BAT（百度、阿里巴巴、腾讯三家公司英文名称首字母的缩写，代指三家公司）为代表的传统巨头，也有科大讯飞这样的细分领域龙头。从完善产品线到构建平台，这些企业形成了国内 AI 行业的核心力量与关键势力。

1. 英飞拓

英飞拓是一家领先的 AI 企业，它是全球领先的电子安防与光通信设备智能制造

商，也是创新型、信息化、集成化的整体安防制造商，致力于为全球安防市场提供高端、安全、值得信赖的解决方案。英飞拓的产品涵盖全系列视频监控设备，包括摄像机、快球、高速云台、视频矩阵、IP 视频系统，以及门禁控制系统。

英飞拓的技术广泛应用于智能交通系统、城市安全与治安监控系统、智能楼宇、金融系统、电力和水利系统等，成功的案例有重庆平安城市、广州珠江新城 W 酒店、深圳会展中心等。

英飞拓的数字营销业务在全球市场加速扩张。2016 年 8 月 2 日，英飞拓发布公告称拟以 6.5 亿元现金收购北京普菲特广告有限公司 100% 的股权。至此，英飞拓拥有 Infinova、March Networks、“藏愚”三大专业安防业务品牌，以及 Swann 民用安防业务品牌，并通过英飞拓环球、印度英飞拓、美国英飞拓、加拿大英飞拓等子公司覆盖全球安防市场。英飞拓已通过搭建销售团队及建设营销网点的方式在全球市场获得相当的市场地位，积累了显著的全球品牌优势。这次收购实施后，英飞拓利用自身资源优势协助普菲特获取海外的媒体资源（全球范围内的数字营销渠道资源），拓展普菲特的海外业务。同时，普菲特也凭借获取的海外数字营销资源以及自身强大的数字营销体系，助力英飞拓深化全球市场扩张策略。

2. 百度：技术驱动的应用型生态

百度在 AI 领域的布局更侧重于应用型生态，在 BAT 三家中，百度也是最接近完成由专用型应用向通用型应用过渡的公司。

截至 2018 年 1 月，百度研究院由原来的三个实验室变为五个实验室，分别是深度学习实验室（IDL）、大数据实验室（BDL）、硅谷人工智能实验室（SVAIL）、商业智能实验室（BIL）、机器人与自动驾驶实验室（RAL）。百度已经在图像识别、图像搜索、语音识别、自然语言处理、智能语义、机器翻译和精准广告等方面取得了显著进展（超过 500 项国际专利，其中包括 270 余项神经语言程序学领域的专利和 120 余项深度学习领域的专利）。

凭借搜索引擎家的百度拥有强大的数据获取和挖掘能力。百度为外界提供了大数据存储、分析和挖掘技术，促进其在医疗、交通等多领域的具体运用，并在若干领域开放了自己的 AI 生态，发布了多款应用型产品。在语音识别方面，吴恩达及研发团队发明了一种新的语音识别系统，这款基于深度学习的语音识别系统可以在嘈杂环

境下实现将近 81% 的辨识准确率。该语音识别系统用深度学习算法取代了原来的模型，在递归神经网络或者模拟神经元阵列中进行训练，让语音识别系统更加强大。

在图像识别方面，百度也一直在利用深度学习技术来提高图像识别的精度。整合了这些技术的百度形成了强大的数据存储与计算能力，从而可以进行多样化的并行计算，支持生成、配置针对不同应用和场景的网络结构，为 AI 提供了有力的硬件支持。其中，最具代表性的“百度大脑”通过深度学习来模拟人类大脑的神经元，参数规模达到百亿级别，构建了世界上最大规模的深度神经网络。“百度大脑”融合了深度学习算法、数据建模、大规模 GPU 并行计算等技术，实现了实时学习和成长，它拥有 200 亿个参数，构成了一个巨大的深度神经网络。

3. 腾讯：基于用户体系的软硬件服务型生态

相比以技术为导向的百度 AI 生态，依托高频应用和庞大用户群体的腾讯在 AI 领域的布局主要聚焦于基于用户体系的软硬件服务型生态。

随着技术和应用环境的成熟（移动设备的普及、用户使用习惯和认可度的提升、云服务软件形态的成熟），整个人脸识别和图像识别行业都属于朝阳产业，因此，有着中国互联网庞大用户基础的腾讯也选择了在人脸识别和图片识别领域切入。

在这一领域的技术研发和应用落地中，腾讯均处于国际领先地位。在 FDDB 人脸检测、LFW 人脸识别、PASVAL VOC 2012 图片识别大赛中，腾讯优图团队都刷新了识别准确率的国际纪录。依托腾讯内部完善的生态，腾讯优图的人脸识别技术已经与金融、安防与身份识别等领域结合在了一起。其中，既包括与腾讯征信、微众银行、财付通、广点通等内部业务的合作，也涉及与信用评估、全国居民身份证号码查询服务等外部业务的合作。在与内、外部的合作中，相关技术也在业务中完整落地并形成闭环，经过实际业务海量数据的长期实际考验，腾讯也通过持续反馈来优化算法。

在成立了腾讯智能计算与搜索实验室之后，微信还与香港科技大学联合建立了 AI 实验室，主要研究方向包含图像识别、语音识别、语义理解等方面，部分基础研究成果（如语音转文字）已被集成为微信的功能。

值得注意的是，在专业应用领域，腾讯也在不同程度上做出了尝试。无论是由腾讯财经研发的自动写作机器人 Dreamwriter，还是在硬件领域推出的智能球形机器人“微宝”，都反映了腾讯计划在未来更多细分领域推出服务型产品的思路。

未来，随着 TencentOS 的应用以及微信物联等产品的落地，腾讯在智能生活层面的硬件产品研发是非常值得关注的。

4. 阿里巴巴：以阿里云为基础的业务蓝图

与英飞拓、百度、腾讯不同的是，阿里巴巴在 AI 领域的布局主要集中在专业领域的通用型应用和智能家居两个方面，而这一切都离不开以阿里云为基础的大规模分布式云计算和 GPU 集群。

从 2015 年开始，阿里巴巴就开始在 AI 领域发力，其中最具代表性的是首个可视化 AI 平台——DTPAI，它集成了阿里巴巴的核心算法库，包括特征工程、大规模机器学习、深度学习等。在此基础上，阿里巴巴推出了 AI 客服——阿里小蜜，以及基于神经网络、社会计算、情绪感知等原理的智能程序。

2015 年 6 月，阿里巴巴向软银旗下的机器人公司 SBRH 战略注资 145 亿日元。相比机器人制造，阿里巴巴的优势在云计算和大数据领域，这也将成为阿里巴巴在机器制造领域差异化竞争的壁垒。

而通过投资 SBRH，阿里巴巴得以建立起机器人产业的通路，并在机器视觉、语音解析、家庭智能控制、智能网络安全等方面与 SBRH 建立研究合作和沟通机制，帮助其在智能家居方面积累和扩展基础技术。

在智能物联领域，阿里巴巴整合了集团旗下的天猫电器城、阿里智能云、淘宝众筹三个业务部门，在内部调动各类优质资源支持智能产品的推进并加速智能硬件的孵化速度。在 2015 年 4 月，阿里巴巴集团宣布成立阿里巴巴智能生活事业部，全面进军智能生活领域。

1.3.4 AI 相关行业的人才现状

AI 既是计算机科学的一个分支，又是计算机科学、控制论、信息论、语言学、神经生理学、心理学、数学、哲学等多个学科相互渗透而发展起来的综合性学科。基于国际计算机学会的分类，AI 领域主要被分为两级，其中包含 27 个一级概念，44 个二级概念，深度学习、信息检索等是 AI 在国内和国际上比较活跃的子领域。

中国 AI 学者主要分布在以下领域：数据挖掘、机器学习、计算机视觉、数据库、自然语言处理、图像处理、计算机图形学、多媒体技术、虚拟现实、人机交互、信息

检索以及社交网络等。

根据东方头条网 2018 年 1 月的数据显示，国内 AI 相关专业的学者数量如下：数据挖掘（230 人）、机器学习（207 人）、计算机视觉（193 人）和数据库（171 人）领域的人才充足，而机器人、跨学科计算、机器学习应用等领域的人才稍显不足。大部分学者的专业是计算机相关专业（如计算机、电子工程、数学等），但少部分学者的专业为与计算机不明显相关的专业（如经济学、生物、心理学等）。

在中国，从事 AI 相关领域研究的学者共 1073 名，其中有 1045 名来自高校，任职于高校的 AI 学者的数量远远多于在科研机构和企业任职的 AI 专家的数量。AI 学者数量最多的三家单位是清华大学、北京大学和中国科学院，分别有 72 名、66 名、51 名，这三家单位 AI 学者的数量远多于其他单位。

人才到底有多缺呢？

自 2017 年被写进政府报告，AI 在国内的前景彻底明朗化，之后连续颁布的《新一代人工智能发展规划》《促进新一代人工智能产业发展三年行动计划（2018—2020 年）》更是将国内 AI 的发展推向了一个高峰。

随着南京大学等高校相继设立 AI 学院，更多的人意识到一个现象：在国家一路高歌为 AI 产业大开绿灯的时候，“人才缺口”却拖了后腿。

此前，AI 孵化器 ElementAI 基于 LinkedIn 和各大会议的专家数据，对中国、美国、加拿大、德国、日本等主流国家的 AI 人才做了一项调查，结果显示，全世界大约有 2.2 万名拥有博士学位的 AI 从业人员和研究人员，以及 5400 位 AI 专家，而中国以总人数 619 人位列第 7，人数只有美国的 5%。2017 年 7 月，LinkedIn 发布《全球 AI 领域人才报告》，报告内容显示，截至 2017 年一季度，全球 AI 领域技术人才的数量超过 190 万，其中美国的人才总数超过 85 万，而中国仅拥有 5 万多人。腾讯研究院联合 Boss 直聘发布的《2017 全球人工智能人才白皮书》显示，仅在 2017 年的前 10 个月，AI 人才需求量已经达到 2016 年的两倍、2015 年的 5.3 倍，年复合增长率超过 200%……

以上是多个平台关于 AI 人才的统计报告，虽然这些数据的统计方法与统计标准不一样，但是从各个角度来看，有一个事实是无可争议的——中国 AI 人才的拥有量与市场需求之间存在巨大缺口。

仅以国内市场为例，用腾讯研究院的俞点的话来说，累积到现在，国内 AI 领域

虽有几万名毕业生，但是市场的人才缺口是百万级的，所以现在的存量远远不能满足市场的需求。

纵观国内，在大环境层面，AI 已经上升为国家战略；在市场层面，AI 有着极大的应用市场（据中商产业研究院发布的《2018—2023 年中国人工智能行业市场前景及投资机会研究报告》的数据显示，2018 年，中国 AI 的市场规模有望达到 238.2 亿元，增长率达到 56.6%）；在学术层面，AI 论文在近年来的各大学术会议上大放异彩，全球超 4 成的 AI 论文是中国人所写，一些论文甚至在 IEEE 等国际学术会议获奖……

不过，我们却在最重要的硬件设备——人才方面出现了不足，尤其是在高端 AI 人才上。

AI 人才从哪里来？源头一定是学校。为了填补 AI 人才缺口，各高校相继设立了 AI 独立学院。事实上，在国家重点关注 AI 之前，作为计算机学科下面的一个分支，不少学校都开设了 AI 相关专业，譬如中国科学院大学开设的“模式识别与智能系统”“计算机应用技术”专业，北京大学、复旦大学、华南理工大学、北京邮电大学、北京科技大学等开设的“智能科学与技术”专业。

不过，这些专业大多挂在计算机、信息科学与技术等学院下面，用 AI 专家、南京大学计算机科学与技术系主任周志华教授的话来说：“目前（2018 年），计算机专业人才的培养模式，已经不能满足培养 AI 人才的迫切需要。”当然，这种人才培养模式也跟不上当前市场需求急速增长的速度。

为了响应国家的号召，也为了缓解人才缺口现象，自 2017 年开始，一些高校相继设立人工智能学院。

设立人工智能学院的高校

| 高 校 | 时 间 | 招 生 对 象 |
|----------|-----------|----------|
| 中国科学院大学 | 2017.5.25 | 研究生/在职人员 |
| 西安电子科技大学 | 2017.11.2 | 本科生/研究生 |
| 湖南工业大学 | 2018.1.13 | 本科生 |
| 上海交通大学 | 2018.1.18 | 研究生 |
| 重庆邮电大学 | 2018.2.7 | 本科生/研究生 |
| 南京大学 | 2018.3.6 | 本科生/研究生 |
| 清华大学 | 2018.6.2 | 研究生 |

在招生方面，这些高校多是从本科生到硕士生、博士生全面覆盖。截至本书完稿时（2018 年），中国科学院人工智能技术学院已经开始招生，首批研究生已于 2017 年 9 月完成入学注册，其余高校也陆续于 2018 年秋季开始招生。

随着学院的成立，加上从国家到各个行业对 AI 产业的关注，一些疑惑随之而来：谁来教？教什么？学生又能学到些什么？

周志华教授在 2018 年 3 月对澎湃新闻表示：按照高校计算机专业的宽口径人才培养模式，150 个学分中大约有 60 个学分是通识课，15 个学分是毕业双创课，人工智能方面的学习很有限，以至于高度浓缩到了只是给学生做高级科普的程度，难以充分培养学生全面深入地掌握人工智能知识、解决企业关键问题的能力，不能适应智能产业发展的要求，因此开设了人工智能学院。

看着这些课程，笔者突然有了几点疑惑。

首先，与已有的计算机教学相比，作为进一步细化而独立出来的学院，人工智能学院在教学内容等方面与前者有什么本质的不同呢？

高校人工智能学院开设的课程

| 学 校 | 西安电子科技大学 | 湖南工业大学 | 重庆邮电大学 | 南京大学 |
|-----|---------------------------|--|-----------------------|----------------------|
| 专 业 | 智能科学与技术 数据科学与大数据技术（拟建） | 智能科学与技术 数据科学与大数据技术 机器人工程 计算机科学与技术（合并） | 智能科学与技术 数据科学与大数据技术 | 机器学习与数据挖掘 智能系统与应用 |

以西安电子科技大学（以下简称“西电”）的人工智能学院为例，其招生范围覆盖了本科生和研究生。相比于研究生课程的千篇一律，本科生的专业方向更是人们关注的重点。西电只为本科生开设了两个特色专业：智能科学与技术、数据科学与大数据技术（拟建）。前者的目标是培养学生的电子技术、信息处理理论、电子信息系统、计算机与互联网络、智能科学与技术等学科的基本知识，而后者则要求学生能掌握大数据相关的统计、计算机、数学和应用学科的基础知识，以及数据分析、技术开发和应用的基本技能。

从本质上来看，这与当前计算机专业所学的内容似乎是一样的，不同的是，人工智能学院要求学生具备更扎实的数学、计算和程序等学科的基础知识。

但基础知识扎实了，后面又该如何继续呢？进一步考研深造？还是凭着扎实的基础知识踏入社会？

为此，笔者也向几家公司了解了一下技术团队的学历结构。在专注于智能语音技术的云知声团队中，具有博士学位的人员占比为 45%；而在以智能交互为核心的极限元团队中，拥有硕博背景的人员占比为 60%，剩下的几乎都是本科毕业生。

在 AI 专业人才缺乏的前提下，“应用数学、物理、统计学方面的（研究生）也会考虑。”图普科技创始人兼 CEO 李明强表示。

既然如此，那就意味着，在产业界看来，AI 人才是急需的，但也存在暂时性的解决方案；高素质 AI 人才才是最为紧缺的，这才是需要人工智能学院解决的问题。那么，如果人工智能学院只是对学生的数学、计算机科学、程序设计基础的相关知识进行更有针对性的“加固”，那么，这些本科生在毕业后还需要继续考研深造，这与现在计算机等相关专业的本科毕业生又有什么本质的区别呢？

到了这里，笔者产生了第二个疑惑，在课程大致体系尚未构建好的时候，人工智能学院面向本科招生真的合适吗？

在众多高校放出“设立人工智能学院”的消息后，就有人质疑了：AI 一级学科还没有建起来，暂时还不能直接招本科生。

“在现有的学科体系下，计算机科学与技术是一级学科，往下是计算机软件与理论、计算机系统与结构、计算机应用等二级学科，二级学科再往下才是 AI 方向。”针对高校的学科划分，四川大学计算机学院院长章毅曾如此表示。

如今，人工智能学院已经成了一个与计算机等学院比肩而立的独立学院，但是，作为一个“新生儿”，其课程体系更多的还是在摸索中，尤其是面向本科生的课程体系。

都说术业有专攻，但对于 AI 产业，面对“人才荒”问题，浙江大学计算机学院副院长、AI 研究所所长吴飞提出，我们需要充分挖掘现有的人才存量资源，计算机科学、自动化、统计学等专业的人才皆可从事 AI 工作。

在解决 AI 产业“用工荒”问题的同时，也折射出一个“尴尬”的现象：一方面，因为专业性与深度的问题，一些人认为当前计算机专业培养的毕业生不能满足市场对 AI 人才的需求；另一方面，在人才需求不能被满足的当下，一些非 AI 专业的毕业生

被“抓”来救急，或是在企业内通过再培训来提升其 AI 素质。最终来讲，AI 领域所缺少的还是那些专业的、高素质的 AI 人才。但是，现在还处于摸索初期的高校教学体系是否真的能将本科生培育成高素质的 AI 人才？

此前，因为 AI 相关课程散落在各个学科里面，被迫碎片化、离散化、简单化，这导致了高素质的 AI 人才难以培养。虽说现在人工智能学院已经是一个独立的存在，但从一些高校设立的课程来看，似乎并没有一级科目、二级科目之分，尤其是对本科学生。

如此一来，在人工智能学院的学习中，相比已经掌握了基本知识与技术，并对未来研究方向有了决断的研究生，学识与技能还是一片空白，且具体研究方向也不定的本科生或许会产生一定的疑惑。由此来看，在一级科目尚未建立的情况下，人工智能学院一上手就直接招收本科生，对学生而言似乎有点“超纲”了。

最后，笔者确定的是，不管是从当前的意义，还是从未来的角度来看，人工智能学院的设立都是一件极具意义的事情。其中，面向本科招生是培养高素质 AI 人才的基础建设。不过，从当前的情况来看，在教学体系等基础建设尚未完善的情况下，高素质 AI 人才的培养也许应该先从研究生阶段开始，至于“数量”的缺口，则可以依靠不需要知晓太多专业性技术原理的 AI 应用型人才稍加弥补。

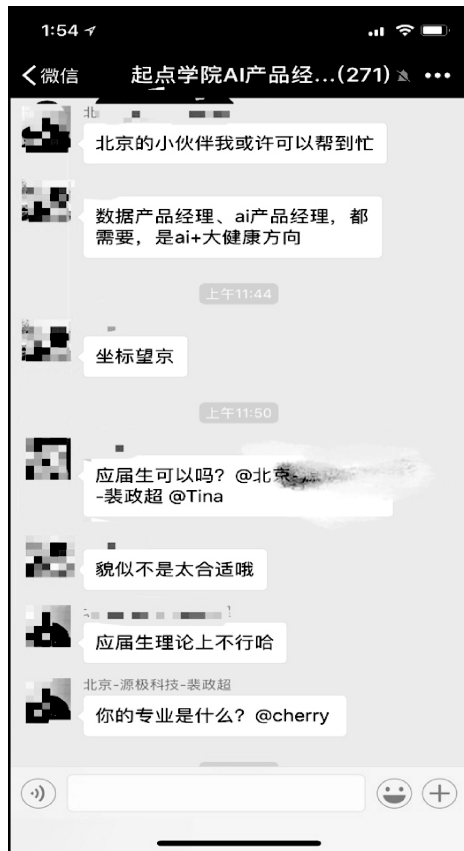
1.4 AI 产品经理的现状

笔者在前面分析了 AI 行业的人才现状，鉴于本书面向的主要是应用型人才，所以有必要单独讲一讲制作应用的人——AI 产品经理的现状。

1.4.1 岗位需求与人员素质不匹配

2018 年春节前夕，人人都是产品经理社区的创始人老曹（曹成明）找到笔者，说能不能给他平台上的同学讲一下打造 AI 产品的方法和 AI 产品经理应该具备的知识、技术。笔者明白，在当下（2018 年），企业对 AI 是非常拥抱的，但苦于没有人才，没有具备实战经验的 AI 产品经理人才，所以就答应老曹做这件事。笔者在跟起点学院联合开发课程的过程中发现这样一个问题：普通产品的产品经理和市场上其他岗位的

人才是非常愿意了解、学习并从事 AI 产品经理工作的，但是自己的能力、经验与岗位需求不匹配。

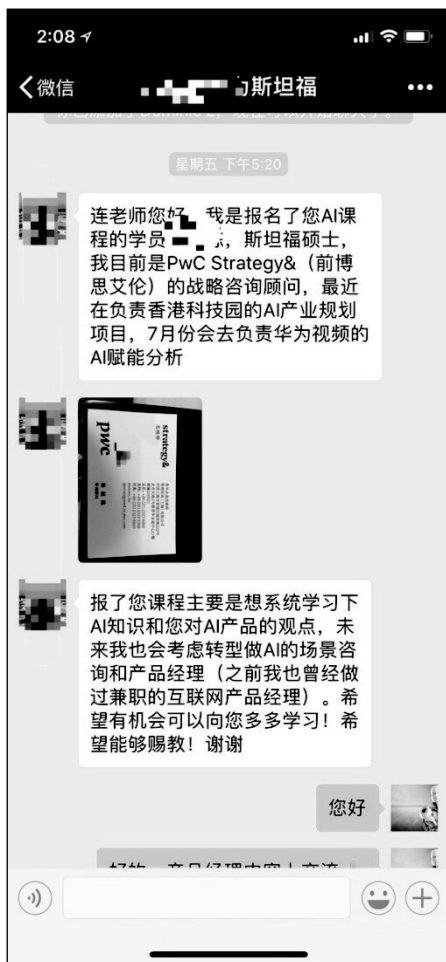


AI 产品经理的岗位要求

1.4.2 AI 产品经理来源广泛

AI 产品经理的人员来源十分广泛，在传统时代，产品经理的来源就已经很广泛了，但是在 AI 时代，产品经理的来源更加广泛，这是因为 AI 的魅力吸引了大量的优秀人才。

例如，笔者的 AI 产品经理入门课程就吸引了毕业于斯坦福大学的高才生，以及来自新加坡等地的优秀人才。



AI 吸引了大量的优秀人才

当下，AI 产品经理主要来自以下几个渠道。

(1) 技术人员转岗。一些毕业于计算机、软件工程等专业，做了一段时间编码工作的人，逐渐发现编码不是自己的“真爱”，但往架构师、技术总监方向发展却苦于相关经验比较少，于是特别希望找一份既需要了解技术，又不需要编程的工作，转到产品经理岗位成了很合适的选择。而传统产品的产品经理岗位显然不是时下的趋势，所以就转到 AI 产品经理岗位上来了。

(2) 市场、销售人员转岗。在传统产品的时代，市场、销售人员转型成为产品经理的先例早已有之。在 AI 时代，市场、销售人员在营销过程中通过对用户进行调研，

会发现已有的传统产品几乎没有生存空间，因此，根据市场需求，自然而然转到了 AI 产品经理的岗位上来。

(3) 运营、设计人员转岗。运营、设计人员在与消费者、用户打交道的时候，能够第一时间发现时下最紧缺的人才就是 AI 产品经理，加之运营、设计人员在日常工作中经常与产品经理打交道，比较熟悉打造产品的流程和环节。两方面的因素导致运营、设计人员转岗成为 AI 产品经理是一件自然而然的事。

(4) 应届毕业生。笔者在打造 AI 产品时，经常会遇到应届毕业生和尚未毕业的学生，这部分准备成为 AI 产品经理的同学，往往因为没有工作经验而感到迷茫，但同时对 AI 能够实现的产品效果存在过度的期待。

(5) 创业者。在 AI 领域的创业者中，有一类是 AI 专业毕业的，他们发现一个需求点（如机器视觉、人脸识别）后，直接投入 AI 的创业大潮中来了；还有一类是行业专家，这类专家在原来的业务发展过程中发现人工服务无法满足消费者的需求，市场的需求倒逼着他们做 AI 产品。

1.4.3 成为 AI 产品经理可能遇到的主要问题及解决方案

上文分析了几种成为 AI 产品经理的渠道，下面重点分析一下典型人群的典型问题，并给出笔者的建议。

1. 技术人员转岗

经常遇到的问题：困于技术。

解决办法：技术人员转岗成为 AI 产品经理通常是两种情况，一种是不希望再对技术进行深造，另一种是技术已入臻化。对前一种人，笔者建议不要再“修炼”技术，而是从 AI 产品的需求出发，从当下 AI 技术的成熟度出发打磨产品。对第二种人，笔者认为，人到了接近真相的境界后会有一个共性——求是，即看到事物的本质，希望第二种产品经理能够平衡好产品的需求和技术思维。

2. 市场、销售人员转岗

经常遇到的问题：功能堆叠。

解决办法：笔者在 2015 年至 2017 年带过不少从市场、销售岗位转型成为 AI 产品经理的人，由于他们经常联系用户，频繁与用户交流需求，同时也会参加很多展会，

因此天生具备调研用户、感知市场前沿需求的能力，这是这类 AI 产品经理的优势。但是，他们也有短处，市场、销售岗位转岗过来的 AI 产品经理的技术知识相对匮乏，经常会把展会上看到的、新的、尚未成熟的样机、展品甚至概念带到产品中来。例如，有一位 AI 产品经理参加 2016 年的 CES 后看到智能试衣镜，就仓促决定做一款家用智能试衣镜产品，而忽视了人体建模算法、人体本身非规则等客观情况。笔者建议从市场、销售岗位转岗成为产品经理的人员注重补上 AI 技术的相关知识。

3. 设计、运营人员转岗

经常遇到的问题：过度重视界面设计，唯用户说法论。

解决办法：笔者还记得 10 多年前帮别人建网站时，用 Photoshop 就可以建一个站点，现在回想起来确实需要再强化一下界面设计。不过设计、运营人员转岗到 AI 产品经理岗位以后，不能再“唯设计论”。毕竟，AI 产品的核心是用 AI 技术赋能产品，所以在传统时期我们强调的界面降噪处理、界面交互设计等内容的重要性要下降。不是说 AI 时代就不需要重视设计，而是相比 AI 的核心流程来说，设计要退居第二位。此外，设计、运营岗位出身的 AI 产品经理需要避免“唯用户说法论”，因为在 AI 时代，有时候用户并不知道产品使用了哪些 AI 技术，即 AI 技术是在产品背后支撑着产品。

4. 应届毕业生

经常遇到的问题：经验匮乏。

解决办法：应届毕业生的天生优势是年轻，犯错的成本较低；劣势是没有经验。应届毕业生应努力加入一个以 AI 技术为核心的产品团队，以便学习 AI 产品的需求趋势、AI 产品的打造流程。

5. 创业者

经常遇到的问题：精力不足。

解决办法：大多数创业者遇到的主要问题是精力不足。因为创业者或多或少都有一种能力或者某些资源，否则是无法启动的，所以笔者建议创业者亲自打磨第一个版本、第一代产品，之后要调整自己的定位，把制定产品的盈利方向作为主要工作，把寻找需求、满足用户需求交给招聘来的 AI 产品经理。

1.5 AI 产品经理认识 AI 的 3 个阶段、9 个过程

AI 产品经理在团队中是理性的，同时也是最懂人性的。AI 产品经理一方面要用符号化的逻辑赋能产品，另一方面要顾及用户无法量化的感受。

根据笔者的经验，AI 产品经理认识 AI 一般会经历 3 个阶段。

(1) 昨夜西风凋碧树，独上高楼，望尽天涯路。这个阶段是 AI 产品经理的迷茫焦虑期，找不到入口且担心被市场淘汰。例如，某位产品负责人因为不知道怎么操作，在北京、上海、广州、深圳等地四处寻找 AI 方案，深感迷茫和焦虑；笔者了解的一家智能工厂，公司创始人也是首席产品经理，经常跟笔者讨论某技术是否成熟、是否还能赶上潮流等问题。

(2) 衣带渐宽终不悔，为伊消得人憔悴。这个阶段的产品经理为了产品场景落地而苦苦打磨产品，综合运用 AI 算法、算力、数据来赋能所操刀的产品。例如，木蚂蚁物流机器人利用数据、算力和算法打造的生产线物流机器人，旷世科技用人脸识别赋能新零售。

(3) 众里寻他千百度，蓦然回首，那人却在灯火阑珊处。这个阶段的产品经理恍然大悟，发现 AI 产品其实是数据、算力和算法到达一定阶段后，需要对用户所需要的商品、服务进行升级，而产品经理需要做的就是利用 AI 技术让用户满意、开心。例如，抖音借助机器学习推荐的视频都是用户感兴趣的、乐意看的。

AI 是一种思维，就像 10 年前的互联网思维一样；AI 还是一种哲学，更是一种驱动市场进化的力量，即为市场创造更大的新价值。产品经理需要经过 9 个过程来成长为优秀的 AI 产品经理。

过程 1：AI 只不过是工具。

过程 2：AI 完全可以具备自我意识。

过程 3：具备自我意识并不等于具备综合生存能力。

过程 4：综合生存能力完全可以被复制。

过程 5：人类生存发展的综合能力尚不完全清楚。

过程 6：AI 不必模拟人类，AI 完全可以超越人类。

过程 7：AI 要为人服务，这应是我们的追求，而且在这方面容易产生意外，我

们要防范意外。

过程 8：人类本身需要觉醒与升华，为 AI 的发展指引方向。

过程 9：AI 在发展与控制方面会给人类带来新的课题，甚至伦理问题，探索的任务任重道远。

人们都说产品经理是离 CEO 最近的人，AI 时代的产品经理不仅是离 CEO 最近的人，更是一个需要具有更高阶综合素质的岗位。AI 时代的产品经理通过调用 AI 技术，寻找行业的切入点，综合市场、用户的需求来打造产品。

1.6 AI 产品经理应该了解的算法

算法包括基本算法、数据结构的算法、数论与代数算法、计算几何的算法、图论的算法、动态规划以及数值分析算法、加密算法、排序算法、检索算法、随机化算法、并行算法、厄米变形模型、随机森林算法等。

算法可以分为三类。

(1) 有限的、确定的算法。这类算法可在有限的一段时间内终止。它们可能要花很长时间来执行指定的任务，但仍将在一定的时间内终止。这类算法得出的结果常取决于输入值。

(2) 有限的、非确定的算法。这类算法可在有限的时间内终止。然而，对于一个（或一些）给定的数值，算法的结果并不是唯一的或确定的。

(3) 无限的算法。无限的算法是那些由于没有定义终止条件，或定义的条件无法由输入的数据满足而不终止运行的算法。通常，产生无限的算法是由于没有明确定义终止条件。

对 AI 产品经理来说，算法只有两种：传统算法和 AI 算法。

1.6.1 传统算法

传统算法是一种逻辑算法，是处理决策，而不是模拟人脑，所以不能叫 AI，可以称之为逻辑运算。例如：加、减、乘、除都是逻辑运算，也是一种算法。排序算法、二分树算法、递归算法都叫算法，虽然是算法，但是它们不智能，智能算法更突出的特性在于解决复杂的问题，在于模拟人脑的思维方式，所以叫 AI 算法。

传统算法是不智能的，没有自主学习和进化的能力，属于逻辑运算且是 AI 算法的一部分。智能算法在利用数据进行训练的基础上才有深度学习的能力。

1.6.2 AI 算法

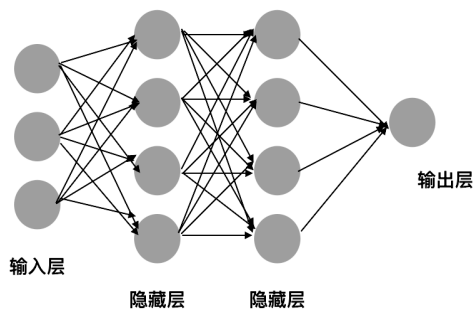
AI 算法首先能模拟人脑，其次能做处理决策，再次则是智能的。

一个机器嵌入了 AI 算法后，这个机器就拥有了人所具有的基本能力，比如观察、思考、学习、创造等。这部分要介绍的就是这种算法，实际上目前（2018 年）的算法还达不到全面超越人类的水平。

AI 算法主要由两部分组成：深度学习（Deep Learning）和强化学习（Reinforce Learning）。

1. 深度学习

深度学习就是多层人工神经网络，如下图所示。

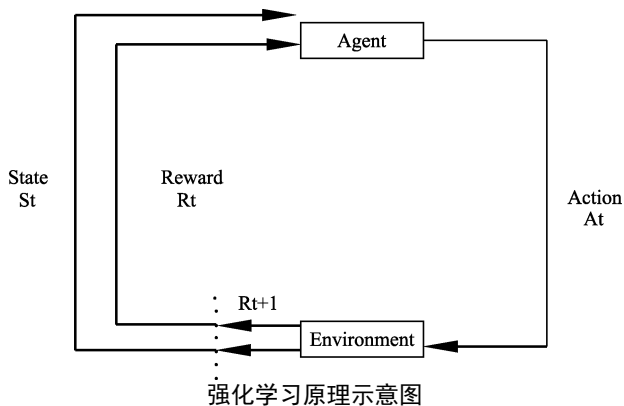


深度学习的示意图

它主要包括三个部分：输入层、隐藏层、输出层。输入层就是向机器输入的数据，比如我们问它“你吃了吗”；隐藏层就是对这句话进行特征提取和分析的过程，机器思考“我今天到底有没有吃饭”；输出层就是输出的结果，比如机器人说“吃了”。

2. 强化学习

强化学习相当于建立了一个机器与环境交互的过程。



强化学习主要包括两个部分：一个是 Agent，即机器；一个是 Environment，即环境。对于机器（Agent）来说，外面的任何事物，如天气、光线、张三、李四，都是环境（Environment）。比如，今天的环境（Environment）很冷，就是环境（Environment）当前时刻的状态（State）是“冷”，现在机器（Agent）知道了有点冷，就执行了一个动作（Action）——把衣服脱了。环境（Environment）知道了机器（Agent）把衣服脱了，天冷了脱衣服不对啊，所以环境（Environment）就给了机器（Agent）一个负的反馈（Reward），因为它的动作（Action）不对。然后环境（Environment）又给了机器（Agent）当前时刻的状态（State）“冷”，这时候机器（Agent）吸取上次的教训，就把衣服穿上了。环境（Environment）知道机器（Agent）把衣服穿上了，这是个正确的行为，所以就给了机器（Agent）正的反馈（Reward）。以后环境（Environment）冷的时候，机器（Agent）就知道要穿衣服了。

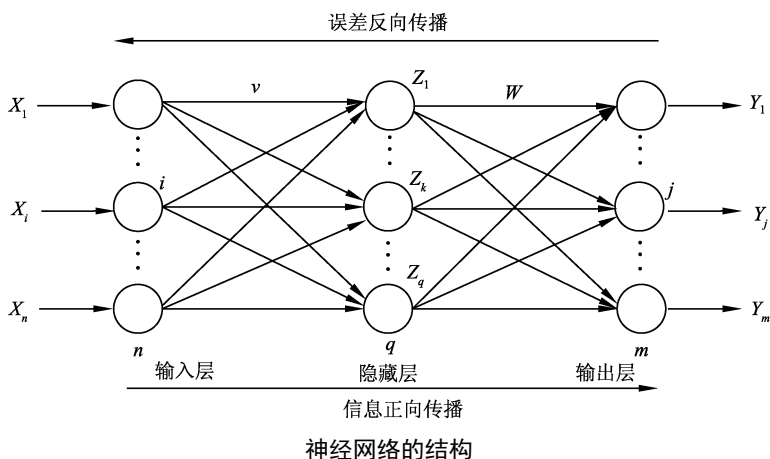
3. 神经网络的基本实现

神经网络的基本单元是神经元，第一层的神经元是我们输入的数据，是已知的。第二层的每个神经元相当于一个函数，它接收输入值 x ，经过函数映射得到输出值 y ，然后把 y 作为下一层的 x 。以此类推，到最后一层，得到最后一层的值。

神经网络的本质是一个分类算法，所谓分类算法就是一个公式，根据不同的输入 x ，经分类算法映射得到的 y 就是分类结果。就像我们输入猫的图片，得到的就是一个猫的标签或者狗的预测，这就是分类。

神经网络的简化表达： $y = ax + b$ ， y 是分类结果，同时也被称为预测值， x 是输入值， a 和 b 是公式参数。

神经网络的目标是调整 a 、 b 的值，使预测值和真值越来越接近，直到它们的差值小于等于我们设定的阈值。其中真值就是上面说的图片的标签，预测值就是公式的结果 y 。神经网络的结构如下图所示。



神经网络的实现主要有两个步骤：网络的搭建，网络的训练。

搭建网络时要考虑神经网络设计成几层，每层有多少个神经元，层与层之间是怎样连接的。层数越多、神经元越多，这个神经网络就越强大。但是当层和神经元太多的时候会出现一个过拟合的现象，就是随着层数的增多，拟合效果不一定越好，层的数量要根据激活函数、正则项、层与层之间的连接方式等因素综合考虑。

怎么训练网络呢？我们拿已经贴好标签的猫、狗的图片来训练。猫的图片我们就贴标签“猫”，狗的图片我们就贴标签“狗”。首先，我们把猫的图片输入神经网络，如果神经网络判断这个图片是猫，那就不去管神经网络里面的参数，如果神经网络判断猫的图片是狗，那就使用反向传播算法更新神经网络的参数。更新原则就是最小化真值和预测值之间的差值。

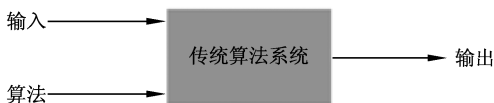
网络搭建完成，神经网络的模型训练好之后，就可以拿模型来识别猫、狗的图片了。

笔者的观点是，深度学习是在强化学习范畴内的。现阶段来看，强化学习仍然是 AI 的 Top-Level 算法。对 AI 其他算法有兴趣的读者也可以结合书中的引导再专门找一些 AI 算法资料来研究。

1.6.3 传统算法和 AI 算法的区别、联系

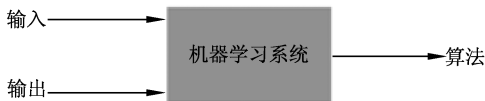
我们知道 AI 算法包含传统算法中解决复杂问题的部分。

先看传统算法系统。传统算法是已经给定了一个算法处理逻辑的系统，当用户给系统一个输入时，系统按照这个算法处理逻辑进行处理，就能得到相应的输出，如下图所示。



传统算法系统原理示意图

再看 AI 算法中的机器学习系统。机器学习系统是给定输入、输出的（这里先介绍监督学习），让系统根据这些数据学习算法，以后，通过这个学习到的算法模型对给定输入进行判断，得到输出结果，如下图所示。



机器学习系统原理示意图

例如，在传统的算法系统中，我们的要求是，给定一个数 n ，输出 $1+2+\dots+n$ 的结果。我们很容易由等差数列求和公式得到算法程序，如下图所示。

```
# include <stdio.h>

double get_sum(int n)
{
    double ans;
    ans = n*(1+n)/2.0;    //等差数列求和公式
    return ans;
}

int main(void)
{
    double res;
    int n;
    printf("please input an integer:\n");
    scanf("%d",&n);
    res = get_sum(n);
    printf("%lf\n",res);
}
```

求 $1+2+\dots+n$ 的结果的算法

根据上面的算法流程，我们输入一个数 n ，直接用程序“跑”一下，就能得到结果。

机器学习系统是否也能够直接编写判断逻辑算法呢？为此，笔者用一个机器学习最经典的入门例子——手写数字识别来解释。

比如我们要判断一个手写数字是几。按照我们的思维逻辑，需要写 10 个 if...else... 来分别判断是不是 0、是不是 1、是不是 2.....

比如判断是不是 3，就会有下图所示的判断过程。

```
if(满足是3的条件)
{
    printf(这个手写数字是3);
}
else
{
    printf(这个手写数字不是3);
}
```

判断是不是 3 的代码

看起来我们好像完成任务了。但是，“if(满足是 3 的条件)”这个“满足是 3”的条件怎么编写呢？我们先看看标准的 3 是什么样子的，如右图所示。



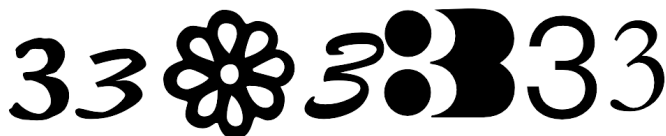
标准的 3

我们可以根据这个标准的 3 和其他数字的不同特征来编写，比如 3 开始是向右上提，最后是向左上提，中间有一条短横线，等等。

其他的数字类似，只要找到每个数字和其他数字的不同，就能编写出 if() 中“满足是数字几”的条件。

这样不就满足要求了？为什么还需要机器学习算法呢？

问题没有这么简单。我们真正碰到的手写的数字 3，根本不会有上图所示这么标准，而是各种各样的都有，如下图所示。



不同种类的“3”

对于上图这些数字，我们通过人眼观察还是能够认出大部分是 3，但是它们都满

足前面刚刚说的 3 的特征吗?上面好几个 3 都没有满足,我们无法一一列出现实生活中的各种情况。

这时,我们就需要机器学习算法自动地学习相关的特征,构建出一个机器学习模型,进而帮助我们判断一个手写数字是 3 还是其他数字。

其实从某种程度上来说,上面这两个例子都能反映两种系统的本质区别。

1.7 AI=数据+算法+算力

AI 是对人的智能的模拟,为了使得 AI 产品具有人的能力,需要收集大量的、正确的人的语言、行为、情感等数据。数据是基础,然后用这些数据来训练 AI 算法。用 AI 算法计算数据的规则时需要大量的计算资源,即算力。

因此,在某种程度上来说,AI=数据+算法+算力,算法、算力、数据也是 AI 产业爆发的三大条件。

例如,平安金融拥有大量的用户办理金融业务的数据,基于此,平安金融的 AI 产品经理协同 AI 算法工程师搭建了 AI 算法模型,即金融 AI 风控模型,利用计算机资源不停地进行训练。平安金融的 AI+金融产品已经能够将原来用户需要到现场才能办理的业务转到线上,这就是通过“数据(庞大的用户数据)+算法(生物识别算法等 AI 算法)+算力”实现的 AI 产品。

进一步来说,AI 从概念到产品的爆发需要具备三个条件,即算法、算力、数据。

在算法方面,随着深度学习理论和工程技术体系的成熟,包括通过云服务或者开源的方式向行业输出技术,先进的算法被封装为易于使用的产品和服务,越来越多的人和公司能够使用这些算法。AI 相关的技术包括了水平层和垂直层的技术,水平层的技术主要体现在算法方面。这两个层面都有很多大厂商在做,谷歌、亚马逊、微软等都试图搭建通用的机器学习和深度学习底层计算平台。而上层的应用,比如语音识别、文字识别、图像识别、即时定位等,也有很多开源的框架,特别是在图像处理领域,很多框架都应用了机器学习和深度学习技术。从这方面来看,笔者认为 AI 的产业化已经开始了,各个层面的企业都已参与其中。

微软在 2016 年 9 月份进行了重组,成立了微软人工智能及微软研究事业部。2018 年 4 月,微软人工智能及微软研究事业部的重要性明显增加,与体验和设备部门、云

计算与人工智能部门成为微软旗下三大工程部门，拥有 5000 多位科学家和工程师，这是一项巨大的投入。再比如，IBM 在 2018 年年初提出转型到认知计算（Cognitive Computing）和云平台，高调进入 AI 领域。除此之外，国内的百度、科大讯飞等公司也在 AI 上投入了大量的资源。

所以，从趋势上来讲，AI 时代已经到来是明确的。

在算力方面，云计算的兴起发挥了非常关键的作用。深度学习是极其消耗计算资源的，而云计算可以以低成本获取大规模的算力，动态地获取几千个 CPU，甚至上万个 CPU 的算力都不是难事。除了云计算，GPU 计算能力的进步对深度学习也有很大的推动作用，它能够加速深度学习中的计算速度，在某些情况下甚至能实现成百上千倍的提高。例如，现在深度学习的算法涉及大量可以并行化的矩阵运算，而 GPU 的工作方式就是多核并行计算，这个特点特别适合 AI 领域中的计算。此外，一些面向 AI 的专用硬件架构也开始出现，比如用 FPGA 去做专用的 AI 加速芯片和基础的加速设施，微软的数据中心就大量运用了 FPGA。

在算力相关的技术领域中，英伟达是较早意识到 AI 潜力的公司。早在 2012 年，英伟达就开始在 AI 上投入大量资源，陆续发布了针对服务器端的 DGX-1 以大幅加快训练速度、基于 Pascal 架构的 Tesla P100 以支持数据中心加速、针对嵌入式产品的 Jetson TX1，获得了市场的高度关注。从总体来看，英伟达的产品线体系已经比较完备，并且英伟达并不只是提供一个空的计算架构，还给软件开发人员提供了很多开发的库，如为实现汽车的自动驾驶提供从模组到软件的一整套解决方案。英伟达的 AI 战略是很清晰的，有很多实质性的产品落地。比如，已经有多家公司将英伟达技术嵌入摄像头中进行视频数据智能分析。

此外，英特尔也不容忽视，虽然它意识到 AI 爆发稍微滞后一点，但是它的补救动作是很及时的。在 2016 年，英特尔收购了包括 Itseez、Nervana System、Movidius 在内的众多 AI 技术公司，加上 Altera 和 Saffron Technology 等，英特尔在计算机视觉技术、数据中心 AI 计算加速、FPGA 芯片等方面快速完成了布局。尽管它的发力晚了一点，但展示出了坚定的决心和强大的实力。笔者认为，推出面向 AI 的 CPU、GPU、FPGA 的融合计算架构，一定是它未来的发展方向。

从总体来看，两家各有优势，如果从收入规模来看，英伟达比英特尔小得多，正是因为它的体量小，所以在 AI 领域的收入比例显得大很多，这是它引起媒体持续关

注的一大原因。

除这两家之外，高通也是一个很有潜力的参与者。高通的强项是移动芯片，而移动芯片的低功耗要求使得运算性能很难做得特别高。因此，数据训练显然不是高通的强项，因为数据训练是性能驱动的。但是高通专注设备端所需的推理计算，在移动端功耗很低、算力不是那么强的场景，高通能够满足场景对 AI 能力的需求。高通一直试图打入一些终端市场，例如车用芯片市场，以实现业务多元化，获得这样的能力有两种途径：一种是加强通用计算的能力，比如 CPU 计算单元应该做一什么样的改进才能适合 AI 的工作；另一种是针对特定领域的问题，比如针对视觉、语音等领域的问题，开发专用硬件并提供给设备端。现在很多公司的发展方向都是把专用领域里的一些能力放到端上去。

一些 AI 应用领域（如 AR、VR、MR）对智能能力有很强需求，如对外界环境的智能感知能力、基本的几何识别能力、语意理解能力的需求。而 AR、VR、MR 所需要的 AI 很多是要在端上实现的，不能完全依赖于云。所以我们可以看到，未来会有很多类似高通的厂商进入设备端智能能力的提供领域。AI 在设备端的需求量是相当大的，预计到 2021 年，我们就能够看到大量设备端 AI 产品出来。

当然，虽然我们谈了很多硬件基础设施对 AI 发展的支撑，但是影响 AI 算力需求最重要的还是软件和算法，因为算法决定了计算的复杂度，也就决定了对算力资源的需求。在另一方面，无论是 CPU、GPU 还是 FPGA 的计算架构，规模庞大的算力将主要通过云服务的方式向广泛的公众用户提供。谷歌、脸书、亚马逊、微软等行业龙头，都在大力开发云服务，积累了大量的算力资源，期望能够成为智能时代的计算平台“霸主”。

在数据方面，近年来移动互联网的爆发使得我们积累了大量的数据，同时物联网也极大地扩展了 AI 获取数据的数量和类型。事实上，相较于算法和算力，数据的获取会更难一点，因为它是建立在已有业务基础上的。以往我们都是先通过非 AI 的方式积累大量的数据，而现在初创企业要去获得数据需要一些“巧劲”。每个时代都要解决不同的问题，今天 AI 公司也一样需要去解决问题，企业要竭力避免打着 AI 的旗号，做一些不接地气的事，要为用户解决实质性问题。

谈到数据，笔者认为这是国外厂商要实现 AI 技术在中国落地需要解决的一个重要问题。在数据管控方面，我们国家对数据主权的保护越来越重视，这对国内的企业

来讲是一个利好。但是，长期来讲，数据的流动是难以避免的。对此，国内企业要及时抓住时间窗口，快速建立起竞争壁垒。

笔者认为，在 2018 年，以上三方面都已经做好启动准备，AI 时代已经真正地到来了。

AI 的发展路径一定是以创造价值为导向，而不是以技术先进为导向的。对不同的产品来说，在某些情况下 AI 是优化产品的支撑手段，而在另外一些情况下则是创造差异化竞争优势的初始核心能力。在创业时一定要明白自己产品的属性到底是什么，以及 AI 在产品不同发展阶段的不同作用。

第 2 章 从零开始成为 AI 产品经理

2.1 AI 带来的改变与挑战

2.1.1 AI 带来的改变

随着 AI 技术的不断突破 ,AI 产品生产成本的不断下降 ,AI 领域的前沿阵地 ,如语音识别、人脸识别、自动翻译等技术 ,开始从实验室走进普通民众的日常生活。在普通人看得见的领域 ,扫地机器人、无人机已经量产销售。2017 年 ,Lady Gaga 在美国超级碗中场秀中表演时 ,全世界都看到了数百架无人机精彩震撼的助演 ; iPhone 手机的语音识别和双摄像头双焦距运用了 AI 的语音识别和机器学习技术 ; 滴滴出行和高德地图背后都是 AI 根据大数据为用户选择司机、线路 ; 申通快递公司已经大规模使用快递机器人分拣快件。位于日本长崎豪斯登堡的世界第一家机器人酒店——奇怪酒店 ,几乎看不到人类员工的身影 ,入住咨询、行李抬送、打扫卫生 ,甚至餐厅备餐的工作都由各类机器人承担。

除了日常生活 ,在工业生产、城市管理领域也都能看到 AI 技术的使用。富士康公司的生产线上已经有 4 万台机器人在工作 ; 杭州已经利用物联网、大数据等技术来实现城市管理 ; 谷歌公司整合 AI、云计算、大数据 ,向业界提供决策参考。

在普通人看不见的领域 ,AI 同样有着润物细无声的影响力。未来 ,从电视、冰箱、洗衣机到一个扫把 ,都可以通过智能交互技术实现连接 ,我们每个人的生活方式都可能随之而改变。



扫地机器人



机器人在分拣快递



日本的机器人酒店



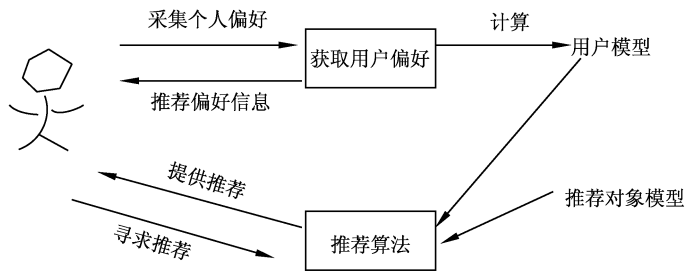
富士康生产线上的机器人

除前文提到的应用外，AI 技术对生活、工作的影响还体现在如下几个方面。

(1) 推荐系统。个性化推荐是指根据用户的兴趣特点和购买行为,向用户推荐可能会感兴趣的信息和商品,帮助用户决定应该购买什么产品,模拟销售人员帮助客户完成购买过程。

随着电子商务规模的不断扩大,商品品种数和种类快速增长,顾客需要花费大量时间才能找到自己想买的商品。这种浏览大量无关信息和产品的过程无疑会使淹没在信息过载问题中的消费者不断流失。为了解决这些问题,个性化推荐系统应运而生。个性化推荐系统是建立在海量数据挖掘基础上的一种高级商务智能平台,用来帮助电子商务网站为其顾客提供完全个性化的决策支持和信息服务。

推荐系统有 3 个重要的模块:用户建模模块、推荐对象建模模块、推荐算法模块。通用的推荐系统模型如下图所示。



推荐系统通用模型

通用的推荐系统模型

推荐系统对用户模型中的兴趣需求信息和推荐对象模型中的特征信息进行匹配,同时使用相应的推荐算法进行计算和筛选,找到用户可能感兴趣的推荐对象,然后推荐给用户。

(2) 新硬件。包含智能穿戴类、智能家居类、智能出行类等硬件,例如语音识别技术赋能的天猫精灵、小米 AI 音箱等。

(3) 新零售商店的智能导购。例如,美的智能导购结合了 4 大 AI 技术:自然语言理解、人脸识别、大数据分析、智能推荐算法。其中,利用自然语言理解技术与消费者进行一对一的会话交流,深度了解用户的需求;通过人脸识别及 107 种表情分析技术,实时理解消费者对产品和对话的即时反馈;利用大数据分析技术,对用户购买历史、服务记录等进行分析,构建客户画像;智能推荐系统融合成百上千的 AI 模型和算法,进一步理解、判断消费者实际的购物需求,分析出他们的购

物习惯和产品偏好，以进行精准的产品、套系、折扣等内容的个性化推荐。消费者还可以在全数字化的环境中，了解产品的每一个细节，同时体验美的产品与不同装修风格搭配的效果。

美的集团全球 AI 负责人、硅谷未来技术中心和深圳 AI 研究所负责人王冬岩博士表示：“美的智能新零售不追求做无人商店，并非简单的‘线上+线下’，而是以消费者需求为中心，以 AI 为引擎，为美的消费者提供全新、便捷、高性价比的购物体验。对消费者而言，智能导购以用户的需求、利益为出发点，方便用户快速找到自己喜欢的产品，提供个性化的贴心服务，实时接收用户反馈，精准推荐，提供高标准、高质量的个性化导购服务；智能导购能让美的更懂客户，让客户更懂美的产品，实现人机和谐共存。”

美的已经建成了异构深度学习 GPU 集群平台、中国成品菜数据集，并开发出了多款 AI 产品及应用，包括 AI 烤箱、AI 米饭机器人、未来厨房等。

(4) 智能安防系统。机场、车站的安检系统借助人脸识别、图像识别等 AI 技术，将会变得更加精准、快捷，今后乘客可能完全感觉不到安检流程。

(5) 城市管理系统。谷歌设计人员认为，人们输入的搜索关键词代表了他们的即时需要，反映了用户的情况。为便于建立关联，设计人员编入“一揽子”流感关键词，包括“温度计”“流感症状”“肌肉疼痛”“胸闷”等。只要用户输入这些关键词，系统就会展开智能跟踪分析，创建地区流感图表和流感地图。为验证“谷歌流感趋势预警系统”的正确性，谷歌多次把测试结果与美国疾病控制和预防中心的报告做比对，结果证实系统的准确性非常高。

2.1.2 AI 带来的挑战

AI 带来的挑战主要存在以下几个方面。

(1) 对语言的挑战。语言是认知的基础，语言是人类智能真正的起点，AI 发展的关键也是语言的运用。从逻辑推理、知识学习、语言理解到认知智能，当 AI 掌握了语言，就能用机器语言描述整个社会、描述周边的环境，并且能够让信息得到提升、交换。比如，当一个机器人走进一个环境的时候，它可以看到各种生物，知道大家在做什么，并用语言表达出来，这个时候我们要学会用人类语言理解机器语言。

(2) 对思维方式的挑战。对事物进行判断时，人类的思维方式除了依赖常识，有

很多时候还夹杂着个人喜好和情感；而 AI 则是依靠人类的历史评判数据做出正确或错误的判断，尽量保持客观和公正。

(3) 对心理的挑战。人类会越发恐惧自己被机器替代，有朝一日会成为机器的“奴隶”。

(4) 对失业的挑战。AI 增强了社会各方面的自动化，提升了工作效率，将会导致结构性失业。

(5) 对安全的挑战。2015 年 7 月 2 日，据英国《金融时报》网络版报道，德国大众公司位于卡塞尔附近的一家工厂发生一起悲剧，一名技术人员因突遭机器人攻击而不幸丧生。

(6) 责任和义务。对 AI 时代，存在截然相反的两种图景：一种是乐观的，比如 AI 可以把人类从繁重的体力或脑力劳动中解放出来，变成无须工作的自由发展的人；另一种则是悲观的，被剥夺了工作机会的人类将变得一文不值，成为真正的“一无所有”者。究竟是该悲观还是该乐观，关键在于 AI 能否带来真正的产品价值革命，让大多数人或者代表大多数人的组织形式成为操纵、使用 AI 的主体，而不是让 AI 成为少数大企业垄断的产业。在这个意义上，笔者认为机会大于危机，AI 将会创造一个更加美好的社会。

2.2 AI 产品经理与传统产品经理的区别

2.2.1 面向的用户不同

AI 产品经理与传统产品经理面向的用户不同，传统互联网产品集中于 ToC 端，AI 产品现阶段则重点面向 ToB 端。例如，旷视科技、科大讯飞等企业的产品面向的主要群体是 B 端客户，B 端客户通过集成其应用，然后赋能行业。现阶段，AI 产品经理与传统产品经理的最大区别是能否把 AI 应用穿透到 C 端用户的具体应用场景中，并提供一个具体的产品和服务。

当下 AI 面向 B 端是理性的选择，原因如下。

第一，数据。AI 公司往往具有自己的算法，但是没有数据，选择 ToB，跟拥有数据积累的企业合作是明智的选择。

第二，获客成本。AI 时代的创业不再是之前“烧钱”获取用户的时代了。

第三，与其从零启动，不如站在巨人的肩膀上。例如，旷世科技在创业之初选择了蚂蚁金服的战略投资，获得蚂蚁金服数据和算力的支持，这比自己从零开始建设这样一套系统来得更快。

下面以智能投资顾问为例进一步来看 AI 在当前阶段适合应用于 ToB 的缘由。

对于大多数行业来说，谁先占得了市场先机，谁就容易在潮流中脱颖而出。但这样的论断有个前提，那就是“市场”是业已存在的，只是谁能“占山为王”的问题。然而，就像早先“火了一把就销声匿迹”的 VR 一样，作为核心技术元素的 AI，稍不留神也容易重蹈覆辙。

截至 2018 年，大多数 C 端的智能投资顾问产品的界面都过于简陋，对用户进行风险评测后即要求绑卡、开户，体验上过于直接、突兀，缺乏应有的引导过程，这显然是在注重技术的同时，产品体验的逻辑没有跟上。

几乎所有面向 C 端的产品都需要从零开始集聚人气，难以逃脱前沿科技产品“冷启动”的宿命。但智能投资顾问难以走通这条道路，此时，面向 B 端的优势就显现出来了。ToB 既能够在内核算法技术上获得充分的依靠，也能够凭借充足的客服经验在产品体验上满足用户的需求。ToB 给了智能投资顾问一个更好的选择，机构原有的用户基础甚至线下的实体门店都给了互联网化的智能投资顾问一个“从巨人肩膀上开始”的机会。最关键的是，除了数量上的“热启动”，B 端自身的基础、积累还能有效治疗 AI 这种前沿科技与传统理念不匹配带来的智能投资顾问顽疾——获客成本高。在用户规模足够大的基础上，一方面投入的获客边际成本会越来越低，另一方面，新用户的进入意愿也相对会更高。

AI 作为 AI 产品甚至母公司的核心，技术内容无疑最为关键。但过分注重技术往往会成为创业公司的弊端，一门心思钻研技术而被市场淘汰的案例比比皆是。在商业化浪潮下，技术极客需要抬起头看世界，思考一下围绕 AI 技术的发展策略同样重要。毕竟，越是前沿的科技，对市场的放大效应越强，在发展上越应当小心翼翼，任何失误都可能是致命的，错失任何有益尝试带来的机会损失都可能是巨大的，推及智能投资顾问，就是同 B 端合作时应当注重谋略、避开禁忌。

小结：纵观金融市场的演化过程，技术一直都扮演着重要的驱动角色，即便看跌的再多，智能投资顾问的价值也无法被否定。在未来市场更趋成熟的时候，AI 将会与

传统金融诸多功能紧密结合，为理性投资者们创造新的效益与价值。而在这之前，基于 AI 发展特性与市场现实，智能投资顾问最佳的发展方式可能是先面向 B 端，谋定而后动，最终征服 B 端和 C 端。

2.2.2 工作内容不同

传统产品的需求大多来自老板，AI 时代的产品需求诞生于数据，来自产品经理对数据的感悟。

我们都知道，AI 是一门前沿综合性的学科，最终形成的模型（相当于现在的程序）可以理解为中间件，是不能直接应用到用户交互操作中的。产品经理应当想办法利用这个“中间件”的性能优势，为用户提供体验更好的产品。那么，产品经理的核心任务会不会就是将 AI 技术应用在有商业价值的需求场景中？

下面笔者从行业+AI 的角度提出对 AI 产品经理工作内容的几点思考。

1. AI 产品经理需要了解 AI 技术的成熟度

笔者认为，AI 技术处于快速发展时期，但是都不算成熟，了解技术的能力范围、技术当前存在的瓶颈、技术本身所需要的成本等，可能是 AI 产品经理需要具备的基础能力。按照以往的经验，如果在最初使用技术时并不了解技术的相关特性，那么当技术产生问题的时候一定会使你浪费掉更多的精力。本书先介绍两种了解 AI 技术现状的渠道。通过公司内部的 AI 技术团队。这种渠道主要针对大公司，当公司自有 AI 技术团队时，产品经理可以很方便地了解现有技术的情况。多与公司的科学家、工程师们沟通，AI 产品经理会得到飞速的进步。通过第三方提供的 AI 平台或 AI 框架了解技术现状。这种渠道主要针对一些中小型公司。中小型公司可能没有 AI 技术团队，这时充分利用第三方平台提供的 AI 技术是一个理想的选择。

2. AI 产品经理需要挖掘并抓住 AI 业务场景的特点

技术处于发展之中，还不够成熟，这就要求 AI 产品经理需要在深度理解技术优/劣势的基础上，寻找可以提供更好用户体验的业务场景。也就是思考，如何结合现有 AI 技术与业务场景，为用户提供更好、更便捷的产品。这时，AI 产品经理的主要工作是：完成当前技术与现有市场需求的映射。这一点笔者要特别说明的是，根据以往的经验，在技术能力不足时，很多场景的业务需求都不需要我们 100%地解决，只需要利用现有技术为用户解决最为关键的问题即可，即抓住主要场景的痛点来解决问题。

只不过，我们需要评估这样是否真的能够提升用户体验，而不是带来了新的麻烦。

另外，如果 AI 产品经理真的发现了具有良好用户体验的智能化场景，还需要充分评估产品定位、用户价值、成本、商业价值等因素是否相匹配，毕竟当前 AI 技术的实现成本相对还是很高的。不是你有想法，技术就能实现；也不是技术能实现，公司就能付得起成本。即使条件都能满足，也存在一个巨大的风险，那就是做出一个没有市场价值的产品。

最后，我们来思考一下什么才是好的 AI 业务场景。笔者了解的几个要点如下。

有大量数据的业务场景。

充分体现用户个性化需求的业务场景。

在进行交互时用户能够反馈结果的业务场景。

强调为用户服务的业务场景。

3. AI 产品经理需要推进产品研发

在项目研发过程中，为了最大限度地发挥机器学习对业务的价值，尽量避免研发过程的差错导致的成本浪费，需要产品经理、研发工程师与数据科学家之间持续保持密切协作，并且，产品经理有责任确保工作的成果对产品是有正面价值的。

4. AI 产品经理需要挖掘产品的市场价值

笔者认为任何产品经理的核心职责都是让产品具备更好的市场价值，AI 产品经理也不应该例外。AI 产品经理应当理解产品，理解用户群体，理解业务场景，这些都是 AI 产品经理应当具备的基本素质。并且在市场逐步由技术竞争转变为产品竞争的时候，AI 产品经理让现有技术所定义的产品价值能够转化为商业价值是其核心的工作。

5. 小结

AI 产品经理比互联网产品经理更关注数据：机器学习模型是使用数据“喂养”出来的成果，AI 产品需要使用数据指导交互设计，即 AI 产品具有更强的数据需求。

2.2.3 技术方案的维度不同

传统产品经理一般只考虑前端、后台，而因为 AI 是一门前沿综合性学科，所以 AI 产品经理需要综合考虑前端、后台、软件、硬件、系统驱动等多个方面。

常见的 AI 产品经理有三种：技术要求高。核心 AI 公司往往在技术层面对产品经理的要求最高，例如商汤科技、科大讯飞等应用型 AI 公司的产品经理。

对“软硬一体”理解较深。例如小米 AI 音箱、小蛋机器人等垂直应用型产品的产品经理。对产品、业务、行业有深入的了解。例如，意能通的客服机器人的产品经理。

那么，产品经理应该如何找到自己当前的定位？“懂技术”该怎样衡量？

无论你身处三个阵营中的哪一个，你的 AI 技术知识都需要帮助你回答下面几个问题。

(1) AI 技术会给你的产品带来多大的价值？产品永远是需求驱动的，而非技术驱动的，再前沿的技术，从理论到产品落地也是需要巨大的投入的。

(2) 从技术角度来看，将 AI 技术应用到你的产品中需要哪些资源或准备，是更多的数据，还是更完善的算法模型？尽管很难量化这样的需求，你还是要尽可能掌握更多的信息。

(3) 从技术角度来识别 AI 领域中的哪些理论已经有了最佳实践，即需要判断技术的成熟度。

当你将 AI 技术应用到产品中时，应该可以对以下问题做出回答。

(1) AI 带来的价值是否真的被客户认可？这样的技术真的比传统技术更好吗？需要多长时间或多少样例数据来验证你的 AI 产品已经站住脚了？

(2) 一旦产品上线后的效果没有预期好，你是否有备用计划？

(3) 任何一个机器学习功能的上线都需要占用研发工作 80% 或更多的时间来完成数据的准备，你是否已经和研发部门充分沟通并达成一致？

2.3 打造 AI 产品的流程

2.3.1 打造传统软件的开发流程

传统软件产品或者传统系统的开发流程如下。

(1) 问题的定义及规划。在此阶段，软件开发者与需求方共同讨论，主要确定软件的开发目标及可行性。文档包括可行性研究报告和项目开发计划。

(2) 需求分析。在确定软件开发可行性的情况下，对软件需要实现的各个功能进行详细的需求分析。文档包括软件需求的说明书、数据要求的说明书。

(3) 软件设计。在此阶段要根据需求分析的结果，对整个软件系统进行设计，如系统框架设计、数据库设计等。软件设计一般分为总体设计和详细设计。文档包括总体设计说明书、详细设计说明书、数据库设计说明书。

(4) 程序编码。在此阶段，将软件设计的结果转化为计算机可运行的程序代码。文档包括用户手册、操作手册、模块开发卷宗。

(5) 软件测试。在软件设计完成之后要进行严密的测试，以发现软件在整个设计过程中存在的问题并加以纠正。整个测试阶段分为单元测试、组装测试、系统测试三个阶段。文档包括测试计划、测试分析报告、项目开发总结报告。

下面，笔者先讲传统产品经理所面对的开发流程。

(1) 需求整理。

需求整理的第一步是确定产品功能。在拿到一个新的产品需求进行产品方案设计以前，产品经理应该对需求进行一次整理，看哪些是真需求，哪些是边缘需求，哪些是伪需求，进而把功能点、产品卖点整理清晰，即先确定产品功能，在有条件的情况下做一份市场需求文档（Market Requirement Document，后文简称 MRD）。

需求整理的第二步是编写产品需求文档（Product Requirement Document，后文简称 PRD），具体内容由产品经理产出。PRD 主要给软件开发与测试人员、其他相关部门人员查看。

需求整理的第三步是编写设计文档，主要用于设计人员在画图之前对整个产品进行把握。设计文档由产品经理主写，交互人员辅助。因为产品经理在写文档的时候是以面向程序员的角度写的，所以设计人员要把交互原型所需要的东西提交给产品经理，辅助完成设计文档。PC 端和移动端的设计文档要分开编写。

那么，设计人员需要做哪些工作呢？设计人员要把产品原型图需要的内容提交给产品经理，并辅助产品经理完成设计文档的编写。根据设计人员提供的要求，产品经理需要提供以下内容：功能类内容，如功能流程图、产品的功能点、主要功能卖点；

模块类内容，如文案、各功能描述、弹出框内容，各模块之间的关系；逻辑类内容，各模块的业务逻辑、结构框架图。

当存在一些比较复杂的流程的时候，为了避免开发人员或设计人员不能正确理解，需要把流程图画出来。但是一个完整的项目只用一个流程图是不可能表达清楚的，所以需要按照模块的流程逐个进行分类，分类以后按照功能的不同、简易程度，分别画出流程图。

（2）交互原型。

在设计文档评审通过后，就可以开始安排页面上每个功能的具体位置、交互方式和跳转方式了，一般由产品经理输出全部交互原型图、原型图解释文案，如没有产品经理，可由设计部门完成。交互原型一样需要评估，参与人员由产品经理指定，如没有产品经理可由设计部门指定。

交互原型图的具体要求：不要加入视觉设计的元素。着眼于大局，不要纠结于细枝末节。不限工具。产品经理最好用 Axure 完成动态设计，如没有，需设计人员用 AI 或 PS 完成。页面流程中所考虑的交互过程需要体现在原型上面，出错、提示等交互细节也应体现。原型图解释文案在画交互图的开始就要记录下来。每个页面之间的联系、承接关系，页面当中的跳转细节、按钮的不同状态、弹框的提示文字、日历的跨年显示效果、小数点是否保留（若保留，保留到小数点后几位）等内容是解释文案的重点。

（3）视觉设计。

在交互原型评审通过后，设计负责人把控 UI 整体风格和所有视觉效果的设计，主要输出 VI（Visual Identity，通常译为视觉识别系统，后文简称 VI）备选方案、所有视觉效果图、资源包等。

之后，选定一个时间节点，组织人员评估，参与人员为 UI 设计人员。在交互原型图评审通过后，按照项目视觉设计的复杂程度再来安排具体时间节点，然后预估工作量，按照项目复杂程度和具体的完成时间来评估工作量。

一般公司的视觉识别系统在一个项目的最开始就要开始做。视觉识别系统一般由 UI 设计人员做 3 套以上的 UI 设计方案，包括颜色、Logo、整体风格、图标等的效果图，然后开会评审，选定一套方案。确认视觉识别方案后方可进行接下来的整体界面设计。提前做好视觉识别方案，一方面能避免因为工作交接而在视觉上产生偏差；另一方面能为项目后期的一些工作做好准备，比如确定产品手册、营销推广页、名片的

风格与颜色等。

颜色的设定要有理有据，开会的时候会有很多人问你为什么要用这个颜色，这个颜色代表什么，这些最好在做设计的时候就想到。

一般情况下公司不需要在这时做 Logo，但是如果有需要的话就要按照 Logo 的制作流程进行设计。主要内容包括 Logo 的设计思路、Logo 的颜色、Logo 的应用场景等。

俗话说细节决定成败，在设计图标时要充分考虑细节。每一个图标的风格要跟页面整体的风格保持一致，这样才能保证所有页面设计完成以后风格是统一的。

除了上述内容，我们还要进行主要页面的视觉评审。主要页面的视觉评审对于开展后面的工作十分重要，设计师需要在评审的时候讲述自己的设计思路，包括颜色的设计理由、风格的设计理由等，要让参与人员了解你做设计时思考的内容，而不单单是作图。

页面整体的设计，具体要求如下。

在视觉风格已经确定的情况下，按照交互原型图进行逐个页面的设计。

在设计中对于细节的把握要准确。

在视觉设计中要考虑不同平台的特性，根据不同平台的特点进行设计，切勿混淆各平台的控件。

对于某些需要适配的移动端页面或是 H5，要考虑尺寸的问题。

在设计中，一些容易忽视的页面也要做出来，比如断网、连不上 Wi-Fi、404 错误、正在加载等页面的展示效果。

在视觉设计评审通过后，接下来会进行一个比较重要的环节——与前端人员合作实现效果。这个时候设计人员通常需要输出 5 个文件资源包，分别是：

视觉效果图的尺寸标注资源包。

所有尺寸的切图资源包。

解释文案资源包。

效果图源文件资源包。

全部视觉效果图资源包。

在完成一款产品后，设计人员需要做一份设计总结，就是针对现在这个版本编写

视觉设计规范。这是对设计工作的一个总结，对下一版本的开发也很有帮助，可以大大提高开发的效率。

（4）整理可以优化的体验点。

在项目第一版已经成功上线，第二版还没开始开发前，产品经理需要整理可以优化的体验点，要对现有的这个版本的体验点进行优化。这部分工作需要开发人员和用户研究人员一起参与。

整理优化体验点的具体工作内容有：

对现有产品的一些功能需要优化的点进行记录，记录有哪些不足之处。

基于用户的反馈，判断哪些功能是可以省略的，哪些功能是需要加强的。

有哪些更好的想法可以在下一个版本中实现。

以上是传统产品经理打造传统产品的流程，介绍上述内容的目的的一方面是帮助大家回顾一下传统产品的打造流程，另外一方面是想跟后面的 AI 产品打造流程做一个比较。

2.3.2 打造硬件产品的流程

大家可能觉得做硬件产品很简单，但事实并不是这样。实际上，一个硬件产品项目最基本的流程就包括：前期的市场调研，产品定义，需求分析，方案设计，外观结构的设计，软、硬件的设计，物料的采购，试生产，测试和整改，生产管控，质量控制，量产出货，售后跟踪等。这是一个很繁杂的链条，这里只讲了十几个大的方向，每个点下面还有很多细节。

有人说：“做硬件产品很简单，设计一个电路板就解决了。”真的是设计一个电路板那么简单吗？如果真有那么简单，硬件工程师就不可能拿那么高的薪水了。设计一个电路板需要考虑的东西很多，如电源的设计和走线是强相关的，整个电路板需要进行抗干扰设计，高速信号需要对仿真、射频、音频之类的弱信号进行保护，需要考虑电路板的防静电能力，等等。这些内容不仅跟工程师的水平有关系，也体现了一个公司的设计水平。

还有人说：“做硬件产品很容易，做一套外壳就解决了。”我们先不谈做外壳需要开模，开模需要花钱。单从设计上来讲，结构设计要满足外观效果的要求，要能降低

模具的成本，要能够实现较高的生产效率，同时结构强度要高，长期使用的话，质量可靠性也需要高，这些都是跟结构设计息息相关的。你找一个初级工程师，他肯定能给你画出来，但他画出来的会是什么质量的？为什么资深工程师比初级工程师的待遇高很多？因为他们的经验比初级工程师丰富很多。这就是为什么有些公司设计费用很低，而有些公司设计费用很高的原因。如果说你做一个外壳，什么都不考虑，只要能把它拼起来就行了，任何可靠性和可生产性都不考虑的话，产品肯定也是能做出来的，但你就等着产品的可靠性出问题吧。

这里还有个重要流程，也是产品经理初次接触硬件产品时一般不会注意到的流程：测试。测试包含硬件电路的测试、软件的测试、结构的测试、整机的可靠性测试，还有大量的压力测试。举个例子，针对手持设备有一个测试叫“微跌落”，即拿着一个产品，从 10cm 左右的高度摔到一块钢板上去。需要摔多少次？国内一线品牌的“微跌落”测试的次数一般在两到三万次！小一些的品牌，至少也要做几千到一万次。这样才能保证产品在一到两年的产品生命周期里，不会因为经常的振动和晃动而损坏。这些测试跟产品的设计有很大的关系，尤其跟结构设计有非常大的关系。如果不知道这些话，做出来的产品在刚出厂的时候是好的，但使用一个月之后，大概率会出现故障。

在介绍硬件产品的开发流程之前，我们先来了解一些硬件的基础知识。

常见的硬件系统主要有小型的单片机硬件系统和大型的 Android 硬件系统。

单片机硬件系统。单片机在小微型智能硬件领域的应用很广泛，其成品价格低、开发周期短，适合运算量小、通信数据量小的应用场景。单片机在智能小家电领域的应用有智能电饭煲、智能花盆、空气净化器、智能台灯等，在智能工业领域的应用有环境温度监测、空气质量监测、水质监控、农业喷灌控制等。随着 BLE、ZigBee、GPRS、NB-IoT 等众多无线传输技术的普及，单片机+云服务架构的产品越来越多。设备的智能化、互联化都离不开智能单片机。丰富的单片机系统开发经验，能让你快速实现从非智能到智能、从单体到组网的产品升级。单片机只能实现简单的数据处理，如果需要处理复杂的数据，例如处理视频、识别语音等，就需要 Android 硬件系统了。

Android 硬件系统。主流的 Android 硬件系统使用 4 核~8 核 ARM Cortex A7 或更强的处理器,集成 GPU,很多还集成 LTE 通信模块,其运算能力强、通信数据量大、软件扩展性好、UI 界面漂亮、人机交互便捷。Android 智能硬件系统已在逐渐取代传统嵌入式 Linux 类硬件和嵌入式 Windows 类硬件。智能车载系统、智能手表、智能家居网关、智能电视、智能工控主机、智能导购屏这些产品,几乎都采用了 Android 硬件系统。

传统的 PC 系统因结构单一、硬件尺寸大,在智能硬件领域的应用不多。嵌入式 Linux 类硬件因开发资源和第三方资源远不如 Android 多,硬件成本也比 Android 智能硬件系统高,因此逐渐被 Android 智能硬件系统取代。

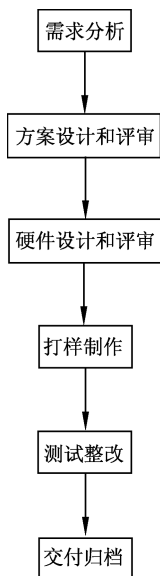
智能硬件的开发流程如右图所示。

(1) 需求分析。

在所有的流程当中,需求分析最为关键。很多创业型产品就是“倒”在不断地修改功能需求上。软件迭代相对快一些,但硬件迭代一次少则一个月,多则两三个月;软件迭代几乎不影响整机,但硬件迭代很有可能导致整机结构发生变化,如果不停地修改产品的功能需求,产品上市就遥遥无期了。需求分析准不准直接关系到产品的上市时间、成本、质量。笔者建议多向客户问问题,以准确分析需求,少走弯路,选择合适的技术路线。

(2) 方案设计和评审。

智能硬件产品往往涉及一些新技术或非常规技术,项目风险会比较大。需求分析结束后直接开始做硬件设计的话,很容易遇到特殊物料买不到、芯片选型不满足指标、电路设计有缺陷等问题,导致项目延期、客户流失。因此,在正式设计之前要做好大量准备工作,包括“关键器件选型”“关键技术验证”“系统框架设计”“产品风险评估”“功能交互设计”“拟定产品测试大纲”等一系列步骤。以上内容需要 CTO 组织人员对每个项目的方案设计做详细的评审,评审通过后才能正式开始设计工作。评审能够极大地提升产品的质量,降低研发风险。



智能硬件的开发流程

（3）硬件设计和评审。

硬件设计主要包括原理图设计和 PCB（Printed Circuit Board，印刷电路板）图设计，看似很简单。很多小公司安排一个工程师就把画原理图、画 PCB 图、写代码、调软件全包了。但你相信一个人能对各个环节都精通吗？硬件设计不是把线路连通就算完成了，还需要考虑功耗、散热、抗辐射、防静电、高速信号走线设计、射频性能等一大堆问题。如果设计不合理，一般功能不会出现太大的问题，但是性能就完全没办法保证了，产品肯定是通不过各项测试的。

硬件设计完成后，需要进行内部评审，包括原理图评审、PCB 图评审、结构评审等。每个评审表格都有几百个项目，通过评审可以检查设计错误，并将常见错误消灭在设计阶段。在设计阶段修改一次只需要一两天的时间，如果已经把 PCB 做出来了，再来修改则至少需要半个月的时间，而且还会带来极大的物料浪费。

（4）打样制作。

硬件设计完成后，硬件工程师就能稍微松一口气了。PCB 的生产是需要一定的周期的，4 层板一般需要一周多，8~10 层板需要两到三周。在板厂制板的这段时间里，采购、资源调度和生产管理部门需要去做元器件备料和 SMT（Surface Mount Technology，表面贴装技术）产线预约。对于一些超长周期的物料，早在设计阶段，甚至在方案评审阶段，就需要下单采购。等所有元器件都到齐了，硬件工程师也已经把生产资料准备好了，就可以上 SMT 线贴片生产了。通常第一次上线生产都会暴露出一些问题，有物料问题、生产制程问题，也有硬件问题、设计问题，工程师需要记录这些问题，在后续设计整改时及时改进。

（5）测试整改。

PCBA（Printed Circuit Board + Assembly，即 PCB 空板经 SMT 上件后，再 DIP 插件）完成后，硬件工程师、软件工程师、测试工程师就开始紧张地调试和测试了。看一款产品设计得好不好，只需要看测试报告就足够了。优秀的测试工程师会结合产品的使用场景，设计出很全面的测试用例，这些用例能够覆盖各种常见和不常见的场景，不断地“折磨”产品，直到它出问题为止。一款经过千锤百炼的产品，品质才有把握。像华为的手机，光试产测试样机都要做上千台，不管是大的问题还是小的问题，都被消灭在研发阶段，以保障质量、口碑。

经过一轮测试后，项目经理会组织项目组成员汇总测试问题，提出并验证解决方法，然后整改到下一个硬件版本中去。如此反复，才能打造出一款优秀的硬件产品。

(6) 交付归档。

项目完结后，所有的资料都会归档保存，除基本的设计资料外，还有评审资料、问题记录资料、测试用例资料等，以供后续查阅。很多小公司不重视资料归档和资料保存，一旦某个项目暂停几个月后再重启，很容易出现资料不全、不知道哪个版本才是正确的、老问题又重新出现等各种乱象。

一定要记得，创业型公司切不可贪大求全。任何一款产品，我们想的都是很美好的，从软件、硬件到平台、生态链，我们都可以规划出来，但如果前期就这么做的话，比如，做一个手环就想把生态链都打通，那投入的资金肯定不止几十万元了，至少要到百万元的级别。

所以，前期尽量把不太需要的功能都砍掉，只保留核心功能。省钱、快速地出产品、出原型机才是王道。现在的风投都很精明，已经不会在你只有一个 PPT 的情况下就投很多钱了。必须拿出来一个实物，来证明你能做出来，风投才敢继续往下投资。

这部分是笔者用血泪总结出来的经验，是经过上亿元投资打磨出来的教训。如果你准备创业做硬件，笔者给你一个可能节省千万元的方法：做原型机。原型机可以不考虑可靠性问题（也可以不考虑结构、外观等），其唯一的作用是让你验证核心业务和核心功能。等你把核心业务、核心功能跑通了之后，就可以用原型机去招商、拉风投。如果招商和拉风投都比较顺利，那么钱的问题自然就解决了。大家都看好你的产品，也相信你能做出来，自然会有很多人想来投资，不管是众筹还是风投。这会支撑你后面拿到更多的资金，做更好的硬件产品。

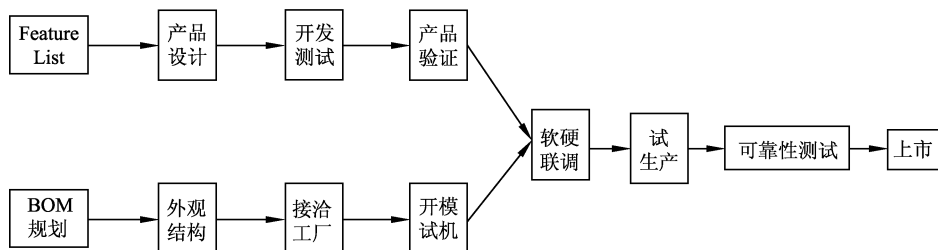
2.3.3 打造 AI 产品的流程

AI 产品有软件产品，如苹果的 Siri 智能语音助手；也有硬件产品，如智能手机、智能音箱等。

但是笔者认为，看 AI 产品不能仅仅从软和硬的单一维度来看（实际上一款 AI 产品一定既有“软”也有“硬”），应该从整体角度、用更高层次的眼光来看 AI 产品，一方面因为 AI 产品离不开数据，而数据的采集离不开传感器，如 NB-IoT、GPRS 等，

这些传感器属于硬件的范畴。另一方面,打造一款普通桌子肯定不是本书讨论的范畴,但如果桌子有操作系统和软件,那么它就是笔者定义的 AI 产品。

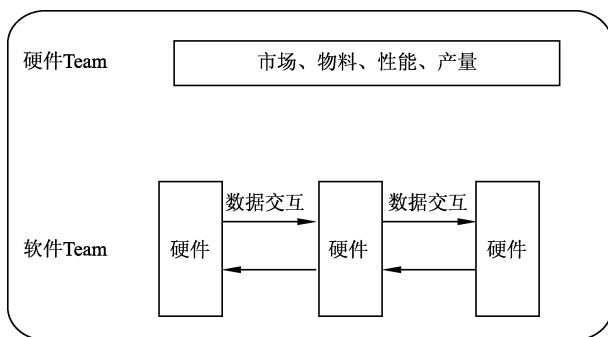
一般来讲,打造一个整体级别的 AI 产品是包含软件、硬件双流程的,如下图所示。



打造 AI 产品的软件、硬件双流程

2.3.4 打造 AI 产品的实际案例

这里以微鲸智能电视为例:产品每个版本的打磨,无论是从 0 到 0.1,还是从 1.0 到 2.0,一方面需要硬件团队对市场物料的性能和产量有稳妥、精准地了解,另一方面需要软件团队熟悉每两个硬件衔接的驱动是否匹配,是否能够相互交换数据,一方能否调用另外一方的数据,产品经理要把握两个团队的流程。



产品经理需要把握打造 AI 产品的双流程

2.3.5 打造 AI 产品的创新因素

融合和跨界是 AI 自带的 DNA,而创新大多数来自跨界和融合,所以 AI 给产品创新带来了好机会。又因为 AI 是个多学科的技术,因此,我们有必要了解一下 AI 产品中哪些因素对产品创新的影响比较大。

对于线性架构,影响创新的因素包括制造设备、制造工艺、制造物料,这些因素

都是很难发生变化的，有的制造工艺甚至从开始到现在都没怎么变化过，例如茅台酒的制造工艺就没怎么变化过，而物料和设备在一定的周期内也是不变的因素。

所以，我们只能将创新的希望寄托在非线性、非结构化的因素上，如制造环境、参与人员、管理水平、实施标准、制造目标、制造节奏等因素。

一方面，我们可以通过改善生产制造环境，如采用富士康代工智能手机的无尘车间；另一方面，我们也可以通过培养员工的生产技能来提升 AI 产品的创新效率，同时还可以改善管理方法，如激励竞争，引入符合创新环境的 KPI 管理机制来激发创新。

笔者建议：我们一定要抓住可以利用的打造 AI 产品的创新因素，要重点强调创新性人才的价值。

2.4 AI 产品经理应该具备的思维

2.4.1 多样化思维

多样化思维不是 AI 产品经理独有的，但是多样化思维与 AI 技术前沿、综合性学科的气质相吻合。多样化思维是与单一性思维相对的，是从多方面考察需求的综合性思维过程。多样化思维从思维的各个层次出发，对需求进行多角度、多方面、多因素、多变量的系统考察。多样化思维是客观世界普遍联系的客观现实在思维世界的反映，它从不同的角度想痛点，在同一个痛点面前，尽量提出多种设想、多种答案，以扩大选择的余地；能灵活地变换影响需求的质和量的某种因素，从而产生新的思路；思维在一个方向上受阻时，能马上转向另一个方向；能用心寻找最优答案，确定实现需求的最佳解决方案。

多样化思维具有单一性思维所不具有的特点，具体包括：（1）多角度性。多样化思维从众多方面考察需求，把需求的相互联系看作多样性的统一，因而要求思维的多角度性。（2）多层次性。多样化思维是一种协同性思维，具有多层次性。思维的不同层次，如经验思维和理论思维、形象思维和抽象思维等，是思维多样化的表现。只有将不同类型、不同层次的思维相互融合，才能更为深刻地认识需求及其发展规律。（3）互为中介性。在用多样化思维认识世界的过程中，各种类型、各种形式的思维活动互相联系、互为中介、相互渗透和转化，使任何事物之间都不存在绝对的界限，都可以

通过一定的条件彼此过渡。多样化思维与单一性思维是对立的统一，二者在发挥各自长处过程中相互补充、密切配合，使思维能力不断进步，帮助决策者不断提高决策能力和决策水平。

例如：在创业初期，马云为了使得团队具有多样化思维，招聘人才时特别注意从多个地域、多个学历、多个背景来招人，在称呼阿里巴巴时，戏称之为“动物园”，因为团队中每个人都不相同，每个人都自带特点。

2.4.2 差异化思维

差异化思维也被称为特色优势思维，是指产品在用户广泛重视的一些方面，努力做到行业内独树一帜。选择许多用户重视的一种或多种特质，并赋予其独特的地位，以满足顾客的要求。特质既可以是先发的、令人兴奋的需求，也可以是后发的、超出用户期待的需求。

产品差异化可以带来较高的收益，以应付来自供应链的压力，同时可以缓解来自消费者的压力。当用户缺乏选择余地时，其对价格不会特别敏感。采取差异化战略赢得用户和客户忠诚的公司，在面对替代品威胁时，其所处的地位比竞争对手更为有利。

有时，实现产品差异化与占领更大的市场份额相矛盾，差异化往往要求公司对这个需求的排他性有思想准备，即这个需求与提高市场份额两者不可兼顾。较为普遍的情况是：如果实现差异化的准备活动总是成本高昂，如需要广泛的研究、复杂的产品设计、高质量的材料、周密的用户服务等，那么实现产品差异化将意味着以提高成本为代价。

有时即便全产业的用户都了解公司的独特优点，也并不意味着所有用户都将愿意或有能力购买公司的高价格产品。

差异化思维的典型例子是苹果公司，其创造的“Think Different”充分体现了苹果公司的差异化思维。

2.4.3 穿透思维

穿透思维也存在于其他行业，但是对于泛 AI 领域来说，穿透思维是赋能发展的一种“刚需”，因为 AI 技术具有跨多学科、多领域的特点，且许多 AI 技术都处于猛烈增长的阶段。

AI 产品经理对穿透思维应该有两方面的认知。

1. 理性的认知

理性的认知又分横向的和纵向的。

横向的理性的认知是指依靠 AI 技术的公司在打磨好赖以起家的 AI 技术之后，横向贯通到供应链领域。例如，一家做语音识别的公司应该如何进行横向拓展？首先，这家公司的语音识别产品的识别率应该已经非常接近或者到达百分之百，但是这家公司的产品依然不理解语义，这样的产品对用户而言是没用的，这时需要横向穿透到供应链领域的语义分析领域，需要自然语言理解技术。

纵向的理性的认知是指以 AI 技术起家的企业，因为需要打磨技术、需要训练模型，所以可以将技术迁移到其他领域，例如商汤科技就是把人脸识别技术迁移到物品识别领域，做人脸识别的同时瞄准安防、银行、新零售等领域“做深”，为行业赋能。

2. 感性的认知

单纯依靠技术建立门槛的企业并不多，技术能够领先市场一小段时间、为企业赢得市场先机已经很不错了。如果想要建立门槛，AI 产品经理需要热爱某个行业，深挖核心 AI 技术，持续赋能，以图领先市场一小步，即“感情不变，能力要变”。

这里举一个穿透思维的例子：某智能电视先做语音识别、语义分析遥控器，紧接着要上 MEMC (Motion Estimate and Motion Compensation，运动估计和运动补偿，液晶电视中常用的画质补偿技术)，之后做老年戏曲节目优化，因为白天在家看电视的大多是老年人。

2.4.4 大道思维

大道思维是指产品选择的“赛道”要足够长、足够宽。

一个不被人看好，但是活得却很好的产品方向是生意。有人说开一家杂货店，这个叫不叫创业？这个当然叫创业，但是这种创业存在非常大的问题，因为它的创新性不强，它的可复制性很强，这只是一个生意。

那什么叫大道思维呢？就是我们不要去某个非常小的生意，我们要做一个能够改变社会形态的、能够影响非常多人生活方式的、创新的产品，这样的产品才是最受投资人欢迎的。就像你说要去开一家店，笔者相信基本没有投资人愿意投资，但是如

果你说要开一千家店，投资人就会很兴奋，这就是区别，大道思维就是看你的野心和梦想有多大。

产品是所有创业公司最核心的竞争力。只有你的产品真正解决了用户的某种核心需求，能够用先进的技术去创造不同的体验，这样你才能够快速获得成功

其实，将大道思维用在产品的商业模式上有很多可以讲的东西。很多人在创业的过程中老是“盯”着别人怎样赚钱，琢磨着自己也要这样赚钱，这是不对的。最优秀的商业模式是让你的竞争对手最赚钱的东西，到了你这里没办法赚钱，你去赚取他想象不到的钱，这个时候你的竞争对手会非常痛苦，因为他不知道怎样跟你竞争。

所以，大多数创业者在选方向的时候，还是需要能落地，眼光放长一点就会发现，迟早要考虑盈利的问题。如果考虑盈利的话，就需要尽量根据现有的已知条件，去测算未来的回报预期。是做一个方向正确的大市场呢？还是挑一个自己有积累的、偏门的小方向呢？抑或是挑一个自己喜欢但盈利空间不大的方向呢？

有调研机构通过调查发现：年轻的创业者在选择创业方向时，首选几乎都是泛 AI 领域。通过 AI 技术起家，通过 AI 技术赋能；或者做一个 AI 应用，让 AI 赋能行业；找准“刚需”，站在巨人的肩膀上提出更好的方案。这都是比较容易成就事业发展、实现财富自由的大道思维。

大道思维的例子比比皆是，比如爱奇艺的奇秀在 2014 年主要运营 Live house 而不是像某些直播平台一样运营限制级内容；今日头条关闭内涵段子等内容平台，专注运营产生高质量内容的短视频平台——抖音。

2.5 成为 AI 产品经理的方法

2.5.1 非技术背景的人员如何转型成为 AI 产品经理

产品经理是一个非常重要的岗位，但是高校并没有培养产品经理特别是 AI 产品经理的专业课程，这方面的书籍也非常稀缺。非技术背景的人员转型成为 AI 产品经理是未来一段时间内的一种趋势，这里主要介绍非技术背景的人员如何转型成为 AI 产品经理，非技术背景的人员包括市场营销、运营、销售、BD（Business Development，商务拓展）、行政等岗位的人员。

首先，非技术背景的人员转型成为做 AI 产品经理需要收集市场上的一些信息。

笔者在为本书收集资料的过程中收集了以下信息,我们先以问答的形式来看一下。

1. AI 产品经理和传统产品经理的工作流程有何不同?面临的问题和困难有哪些?

回答:AI 产品经理和传统产品经理的工作流程基本上是一致的,不同之处在于以下几点。

(1) 需求把握。AI 产品还处于探索期(找刚需),ToC 产品的产品形态甚至典型用户群体(用户画像)都还不明确,所以信息(行业信息、竞品信息、用户信息)收集、创意思考、产品验证的工作会更加困难。

(2) 闭环验证。在进行产品核心价值的设计和验证工作时,AI 产品经理除了以数据分析为驱动,还需要有大胆的思路、敏锐的洞见。首先,从手机场景升级到 AI 场景,产品使用场景发生了翻天覆地的变化;其次,AI 产品的用户门槛远高于互联网产品,用户量和用户数据的规模远比不上互联网产品的量级;再次,AI 产品与硬件关系紧密,产品迭代周期更长,更难收集有效数据。

(3) 交互设计。场景巨变使得交互方式从纯软件(界面、触摸屏)、纯硬件的形式,升级到多模态交互等更复杂的人机交互形式,截至本书完稿时(2018年),AI 场景还没有形成清晰的交互体系。很多人没有意识到新时代不仅是新技术驱动的,更关键的标志是新交互(还可能是新硬件)。

(4) 功能设计。比如,搜索产品和 AI 问答产品是很不同的东西,同时,产品一旦和硬件相关,难度就会陡增。

(5) 数据分析。语音交互产生的数据分析难度远高于触摸交互。因为触摸屏有效表达用户意图的概率非常高,而通过语音识别得到的信息往往和用户意图有很大的偏差。

2. AI 产品经理需要具备的能力、素质和传统产品经理有何不同?

回答:各方面的能力和素质都需要跨越式的升级。

(1) AI 产品经理需要更广的知识面、更深的知识(经验)积累,不仅需要常规的 App 开发、网站建设的相关知识,还需要硬件、操作系统、生产的相关知识,甚至认知新的用户群体(小孩、老年人、外国人……)的相关知识。

(2) AI 产品经理需要更强的技术理解能力。对这一点存在两种观点:必须有技术背景,真的懂软件、硬件开发到底是怎么回事。清楚“什么能做,什么不能做”

即可，甚至只要整个产品团队有人能做到这点，能够进行配合就行。无论如何，对 AI 产品经理的要求比对传统产品经理的要求更高。如果学生时代就接触神经网络、深度学习、自动化等内容，会存在优势。

(3) AI 产品经理需要更强的学习能力（悟性）。不仅需要学习传统产品的相关知识，还可能需要随时去学习新出现的知识（甚至是新出现的技术知识）。

(4) AI 产品经理需要重新认识人的感知和交互方式。

(5) AI 产品经理需要具备更强的洞察力、创造力，以及想象力。

(6) AI 产品经理需要对行业有正确的认识，对趋势有准确的判断。行业周期性如何？AI 的发展方向如何？新的技术、新的产品形态，哪种能最终胜出？如何组织这些新的技术、新的产品形态？时间窗（时机）如何？政府、大学等机构对行业的影响该如何理解、应对和借势……

(7) AI 产品经理需要更深厚的人文素养和更高的灵魂境界。笔者认为，真正的 AI 产品，需要超越纯逻辑性的思维和内涵。比如，有人认为，对于机器人产品而言，把功能做好（有用）就可以了，但笔者认为，机器人和人交互时，一定会伴随着情感等非理性因素的影响，这不是设计者想规避就能规避的。当笔者第一次近距离看到超大的工业机器人时完全被震撼到了，因为那种协调的动作、和谐的节奏，会让观察者本能地觉得“像人一样”。

更深入地说，一个产品，本质是其公司、设计者灵魂层次的外化，一个精神层次不高的团队，不可能做出一款跨时代的 AI 产品。

3. 如何判断自己是否对 AI 领域真的感兴趣？如何判断自己是否适合 AI 领域？

回答：“感兴趣”不是一个形容词，而是一个动词。关键不是你多么兴奋，而是在兴奋过后，你具体做了什么。

如果做完了（至少做了一部分）后面第 4 个问题提到的任务，笔者才认为你是真正的“感兴趣”。也只有基于此，并且真的去做了 1、2 个产品特征设计，才可以判断你是否合适从事 AI 产品经理工作。

另外，如果讨论的还是下列问题（无营养、无逻辑、无边界、无方案），那就不适合从事 AI 相关的工作。

(1) 对于“强 AI”的好奇心与普通群众或科幻小说家并无二致。

(2) AI 再厉害也摆脱不了“吹灯拔蜡(断电)”的尴尬,根本不智能。

(3) 我想和 AI 谈恋爱。

(4) 大自然到处存在着人类无法理解的算法, AI 再厉害,也只是大自然的一部分。

(5) 智能家居会随时监控我的生活状态,这种可能产生危险的东西,除非能够证明其可靠性,否则我不会使用。

4. 要成为一名合格的 AI 产品经理,需要做什么准备工作?

回答:对于这个问题,笔者有以下建议。

(1) 查看知乎上所有和 AI、机器人相关的问题,看相关书籍或课程。

(2) 体验各种 AI 产品,了解产品、技术的发展现状(比如与语音交互相关的各种传感器及其作用),收集行业、公司、团队的相关信息。

(3) 整理出自己的独到见解,认真思考人类行为,研究自我。

(4) 筛选出几个你真正认可的创业团队,想尽办法结识其中的牛人(带上你的独到见解去交流)。

非技术背景的人员转型成为 AI 产品经理还需要补充一些必需的技术知识。

细数全世界优秀的产品经理。乔布斯是极致的代言人,他定义并设计的苹果产品改变了一个时代,引领了潮流。他的苛刻、极致、改变世界的初心影响着如今科技行业的产品经理们。张一鸣,抖音之父,深谙人性、引领潮流,能把一款产品做到人们的生活中,让几亿人都为之买单,实属境界。相信每一个产品经理都有改变世界的梦想,也都在这条不归路上蹒跚前行,但改变世界的毕竟是少数人,更容易改变的是自己,在产品之路上修炼自己、完善自己,也许在不经意间会发现自己已经做了一件了不起的事,产品经理要有宇宙般宽广的胸怀,同时也要有蚂蚁般的勤奋。

产品经理成了互联网公司的标配,但产品经理这一职能是没有体系化的培养系统的,现有的产品经理更多的是从其他岗位转岗过来的。有从技术岗位转岗的,有从设计岗位转岗的,也有从运营岗位转岗的,还有从销售岗位转岗的,产品经理的背景差异很大,不同背景的人做产品经理的风格也各有差异。但产品经理这一职能的根本作用是挖掘并分析用户需求,通过 IT 技术打造一款产品满足用户的需求。所以,对于产品经理来说,工作就是一个创造的过程,这个过程需要具备的首要思维

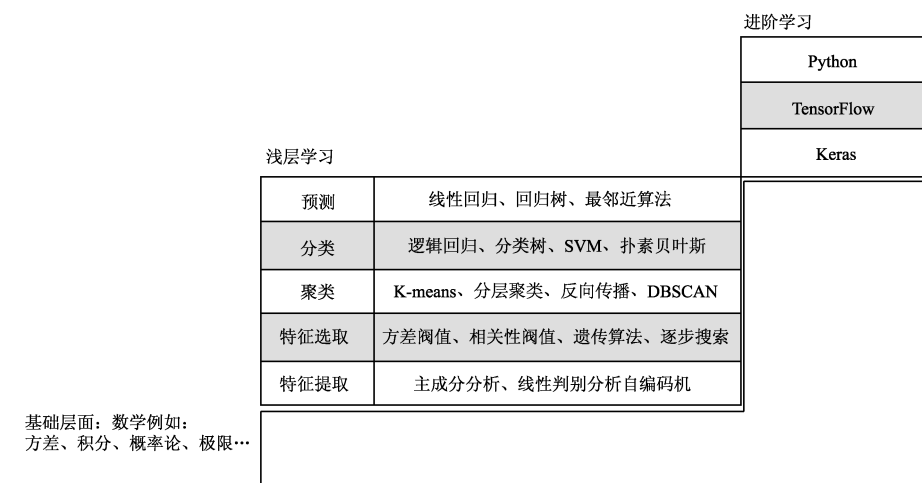
是产品思维，或者说用户思维；其次，产品经理要通过 AI 技术工具去实现产品，去满足用户，非技术背景的产品经理要想做好 AI 产品，对使用 AI 技术工具去创造的过程也要有全面的了解。

如今的产品经理基本分为两大类，一类是技术背景型，也就是从程序员转型过来的；另一类是非技术背景型，也就是从设计、运营或其他职能转型过来的。乔布斯是典型的非技术背景型的产品经理，张小龙是典型的技术背景型的产品经理。但最终，他们的产品都是以优秀的用户体验为导向。

产品经理需要与各种职能的人打交道，其中，产品经理与工程师的互动最频繁，矛盾也最多。在工作流程中，产品经理的工作属于流程的上游，也就是设计者，而工程师的工作属于流程的下游，也就是实施者。简单来说，工程师是帮助产品经理实现想法的人。要知道，人是很难理解别人的。技术思维的工程师和产品思维的产品经理，天然就存在一些离心力，这也是为什么在互联网行业里，产品经理和工程师一直存在一种对立，这种现象本质上是由这两种职能的工作内容和思维方式决定的。对于产品经理来说，特别是对于非技术背景的产品经理来说，如何学习并使用技术思维，做到与工程师顺利配合和沟通，对推动双方合作和产品发展都有非常重要的意义。

如果按上述两大类来区分产品经理，技术背景型的产品经理在与工程师合作的过程中具备天然的优势，因为“本是同根生”，所以在合作和沟通过程中有共同语言，产品经理会用技术思维去思考问题。但对于非技术背景型的产品经理来说，在与工程师的沟通中会存在“语言”差异，当工程师用技术语言和非技术背景型的产品经理讨论一个产品实现方案时，对于非技术背景型的产品经理来说，很可能就处于信息输入无效的状态，而产品经理用自身原有思维去与工程师沟通时，也会出现“语言”不对的情形，最终双方很难达成一致，致使问题得不到解决。

对于非技术背景型的产品经理来说，掌握一些基础的技术知识尤为重要，对于 AI 技术的理解可以不用深入到实现层面，但对于 AI 的基本原理以及产品背后的整体技术架构和 AI 技术的成熟度要做到心中有数。产品经理处于信息流的上游，在与工程师沟通自己的产品想法和推动产品实施的过程中，对技术有一定的了解，就好比手上多了一把“好武器”，能让问题顺利解决，让产品不断向前发展，不管是在与工程师合作环节还是在产品设计环节都能给自己带来非常大的帮助。



成为 AI 产品经理需要抓住的技术点

2.5.2 有技术背景的人员如何转型成为 AI 产品经理

上文提到了非技术背景的人员转型成为 AI 产品经理的一些方法论，本节主要讲有技术背景的人员转型成为 AI 产品经理的方法，笔者建议有技术背景的人员转型成为 AI 产品经理需要重点抓所应用的 AI 技术适用的场景。

比如，技术人员在思维模式上可以突破技术的限制，如意识到图像识别技术除应用于金融窗口、酒店窗口外，还能被应用在路边停车员手持的车牌识别设备。再比如，语音识别技术除了应用在智能音箱类产品，还能被应用在陪老人、儿童进行聊天的产品中，对应的落地产品有日本的 Pepper 情感关怀机器人。

再看一个笔者在工作中接触的案例。某位程序员拥有 12 年的技术背景，在一家中型（300 多人）互联网公司担任技术总监，但是也需要写一些代码，确定一些架构，引进新的 AI 技术。他自己也在不断完善自己的一套开发框架，对技术感兴趣，爱钻研，他个人的规划原本是再积累 2~3 年，等自己的框架再成熟一些的时候到大公司做架构师，甚至到一些不算太大的企业担任 CTO。

其性格比较开朗，所以现在的职位里的一部分沟通、协调、管理的工作由他负责。当一位产品经理要离职时，老板找这位技术总监谈话，希望他尝试一下产品经理职位。老板认为他平时对产品比较有想法，几个高层也都觉得他有做产品的潜力。他很纠结，倒不是排斥产品经理这个职位，但是放不下技术，于是找到笔者进行交流。

笔者的建议概括如下：产品经理和技术工程师工作内容的区别是很大的，产品经理对研发、运营、销售、服务、管理等工作都要了解，对上要给公司高层做产品策划，对外要对接市场和渠道的相关人员，会议多，出差也多，如果喜欢沟通，想法比较多，建议从事产品经理工作，发展空间会很大。如果喜欢钻研技术，个人成就感都来源于技术，那笔者不建议他从事产品经理工作，因为产品经理一般都不会聚焦在某个技术上，要耐得住产品“火”起来前打磨产品的寂寞。产品经理相当于军师，技术人员相当于战将，两者的侧重点不同。

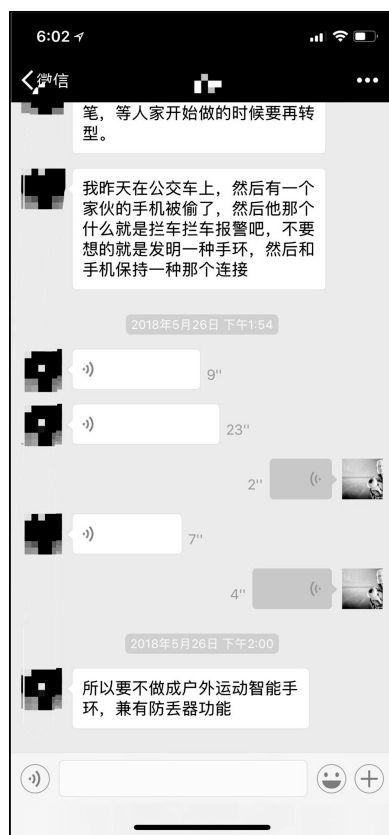
但是无论是钻研技术还是转型成为 AI 产品经理，目标都要明确，AI 产品是一个系统，随着经历的增长，不论哪个岗位都要有系统思维。

建议技术人员转型成为 AI 产品经理后要主攻应用场景，即要能够将技术结合进场景。

2.5.3 有技术背景的人员转型成为 AI 产品经理的案例

这里举一个拥有 AI 技术背景且很擅长机器学习的技术人员转型成为 AI 产品经理的例子。这位转型者拥有 7 年的 IT 技术经验，4 年的 AI 技术经验，现在刚转型成为 AI 产品经理 6 个月，因为之前一直从事 IT 产品开发工作，所以他在记得产品的痛点、需求、市场先机是什么。基于上述背景我们来分析一下右图这个案例。

首先，他的产品意识和想做产品的心态是积极的、正面的，但是他没有意识到“做一款智能硬件是一件高成本的事，耗时费力，当发现一个需求的时候要仔细琢磨，千万不能‘病急乱投医’，如果确实要做也要先从原型机开始，等原型机获得市场和投资者的认可后再生产样机。”所以，不能发现一个需求就全部投入，要思考这个需求本



有技术背景的人员转型成为 AI 产品经理的案例

身的价值,要思考这款产品用到了哪些技术,每个技术的成熟度等,最重要的是通过原型机验证这个需求的大小、强弱。

2.6 AI 产品经理的创新方法

许多 AI 技术在飞速发展之中,有些 AI 技术是自带概率属性的,所以 AI 创新是一场无限逼近完美的过程。

笔者对 AI 产品创新实战的要求汇总如下。

(1) 多做。工作量越大,找到创新方法的可能性越高,想出 200 个以上的点子,实现从量变到质变的跨越。

(2) 不盲从,打破惯性。

(3) 具备向竞争对手竞争的思维。

(4) 学会书面头脑风暴,挖掘个体的需求。

(5) 及时激励最佳创意的提出者。

(6) 具备鉴别好想法的能力。

(7) 产品总监要营造竞争氛围。

(8) 要将产品趣味化。

(9) 要将团队敏捷化。

此外,笔者再深度解读一下协同打造 AI 产品的方法。

在产品的市场竞争压力不断增大、产品的创新成本快速上升、用户需求个性多样的背景下,越来越多的企业将有限的资源投入到能够更好地满足用户关键需求的“点”来进行产品创新,以实现产品差异化发展,达到提高产品竞争力、附加值、创新效率和客户满意度的目的。用户协同产品创新(Customer Collaborative Product Innovation, 简称 CCPI)作为一种新兴的、能够最大限度发挥用户潜力的产品创新模式,得到了 AI 产品创新设计、研究领域的专家及学者的普遍关注。然而,撰写本书时笔者未检索到 CCPI 中“面向用户关键需求的产品协同创新方法”的系统性研究成果。

因此,笔者在这里立足于“对用户需求进行排序”等相关理论和技术,重点对 CCPI 中“面向用户关键需求进行协同创新”的原理及形成机制、影响协同创新的因

素及识别企业内部协同创新因素、产品创新方案生成法进行深入研究。

研究开发、工程设计等专业人员的主要任务是创新。潜在的创造力和解决问题的能力是每个人都有的，单个人或组织的创造力是有限的、渐进的，但活动累加的影响是深远的。大多数企业的创新活动建立在个别技术的基础之上，依靠 AI 技术突破取得成功的企业尚未形成一套系统、全面、驱动全员参与的创新流程。基于用户需求的变化，引导并驱动创新活动的开展，调动满足用户需求的积极性和有利因素，是协同创新的关键。

2.6.1 用户协同创新

1. 用户协同创新的形成机制

首先回顾一下用户协同创新的知识内容。用户协同创新是适应信息技术和信息经济发展的一种新的创新模式。通过利用用户协同创新模式，挖掘用户不同的需求模式和需求信息，区分不同类型的用户需求，是实现基于用户需求的协同创新的关键。任何类型的项目或产品，其复杂性和新颖性都是影响其开发成功的关键因素。

高新颖度产品最常用的创新方法是市场细分、原型设计和成为产业专家。低新颖度产品常用的创新方法是消费者参与、市场细分和用户开发。常规产品常用的创新方法是市场细分和消费者参与。

对产品的市场成熟度、技术成熟度有清晰的认识是实施产品创新的前提。不同的创新驱动着不同技术成熟度、市场成熟度的产品的开发和商业化过程。当技术和市场都处于低新颖度时，创新的重点在于改进现有技术的使用以满足用户的特定需求，通过与用户协同，实现产品和服务在定价、包装和售后服务等方面的差异化创新。当使用原有技术去开发新的市场，或为原有市场提供新的使用方式时，企业主要通过原有技术的新颖组合服务于特定市场，此时企业必将与顾客保持密切的关系，创新的主要来源在于潜在用户的参与。当开发新技术以满足成熟市场的顾客需求时，企业竞争的焦点主要在于产品的性能和新颖性，即为现有市场提供性能更高或者更新颖的产品，而不在于产品的质量和价格。创新的源泉是开发者，但了解消费者的潜在需求和欲望是创新的关键。当技术和市场的新颖度都较高时，由于新技术的应用和新市场的开发都没有明确的界定，所以新产品在多数情况下是由开发者和领先用户合作来进行的。

2. 用户协同创新的流程

作为 AI 产品经理,再来看一下用户协同产品创新的流程。由于不断接触新产品,领先用户会不断给销售或服务人员提出他们的需求愿望,这种愿望通过销售或服务人员的描述将在专业设计人员那里转变成产品构思,产品设计人员依据该产品构思,利用自己的专业知识,最终形成满足客户需求的成熟的创意方案。

形成创意方案后,进入不同的协同创新流程。技术和市场比较成熟的项目直接进入协同开发流程,以快速满足用户需求;技术成熟而市场新颖度较高的项目,需要对现有技术进行重新组合,在制定项目任务书之后进入协同创新流程;市场比较成熟而技术比较新颖的项目,将制订技术开发路线图,与用户协同开发新技术,形成技术开发路标,并最终形成技术开发的任务书,进入产品协同创新流程;市场和技术都比较新颖的项目,进入市场管理流程,形成产品开发计划、市场和技术开发计划,形成技术与市场开发路标后制订项目任务书,再进入用户协同产品创新流程。

实现用户创意包括三个渐进的过程: 用户创意的形成,由于现有产品可能在性能、款式、颜色或者其他满足用户价值的因素上不能完全满足用户的要求,或是信息黏滞驱使用户产生产品创新的动机。 用户创意的实现,用户创意经过市场营销人员的加工,变得更加清晰而明了,但要真正实现用户的创意,产品经理必须加强与专业设计人员的沟通与联系。 形成成熟的产品创意方案,用户、营销人员和产品设计人员的良好沟通是形成产品创意方案的基础,设计人员通过沟通不断学习和积累知识,并将用户的创意不断融入产品创新的过程,设计人员的知识将得到更新和传播,随着知识的更新和传播,设计人员可能会被激发产生新的知识,从而保证用户最初的创意演化成完全可行的创新产品。当整个协同创新过程完成之后,可以通过对整个产品协同创新过程进行评价,更新产品设计库,形成企业知识库,以便让企业更好地满足客户的需求。

2.6.2 影响用户协同创新的因素

上面回顾了用户协同创新的相关知识,我们再来看一下影响用户协同创新的因素,因为这些因素决定了用户协同创新的深度。

1. 基于用户需求的产品协同创新的特征

(1) 用户深度参与。在基于用户需求的协同创新过程中,用户不仅仅是产品的消

费者，更是产品创新的参与者。用户对于企业产品创新不再是通过定制或者是选型、配型来参与产品的设计、生产活动，而是与企业的市场营销、研发、生产等人员一道协同进行产品的创新。

(2) 协同技术网络化。市场信息化和知识网络化的快速发展为用户协同创新提供了更为便利的条件，尽管用户、营销、研发、生产等人员分布在世界各地，但基于信息技术和互联网技术的协同创新环境可以利用各种先进的创新工具和创新知识来进行用户所需的产品创新，满足用户的个性化需要。

(3) 目标一体化。用户协同产品创新始终以满足客户需求为特征，创新团队成员通过相互交流、学习，他们的知识、经验和创新能力出现螺旋式上升，快速实现新知识的创造，并确保创新知识或创意转化为最终的创新产品，而客户在此创新过程中进一步理解自己的需求。

基于用户需求的产品协同创新是网络环境下的、以满足用户需求为目标的、用户深度参与的、复杂的协同创新过程，主要受到企业外部的协同创新因素、企业内部的协同创新因素、用户因素的影响。

2. 企业外部的协同创新因素

好的协同创新环境可以激发创新组织的主动性与积极性，加速产品创新的进度。外部的协同创新因素主要包括市场环境、政策环境、技术环境等。

(1) 市场环境。企业需要对市场环境进行动态的把握和理解，识别市场要素的各种变化所蕴含的潜在机会；充分利用各种预测工具获得有价值的信息，进而获得工艺创新的目标。市场环境调研是与用户合作，由用户提供所需要的产品或服务的信息，这些信息可以通过小组调查、常规调查和其他相关的调查获得，以快速响应顾客、提高产品创新的针对性和成功率、降低创新成本。成功的协同创新强调在供应商和顾客之间维持长期的互动，企业可以通过邀请客户加入协同创新小组或进行双向拜访的方式维持与用户的持续沟通，将用户理念不断灌输到企业所有职能部门，而不是简单地将其作为一个市场信号传递给产品研发人员，牢固树立“全体员工为用户服务”的理念。

(2) 政策环境。国家创新系统环境，国家创新系统环境在一定程度上决定了企业面对机会和威胁时的选择范围，多数企业是在一个或者两个国家内培育、制定和发

展其执行创新战略所需要的技能和知识,各国都有不同的领先技术领域。国家企业治理系统,即企业所有权和控制权的实施和改善体系。不同治理模式决定着不同的创新激励机会和创新动力。他国创新系统环境,密切关注他国创新系统在组织、技术和生产能力方面的优势,学习那些发展势头“猛”的新型国家的创新系统。因为这些创新系统可能培养出一大批具有较强创新竞争力的新企业,本国企业也可以从国外创新系统产生的技术获利。

(3) 技术环境。新产品创新的机会大多来源于知识的不断积累和发展,新产品、新工艺和新服务的很多创意和构思之所以会成为一种可能,是因为明确了一种技术在未来的发展轨迹之后,企业会积极寻求信息或技术供应商,并建立联系,联系的模式可以通过建立“虚拟大学”等形式将各方面的专家联系起来,为公司提供新的技术信息。技术信息的搜寻和组合还必须通过开展标杆比较去探索技术信号最有力的变化,从而去发现更有竞争力的新产品、新工艺或者新的市场。

3. 企业内部的协同创新因素

(1) 精益生产。对任何企业来说,如果没有建立能够对变革做出快速反应的组织结构,将很难获得成功。在当前的市场竞争条件下,创新已不仅仅是研发部门的事情,而日益变成整个组织的任务,组织创新的重要性日益增长和组织制度的不断变革,给传统的稳定而机械的组织结构带来了严峻的挑战。精益生产就是一种解决创新问题的生产组织模式,这种组织模式通过团队协作、共同参与、灵活的工作单元和打破层级壁垒,建立了一种更加松散的组织结构,将非常有利于协同创新活动的开展。

(2) 核心人物。创新的复杂性和不确定性往往会造成很多创新的天折,解决这一问题的有效途径是确定创新活动的灵魂人物,他支持创新,并以自己的极大热情和精力去组织创新、推动创新。灵魂人物的作用在于能够理解关键技术、核心知识,能够解决从实验到实践过程中出现的大量问题;对创新成果拥有的潜力有极大的自信,而这种自信可以有效地说服企业决策者为之提供有效的支持。创新的成功离不开高质量的信息交流,仅仅通过先进的IT技术和软件是不可能实现有效的沟通和知识共享的,信息交流更重要的是通过人与人之间的交流来实现,核心人物在知识管理领域的作用也日益凸显。

(3) 培训、发展。通过员工培训和员工发展可以让员工运用新设备、掌握新工艺、理解新理念,全面提高员工素质,同时也能够增强员工的责任感,调动员工的工作积

极性。当组织进行一项重要创新时，不少员工也许会感到创新对技术能力的要求是他所不具备的，由此会产生焦虑，这种焦虑和不安使得员工可能会抵制创新。培训使他们明白了创新战略的合理性，使得创新项目得到顺利推动。培养组织、团队良好的学习习惯和学习方法是培训的另一个重要作用，培训有助于建立一个良好的“学习型组织”。“学习型组织”的核心在于其学习的持续性和团队性，但要让员工做到这一点并不容易，所以，我们在制订员工技能培训计划时还要考虑如何培养员工养成良好的团队学习习惯。

(4) 团队协作。有效开展协同创新的一个最佳组织方式是团队。融合不同的观点来解决现实问题是协同创新解决问题的主要方法，团队工作模式为这种融合创造了可能。衡量一个团队高效与否的重要标志是团队的任务和目标是否明确；团队的领导是否高效；团队内个人角色与能力的匹配是否合理；团队内部的争端处理机制是否健全；团队与外部组织的持续联络是否通畅。

(5) 创新氛围。创新不仅是新事物的发明，更关键的是新事物的应用。许多伟大的发明都来源于幸运的偶然发现。虽然创意的形成需要灵感的闪现，但在实现和应用创意的过程中需要不断发现和解决一系列的小问题。不管是灯泡、即时贴、便签条的发明，还是其他一些成功的创新活动，都包含着各个时代、各种人的创造性劳动。建立创新性文化氛围涉及组织结构、沟通渠道、奖惩系统、培训政策、会计和评估系统以及战略展开这一整套内容。

(6) 关联业务。业务联系密切的部门之间的关系协调与否会对协同创新产生重要的影响。协同创新尤其需要两个业务往来频繁的部门间进行有效的协同。如在一些大型企业中，由于部门相对较多，业务处理复杂，涉及人员多，如果部门业务衔接失利，可能会导致企业处理问题的周期过长，造成企业无法有效对用户的需求产生快速反应，从而失去用户、失去市场。因此，我们可以通过制定合理的利益分配方案，促使不同部门在同一项任务中分工明确，既各司其职又精诚合作。

4. 用户因素

用户不仅提供需求信息和形态各异创新思维，更将自己在产品使用过程及其他方面获得的与产品创新有关的经验、知识和创新技能及时有效地和营销人员、设计人员进行沟通和共享。影响用户协同创新的用户因素主要是用户的知识类型和知识结构。用户的知识有以下几种类型：用户的基本知识，主要是指用户的基本信息以及由于

用户与企业之间相互作用而产生的大量信息；用户的创新知识，主要是指用户在现有知识的基础上，由于参与创新过程而学习并创造的新知识；用户的技能知识，主要是指创新用户所具备的技能；用户的经验知识，主要是指用户在与团队成员不断交流以及不断体验产品所获得的经验。用户协同创新还需要研究用户行为及用户与制造商的关系，具体包括识别领先用户、创新用户群体对创新的贡献、用户参与创新的心理及行为特点等，对这些因素的认知和了解可以帮助企业构建良好的用户协同创新系统。

笔者对用户需求协同创新方法进行整理和研究后，结合 AI 产品的实战，发现 AI 产品的创新方法与传统产品的创新方法的原则是一致的，AI 产品的创新方法包含传统产品的创新方法，且因为借助 AI 进行创新的灵活性更高，借助 AI 进行创新的方法为 AI 产品所独有。

2.6.3 AI 产品的创新方法

这一部分笔者将带读者一同了解“框定熟悉感”和“建立新奇感”两大 AI 产品创新方法。

如果你刷抖音，你一定在抖音上看到过海底捞的各种“网红”吃法；如果你经常去肯德基、必胜客等快餐店，不难发现它们每段时间都会推出新品；如果你看综艺节目，肯定知道“我是歌手”“中国有嘻哈”“火星情报局”等；如果你玩《王者荣耀》《英雄联盟》，会发现隔段时间就会有新版本（新英雄、新属性、新时装）出现……

还记得健力宝、大宝 SOD 蜜吗？还记得天涯社区、猫扑网、西祠胡同吗？这些产品是不是已经被你逐渐淡忘了。

上述产品的兴衰，其实大多都是因为一件事：新奇感。PF(Parapsychology Foundation) 等国际顶级研究机构对“新奇感”进行了研究，不论是从商业、记忆、历史等方向，还是从其他方向进行研究，得出的结论都是“人类喜欢新奇的事物”。

但是，我们不断更新产品、不断创新就能刺激用户来体验吗？

可口可乐的咖啡味饮料、高露洁的牛肉宽面条、谷歌眼镜等产品都做了创新，为什么没有成功呢？其实，这些产品在做创新时少考虑了一件事：熟悉感。

一个产品通过创新刺激用户应该遵守一个原则——保持熟悉的**新奇感**。这其实与

旅游行业的底层原理相似，人类喜欢熟悉且新奇的事物。旅游满足了用户对于新事物的好奇心，让用户体验新的美景、新的食物、新的生活方式。但是如果携程给你推送一条“博茨瓦纳”的旅行信息，而你对这个地方完全不了解，此时恐怕文案再深刻，图片再绚丽，也不会马上打动你，因为你完全不了解这个地方，这会让你产生“对未知的恐惧”。

人类恐惧未知的事物，保险行业就遵循了这个底层原理——满足人类，提供确定性和安全感。所以，你选择的旅行目的地肯定是你从某个信息源（音乐、电影、文章、媒体、社群等），有过一定的了解的地方。

问题来了，如何掌握“熟悉的新奇感”，完成产品创新？

笔者整理了2个方法供读者参考。

1. 携带熟悉感

(1) 携带场景。

目标用户对这个场景的熟悉程度如何？会不会因为完全熟悉而无法产生想要体验的欲望？会不会因为完全陌生而需要花大量时间来培养习惯？

场景熟悉感弱：下午茶源于英国，我国是没有下午茶这个概念的。虽然有商业刺激（外卖平台众多的下午茶品类）和文娱内容（电视剧、音乐中对下午茶的呈现），但在下午茶这个场景中，相关产品的体量并不大，这种情况在一线城市以外的地区尤为明显。所以，如果一个蛋糕店针对下午茶这个场景做产品创新，其性价比肯定不高。

在外卖行业发展初期，大多数用户对“点外卖”这个场景并不熟悉，所以饿了么和美团最开始针对的是大学生市场，因为大学生对新事物的接受度高，后来随着资本的介入，企业通过大量的补贴来培养白领用户，外卖场景才逐渐被大多数人所熟悉，滴滴打车的发展也是如此。

场景熟悉感强：肯德基和麦当劳在早餐中推出粥、油条、卷饼、春卷等产品，因为这些都是中式早餐非常常见的产品，用户对粥、油条、卷饼等产品的熟悉感被携带了过来，再通过品牌加持，产品自然很容易就被用户接受。

(2) 建立新奇感。

产品熟悉感弱：对爱奇艺的三款综艺节目“中国有嘻哈”“街舞团”“机器人争霸”

进行分析会发现，普通用户能分别评判节目并给出如下评价。

中国有嘻哈：押韵、节奏感强、好看；

街舞团：节奏感强、好看；

机器人争霸：好看。

所以，三个综艺节目的主要元素在大众的认知范围内的排名是：中国有嘻哈>街舞团>机器人争霸。对于元素的了解程度是能否让用户沉浸其中的核心指标。所以，抛开其他因素（明星、剪辑等），“机器人争霸”定位的目标受众自然就少于其他两个节目。

产品熟悉感强：小米之家。在小米之家出售的产品（路由器、插线板、旅行箱等）基本都是用户熟悉的产品，所以不需要再去培养用户习惯，而小米通过性价比给用户带来新意。② 欧丽薇兰橄榄油。橄榄油是用户熟悉的品类，具有“健康”属性，由于用户对产品概念熟悉，所以不需要再培育市场。欧丽薇兰基于熟悉感，通过“品牌营销+移动营销”获取了市场成功。

2. 建立新奇感

更新产品“感觉”是指：产品本质功能不变，利用外部因素（广告语、包装、场景、用户等）使产品产生“新奇感”。

（1）广告语。加多宝就是通过不同的广告语传达不同的感觉，从而赋予产品“新奇感”。正宗好凉茶，正宗好声音（传达加多宝为正宗，XX 不正宗的感觉）；全国销量领先的红罐凉茶，改名加多宝（传达 XX 改名为加多宝的感觉）；对不起系列（传达加多宝受到不公平待遇的信息，激发人同情弱者的本性）。

（2）包装。可口可乐的“歌词瓶”和“城市套罐”将外部感觉（歌词的情绪、城市调性）赋予可口可乐，使用户看到产品包装时候产生“新奇感”。

（3）场景。喜茶将北京分店开在三里屯，将三里屯的时尚、潮流等元素赋给“喜茶”品牌，当用户收到三里屯和喜茶两个品牌的相关信息时，可能觉得喝“喜茶”是一件很“潮”的事。

（4）用户。“特殊”用户促进连带效应。比如，Supreme 作为年轻潮牌，在我国引爆于“中国有嘻哈”，基于“中国有嘻哈”的流量，大批用户看到很多玩嘻哈的人都穿这个牌子，会产生一种感觉：穿 Supreme 才是最“潮”的事。

创新产品功能是指：基于用户的痛点、爽点、痒点更新产品的本质功能，使产品产生“新奇感”。以下是几个典型案例。

零度可乐、减脂雪碧：解决喝可乐的人怕影响身体健康的痛点。

小米插线板：在插线板中加入 USB 插孔功能，解决大插排不好看，小插排插孔不够用的痛点。

偶像练习生：在大家都了解的选秀模式中，更换最主要的元素（把路人替换成练习生），解决选秀节目质量较差，用户观感不强的痛点。

抖音：区别于之前的内涵段子，极大地简化用户操作流程，且重视用户“好看”的需求（智能推送、内容演绎等），让用户使用抖音的时候特别“爽”，为用户制造爽点。

足迹：让每个人都能拍出电影大片一样的照片，满足用户体现自我拍照能力的这个痒点。

2.7 成为 AI 产品经理的核心路径

这一节我们看看成为 AI 产品经理的几个阶段。

首先看入行前的心理认知。笔者建议读者不要盲目崇拜 AI，同时也绝不能因为迷茫而轻视 AI。比如，很多人认为 AI 很快就要改变我们的生活，在不久的将来我们将不再需要人工驾驶汽车，汽车可以自动驾驶，因此可以不学车了；很多人还认为在不久的将来也许不需要上网去社交，因为出现了 AI 生成的男女朋友。笔者觉得要实现以上两个场景至少还需要 15 年以上的时间。

大家对 AI 在消费领域带来的产品创新的期望过高，因此投资过多。笔者觉得对消费直接产生影响的 AI 产品的质和量都比较有限。

成为 AI 产品经理之前在心理上应该认知到：在生产领域，AI 已经改变了我们的生产方式；在商业决策领域，AI 正在改变一切。

以前，产品经理是科技网络行业里最没有标准的职位。无论是谁，无论是大公司还是小公司，所描述的产品经理可能都是不一样的，产品经理在实际工作中所做的工作也千差万别。转行成为 AI 产品经理时也面临同样的问题，因此笔者在这一节总结了转行成为 AI 产品经理的特色方法。

1. 第一个阶段：入行前

也许你在各种各样的 AI 会议、论坛上听说过 AI 产品经理这个职位，也了解过一些相关的信息。这里面有的把 AI 产品经理夸成未来世界的设计师，离 CEO 最近的人；有的说 AI 产品经理是打杂的；有的觉得 AI 产品经理的工作成就感爆棚、专业性极强；有的又说 AI 产品经理是样样都懂、样样不行……在了解了一些信息后，你可能觉得自己还是想要从事 AI 产品经理这个工作。于是你开始找各种各样的资料来学习，也许是报各种成为 AI 产品经理的辅导班；也许是阅读产品经理写的书，笔者之前就写过一本《产品经理进化论：AI+时代产品经理的思维方法》；也许是……你可以学习的东西太多，根本不可能学完。不过，这些学习还是有好处的。通过这些杂乱无章的学习，你开始有了一点信心，觉得自己对做 AI 产品有了一些了解。终于，你开始投递简历，或者在公司内部要求转岗，开始走上 AI 产品经理这条路。

入行前的这个阶段，需要你具备自我学习的能力，否则很难有机会获得进入行业的机会，这个行业的人太“拼”了。

2. 第二个阶段：做协助

经过面试或公司内部转岗，你终于成为了一名 AI 产品经理。有的公司可能不会给这个头衔，而是 AI 产品助理，不过都是一回事，你做的工作应该都是差不多的。

如果你看过产品经理所要做的工作，会觉得自己一上来就应该大杀四方、独当一面。事实上，这个阶段你基本上是没有这样的机会的，因为你对 AI 产品的认识还不够。多数时候，公司会安排你做一些杂事。比如：录入、整理反馈意见；收集、整理各种渠道汇集过来的需求；做一些产品会议记录；做一些数据统计工作；挑一挑公司产品细节上的毛病……你会发现，你做的事情都是杂事。你也许会抱怨，为什么不让我做一款产品或一个功能呢？别急，虽然你前面做的这些事情都是非常基础的工作，但是很多厉害的产品经理现在为了能够真正了解一线的情况，还在亲自做这些事情。所以，公司让你做这些事儿其实是在为你打基础，很多 MIT 和 UIUC 的 AI 专业的高才生也会做基础工作，如数据的标注工作。毕竟，你刚刚成为一名 AI 产品经理。

在这个阶段，你所需要具备的核心能力也许就是吃苦的精神、永不放弃的精神了。

3. 第三个阶段：做 AI 功能

如果你顺利地适应了协助 AI 产品经理的工作，而且积累了一些自己的心得。你

会发现公司或你的上级会开始给你一些小需求练手。刚开始，真的是非常小的功能，比如：改版一下异常交互、设计一个产品后台的界面、做一个简单的在线反馈功能……可别小看这些小功能，这些功能同样有各种各样需要考虑的地方。你也许会觉得做这些小功能是不是太小了，对于老手来说，可能会有一点，但对于新手，完全不会。首先，这些小功能会让你熟悉 AI 产品的整个开发流程，尤其是在 AI 技术发展时期，这些功能是完善产品的重要组成部分，能让你和设计、研发、测试的同事们逐渐熟悉起来，以便后续工作的开展。其次，这些小功能也会让你对做产品有一个真实的感觉。很多东西，知道是一回事儿，自己真正做是另外一回事。一个异常为什么放这里而不是放在别的地方可能就够你思考一阵子了。

当你扎扎实实地做一个功能时，随着经验的积累，你会发现你逐渐接触到了公司产品的核心功能。

在这个阶段，一名 AI 产品经理最需要的能力是缜密的逻辑能力。软件部分和硬件部分都要思考，你必须考虑清楚每个功能的流程及各种状态，才能将具体的功能做好。

4. 第四个阶段：做模块

如果你 AI 功能做得不错，你会被委以重任——做 AI 模块。AI 模块所包含的东西就不仅仅是单个功能了，它是一个能够独立运行的有自己定义的点或者一系列功能的组合。在这时，你可能发现你在做功能阶段所考虑的那些东西在这里更加复杂了。以前是一个功能内部的逻辑，现在要考虑多个功能之间彼此的关系。谁重谁轻，谁先谁后，谁和谁彼此影响，谁和谁又互不干扰，这些都是你需要考虑的。

如果说一个小的功能还不能被称为完整的产品的话（小功能很多时候不能让用户完成一个真正的任务），一个模块完全称得上是一个完全的产品。用户使用一个模块往往能够完成一项或多项任务。如果你在上一个阶段的基础不够扎实，这个阶段可能就会比较辛苦。毕竟，你要考虑的是比以前复杂很多倍的东西。

在这个阶段，除了以前要求的能力，你最需要的可能是归类或分类的能力，轻重缓急、彼此关系都需要你自己来判断并组织。

5. 第五个阶段：做 AI 产品

如果你有机会的话，你也许能够做到这个阶段。在很多大公司里，能够真正独

立负责一款产品的机会是很少的，大公司的产品经理大多数是做一个产品里的一个模块。而这时候要想独立做一款产品，要么跳槽，要么内部创业。

做产品是做模块的升级版。虽说是升级，但有不同的“风景”。如果说一个功能做“坏”的影响为 1%的话，一个模块做“坏”的影响可能是 10%，而在做产品的过程中，对整体产品的判断如果做得不好，可能造成几个或一系列模块的失败，影响可能是 100% 以上。

在这个阶段，真正重要的就是那句老话“重要的不是做什么，而是不做什么”。一款 AI 产品必然有它自己的边界，不可能什么都做，此时，“做什么，不做什么”就是负责整个产品的 AI 产品经理需要考虑的问题。不能说一个决策定生死，但这时需要考虑的因素和决策对产品的影响无疑是巨大的。

能不能做好产品的整体架构，是这个阶段的 AI 产品经理的核心素质。

笔者强烈建议在一个迭代周期内尽量少做功能模块。

6. 第六个阶段：做 AI 产品的商业模式

ToB 产品是当下 AI 产品落地的主要形式，并且对提升工业制造的效率有重要的影响。

如果说之前还是在做产品的范畴之内辗转腾挪。那么，做 AI 产品的商业模式就是在公司层面上对产品的方方面面进行组织或判断。

比如：如果做内容型产品，那么你应该考虑的就不仅是一个网站或一个 App 的产品细节，可能还需要考虑如何源源不断地产生内容，如何产生高质量的内容……然后，你要考虑如何让用户发现自己想要的内容，如何引导用户消费更多内容……你可能还需要考虑用户消费完内容后怎么办，让用户直接离开肯定是不行的。你要让整个产品的方方面面有机地组合在一起，让整个系统运转起来，生生不息。例如，如果你是抖音机器学习产品的产品经理，你做的产品策略是自上而下，那么，你在设计这款 AI 产品时要考虑产品在“冷启动”时期（一方面，优秀的短视频内容生产者少；另一方面，用户因为内容少而很少点赞，点赞量少又导致头部内容生产者动力不足、成就感不强的时期）的运营策略。

如果你做的是社交型产品，你得考虑怎样吸引初始用户，用户来了做什么，用户做的这些事情是否能够吸引更多的用户，更多的用户来了又能否有更多的事情可做，

用户多了怎样不影响产品的初衷，有没有可能让用户彼此之间能够产生紧密的联系从而吸引用户不断地使用产品……

在这个阶段，AI 产品经理最重要的是给出一个自洽的循环，告诉大家怎样做才可以让整个产品运转起来，让用户不会到某个地方突然“卡壳”。

7. 第七个阶段：做定位

产品定位一般针对高阶的 AI 产品经理，这个时候的你需要协助 CEO 和总裁等公司核心团队制定产品的发展方向、产品针对的用户群体。如果你自己创业了，自己做自己的产品，那么产品定位可能是你的核心工作。

8. 第八个阶段：明星产品经理

第八个阶段之所以放在第七个阶段的后面，是因为创业产品在市场知名度和用户数量上都不如大公司的产品。例如，笔者曾经负责爱奇艺旗下的奇秀产品线，当时很有成就感，因为产品的名气大、用户量大，更主要的是学习数据分析以及视频硬解码、软解码的过程很“过瘾”。

9. 第九个阶段：做出自己的明星产品

这个阶段是笔者当下的阶段。

第3章 非技术出身的 AI 产品经理应该了解的 AI 技术

3.1 常见的 AI 技术

在本书的第1章，笔者提到了 AI 的概念、定义，本章主要对当下和未来一段时间内主流的 AI 技术进行概括性介绍和技术点解析，如果读者觉得某些技术点对自己有帮助，笔者建议读者结合自己的工作进一步搜集这方面的技术实战，并进行落地研究。

在笔者个人的方法论中，一个带有完整上下文且结构良好的知识框架在某种程度上比一些细枝末节的详细知识点要重要得多，一旦构建了一个完备的知识结构框架，那么，剩下要做的就是将一些零零碎碎的细节填补进框架而已；而反过来却很难，知识堆砌只会让你的思维混乱。

我们先看一下 AI 技术框架有哪些。

- (1) 自然语言生成。
- (2) 语音识别。
- (3) 虚拟代理。
- (4) 机器学习平台。
- (5) 针对 AI 优化的硬件和芯片：GPU 和其他设备。
- (6) 决策管理技术。
- (7) 深度学习技术。
- (8) 生物特征识别技术。

(9) 机器人技术。

(10) 文本分析和自然语言处理。

下面笔者对以上技术进行深入分析。

3.1.1 自然语言生成

自然语言生成是人工智能和计算语言学的分支，语言生成系统是基于语言信息处理的计算机模型，其工作过程与自然语言分析相反，是从抽象的概念层次开始，通过选择并执行一定的语义和语法规则来生成文本的。开发 NLG 系统主要有两个目的。

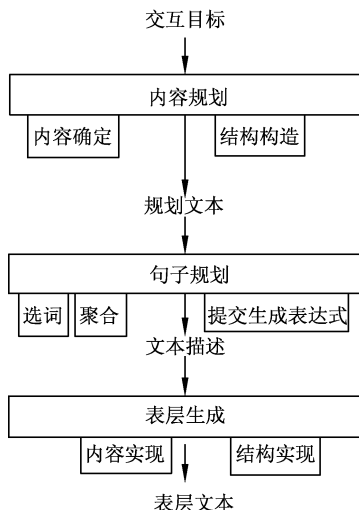
第一，作为人们生活中的交流工具。这主要是从经济角度考虑的，借助生成系统在生产速度快、错误率低、能进行多语言生成等方面的优势，可以利用语言知识和领域知识快速生成分析报告、帮助消息等。

第二，作为检验特定语言理论的一种技术手段。从这一角度来看，无论是在理论上还是在描述上，其工作过程都与研究自然语言本身有着紧密的联系，涉及语言理论等方面的内容。自然语言生成作为理论语言学与人工智能语言学共同的研究课题，近年来发展迅速。国内在内容规划、系统复用性、人机接口等方面的发展尤为迅速，但在基本理论方面的关注较少，因此，这一部分将详细介绍自然语言生成系统的体系结构及其发展现状，希望对读者有所帮助。

1. NLG 系统的体系结构

由于经典管道模型具有较好的鲁棒性、复用性和独立性，且具有广泛的运用，下面以其为基础介绍 NLG 系统的体系结构，如下图所示。

NLG 系统的体系结构包括内容规划（也称宏观规划）、句子规划（也称微观规划）和表层生成三个基本功能模块。在生成过程中，系统根据应用目标和用户模式，通过相应的语义表示、语法分析、话语结构实现来生成文本，多数 NLG 系统的体系结构随具体应用的不同而有所不同。对于输出结果而言，多数 NLG 系统并不关心输出数据的格式、显示形式等细节问题，而仅仅关心 ASCII 文本，对有特定数据格式和多语言要求的输出而言，此部分工作将交由文档表示系统（也称后处理器）完成。



NLG 系统的体系结构

(1) 内容规划。

在 NLG 系统的体系结构中，内容规划器作为关键部件，主要完成内容确定和结构构造两项任务。顾名思义，内容确定主要解决生成文本表达什么内容的问题，它非常依赖具体的应用，需要大量领域知识的支持；结构构造描述已确定内容的结构，即用一定的结构将所要表达的内容组织起来，使其更加符合人类阅读和理解的习惯。从数据处理的流程来看，内容规划器的结构模型一般有两种设计方法：串行设计方法。内容确定组件首先标记确定的内容消息，然后结构构造器在此基础上再构造结构树。

并行设计方法，即同时对内容和结构进行规划，两者同步进行。串行设计方法以数据模型为驱动，将内容确定和结构构造两个组件分离，以保证修改数据时二者相互独立。采用并行设计方法时，由于两个组件同步工作，得到的数据必然相互影响。

内容确定的域建模和消息定义。从存在论的角度分析，内容确定的关键是域建模，文本作为信息的载体，信息是实体属性的说明及实体间相互关系的描述，故讨论一个问题域要从域中实体的概念、类型、属性和相互关系四个方面来进行分析和判断。域建模主要包括两项工作：分析大量应用域的范文集，从中抽取实体集，构造文档框架原型。文档框架应覆盖所需表达的全部内容，尤其是边界内容及非正常实例。规划文档框架的内容成分，即按固定成分、输入成分和推导成分来确定内容。固定成分对应多数应用都包含的固定内容；输入成分来自特定应用的实际输入数据或知识库的

有关知识，与具体应用密切相关；推导成分对应对输入数据和历史数据进行概括推理之后得到的数据，为具体应用服务。

域建模和消息形式化表示的核心是它能表示语言。在实际应用中，句子或子句表达的基本概念用消息表示，消息的逻辑形式定义有抽象语义结构、词格框架表示、罐装文本等多种方法，相应的粒度取决于输出文本的类型，没有必要要求所有消息的粒度都相同。通常，消息的形式定义主要从专家处或从成功案例中获得，是“修改—请教—更正”的反复过程，包括以下五个步骤。

① 选择对应于目标文本的范文集。

在对范文集进行分句的基础上进行短语划分，经整理后使短语对应消息。

对消息进行聚合。

确定消息的逻辑形式。

同专家进行交流，反复修改更正。

内容确定的规则和依赖因素。NLG系统的目的是使生成的文本仿佛出自专家之手，故通过与专家一起对范文集进行分析、讨论和校验来形成规则，然后选择一定的规则来确定需要构造的内容。规则的获得过程如下。

对范文进行分析，求出消息全集，并根据需求反复修订消息定义。

使消息关联输入数据，形成面向具体应用的消息集。通常消息关联数据可采用直接选择、概括和推理等方法，所获信息分别对应文档内容的三种成分。

文本可能有很多种类型，每个文本类型可能又使用多种消息。

系统通过调用不同的规则来确定内容，以上内容确定的过程也是求得规则的过程，一旦规则得以确定，相应的内容也就确定了。

同专家进行交流，扩大范文集，重复上述过程。显然，NLG系统的内容确定依赖于交互目标、用户模式、输出约束以及知识库，不同的需求、不同的用户知识背景以及系统知识库的完备与否等都将直接影响内容确定。

内容规划器的输入与输出。对NLG系统而言，系统输入即内容规划器的输入，在多数应用系统中，系统输入通常被抽象描述为四元组 (k, c, u, d) 。其中， k 表示拥有领域知识的一个或多个数据库和知识库，为生成文本提供知识或数据支持； c 表示交互目标，不同目标对应不同的内容和结构，系统能根据不同目标进行响应； u 表

示用户模式,不同的用户对生成文本有不同的要求,通过指定用户类型、适应用户变化,使生成的文本更具个性;d表示历史描述,用于记录用户与系统交互的历史信息。内容规划的处理结果是规划文本,描述文档的基本内容和结构,其典型结构是一棵结构树,树的叶子节点对应用消息表示相应文档的具体内容,内部子节点相应文档的结构,可直接映射到文档的段落、章节等。

结构构造。文档不是句子的随机组合,只有经过一定结构的表示才能让用户更好地阅读和理解全文。对 NLG 系统而言,如果输出语义由单句子表达(如记录类型输出),就不需要考虑比句子更复杂的结构。除此以外,如果要考虑多行文本的输出,就必须考虑文档结构,具体地说,文本能被分析成树状结构。结构构造的主要任务就是构造结构树,用树中叶子节点表达具体消息。内部节点描述消息如何按照相应主题聚合、消息顺序、消息组之间的关系、哪些消息对应段落或章节等。有时,结构构造部件仅能指定以上部分信息,其余部分交给句子规划器处理,具体构造方法随不同生成技术而有所不同,通常包括自上而下和由下而上两种构造方法。

内容规划技术。在诸多实际应用中,尽管内容规划、句子规划和最后的文本实现有许多差别,但是使用的技术通常是差不多的,开发人员可以根据生成文本的表达能力的强弱来选择不同复杂程度的生成技术。当然,同一生成技术用于不同任务时,它使用的数据结构和实现方法也有所不同,有许多应用系统在不同的生成阶段采用不同的生成技术。一般有四种文本自动生成技术,分别是模板生成技术、模式生成技术、短语/规划扩展技术、属性特征生成技术。

下面主要介绍模式生成技术和短语/规划扩展技术。

模式生成技术(又称 Schema 技术)。模式生成技术是基于语言学中的谓词(Predicate)来表达文本结构的一种方法,它采用谓词来描述文本结构的规律,是文本的骨架表示,在该表示中也明确了文本中主题的表达顺序。文本都是由命题组成的,命题是指一个句子或者一个从句,谓词对文本中的命题进行分类,每个命题都被归纳为特定的谓词。同类型文本存在着一些标准的谓词组合模式来表示文本的结构。

相应结构树中的节点通常分为五种类型:Root、Schema、Predicate、Argument 和 Modifier。其中,Root 是树的根节点,表示一篇文章。每个 Root 下面有若干个 Schema 子节点,一个 Schema 表示一个段落或句群,Schema 下面的子节点可以继续是 Schema,也可以是 Predicate。一棵以 Predicate 为根的子树表示一个句子,它是文

章的基本单位,句子中每一个基本语义成分均是 Predicate 的一个子节点,用 Argument 表示。若 Argument 有修饰成分,则用子节点 Modifier 标志。Argument 或 Modifier 是树的叶子节点,树中每个节点都含有若干个槽,用来标记各种信息以供生成使用。

下面以“中央气象台在 2017 年 10 月 10 日 17 点发布的上海市天气预报”为例来说明内容规划的过程。我们为每个节点定义了六个槽,分别表示节点类型、具体数据、语义特征、语义角色、内部标志以及节点个数,如下图所示。

```
{Schema,HeadIllustration,Paragraph,Empty,Empty}<1>, //此段只有一句构成
({Predicate,发布,HeadIllustration(Main),ReportHead,Empty}<4>, //此句共四个主要成分
 {Argument,2017. 10. 10,Date,Tim,Empty}<0>),
 ({Argument,点钟,Time,Tim,Empty}<1>,
  ({Modifier,17,Number,Num,Field}<0>,)), //表示修饰
 ({Argument,天气预报,ReportType,Obj,Empty}<1>,
  ({Modifier,上海市,PlaceName,Mod,Field}<0>,)), //表示修饰
 ({Argument,气象台,StationName,Age,Empty}<1>,
  ({Modifier,中央,Location,Mod,Field}<0>)) //表示修饰
```

案例语句中的节点及槽

Predicate 的最大特点就是表达简洁,不受文本篇幅的限制,用 Schema 表示的文本结构具有层次,即 Schema 可以递归嵌套。

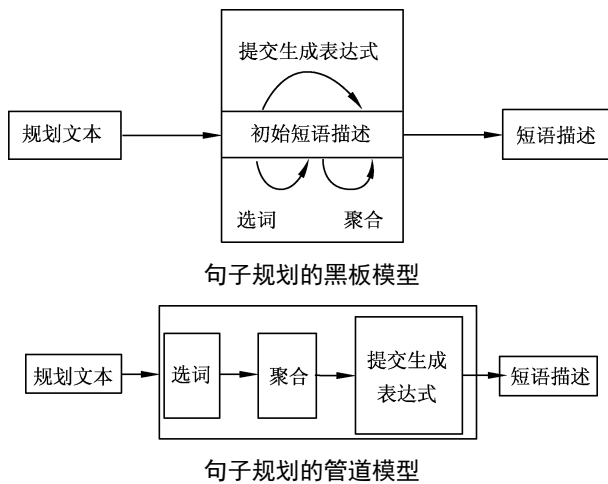
短语/规划扩展技术(又称 RST 生成技术)。RST 生成技术是基于描述文本结构的修辞结构理论。RST 理论认为文章的各部分(无论是子句还是更大的组成单位)都是由一些为数不多、反复出现的关系按照一定的层次聚在一起的。多数 NLG 生成系统包含一个修辞关系集,而与具体应用对应的关系集是其子集。RST 关系有 Nucleus-Satellite 和 Multi-Nucleus 两种模式。其中,Nucleus-Satellite 模式包括核心部分(Nucleus)和附属部分(Satellite),核心部分表达基本命题,附属部分表达一个附属命题,多用于描述目的、因果、转折、背景等关系;Multi-Nucleus 模式涉及一个或多个语段,它没有附属部分,多用于描述顺序、并列等关系。下面的例子简单说明了修辞结构树的节点结构。

(Node<Father> <Sons> <Rhet-Role> <Rhet-Relation>)

其中,<Father>代表父节点,<Sons>代表子节点,均对应于句子或子句。<Rhet-Role>是<Father>中与直接父节点具有修辞关系的节点,起修辞作用;<Rhet-Relation>是<Sons>节点间的修辞关系。

(2) 句子规划。

通常,内容规划并没有完全指定输出文本的内容和结构,句子规划的任务就是进一步明确定义规划文本的细节,具体包括选词(Lexicalization)、聚合(Aggregation)、提交生成表达式(Referring Expression Generation)等工作。句子规划的输出是文本描述(TextSpecification),但其仍然不是最终输出文本,仍有句法、词法等特征需要进一步处理。一般文本描述的层次结构仍然对应逻辑结构,须经过文本表示系统实现从逻辑结构向物理结构(段落章节)的映射,才能最终生成文本。如下面两张图所示。



上面两张图分别给出了句子规划的黑板模型和管道模型,其中管道模型运用较多,也易于实现。

选词 (Lexicalization): 在应用中,特定信息必须根据上下文环境、交互目标和实际因素(如对人的知识背景做强调等特定处理)用词或短语来表示。选择特定的词、语法结构以表示规划文本的信息意味着对规划文本进行消息映射,有时只有一种选词方法用来表示信息或信息片段,在多数系统中允许有多种选词方法。如 Lexicalize (Spell) 是递归过程,其中 Spell 是时间间隔参数,算法必须对 Spell 进行分析判断,根据 Spell 的长度来调用不同的类(词)以描述 Spell。

聚合 (Aggregation): 为了能够消除句子间的冗余信息,增加可读性以及能从子句构造更复杂的句子,在句子规划中应用了聚合技术,通常按照粒度区分句子、词汇、语义、修辞和概念聚合等。从语言学角度来看,最好理解的聚合类型是句子聚合。

实现句子聚合有四种方法。

- 简单连接。直接利用连接词对多个句子（或子句）进行组合。连接词的选择将依赖信息间的修辞关系，因简单连接没有改变内部成分的语法和词汇等内容，故其非常适用于在离表层生成较近的层次运用。
- 经共享成分连接。如主语相同、宾语相同或行为相同，可以保留相同部分且仅出现一次，而将其他成分进行聚合。是不是具有相同成分就可以聚合呢？不是。聚合操作必须了解语义，仅仅在字面上相同是不够的。
- 经共享结构连接。子句具有相同的结构，且仅个别成分不同，此时可以通过插入连接词连接不相同成分进行聚合。
- 句法修饰。这是最复杂的聚合形式，主要是对语义进行概念上的有界聚合和无界聚合。聚合是非常有目的地缩短文本的机制，因此语义简明和朴素易懂是选择聚合机制的基本要求。在许多应用中，聚合使句子语义更加复杂，但实际上有时并不需要，所以聚合应针对系统的具体目标，随应用而定。

提交生成表达式：域是实体的集合，而实体是具体或抽象的讨论对象，用名词短语表示。在实现选词和聚合之后，提交生成表达式的工作就是让句子的表达更具语言色彩，对已经描述的对象进行指代以增加文本的可读性。重复性称谓是利用代词、固定名词、完整或缩略名词短语来指代实体。名词短语通常包括确指名词短语和非确指名词短语，固定名词和代词都是确指名词短语的子类。通常，非确指名词短语用来描述首次提及的实体，而随后提及的实体用确指名字短语表示。在生成系统实现指代的处理中，一方面应注意避免歧义和指代不当，过度“咬文嚼字”往往不实用；另一方面应尽量避免信息冗余。

（3）表层生成。

表层生成主要涉及结构实现和语言实现两部分。具体来讲就是将经句子规划后的文本描述映射至由文字、标点符号和结构注解信息组成的表层文本。生成算法首先按主谓宾的形式进行语法分析，并决定动词的时态和形态，再完成遍历、输出。其中，结构实现将结构注解信息映射至文本实际段落、章节等结构；语言实现将短语描述映射到实际表层的句子或句子片段。一般而言，NLG系统并不关心输出数据的格式等细节，而直接由文档处理系统将逻辑结构映射至表示设备。换句话说，我们能利用文档

处理系统，如 Word 或其他网页浏览器作为后处理器来生成文档。

2. 自然语言生成的发展与现状

自然语言生成已经发展了近六十年，在这段时间里，从事该领域研究的专家不断提出新的理论和方法，设计出新的生成模型，使语言生成的研究不断取得新的进展。语言生成的研究侧重于以下几个方面。

在特定的语法理论框架内更加广泛、更加深入地处理语言现象，如 Fawcett 的 GeneSys 生成系统。

在同一语法环境下生成多种语言，例如英国斯特林大学的 Nigel 多语言生成系统；上海交通大学的多语言天气预报发布系统。

面向实际应用的开发，如颐和园的导游系统；中国科技大学的足球现场解说系统。

在生成过程中对所要表达的信息进行语义和句法方面的聚合也是研究重点之一。语言生成的研究方向主要在语言表示形式、信息内容规划以及语言生成模型等方面。

3.1.2 语音识别

语音识别是指将人类语言转录和转换成对计算机应用软件来说有用的内容。

“与机器进行语音交流，让机器明白你说的是什么”是人们长期以来梦寐以求的事情。中国 IoT 校企联盟把语音识别形象地比作“机器的听觉系统”。语音识别技术就是让机器通过识别和理解过程把语音信号转变为相应的文本或命令的 AI 技术。

语音识别技术主要包括特征提取技术、模式匹配准则及模型训练技术三个方面。语音识别技术在车联网领域得到了充分的应用，例如：在奔驰的车联网应用中，只需按一个键即可直通客服人员，车主通过口述即可设置目的地，交互变得安全、便捷。

语言是人类相互交流最常用、最有效、最重要和最方便的通信形式，语音是语言的声学表现。随着 AI 技术整体的飞速发展，语音识别技术也取得了突破性的成就，人与机器用自然语言进行对话的梦想逐步实现。语音识别技术的应用范围极为广泛，不仅涉及日常生活的方方面面，在军事领域也发挥着极其重要的作用。它是信息社会朝着智能化和自动化方向发展的关键技术，它使人们处理和获取信息变得更加便捷，从而提高了人们的工作效率。

(1) 语音识别技术的发展历史。

语音识别技术始于 20 世纪 50 年代。在这一时期，语音识别技术的研究范围主要集中在对元音、辅音、数字以及孤立词的识别上。

在 20 世纪 60 年代，语音识别研究取得实质性进展。线性预测分析和动态规划的提出较好地解决了语音信号模型的产生和语音信号不等长两个问题，并通过语音信号的线性预测编码，有效解决了语音信号的特征提取等问题。

在 20 世纪 70 年代，语音识别技术取得突破性进展。基于模式匹配的动态时间规整 (Dynamic Time Warping, 后文简称 DTW) 技术基本成熟，出现了矢量量化 (Vector Quantization, 后文简称 VQ) 理论和隐马尔可夫模型 (Hidden Markov Model, 后文简称 HMM)。

在 20 世纪 80 年代，语音识别任务开始从对孤立词、连接词的识别转向对大词汇量、非特定人、连续语音的识别，识别算法也从传统的基于标准模板匹配的方法转向基于统计模型的方法。在声学模型方面，由于 HMM 模型能够很好地描述语音的时变性和平稳性，开始被广泛应用于大词汇量连续语音识别 (Large Vocabulary Continuous Speech Recognition, 后文简称 LVCSR) 声学建模。在语言模型方面，以 N 元文法为代表的统计语言模型开始广泛应用于语音识别系统。在这一阶段，基于 HMM/VQ、HMM/高斯混合模型、HMM/人工神经网络的语音建模方法开始广泛应用于 LVCSR 系统，语音识别技术取得新突破。

20 世纪 90 年代以后，伴随着语音识别系统走向实用化，语音识别在细化模型设计、提取和优化参数、系统的自适应方面取得较大进展。同时，人们更多地关注话者自适应、听觉模型、快速搜索识别算法以及新的语言模型研究等课题。此外，语音识别技术开始与其他领域的相关技术相结合，以提高识别的准确率，实现语音识别技术的产品化。

近年来，科大讯飞的语音识别系统在语音平台、通信增值、嵌入式、教育、行业软件和系统集成等方向取得突破。

(2) 语音识别的基础概念。

语音识别是将人类的声音信号转化为文字或者指令的过程。语音识别以语音为研究对象，它是语音信号处理的一个重要研究方向，是模式识别的一个分支。

语音识别涉及计算机、人工智能、数字信号处理、模式识别、声学、语言学和认

知科学等许多学科，是一个多学科综合性的研究领域。

根据不同限制条件下的研究任务，语音识别产生了不同的研究领域。这些领域包括：根据说话人的说话方式，可分为孤立字（词）、连接词和连续语音识别系统；根据对说话人的依赖程度，可分为特定人和非特定人语音识别系统；根据词汇量的大小，可分为小词汇量、中等词汇量、大词汇量以及无限词汇量语音识别系统。

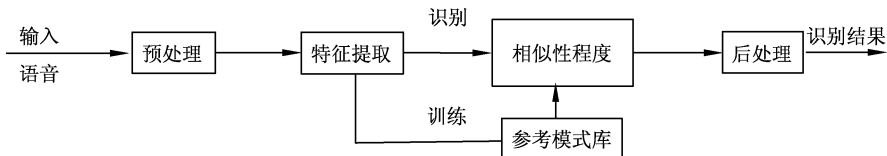
（3）语音识别的基本原理。

从语音识别模型的角度来讲，主流的语音识别理论是建立在统计模式识别基础之上的。语音识别的目标是利用语音学与语言学的相关知识，把输入的语音特征矢量序列 $X = X_1, X_2, \dots, X_T$ 转化成词序列 $W = X_1, X_2, \dots, W_N$ 并输出。基于最大后验概率的语音识别模型如下式所示。

$$\begin{aligned}\hat{W} &= \arg \max_W \{P(W|X)\} \\ &= \arg \max_W \{P(X|W)P(W)/P(X)\} \\ &= \arg \max_W \{P(X|W)P(W)\} \\ &= \arg \max_W \{\log P(X|W) + \lambda \log P(W)\}\end{aligned}$$

上式表明，要寻找最可能的词序列 W ，应该使 $P(X|W)$ 与 $P(W)$ 的乘积达到最大。其中， $P(X|W)$ 是特征矢量序列 X 在给定 W 条件下的条件概率，由声学模型决定。 $P(W)$ 是 W 独立于语音特征矢量的先验概率，由语言模型决定。由于将概率取对数不影响 W 的选取，第四个等式成立。 $\log P(X|W)$ 与 $\log P(W)$ 分别表示声学得分与语言学得分，且分别通过声学模型与语言模型计算得到。 λ 是平衡声学模型与语言模型的权重。从语音识别系统构成的角度讲，一个完整的语音识别系统包括特征提取、声学模型、语言模型、搜索算法等模块。语音识别系统本质上是一种多维模式识别系统，对于不同的语音识别系统，人们所采用的具体识别方法及技术不同，但其基本原理是相同的，即将采集到的语音信号送到特征提取模块进行处理，将得到的语音特征参数送入模式库模块进行处理，由声音模式匹配模块根据模式库对该段语音进行识别，最后得出识别结果。

语音识别系统的基本原理如下图所示。



语音识别系统的基本原理

预处理模块去除原始语音信号中的次要信息及背景噪声等，包括抗混叠滤波、预加重、模/数转换、自动增益控制等处理过程，将语音信号数字化。特征提取模块对语音的声学参数进行分析后提取出语音特征参数，形成特征矢量序列。语音识别系统常用的特征参数有短时平均幅度、短时平均能量、线性预测编码系数、短时频谱等。特征提取和特征选择是构建系统的关键，对改善识别效果极为重要。

由于语音信号本质上属于非平稳信号，对语音信号的分析是建立在短时平稳性假设之上的。在对语音信号作短时平稳假设以后，通过对语音信号进行加强，实现短时语音片段上的特征提取。这些短时片段被称为帧，以帧为单位的特征矢量序列构成语音识别系统的输入。由于梅尔倒谱系数及感知线性预测系数能够从人耳听觉特性的角度准确刻画语音信号，其已经成为主流的语音特征。人们在使用梅尔倒谱系数及感知线性预测系数时，通常加上它们的一阶、二阶差分，以引入信号特征的动态特征。声学模型是语音识别系统中最为重要的部分之一，声学建模涉及建模单元选取、模型状态聚类、模型参数估计等很多方面。在 LVCSR 系统中，普遍采用上下文相关的模型作为基本建模单元，以刻画连续语音的协同发音现象。在考虑了语境的影响后，声学模型的数量急剧增加，LVCSR 系统通常采用状态聚类的方法压缩声学参数的数量，以简化模型的训练。在训练过程中，系统对若干次训练语音进行预处理，并通过特征提取得到特征矢量序列，然后由特征建模模块建立训练语音的参考模式库。

搜索是在指定的空间中，按照一定的优化准则，寻找最优词序列的过程。搜索的本质是问题求解，其广泛应用于语音识别、机器翻译等人工智能和模式识别的各个领域。它通过利用已掌握的知识（声学知识、语音学知识、词典知识、语言模型知识等），在状态空间中找到最优的状态序列。最终的词序列是对输入的语音信号在一定准则下的一个最优描述。在识别阶段，将输入语音的特征矢量参数同训练得到的参考模式库中的模式进行相似度比较，将相似度最高的模式所属的类别作为识别中间的候选结果输出。为了提高识别的正确率，在后处理模块中对得到的候选识别结果继续进行处理，包括通过 Lattice 重打分以融合更高元的语言模型，通过置信度量得到识别结果的可靠程度等。最终通过增加约束，得到更可靠的识别结果。

（4）声学建模方法。

常用的声学建模方法包含以下三种：基于模式匹配的动态时间规整法（DTW 法）；

隐马尔可夫模型法 (HMM 法); 基于人工神经网络的识别法 (ANN 法)。

DTW 法是较早的一种模式匹配的方法。它基于动态规划的思想, 解决孤立词语音识别中语音信号特征参数序列长度不一致的模式匹配问题。在实际应用中, DTW 法通过计算已预处理、分帧的语音信号与参考模式之间的相似度, 再按照某种距离测度计算出模式间的相似度并选择最佳路径。

HMM 法是针对语音信号的时间序列结构建立的统计模型, 是在马尔可夫链的基础上发展起来的一种统计识别方法。HMM 法可模仿人类的言语过程, 可视作一个双重随机过程: 一个是用具有有限状态数的马尔可夫链来模拟语音信号统计特性变化的随机过程, 另一个是观测与马尔可夫链的每一个状态相关联的序列的随机过程。

ANN 法以数学模型模拟神经元活动, 将人工神经网络中大量神经元并行分布运算的原理、高效的学习算法以及对人的认知系统的模仿能力充分运用到语音识别领域, 并结合神经网络和隐马尔可夫模型的识别算法, 克服了 ANN 在描述语音信号时间动态特性方面的缺点, 进一步提高了语音识别系统的鲁棒性和准确率。其中成功的方法就是在混合模型中用 ANN 替代高斯混合模型估计音素或状态的后验概率。2017 年, 微软 Cortana 以深度神经网络代替多层感知机形成的混合模型系统大大提高了语音识别的准确率。

(5) 语音识别的应用范围。

语音输入使得人们可以甩掉键盘, 通过识别语音中的要求、请求、命令或询问来做出正确的响应, 这样既可以克服键盘输入速度慢、极易出差错的缺点, 又有利于缩短系统的反应时间, 使人机交流变得简便易行, 比如声控语音拨号系统、声控智能玩具、声控智能家电等。在智能对话查询系统中, 人们通过语音命令, 可以从远端的数据库系统查询与提取有关信息, 享受自然、友好的数据库检索服务。语音识别技术还可以应用于自动口语翻译, 即通过将口语识别技术、机器翻译技术、语音合成技术相结合, 将一种语言的语音输入翻译为另一种语言的语音输出, 实现跨语言交流。

语音识别技术在军事领域里也有着极为重要的应用价值和极其广阔的应用空间。一些语音识别技术就是为军事活动而研发, 并在军事领域率先应用、首获成效的, 军事应用对语音识别系统的识别精度、响应时间、恶劣环境下的鲁棒性都提出了更高的要求。语音识别技术已经在军事指挥和控制自动化方面得以应用。比如, 将语

音识别技术应用于战斗机的飞行控制，可快速提高作战效率、减轻飞行员的工作负担。飞行员利用语音输入代替传统的手动操作，控制各种开关和设备，把时间和精力集中于对攻击目标的判断和完成其他操作上，可更快获得信息来发挥战术优势。

小结：语音识别是一门交叉学科。近年来，语音识别技术取得显著进步，早已经开始从实验室走向市场。语音识别技术已经进入工业、家电、通信、汽车电子、医疗、家庭服务、消费电子产品等领域。语音识别的研究工作对于信息化社会的发展，提高人民生活水平等有着深远的影响。随着计算机信息技术的不断发展，语音识别技术将取得更多重大突破，语音识别系统的研究将会更加深入，语音识别有着更加广阔的发展空间。

3.1.3 虚拟代理

虚拟代理主要包括智能客服及智能家居管理器。

在客户关系管理里，虚拟代理（Virtual Agent，有时候也被称作智能代理）是自动答话机器人程序，它如同在线客户服务代表那样为组织服务。由于虚拟代理具有类似人类的外表且能正确回答用户的问题，因此适于作为向个人提供服务的自动交互工具。结合图形界面之后，虚拟代理还可以帮助人们完成信息分配、下达命令和预订等工作。使用虚拟代理完成客户应答有着相当好的效果，比如，客户和虚拟代理谈话的时间要比和人工客服谈话的时间长，因为客户会对自动应答感到新奇。虚拟代理通常是用已经设定好的内容来回答不同的问题。如果客户对预定的回答感到不满并且变得粗鲁时，虚拟代理就会像人一样出现“不耐烦”。比如，当客户说虚拟代理很蠢的时候，一个可能的应答是：“我是死的。你在侮辱一个死的东西。我们俩谁更像字典里定义的傻呢？”根据某虚拟代理提供商的看法，自动应答能够有效减少客服成本，鼓励客户自我服务，使得客户更忠实，并可以作为企业的一个品牌营销工具。例如，中国电信公司的一名人工客服一个月的人力成本超过1万元，而虚拟代理只需要2000元就可以完成一个座席的工作。

在广义范畴上，虚拟代理是智能答话机器人的具体应用，属于智能语音技术在客服系统中的应用。

客服系统是企业与客户之间重要的沟通渠道，双方通过信息交互完成服务，从客服行业的发展现状看，占比最大的信息交互方式是语音交互。由于语音信息具有

非结构化的特点,因此,将语音信息转换为结构化信息需要人来完成,人的低效率造成了大量信息不能及时得到处理、分析,因此不能有效提高服务水平。要解决这一问题,需要对语音信息进行自动化、智能化处理,将原来主要由人完成的工作交给系统自动完成,人只需要进行辅助处理。这样既能减少人力资源的投入,又能提高客服系统的处理效率和智能化程度,改善客户体验。客服系统存在的问题随着业务不断发展而不断变大,大的客服系统(如携程的客服系统,飞猪的客服系统以及饿了么的客服系统)的座席客服的数量越来越大,而客户对服务的质量要求也越来越高,传统客服系统在语音信息处理上的局限性制约了系统服务能力的提升。

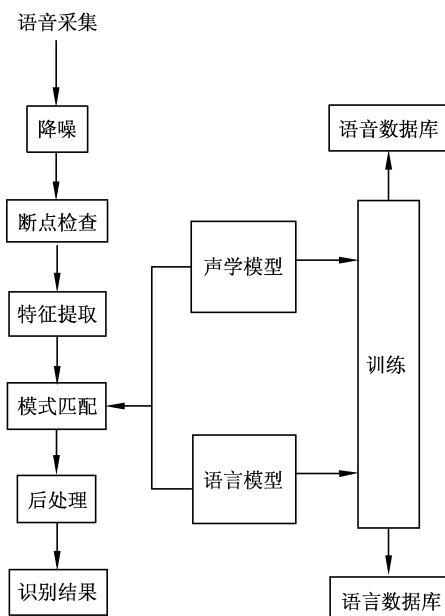
我们先看服务方面的问题。服务方面存在的问题并非是语音信息处理的局限性导致的,而是由于业务的多样性、复杂性带来的海量信息不能得到快速、有效处理导致的。随着移动通信技术的发展,业务的不断增加,如4G、5G等新业务的上线,客服系统的数据量越来越大。加上原有的传统业务,知识库的内容急剧增加,且更新频繁,客服人员与客户对话时需要从庞大的知识库体系中快速搜索、定位知识点。此外,接口众多导致客服人员在服务过程中需要查询包括业务运营支撑系统(BOOS)客户关系管理系统(CRM)在内的多个系统的数据内容才能全面回答客户的问题,多系统、多接口的应用模式严重影响了工作效率。此外,传统客服系统的服务过度依赖人工处理,由于人员流动、培训效果不一致等问题,传统的人工服务带来了服务不一致、客户体验和情绪感知不足等问题,降低了客户的满意度。

除了服务方面,运营管理方面的问题主要是语音信息处理的局限性带来的直接问题。数据单薄,单一客服系统运营分析的主要数据来源是座席代表所标注的来电原因和投诉内容等信息,信息较为有限,不够全面,且受到客服人员素质的制约。数据资源浪费,对话录音文件中包含着客户需求、竞争对手情报等大量信息,但录音文件是非结构化数据,无法对其进行查询、统计和分析,这造成了数据资源的浪费。质量管控的及时性差,传统的座席质量管控模式为事后质检、事后分析,只有等出现了问题才去解决问题,无法在座席服务过程中管控通话质量并预警问题。

虚拟代理技术的基础是智能语音技术,主要包括语音合成技术(TTS)和语音识别技术(ASR)。语音合成技术发展较早,已较为成熟,这里不做具体介绍。语音识别技术是让系统或者机器能够理解人类说的话,其根本目的是创建一种具有听觉功能

的系统或机器，使系统或机器能直接理解人的意图并做出相应的反应，其基本原理是使用含有语音识别技术的智能体接收声音信号，将信号转换成文字，然后根据需要做记录，并进行相应的操作。

语音识别系统本质上是多维模式识别系统，它的识别过程与人对语言的处理过程基本一致，语音识别一般分为两个过程：系统的识别过程和系统的训练过程。在语音识别系统的训练阶段建立对应的模式库，库中每个词都对应一个参考模式，它由这个词重复发音多遍，再经特征提取和某种训练得到。识别阶段是整个系统的核心，它的作用是按照一定的规则求得待测语音的特征参数及语音信息，并与模式库中的模型进行匹配，匹配度最高的结果即为识别结果，语音识别系统的结构如下图所示。



语音识别系统的结构

笔者在后面的章节中会用实际案例分析虚拟代理的使用场景，这里先简单介绍一下智能语音技术的需求场景。要想突破传统语音处理的局限，就要实现系统对语音信息的自动化处理，通过建立具备智能语音分析能力的语音识别平台，实现对客户语音的实时识别及关联处理，同时实现对语音文件的统一分析，提供支撑业务运营管理的有效工具，语音分析引擎对送入的语音进行文字转写、维度结合后生成索引文件，这些结构化的索引文件包含了语音的所有信息，索引文件存储在数据仓库中，通过对海量录音文件、音

频文件进行分析、处理，对语音中的知识进行挖掘，对语音文件进行快速检索。

（1）通话语音实时识别。

通话过程中的语音识别包括通话过程中的关键字识别及基于关键词的整句及全文识别，对于自动语音业务，语音实时识别能够改变主要由 IVR（互动式语音应答）语音导航菜单带动的长时段、低效率的业务模式，通过实时语音识别能快速将客户直接引导至需要或感兴趣的业务。对于人工处理业务，语音实时识别系统通过给每个座席配备一个语音助手，能够在通话过程中对客户信息进行自动查询，对知识进行高效检索，对易错点进行提醒，从而提升座席服务效率，实现来电数据实时分析，同时能为其他分析系统提供数据。借助通话语音实时识别功能，可以在现有事后质检的基础上，增加事中质检的功能，对客服人员实时通话中的服务质量进行实时监控并进行预警。

（2）智能语音质检。

建立了智能化质检系统后，可以在对录音进行全文分析的基础上，将服务禁用语、热点词、空白音、抢话等作为质检索引条件送入客服质检系统，同时为客服质检系统提供看听结合的录音听取页面，扩大质检的覆盖范围，提升质检效率。截至 2018 年，客服中心的质检方式仍较为传统，由管理人员随机抽调录音进行检查。检查过程为播放录音、听取通话内容，这种方式效率低下，且检查范围有限。智能化的质检系统将彻底改变这种低效、片面的工作模式，全面提高管理能力，同时减少人力投入。

录音文件挖掘分析能力将语音转换为结构化文字以后，可以基于业务进行分析，直接了解客户的兴趣点、咨询热点，深入挖掘产品存在的问题，控制风险，快速推动产品和营销方案优化，从根源上提升客户满意度。

一般来说，客户与客服座席的每一次交互都能反映出其当前的消费情况和以后的消费行为趋势，客户对产品的喜好程度、关注焦点、心理预期等都会通过通话向企业反馈，可针对“客户来电原因”“希望解决的问题”“关注的焦点”等进行更深层的数据挖掘，同时将这些客户行为特征与客户年龄段、ARPU（平均每个用户每月贡献的业务收入）值、套餐类型、业务偏好、市场活动等信息综合起来进行运营分析。

通过分析客户行为，预测客户的行为趋势，可以改变被动响应突发的话务高峰和话务低谷的情况，提高话务分流的超前性及准确性，通过更合理的人工分配来节省人

力成本,可以对市场推广的效果、新产品推出及资费变更后的客户反馈进行即时分析,快速、准确地获得客户建议,了解相应策略的效果,改变运营思路和方式。

通过应用智能语音分析技术将录音转化为结构化的索引信息,可以对客户行为数据进行数据挖掘,并将得到的结果应用于市场营销、客户感知、品牌推广、提升首次呼叫解决率等方面。对提升服务水平、提高客户满意度、提升运营管理水平、推进客服中心市场化等都有着重要的意义。

智能语音技术发展迅速,苹果的 Siri、谷歌的 Assistant、微软的 Cortana、小米的小爱同学等产品各有千秋。

通过多年的技术改进,语音识别的准确率有了较大提升,用户通过语言能完成更多操作。这些语音助手均采用在本地进行语音采集、在云端进行语音识别、在本地执行语音指令的解决方案,客服系统需要综合使用实时语音助手及离线语音分析、关联分析等功能,其核心是实现实时语音助手的功能。在实现语音识别的基础上,其他功能(如离线语音分析、关联分析等功能)只是基础功能的扩展,根据需求进行定制开发即可。

截至本书完稿时(2018年),国内的科大讯飞已经能提供智能语音服务,包括智能客服服务和语音分析服务。智能客服服务基于科大讯飞智能语音技术,提供语音合成、语音识别、声纹识别及自然语言理解能力,可以与客户进行基于自然语音或文字的交互,智能引导客户并响应客户需求。语音分析服务能对海量的录音数据进行结构化处理,形成可查询的数据,辅助企业了解客户心声,挖掘客户潜在需求,同时诊断现有管理体系,提升服务水平。

小结:智能代理能够帮助买卖双方协商。所谓的代理、智能代理、软件代理或软件机器人实际上都是指同一种程序,它可以代表个人或组织完成信息收集、过滤和调整等工作。

3.1.4 机器学习平台

机器学习平台不仅提供模型,具备将模型部署到应用软件、流程及其他机器的计算能力,还提供算法、接口、开发工具包和训练工具包。

学习是人类具有的一种重要能力,但究竟什么是机器学习,长期以来却众说纷纭。社会学家、逻辑学家和心理学家都有自己的看法。

Langley 定义的机器学习是：“机器学习是一门人工智能的科学，该领域的主要研究对象是人工智能，特别是如何在经验中学习改善具体算法的性能。”(Machine learning is a science of the artificial. The field's main objects of study are artifacts, specifically algorithms that improve their performance with experience.)

Tom Mitchell 定义的机器学习是：“机器学习是对通过经验自动改进的计算机算法的研究。”(Machine learning is the study of computer algorithms that improve automatically through experience.)

Alpaydin 也提出了自己对机器学习的定义：“机器学习是用数据或以往的经验，优化计算机程序的性能。”(Machine learning is programming computers to optimize a performance criterion using example data or past experience.)

为了便于讨论和估计学科的进展，笔者认为有必要对机器学习给出自己的定义，即使这种定义是不完全的、不充分的。笔者认为：机器学习是研究如何使用机器来模拟人类学习活动的一门学科。稍为严格的提法是：机器学习是一门研究机器获取新知识和新技能，并识别现有知识的学问。这里所说的“机器”，指的就是计算机，包括电子计算机、中子计算机、光子计算机、神经计算机、量子计算机等。

机器能否像人类一样具有学习能力？1959年，美国的塞缪尔设计了一个下棋程序，这个程序具有学习能力，它可以在对弈中提升自己的棋艺。4年后，这个程序战胜了设计者本人。又过了3年，这个程序战胜了美国一位连续8年不曾失败的冠军。这个程序向人们展示了机器学习的能力，提出了许多令人深思的社会问题与哲学问题。

机器的能力能否超过人类？很多人对此持否定态度，他们的一个主要论据是：机器是人造的，其性能和动作完全是由设计者规定的，因此，无论如何，其能力也不会超过设计者本人。这种意见对不具备学习能力的机器来说是对的，但是对具备学习能力的机器来说就不一定了，对于具备学习能力的机器，其能力在应用中会不断提高。

机器学习已经有了十分广泛的应用，如数据挖掘、计算机视觉、自然语言处理、生物特征识别、搜索引擎优化、医学诊断、检测信用卡欺诈、证券市场分析、DNA序列测定、手写文字识别等。

1. 大数据背景下的机器学习

我们知道，机器学习是在海量、复杂多样、变化快的大数据环境下利用机器算力

学习设计好的算法。

随着数据量的爆炸式增长，数据以前所未有的速度积累，大数据概念受到越来越多的关注。大数据正在给数据密集型企业带来丰厚的利润。据统计，谷歌的品牌价值在 2017 年以 1094 亿美元排名北美第一。学术界和产业界关于大数据的认识也在逐步清晰并形成共识。在 2011 年的语义技术信息（Semantic Technologies Information，STI）峰会上，与会的语义网络和数据库学者讨论了语义网络、语义技术及数据库领域在大数据时代所面临的挑战。

Michael L.Brodie 认为在真实、无模式和复杂的大数据或大数据语义网络中进行有意义的数据集成需要多学科、多技术交叉。

Bizer 列举了在大数据时代研究 Web 数据的 3 个挑战：Web 数据的拓扑结构，互联网中存在大量不同形式的数据，各种类型的数据的量都非常大；Web 数据的特点，这对于数据集成和大数据处理来说是一个值得研究的问题；如何利用一些公开的、可用的预先获取的 Web 数据集。

Michael L.Brodie 认为未来的挑战都将围绕数据集成、大规模资源描述框架、数据质量评定展开。

Peter Boncz 认为如果人们想更广泛地使用语义网络，存在两大挑战：缺乏好的使用案例；现有的数据集成方法使得创建链接非常困难。

在产业方面，大数据是现有产业升级与新产业诞生的重要推动力量。随着大数据时代的到来，产业界的需求与关注点发生了重大转变。企业关注的重点转向数据，计算机行业正在转变为真正的信息行业，从追求计算速度转变为关注大数据处理能力，软件也将从编程为主转变为以数据处理为主。大数据处理的兴起也改变着云计算的发展方向，使其进入以分析即服务（Analytics as a Service，AaaS）为主要标志的 Cloud 2.0 时代。

机器学习在学术界和产业界都有巨大的实用价值。由于大数据具有大量、复杂的特性，传统的、在小规模数据上适用的很多机器学习算法已不再适用。因此，这里重点剖析大数据环境下的机器学习算法。并行是处理大数据的主流方法，本部分将单独介绍一些并行算法，并引出大数据环境下机器学习研究的主要问题。

机器学习算法离不开大数据的训练。有关大数据的定义有多种，狭义的定义是：大数据是指不能装载进计算机内部存储器的数据。尽管这是一个非正式的定义，但易

于理解。广义的定义为：大数据是指无法在可容忍的时间内用传统 IT 技术和硬件工具对其进行感知、获取、管理、处理和服务的数据集合。

（1）大数据有多种特点。

从最开始的 3V 模型到扩展的 4V 模型都是以大数据的特点命名的。Larey 的 3V 模型包括体积（Volume）、速度（Velocity）和多样性（Variety）。人们对 4V 模型中的第 4 个 V 有多种解释，如变化性（Variability）、虚拟化（Virtual）、价值（Value）等。针对这些特点，有的大数据专家认为在大数据时代，知识解析、机器与人类共同工作及智能分析系统将会扮演重要角色，需要一种智能分析接口将人类与计算机世界进行连接，否则人类将被淹没在大数据的洪流中。

（2）大数据环境下机器学习的原理。

随着大数据时代的到来，大数据逐渐成为学术界和产业界的热点，已在很多技术和行业广泛应用，从大规模数据库到商业智能和数据挖掘；从搜索引擎到推荐系统，再到最新的语音识别、机器翻译等。大数据算法的设计、分析和工程实现涉及很多方面，包括大规模并行计算、流算法、云技术等。由于大数据存在复杂、高维、多变等特性，如何从真实、凌乱、无模式的大数据中挖掘出人类感兴趣的知识，需要更深刻的机器学习理论进行指导。

传统机器学习的研究内容主要包括如下 4 项：理解并且模拟人类的学习过程；针对计算机系统和人类用户之间的自然语言接口；利用不完全信息进行推理的能力，即自动规划问题；构造可发现新事物的程序。传统机器学习面临的一个新的挑战是如何处理大数据。包含大规模数据的机器学习问题是普遍存在的。但是，由于现有的许多机器学习算法是基于内存的，而大数据却无法装载进计算机内存，故现有的诸多算法不能处理大数据。如何提出新的机器学习算法以适应大数据处理的需求，是大数据时代的热点研究方向之一。

（3）大数据环境下的机器学习算法。

数据、算法、算力是笔者在第 1 章提出来的构成人工智能必不可少的因素。大数据处理策略包括分治策略与抽样。分治策略是一种处理大数据问题的计算范例，在分布式和并行计算有很大发展的情况下，分治策略显得尤为重要。一般来说，数据中不同样本对学习结果的重要程度也不相同。一些冗余和噪声数据不仅耗费大量的存储空

间,降低学习算法的运行效率,还会影响学习精度。因此,应该依据一定的性能标准(如保持样本的分布、拓扑结构及保持分类精度等)选择有代表性的样本形成原样本空间的一个子集,之后在这个子集上构造学习方法,完成学习任务。这样能在不降低甚至提高某方面性能的基础上,最大限度地降低时间、空间的耗费。

在大数据的背景下,样本选取的需求更迫切。但已有的大部分样本选取方法只适用于规模较小的数据集,如传统的压缩最近邻(Condensed Nearest Neighbor, CNN)、约减最近邻(Reduced Nearest Neighbor, RNN)、编辑最近邻(Edited Nearest Neighbor, ENN),它们的核心概念都是最小一致子集,而要找到这个子集需要逐一测试样本。可以基于局部几何和概率分布来选择分类边缘和边界样本,以保持原有数据的空间信息,但须计算每个样本的k近邻。在CNN的基础上,部分文献提出了一种快速压缩最近邻算法(Fast CNN, FCNN),这类文献倾向于选择分类边界的样本。Michael I.Jordan(美国国家科学院、美国国家工程院、美国国家医学院院士,机器学习泰斗)提出了一些关于大数据的统计推理方法。当用分治算法来处理统计推理问题时,需要从庞大的数据集中获取置信区间。Bootstrap理论通过重新采样数据来获取评估值的波动范围,进而获取置信区间,然而这对大数据是不可行的。数据的不完全抽样会导致错误的波动范围,必须进行更正以提供校准的统计推理。Jordan提出的Bag of Little Bootstraps可以避免这一问题,不仅继承Bootstrap的理论性质,并且有许多计算上的优势。Michael I.Jordan讨论的另一个问题是大规模矩阵计算。分治策略是一种启发式策略,在实际应用中会产生较好的效果,但当试图描述分治算法的统计性能时,新的理论问题就会出现。基于此,Michael I.Jordan提出基于随机矩阵的浓度定理。

综上所述,数据分治与并行处理策略是大数据处理的基本策略,但分治与并行处理策略较少利用大数据的分布知识,且影响大数据处理的负载均衡与计算效率。如何用大数据的分布知识优化负载均衡是一个亟待解决的问题。

(4) 大数据的特征选择。

在数据挖掘、文档分类和多媒体索引等新兴领域中,所面临的数据对象往往是大数据集,其中包含的属性数和记录数都很大,这导致处理算法的执行效率低下。通过属性选择可剔除无关属性,增加分析任务的有效性,从而提高模型精度,减少运行时间。大数据处理的一个巨大挑战是如何处理高维、稀疏数据。在大数据环境下,网络

流量、通信记录、大规模社会网络产生大量、多方面的高维数据，张量表示法提供了一种表示大数据的方法，故张量分解成为一种重要的汇总和分析工具。Kolda（基于标签的个性化信息设计推荐模型的专家）提出了一种高效使用内存的 Tucker 分解方法（Memory Efficient Tucker Decomposition，MET），用于解决传统的张量分解算法无法解决的时间和空间利用问题。MET 在分解的过程中基于可用内存自适应，选择正确的执行策略。该算法在充分利用可用内存的前提下最大化计算速度。MET 避免处理计算过程中产生的大量零星中间结果，自适应选择操作顺序，不仅消除了中间溢出问题，而且在不减少精确度的前提下节省了内存。

另一方面，Walid Wahba 探讨了涉及离散、含噪声、不完全相异数据的统计/机器学习模型的设计方法——正则化核估计（Regularized Kernel Estimation，RKE）和鲁棒流形展开（Robust Manifold Unfolding，RMU）。这些方法使用训练集中对象之间相异的信息，得到一个非负的低阶正定矩阵，用于将对象嵌入到一个低维欧几里得空间，其坐标可被用作各种学习模式中的属性。同样，大多数在线学习研究需访问训练实例的所有属性/特征，当获得完整的属性/特征集合的代价很昂贵时，这样的经典场景不再适用。

为了突破此限制，Hoi Kwan Lau 等人通过研究稀疏正则化和截断技术，解决如何利用固定数目的活跃特征来准确预测在线特征，并提出一个有效的算法，评估所提出算法在一些公共数据集上进行在线特征选择的经验性能，并证明用在线特征选择技术解决现实世界的大数据挖掘问题比一些著名的批处理特征选择算法具有更显著的可扩展性。

传统的自组织映射（Self Organizing Map，SOM）可用于特征提取，但当数据集较大时，SOM 速度慢的缺点会变得突出，Alaa Sagheer 等人提出了一种快速自组织映射（Fast SOM，FSOM），主要用于解决该问题。该方法的主要思想是：大数据的主要信息分布于特征空间的某一区域，如果能找到这些区域并直接在这些区域提取信息，而不是在整个数据空间提取信息，则能大幅度减少所用时间。

Quevedo 等人基于输入变量的有用性，采用经典技术（如相关性和正交性）的简单组合，提出一种输入变量排名算法，用于大数据降维和特征提取，取得了良好的效果。

Gheyas 等人结合模拟退火算法、遗传算法、贪心算法及神经网络算法的优点，提出一种模拟退火和遗传算法（Simulated Annealing and Genetic Algorithm，SAGA）。混

合算法用于解决选择最优化特征子集的 NP 时间问题。实验结果表明,该算法能更好地最优化特征子集,并降低时间复杂度。Gheyas 等人在论文结尾之处指出,很少有一种单一算法能解决所有问题,但如果能将多种算法结合,则能较好地弥补算法的不足。

Pal 等人提出了一种基于支持向量机 (Support Vector Machine, SVM) 的用于分类的特征选择方法, SVM 分类算法的准确度与数据集的特征数和数据集大小有关。因此,在分类前对数据进行特征选择有利于提高分类准确度。特征选择方法也很重要,用不同的特征选择方法选出来的特征区别很大。

Sun 等人提出一种用于分类的特征选择算法。该算法利用局部学习理论首先将复杂的非线性问题转换为一组线性问题,然后在最大间隔的框架下学习特征关联性。该算法基于机器学习和数值分析等技术,不需假设有关数据的分布。由于不需任何启发搜索,该算法在含有大量不相关特征的数据集上也具有较好的性能。

Hua 等人对比一些现有的特征选择方法,提出一种特征标签分布式模型,这个模型中可用特征的数量与它们在真实数据中是一样的,模型在一些高维数据上的测试结果显示,不同的特征选择算法在不同的模型条件下性能不一样,它们与样本数量、全局和异构标记的数量有关。该模型不仅适用于论文中提到的特征选择算法,也适用于其他分类算法和特征选择算法。

Song 等人研究降维技术在轨迹聚类中的作用,运用 3 种主流降维方法 (SVD、RP 和 PCA), 结合最通用的轨迹聚类算法,可以提高算法性能。

综上所述,由于大数据存在复杂、高维、多变等特性,如何采用降维和特征选择技术降低处理大数据的难度,是大数据特征选择技术迫切需要解决的问题。

(5) 大数据分类。

监督学习 (分类) 面临的一个新挑战是如何处理大数据。包含大规模数据的分类问题是普遍存在的,但是传统分类算法不能处理大数据。下面介绍几种大数据时代的分类方法。

SVM 分类。传统统计机器学习方法用于大数据分类有两大瓶颈问题: 计算密集型方法几乎不能用于大规模数据集; 鲁棒和非参数的置信区间的拟合模型的预测往往是未知的。针对上述问题, Lau 等人为 SVM 提出一种在线学习算法, 用于处理按顺序逐渐输入的数据的分类问题。该算法速度更快, 所用的支持向量的个数更少, 并

具有更优的泛化能力。Laskov 等人提出一种快速、数值稳定和鲁棒的增量 SVM 学习方法。黄广斌等提出一种大边缘分类器 M4, 与其他大边缘分类器或局部或全局地构建分离超平面不同, 该模型能局部和全局地学习判定边界。SVM 和极大极小概率机 (Minimax Probability Machine, MPM) 与该模型有密切联系, 该模型具有重要理论意义。针对大规模数据的分类问题, 在增量核主成分分析 (Incremental Kernel PCA) 和基于共轭梯度的最小二乘支持向量机 (Least Squares SVM, LS-SVM) 算法基础之上, Kim 等人提出适用于大数据的特征提取和分类算法。该算法所需内存较少, 无须存储较大矩阵, 可更好地解决大规模数据的分类问题。

决策树分类。传统决策树作为一种经典的分类学习算法, 在处理大数据时存在内存开销过大的问题。Franco Arcega 等人提出一种从大规模数据中构造决策树的方法, 解决当前算法中的一些限制条件, 新方法可利用所有的训练集数据, 但不需将它们都保存在内存中。实验表明, 该方法比决策树算法的计算速度更快。杨强 (笔者注: 香港科技大学教授, 机器学习算法专家) 等人提出一种增量优化的快速决策树算法 (Incrementally Optimized Very Fast Decision Tree, iOVFDT) 用于处理带有噪声的大数据, 与传统的挖掘大数据的决策树算法相比, 该算法的主要优势是实时挖掘能力, 这使得当移动数据流是无限时, 它能存储完整的数据用于再训练。Ben Haim 提出一种构建决策树分类器的算法。该算法在分布式环境中运行, 适用于大数据集和流数据, 与串行决策树相比, 该方法在精度误差近似的前提下能提高效率。

神经网络与极端学习机 (Extreme Learning Machine, ELM)。传统前馈神经网络一般采用梯度下降算法调整权值参数, 学习速度慢、泛化性能差等问题是制约前馈神经网络应用的瓶颈。黄广斌等摒弃梯度下降算法的迭代调整策略, 提出 ELM。该方法随机赋值单隐层神经网络的输入权值和偏差项, 并通过一步计算求出网络的输出权值。相比于传统前馈神经网络训练算法需经多次迭代调整才能最终确定网络权值, ELM 的训练速度获得较显著提升。然而, 由于计算能力和复杂性的限制, 在大数据上训练出单一 ELM 是一个困难的问题。解决这一困难通常有两种途径: 训练基于分治策略的 ELM; 在训练单一 ELM 中引入并行机制。单一 ELM 有很强的函数逼近能力, 能否将这种逼近能力延拓到基于分治策略的 ELM 是衡量 ELM 是否适用于大数据学习的一个关键指标。尚未见到基于分治策略的 ELM 逼近的研究报道, 一些相关的研究还包括有效性学习。

应用领域的分类算法。除上述内容外，在一些应用领域，也有针对大数据的分类算法被提出。在计算机辅助诊断领域，机器学习广泛应用于帮助医学专家从已诊断案例中获取先验知识，但大量的已诊断样本很难获取。有机器学习算法专家提出了一种半监督的学习算法——基于随机森林的协同训练，用来估计未诊断样本的标记自信度，能比较容易地得出先验知识。该方法在基准数据集上得到了较好的结果。针对大规模图像数据集的分类性能问题，有机器学习算法专家提出了在特征提取和分类器训练方面提高效率。对于特征提取，利用 Hadoop 架构，在几百个 Mapper 上并行计算；对于训练 SVM，提出并行平均随机梯度下降算法，可处理具有 120 万个图像、1000 个类别的图像数据，并具有较快的收敛速度。

此外，中文网页标记数据稀缺，英文网页标记数据较丰富，有机器学习算法专家用英文网页标记信息来解决跨语言分类问题，提出基于信息瓶颈的方法。该方法首先将中文翻译成英文，然后将所有网页用一个只允许有限信息通过的信息瓶颈来编码。该方法可使跨语言分类更准确，显著提高一些已有的监督与半监督分类器的准确率。

综上所述，传统机器学习的分类方法很难直接运用到大数据环境，不同的分类算法都面临着大数据环境的挑战，如何针对不同分类算法研究并行或改进策略成为大数据环境下分类学习算法研究的主要方向。

(6) 大数据聚类。

聚类学习是最早被用于模式识别及数据挖掘任务的方法之一，并且被用来研究各种应用中的大数据库，因此用于大数据的聚类算法受到越来越多的关注。Havens 等人对比了 3 种扩展的模糊 C 均值 (FCM) 聚类算法对于大数据的执行效率。Havens 等人用可装载的数据集和 VL 数据集进行数值型实验，这些实验对时间复杂度、空间复杂度、速度、处理装载数据的批量 FCM 的近似质量等进行对比。实验结果显示，随机取样可扩展 FCM (Random Sampling Plus Extension FCM)、位减少 FCM (Bit Reduced FCM) 及近似核 FCM (Approximate Kernel FCM) 都是较好的选择，都近似于 FCM。最后，Havens 等人展示含有 50 亿个对象的数据集的大数据算法，并就如何使用不同的大数据 FCM 聚类策略提出一系列建议。

随着数据体积的增大，I/O 瓶颈变成数据分析的一个重要问题。数据压缩能起到缓解作用。以 K-means 为例，Xue 等人提出了一种压缩感知性能提升模型用于大数据聚类。该模型定量分析整个计算过程中与压缩有关的诸多影响因素。在包含上百个计

算核的集群上对大到 1.114TB 的 10 维数据进行聚类实验,实验结果证明压缩能改善 I/O 性能,并且该模型能有效决定何时、如何使用压缩来改善大数据分析中的 I/O 性能。

针对分布式聚类、流数据聚类,Hall 等人研究二次抽样方法以提高聚类算法的可扩展性。实验表明,人们可构造一个好的模型而不必知道所有的数据,如果需要,修改后的算法可应用于 TB 级或更大级别的数据。

为实现数据的大规模并行处理,MapReduce 模型成为学术界和工业界最为流行的工具。从数据库角度看,MapReduce 是一种简单但强大的执行引擎,可与其他数据存储和管理组件有效融合。为解决大规模数据分析难题,有机器学习算法专家提出了基于 MapReduce 的 K-means 算法,并在 speedup、sizeup、scaleup 这 3 个指标上获得较好的并行性能。

Papadimitriou 等人提出了一种利用 MapReduce 模型实现协同聚类的系统框架——分布式协同聚类框架,并引入分布式数据预处理、协同聚类等方法,并在 Hadoop 上实现该系统。实验结果证明分布式协同聚类框架具有良好的可扩展性、高执行效率,能处理几百 GB 的数据。

Ferreira 等人提出了一种利用 MapReduce 开展大规模数据聚类的方法。主要针对问题包括:如何将 I/O 开销最小化;如何降低节点之间的通信开销。Ferreira 所提出的 BoW (Best of both Worlds) 方法能自动发现系统瓶颈并选择应对策略。总体来说,其贡献在于:提出 BoW,并设计开销函数,动态选择最优策略;实验证明 BoW 有明显优势。

Havens 等人分析了在大数据上运行 C-mean 的困难,指出模糊技术在处理大数据上的有效性。

随着信息技术的迅猛发展,聚类所面临的不仅是数据量越来越大的问题,更重要的还是数据的高维问题。由于数据来源的丰富多样,图片、文字、声音甚至视频都逐渐成为聚类处理的目标对象,这些特殊对象的属性信息往往要从数百个甚至成千上万个方面来表现,其每个属性都成为数据对象的一个维度,对高维数据的聚类分析,已成为众多领域的研究方向之一。高维数据的聚类方法包括基于降维的聚类、子空间聚类、基于图的聚类等。在很多需处理高维数据的应用领域,降维是常用的方法之一。直观地讲,降维就是通过把数据点映射到更低维的空间上以寻求数据的紧凑表示的一

种技术，这种低维空间的紧凑表示有利于对数据进行进一步处理。降维代表着数据信息的损失，如何保证降维而不损失聚类性能是传统数据领域仍在研究和讨论的问题。

综上所述，经典的聚类算法在大数据环境下面临数据量大、数据体积过大、复杂度高等众多挑战，如何并行或改进现有聚类算法，进而提出新的聚类算法成为研究的关键。

（7）大数据关联分析。

关联分析的研究源自 Apriori 算法，Apriori 算法的性质要求频繁模式的子模式必须也是频繁的。基于这一启发，出现了一系列的类 Apriori 算法，如 AprioriAll、AprioriSome、DynamicSome、广义序列模式及基于等价类的序列模式。后来，研究者又提出一系列基于数据投影的算法，包括频繁模式投影序列模式挖掘算法（Frequent-Pattern-Projected Sequential Pattern Mining，FreeSpan）和前缀投影序列模式挖掘算法（Prefix-Projected Sequential Pattern Mining，PrefixSpan）。基于等价类的序列模式是基于格点的算法。基于内存索引的序列模式挖掘算法是基于内存索引的方法。而基于正则表达式约束的序列模式挖掘算法通过使用正则表达式将约束整合到一起。实现大数据的关联分析主要有两种途径：并行和增量。

在并行方面，有机器学习算法专家提出了一种基于 MapReduce 的并行 Apriori 算法。Apriori 算法最主要的作用是产生候选项集，该算法将产生候选项集的过程并行化，以提高运行效率，并具有良好的加速比和伸缩性。

一些文献介绍了基于广义序列模式（Generalized Sequential Pattern，GSP）算法和基于 GSP 的频繁序列挖掘（Mining Frequent Sequences，MFS）算法的增量挖掘算法 GSP+和 MFS+。

也有一些文献提出了基于等价类的序列模式的增量序列挖掘算法（Incremental Sequence Mining，ISM），该方法不仅可在数据库更新时维持频繁序列，还提供了一个用户交互接口以方便用户修改约束（如最小支持度等）。ISM 只考虑序列追加，而一些文献提出的增量频繁序列挖掘算法（Incremental Frequent Sequences Mining，ISEM）还考虑插入新序列的情况。

ISEM 只考虑对频繁序列的后缀进行扩充，而一些文献中提出的增量更新序列算法（Incrementally Updating Sequences，IUS）对旧的频繁序列的前缀和后缀都可进行扩充。

有一些文献指出算法可在维持序列模式的同时从数据库中删除一些记录。还有一些文献提出一个实验性方法——性能与差异均衡算法 (Tradeoff between Performance and Difference, TPD), 来确定何时更新序列模式。

以上增量算法都是为了提高算法效果, 在大数据集上的运行效果有待进一步验证。

(8) 大数据并行算法。

如何把传统机器学习算法运用到大数据环境中, 一个典型策略是对现有学习算法并行化。例如, 图形处理器 (Graphic Processing Unit, GPU) 通过并行得到较显著的性能提升。这些 GPU 平台由于采用并行架构, 使用并行编程方法, 使得计算能力呈几何级增长。

有机器学习算法专家提出了一种非平凡的策略用来并行处理一系列数据挖掘与机器学习问题, 包括一类分类 SVM 和两类分类 SVM、非负最小二乘问题、L1 正则化回归问题。由此得到的乘法算法, 可直接在 MapReduce 和通用并行计算架构 (Compute Unified Device Architecture, CUDA) 等并行计算环境中实现。在大数据分类和聚类学习当中, MapReduce 框架被用于并行化的传统机器学习算法以适应大数据处理的需求。

Shim 在 MapReduce 框架下, 讨论如何设计高效的 MapReduce 算法, 对当前一些基于 MapReduce 的机器学习和数据挖掘算法进行归纳总结, 以便进行大数据的分析。

有机器学习算法专家提出了一种大数据挖掘技术, 即利用 MapReduce 实现并行的基于粗糙集的知识获取算法。

Hefeeda 等人提出一种近似算法使基于核的机器学习算法可有效处理大规模数据集。当前的基于核的机器学习算法由于需计算核矩阵而面临可伸缩性问题, 该算法在计算核矩阵时不仅能大幅降低计算和内存开销, 而且没有明显影响结果的精确度。此外, 由于传统的提升算法本身具有串行特点, 不易获得好的扩展性。因此, 需提出并行的提升算法来高效处理大规模数据。Palit 等提出 2 种并行提升算法: ADABOOST.PL 和 LOGITBOOST.PL。它们可使多个计算节点同时参与计算并且可构造出一个提升集成分类器。该方法通过 MapReduce 框架来实现, 在合成数据和真实数据集上的实验表明其在分类准确率、加速比和放大率方面都能取得较好的结果。

Kaiser 等人还利用 MapReduce 框架分布式实现一系列核函数学习机训练, 该方法适用于基于核的分类和回归。Kaiser 还介绍了一种扩展版的从区域到点的建模方法,

适应来自空间区域的大量数据。

Yan 考虑潜在狄利克雷分配模型 (Latent Dirichlet Allocation, LDA) 的两种推理方法——塌缩吉布斯采样 (Collapsed Gibbs Sampling, CGS) 和塌缩变分贝叶斯推理 (Collapsed Variational Bayesian, CVB) 在 GPU 上的并行化问题。为解决 GPU 的内存限制问题, Yan 提出了一种能有效降低内存开销的数据划分方案。这种划分方案也能平衡多重处理器的计算开销, 并能避免内存访问冲突, 使用数据流来处理超大的数据集。

针对在异构云中进行大数据分析服务的并行化问题, 有机器学习算法专家提出了最大覆盖装箱算法来决定系统中有多少节点, 哪些节点应该应用于大数据分析的并行执行。该方法可使大数据得到分配, 使得各个计算节点可同步结束计算, 并且使数据块的传输可以和上一个块的计算重叠, 从而节省时间。实验表明, 该方法相比其他方法可提高大约 60% 的性能。在分布式系统方面, 中国科学院计算技术研究所智能信息处理重点实验室数据挖掘与机器学习组在 2008 年年底与中国移动合作, 开发完成了分布式并行数据挖掘系统 PDMiner, 这是中国最早的基于云计算平台的并行数据挖掘系统之一。该系统提供多种并行数据转换规则和并行数据挖掘算法, 已用于中国移动 TB 级实际数据的挖掘, 并达到商用软件的精度。

还有机器学习算法专家提出了一个可伸缩的分布式系统——广义线性聚合分布式引擎 (Generalized Linear Aggregates Distributed Engine, GLADE)。GLADE 通过用户自定义聚合 (User Defined Aggregate, UDA) 接口并且在输入数据上有效运行来进行数据分析。作者从两个方面来论证系统的有效性: 展示如何使用一系列分析功能来完成数据处理。将 GLADE 与两种不同类型的系统 (用 UDA 进行改良的关系型数据库和 MapReduce) 进行对接。然后从运行结果、伸缩性及运行时间上与不同类型的系统进行对比。

综上所述, 并行策略是将传统机器学习算法运用于大数据的典型策略之一, 并且在一定范围内取得了一些进展, 能处理一定量级的大数据。如何研究高效的并行策略以高效处理大数据也是当今的研究热点之一。

2. 机器学习综述

笔者理解的 AI 是“算法+算力+数据”。也有人认为 AI 是由数据、机器学习和 GPU

硬件驱动的。这两种观点本质上是一致，下面我们从机器学习的历史、常见方法和应用的角度来看 AI 的本质是什么。

（1）机器学习的发展历史与相关要素。

机器学习属于 AI 中一个较为年轻的分支，大致可以分为以下三个发展阶段。

第一阶段：20 世纪 50 年代中期至 20 世纪 60 年代中期，这一时期是机器学习的萌芽时期。人们试图通过软件编程来操控计算机完成一系列的逻辑推理功能，进而使计算机在一定程度上具有类似人类的智能思考能力。然而，这时期计算机推理出的结果远没有达到人们对机器学习的期望。通过进一步研究发现，只具有逻辑推理能力并不能使机器产生智能。研究者认为，使得机器拥有智能的前提是机器必须拥有大量的先验知识。

第二阶段：20 世纪 60 年代中期至 20 世纪 80 年代中期，这一时期的机器学习处于发展时期。人们试图利用根据自身思维提取出来的规则教会计算机执行决策行为，主流是各式各样的专家系统，然而这些系统总会面临“知识稀疏”的问题，即面对无穷无尽的知识与信息，人们无法总结出万无一失的规律。因此，让机器自主学习的设想“浮出水面”。基于 20 世纪 50 年代对于神经网络的研究，人们开始研究如何让机器自主学习。

第三阶段：20 世纪 80 年代至今，机器学习达到了一个繁荣时期。由于这一时期互联网大数据的发展以及 GPU 的出现，机器学习度过了瓶颈期。机器学习呈现爆发式发展，开始成为一门热门学科并且被应用到各个领域。各种机器学习算法不断涌现，而利用深层次神经网络的深度学习也得到进一步发展。同时，机器学习的蓬勃发展还促进了其他分支的出现，如模式识别、数据挖掘、生物信息学和自动驾驶等。

（2）机器学习的常见方法。

监督学习。监督学习表示机器学习的数据是带标记的，这些标记可以是数据类别、数据属性以及特征点位置等。这些标记作为预期效果，不断修正机器的预测结果。具体过程是：首先，通过大量带有标记的数据训练机器，机器将预测结果与期望结果进行比对；然后，根据比对结果修改模型中的参数，再一次输出预测结果；最后，将预测结果与期望结果进行比对……重复多次直至收敛，最终生成具有一定鲁棒性的模型来使机器产生智能决策能力。常见的监督学习有分类、回归。分类是将一些实例

数据分到合适的类别中，它的预测结果是离散的。回归是将数据归到一条“线”上，即为离散数据生产拟合曲线，因此其预测结果是连续的。

无监督学习。无监督学习表示机器学习的数据是没有标记的。机器从无标记的数据中探索并推断出潜在的联系。常见的无监督学习有聚类、降维。在聚类工作中，由于事先不知道数据类别，因此只能通过分析数据样本在特征空间中的分布，例如基于密度或是基于统计学概率模型等，从而将不同的数据分开，把相似的数据聚为一类。降维是将数据的维度降低。例如描述一个西瓜，若只考虑外皮色泽、根蒂形状、敲击回声、纹理、密度以及含糖率这 6 个属性，则西瓜的数据的维度为 6。由于数据本身具有庞大的数量和各种属性特征，若对全部数据信息进行分析，将会增加训练的负担和存储空间。因此可以通过主成分分析等方法，考虑主要影响因素，舍弃次要影响因素，来平衡准确度与效率。

强化学习。强化学习是带激励的学习，具体来说就是，如果机器的行动正确，将施加一定的“正激励”；如果机器的行动错误，会给出一个“惩罚”（也可称为“负激励”）。在这种情况下，机器将会考虑在一个环境中如何行动才能达到激励的最大化。例如，在贪吃蛇游戏中，贪吃蛇需要不断“吃”到“食物”来加分。为了不断提高分数，贪吃蛇需要考虑在自身位置上如何转向才能“吃”到“食物”，这种学习过程便可理解为是一种强化学习。强化学习最为火热的一个应用便是谷歌 AlphaGo 的升级版——AlphaGo Zero。相较于 AlphaGo，AlphaGo Zero 舍弃了先验知识，不再需要人为设计特征，而是直接将棋盘上黑、白棋子的摆放情况作为原始数据输入到模型中，机器使用强化学习来自我博弈，不断提升自己，最终出色地完成下棋任务。AlphaGo Zero 的成功证明了在没有人类的经验和指导下，深度强化学习依然能够出色地完成指定任务。

（3）机器学习的应用。

由于近 20 年科技的迅速发展，机器学习已经成为一门活跃并且充满生命力的学科。机器学习具备了一定的解决实际问题的能力，被当作是一种基础服务技术，并和不同技术进行结合。我们通过几个具体的例子来进一步了解机器学习的优势。

标注问题。标注问题就是为数据标记类别，这些数据可以是图片、声音、文字等。例如，在下面这张图片中，图片包含的动物有猫和狗，因此，可以利用多标注将图片标注为“有猫并且有狗”。在现实中，需要标注的数据往往以百万甚至以亿计，此时人力标注就显得很吃力了；通过机器学习的自动标注可以节省大量的时间，提高

工作效率。



猫和狗的图片

搜索与排序。搜索与排序研究的是如何对对象进行排序。例如，在检索信息时，我们常常关注如何把对象按照与检索目的的相关性进行排序。在 AI 时代，由于各种搜索引擎的流行，我们将目光聚集在如何对网页进行相关性排序。基于机器学习的网页排序可以根据用户喜好，智能并且个性化地呈现搜索结果。

推荐系统。推荐系统与排序问题紧密相连，并被广泛应用于新闻网站、购物网站以及搜索引擎上。推荐系统的目标是将不同用户各自感兴趣的内容推荐给他们。例如 Comcast 公司（美国一家提供宽带网络、有线电视及 IP 电话服务的公司）利用机器学习技术分析了大量用户的历史购买记录，对每一个用户输出对应的品位描述，然后根据品位描述把顾客分成不同类别。对于不同类别的顾客，Comcast 都能实时跟踪并推荐相关的最流行的节目内容。基于机器学习生成了更好的推荐系统，达到了更高的购买率，提高了顾客的满意度。

小结：机器学习是一门多领域交叉学科，涉及概率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科。专门研究计算机怎样模拟或实现人类的学习行为，以获取新的知识或技能，重新组织已有的知识结构使之不断改善自身的性能。

机器学习是人工智能的核心之一，是使计算机具有智能的根本途径，其应用遍及人工智能的各个领域，它主要使用归纳、综合而不是演绎。

3.1.5 GPU 和其他设备

针对 AI 优化的 GPU 和其他设备的架构可以高效运行面向 AI 的计算任务。

图形处理器又称显示核心、显示芯片、视觉处理器，是一种专门在个人电脑、工作站、游戏机和一些移动设备（如平板电脑、智能手机等）上进行图像运算工作的微处理器。

GPU 的用途是将计算机系统所需要的显示信息进行转换，并向显示器提供行扫描信号，控制显示器的正确显示，是连接显示器和个人电脑主板的重要元件，也是“人机对话”的重要设备之一。

GPU 是为了更快处理图像而诞生的。在 20 世纪 90 年代，一批工程师意识到：在屏幕上进行多边形图像渲染，本质上是处理并行的任务——对每个像素点的色彩都可以独立计算，而不需要考虑其他像素点。于是 GPU 诞生，成为比 CPU 更高效的渲染工具。简而言之，由于 CPU 在图像渲染方面的能力不足，GPU 被发明出来分担这部分工作，此后就成了专门的硬件。凭借上千个更加简单的核心，GPU 能高效处理让 CPU 十分吃力的任务。只要配合合适的代码，这些核心就能处理超大规模的数学运算，实现逼真的游戏体验。

GPU 作为显卡的“大脑”，决定了该显卡的档次和大部分性能，同时 GPU 也是 2D 显示卡和 3D 显卡的区别依据。2D 显卡在处理 3D 图像与特效时主要依赖 CPU 的处理能力，被称为“软加速”。3D 显卡把三维图像和特效处理功能集中在 GPU 内，这就是所谓的“硬件加速”功能。GPU 一般是显卡上最大的芯片。时下市场上的显卡大多采用 NVIDIA 和 AMD-ATI 两家公司的 GPU。NVIDIA 公司在 1999 年发布 GeForce 256 图形处理器时首次提出 GPU 的概念。从此 NVIDIA 的图形处理器就用 GPU 这个新名字来称呼。GPU 使显卡减少了对 CPU 的依赖，并承担部分原本属于 CPU 的工作，特别是在处理 3D 图形时。

GPU 能将 3D 模型的信息转换为 2D 的信息并进行表示，同时添加不同的纹理和阴影效果，所以，GPU 在硬件里也是比较特殊的存在。从 3D 建模到最终显示在屏幕上，GPU 渲染场景使用的是流水线操作。在早期，对流水线操作是不能做任何改动的，整个操作从读取三角形的顶点数据开始，GPU 处理完相关数据后将图像信息放入帧缓冲区并准备发送给显示器。GPU 也能对场景的某些特定效果进行处理，不过这些都是由工程师设计好的，能做的修改很少。

GPU 在设计之初并非是针对深度学习而是针对并行计算。并行计算意味着可以同时处理运算，而不是一步一步地进行计算。复杂问题可以被分解为更简单的问题，然

后同时进行处理。并行计算适用于 HPC 和超算领域所涉及的许多问题,比如气象模型、宇宙模型和 DNA 序列。但并不是只有天体物理学家和气象学家才能充分利用并行计算的优点。事实证明,许多企业应用也能从并行计算获得好处。这些应用包括数据库查询、密码学领域的暴力搜索、对不同场景的计算机模拟、机器学习/深度学习、地理信息可视化。

由于 GPU 在设计之初并非针对深度学习,而是针对图形加速,因此在 NVIDIA 推出 CUDA 架构之前, GPU 并无太强的针对深度学习的运算能力。

截至 2019 年, NVIDIA 提供了基于其 GPU 的从后端模型训练到前端推理应用的全套深度学习解决方案,普通开发人员也可以非常容易地上手使用 GPU 进行深度学习开发,或者高性能运算。开发 CUDA 架构耗费了 NVIDIA 巨大的人力、物力,可以说,正是因为有了 CUDA 这个中间层,开发者才真正爱上了 GPU, NVIDIA 胜在软件。但 CUDA 还不能被称为算法,它只是计算硬件与算法之间的桥梁。

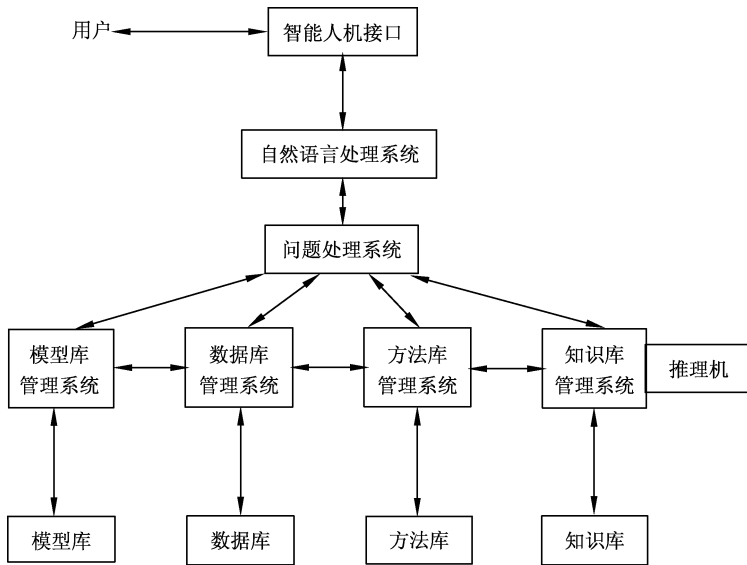
在 2012 年, AlexNet 利用“深度学习+GPU”的方案,一举赢得 ILSVRL,并奇迹般地将识别成功率从 74%提升到 85%。NVIDIA 敏锐地觉察到了这一趋势,并大力优化基于 GPU 的深度学习生态系统,在三年时间内将 GPU 的性能提升了 65 倍,从而奠定了自身的“王者”地位。

3.1.6 决策管理技术

智能决策支持系统(IDSS)将 AI 和决策支持系统(Decision-making Support System, 后文简称 DSS)相结合,应用专家系统(Expert System, 后文简称 ES)技术,使 DSS 能够更充分应用人类的知识(如关于决策问题的描述性知识、决策过程中的过程性知识、求解问题的推理性知识),通过逻辑推理来解决复杂的决策问题。

IDSS 这个概念最早由美国学者于 20 世纪 80 年代提出,它既能处理定量问题,又能处理定性问题。IDSS 的核心思想是将 AI 与其他相关科学成果相结合,使 DSS 具有智能。

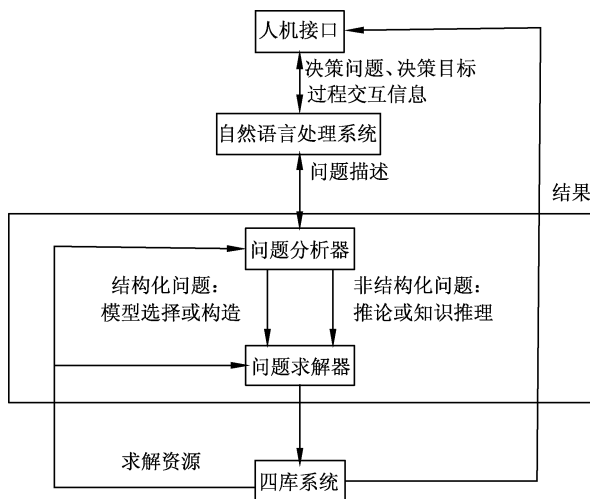
较完整与较典型的 IDSS 的结构是在传统三库 DSS 的基础上增设知识库与推理机,在智能人机接口与子系统之间加入自然语言处理系统(LS),在四库(模型库、数据库、方法库、知识库)与自然语言处理系统之间插入问题处理系统(PSS)而构成新的四库系统结构, IDSS 的结构如下图所示。



智能决策支持系统的结构

(1) 智能人机接口。四库系统的智能人机接口接受使用自然语言或接近自然语言的方式表达决策问题及决策目标，这较大幅度地改变了人机界面的性能。

(2) 问题处理系统。问题处理系统处于 IDSS 的中心位置，是联系人于机器及所存储的求解资源的桥梁，主要由问题分析器与问题求解器两部分组成，其工作流程如下图所示。



问题处理系统的工作流程

问题处理系统是 IDSS 中最活跃的部件，它既要识别与分析问题，设计求解方案，还要为问题求解器调用四库中的数据、模型、方法及知识等资源，对半结构化问题或非结构化问题还要运用推理机作推理。

(3) 自然语言处理系统能转换问题描述，由问题分析器判断问题的结构化程度，对结构化问题选择或构造模型，采用传统的模型进行计算、求解；对半结构化问题或非结构化问题，则采用规则模型与推理机制来求解。

(4) 知识库子系统可分为三个部分：知识库管理系统、知识库及推理机。

知识库管理系统。它的功能主要有两个：一是回应知识库增、删、改知识的请求；二是回应决策过程中对知识的请求。

知识库。知识库是知识库子系统的核心。知识库中存储的是那些既不能用数据表示，也不能用模型方法描述的专家知识和经验，即决策专家的决策知识和经验知识，同时也包括一些针对特定问题的专门知识。知识库中的知识表示是为描述世界所做的一组约定，是知识的符号化。对于同一知识，可以有不同的知识表示形式，知识的表示形式直接影响推理方式，并在很大程度上决定着—个系统的能力和通用性。

知识库包含事实库和规则库两部分。例如：事实库中存放了“任务 A 是紧急订货”“任务 B 是出口任务”这样的事实。规则库中存放着“IF 任务 i 是紧急订货，AND 任务 i 是出口任务，THEN 任务 i 最优先安排计划”“IF 任务 i 是紧急订货，THEN 任务 i 按优先安排计划”这样的规则。

推理机。推理是指从已知事实推出新事实（结论）的过程。推理机是一组程序，它针对用户问题去处理知识库（规则和事实）。推理原理如下：若事实 M 为真，且有一规则“IF M THEN N”存在，则 N 为真。因此，如果事实“任务 A 是紧急订货”为真，且有一规则“IF 任务 i 是紧急订货，THEN 任务 i 按优先安排计划”存在，则任务 A 就应该优先安排。

智能决策支持系统包含以下 6 个特点。

基于成熟的技术，容易构造出实用系统。

充分利用了各层次的信息资源。

基于规则的表达方式，使用户易于掌握、使用。

具有很强的模块化特性，并且模块重用性好，系统的开发成本低。

系统的各部分组合灵活，可实现强大的功能，并且易于维护。

系统可迅速采用先进的支撑技术，如 AI 技术等。

由于在 IDSS 的运行过程中各模块要反复调用上层模块，比起直接调用底层模块的方式，其运行效率要低。但是考虑到 IDSS 只在高层管理者做重大决策时才运行，其运行频率与其他信息系统相比要低得多，况且每次运行时环境条件差异很大，所以牺牲部分运行效率以换取系统维护的效率是完全值得的。

智能决策系统是管理信息系统应用概念的深化，是在管理信息系统的基础上发展起来的系统。DSS 是用来解决非结构化问题、服务于高层决策的管理信息系统，按功能可分为专用 DSS、DSS 工具和 DSS 生成器。专用 DSS 是为解决某一领域问题而设计的 DSS。DSS 工具是指某种语言、某种操作系统、某种数据库系统。DSS 生成器是通用决策支持系统，一般包括数据库（DBMS）、模型库（MBMS）、方法库、知识库和会话部件。DSS 的数据库不同于一般的数据库，它有很高的性能要求，是在原基层数据库的基础上建立起来的专用数据库，一般由数据仓库充当 DSS 数据库，数据库为决策提供数据能力或资料能力。模型库为决策提供分析能力。会话部件又被称为接口部件，它是人和决策支持系统联系的接口。智能决策支持系统（IDSS）在一般 DSS 基础上增加了深度知识库，所以成了人工智能决策系统。

3.1.7 深度学习技术

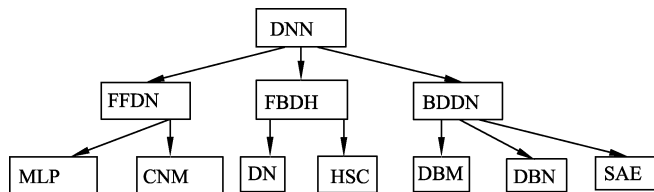
深度学习是机器学习领域一个新的研究方向，近年来在语音识别、图像识别等多类应用中取得突破性进展。其目的在于建立模型模拟人类大脑的神经连接结构，在处理图像、声音和文本这些信号时，通过多个变换阶段对数据特征进行描述，进而给出数据的解释。以图像数据为例，灵长类的视觉系统中对这类信号的处理顺序为：检测边缘和形状；逐步形成更复杂的视觉形状。同样地，深度学习通过组合低层特征形成更加抽象的高层表示、属性类别或特征，给出数据的分层特征表示。

深度学习之所以被称为“深度”，是相对支持向量机、提升方法、最大熵方法等“浅层学习”方法而言的，在深度学习模型中，非线性操作的层级数更多。浅层学习依靠人工抽取样本特征，经网络模型学习后获得的是没有层次结构的单层特征；而深度学习通过对原始信号进行逐层特征变换，将样本在原空间的特征表示变换到新的特征空间，自动学习得到层次化的特征表示，从而更有利于分类或特征的可视化。深度学习理论的另外一个理论动机是：如果一个函数可以用 k 层结构以简洁的形式表达，

那么用 $k - 1$ 层的结构表达则可能需要增加指数级数量的参数 (相对于输入信号), 且泛化能力不足。

1. 从技术角度讲解深度学习技术

深度学习的概念最早由多伦多大学的 Geoffrey E.Hinton 等人于 2006 年提出, 指基于样本数据, 通过一定的训练方法得到包含多个层级的深度网络结构的机器学习过程。传统的神经网络随机初始化网络中的权值, 导致网络很容易收敛到局部最小值, 为解决这一问题, Hinton 提出使用无监督预训练的方法优化网络权值的初值, 再进行权值微调的方法, 拉开了深度学习的序幕。深度学习所得到的深度网络结构包含大量的单一元素 (神经元), 每个神经元与大量其他神经元相连接, 神经元间的连接强度 (权值) 在学习过程中被修改并决定网络的功能。通过深度学习得到的深度网络结构符合神经网络的特征, 因此深度网络就是深层次的神经网络, 即深度神经网络 (Deep Neural Network, DNN)。深度神经网络是由多个单层非线性网络叠加而成的, 常见的单层网络按照编码、解码情况分为 3 类: 只包含编码器部分; 只包含解码器部分; 既包含编码器部分也包含解码器部分。编码器提供从输入到隐含特征空间的自底向上的映射, 解码器以重建结果尽可能接近原始输入为目标, 将隐含特征映射到输入空间。深度神经网络分为以下 3 类, 如下图所示。



深度神经网络的分类

前馈深度网络 (Feed-Forward Deep Network, FFDN), 由多个编码器层叠加而成, 如多层感知机 (Multi-Layer Perception, MLP) 卷积神经网络 (Convolutional Neural Network, CNN) 等。

反馈深度网络 (Feed-Back Deep Network, FBDN), 由多个解码器层叠加而成, 如反卷积网络 (Deconvolutional Network, DN) 层次稀疏编码网络 (Hierarchical Sparse Coding, HSC) 等。

双向深度网络 (Bi-Directional Deep Network, BDDN), 通过叠加多个编码器层和解码器层构成 (每层可能是单独的编码过程或解码过程, 也可能既包含编

码过程也包含解码过程),如深度玻尔兹曼机(Deep Boltzmann Machine, DBM)、深度信念网络(Deep Belief Network, DBN)、栈式自编码器(Stacked Auto Encoder, SAE)等。

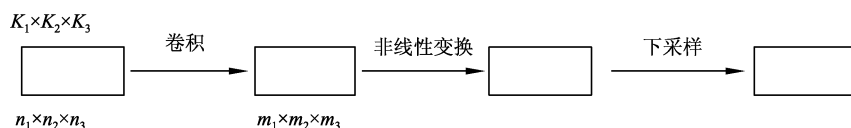
(1) 前馈深度网络。

前馈深度网络是最早的人工神经网络模型之一。在这种网络中,信息只沿一个方向流动,输入单元通过一个或多个隐藏层到达输出单元,在网络中没有封闭环路。典型的前馈深度网络有多层感知机和卷积神经网络等。

F.Rosenblatt 提出的感知机是最简单的单层前向人工神经网络,但随后马文·明斯基等人证明单层感知机无法解决线性不可分问题,这一结论将神经网络的研究引入一个低潮期,直到研究人员认识到多层感知机可解决线性不可分问题,以及反向传播算法与神经网络结合研究的兴起,对神经网络的研究才重新开始成为热点。但是由于传统的反向传播算法具有收敛速度慢、需要大量带标签的训练数据、容易陷入局部最优等缺点,多层感知机的效果并不是十分理想。

1984年,日本学者 K.Fukushima 等人基于感受野(感受器受刺激而兴奋时,通过感受器官中的向心神经元将神经冲动传到上位中枢,一个神经元所支配的刺激区域就叫作神经元的感受野)概念提出的神经认知机可看作卷积神经网络的一种特例, Yann LeCun 等人提出的卷积神经网络是神经认知机的推广形式。卷积神经网络是由多个单层卷积神经网络组成的可训练的多层网络结构。每个单层卷积神经网络包括卷积、非线性变换和下采样3个阶段,其中下采样阶段不是每层都必需的。每层的输入和输出为一组向量构成的特征图(Feature Map),第一层的原始输入信号可以看作一个具有高稀疏度的高维特征图。例如,输入部分是一张彩色图像,每个特征图对应的则是一个包含输入图像的彩色通道的二维数组(对于音频输入,特征图对应的是一维向量;对于视频或立体影像,特征图对应的是三维数组)。每个特征图对应的输出部分表示从输入图片所有位置提取的特定特征。

我们再来看由卷积、非线性变换和下采样3个阶段构成的单层卷积神经网络,如下图所示。



单层卷积神经网络的3个阶段

卷积阶段通过提取信号的不同特征实现对输入信号进行特定模式的观测。其观测模式也被称为卷积核，其定义源于 D.H.Hubel 等人提出的局部感受野概念。每个卷积核检测输入特征图上所有位置的特定特征，实现同一个输入特征图上的权值共享。为了提取输入特征图上不同的特征，使用不同的卷积核进行卷积操作。

卷积阶段的输入是由 n_1 个 $n_2 \times n_3$ 大小的二维特征图构成的三维数组。每个特征图记为 X_i 。该阶段的输出 Y 也是个三维数组，由 m_1 个 $m_2 \times m_3$ 大小的特征图构成。在卷积阶段，连接输入特征图 X_i 和输出特征图 Y_j 的权值记为 W_{ij} ，即可训练的卷积核（局部感受野），卷积核的大小为 $K_2 \times K_3$ 。输出特征图如下。

$$= + \sum *$$

在上式中： $*$ 为二维离散卷积运算符； b_j 是可训练的偏置参数。

在非线性阶段，对卷积阶段得到的特征按照一定的原则进行筛选，筛选原则通常采用非线性变换的方式，以避免线性模型表达能力不足的问题。非线性阶段将卷积阶段提取的特征作为输入，进行非线性映射 $R = h(y)$ 。传统卷积神经网络中非线性操作采用 sigmoid、tanh 或 softsign 等饱和非线性（Saturating Nonlinearities）函数，近几年的卷积神经网络多采用不饱和非线性（Non-Saturating Nonlinearity）函数 ReLU。

在训练梯度下降时，ReLU 比传统的饱和非线性函数有更快的收敛速度，因此在训练整个网络时，训练速度也比传统的方法快很多。4 种非线性操作函数的公式如下。

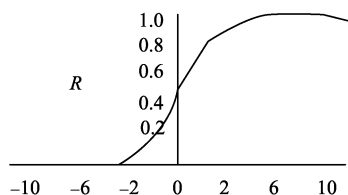
$$\text{sigmoid} : R = 1 / (1 + e^{-y})$$

$$\text{tanh} : R = e^y / (e^y + e^{-y})$$

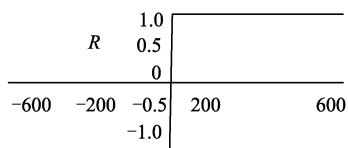
$$\text{softsign} : R = Y / (1 + |Y|)$$

$$\text{ReLU} : R = \max(0, Y)$$

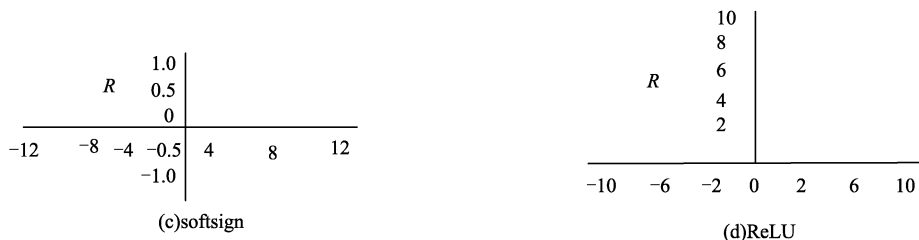
其函数形态如下图所示。



(a) sigmoid



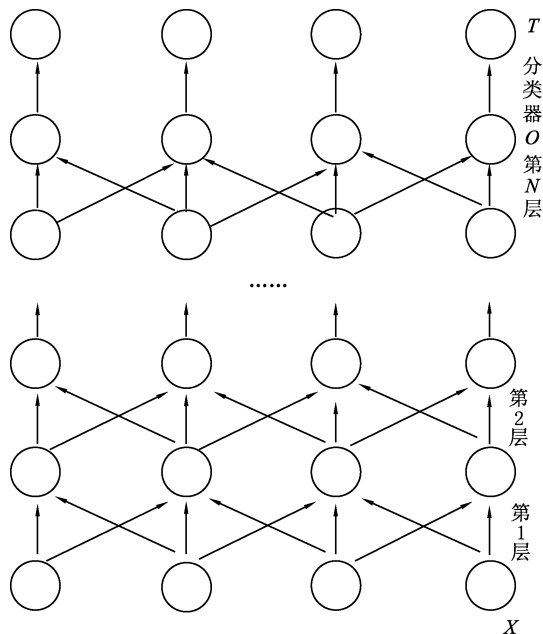
(b) tanh



4 种非线性操作函数

下采样阶段对每个特征图进行独立操作，通常采用平均池化或最大池化的操作。平均池化依据定义的邻域窗口计算特定范围内像素的均值（PA），邻域窗口平移步长大于 1（小于等于池化窗口的大小）；最大池化则将均值替换为最大值（PM）输出到下个阶段。池化操作完成后，输出特征图的分辨率降低，但能较好地保持高分辨率特征图描述的特征。一些卷积神经网络完全去掉下采样阶段，通过在卷积阶段设置卷积核窗口使滑动步长大于 1，以达到降低分辨率的目的。

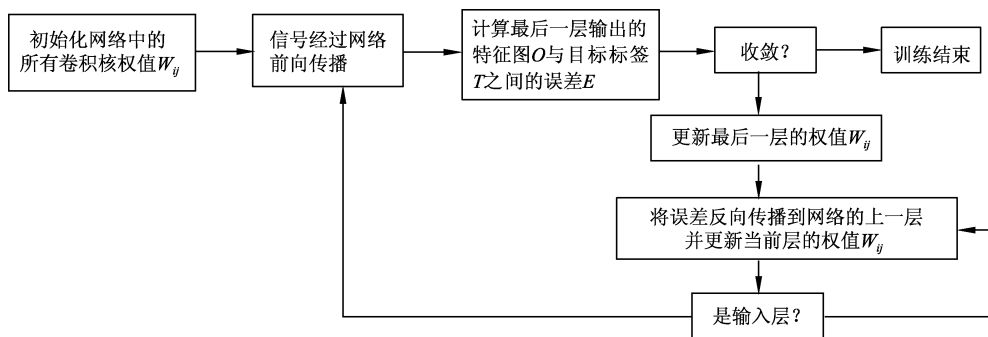
卷积神经网络的结构如下图所示。



卷积神经网络的结构

将单层的卷积神经网络进行多次堆叠，前一层的输出作为后一层的输入，便构成卷积神经网络。其中每 2 个节点间的连线代表输入节点经过卷积、非线性变换、下采

样 3 个阶段变为输出节点,一般最后一层的输出特征图后接一个全连接层和分类器。为了减少数据的过拟合,最近的一些卷积神经网络,在全连接层引入 Dropout 或 DropConnect 方法,即在训练过程中以一定概率 P 将隐藏层节点的输出值(对于 DropConnect 为输入权值)清零;而用反向传播算法更新权值时,不再更新与该节点相连的权值。但是这 2 种方法都会降低训练速度。在训练卷积神经网络时,最常用的方法是采用反向传播算法以及有监督的训练方式,算法流程如下图所示。

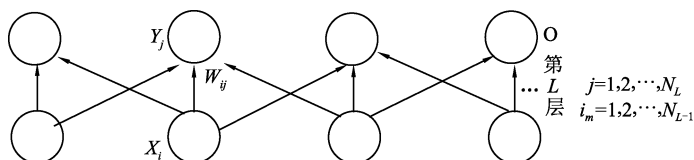


卷积神经网络的训练过程

网络中信号是前向传播的,即从输入特征向输出特征的方向传播,第 1 层的输入 X ,经过多个卷积神经网络层,变成最后一层输出的特征图 O 。将特征图 O 与目标标签 T 进行比较,生成误差 E 。通过遍历网络的反向路径,将误差逐层传递到每个节点,根据权值更新公式,更新相应的卷积核权值 W_{ij} 。在训练过程中,网络中权值的初值通常随机初始化(也可通过无监督的方式进行预训练),网络误差随迭代次数的增加而减少,并且这一过程收敛于一个稳定的权值集合,额外的训练次数呈现出较小的影响。对于卷积神经网络的任意一层 L ,其第 i 个输入特征 X_i 和第 j 个输出特征 Y_j 之间的权值 W_{ij} 的更新公式如下。

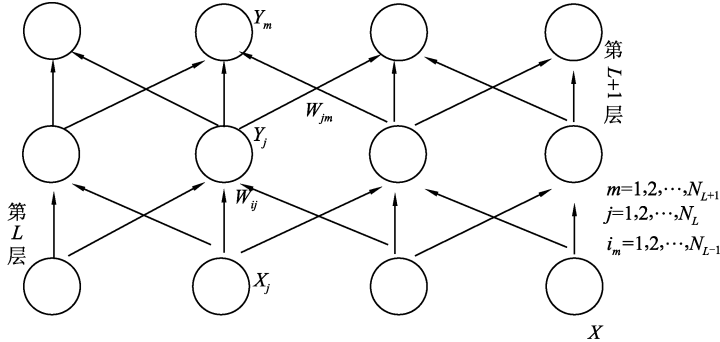
$$\Delta W_{ij} = \alpha \delta_j X_i$$

当第 L 层是卷积神经网络的最后一层时,结构如下图所示。

当第 L 层是最后一层时卷积神经网络的结构

$= (\quad - \quad) h' (\quad)$ 。为第 j 个目标标签； $h'(x)$ 为非线性映射函数的导数； $j = 1, 2, \dots, N_L$ 。

当第 L 层不是卷积神经网络的最后一层时，结构如下图所示。



当第 L 层不是最后一层时卷积神经网络的结构

$$\delta_j = h'_L(X_j) \sum_{m=1}^{N_{L+1}} \delta_m W_{jm}。$$

N_{L+1} 为第 $L+1$ 层输出特征的数目； $m = 1, 2, \dots, N_{L+1}$ ； W_{jm} 为第 L 层的第 j 个输出（作为 $L+1$ 层的第 j 个输入）与 $L+1$ 层第 m 个输出之间的权值。

卷积神经网络的特点在于，采用原始信号（一般为图像）直接作为网络的输入，避免了传统识别算法中复杂的特征提取和图像重建过程。获取的观测特征与平移、缩放和旋转无关。卷积阶段利用权值共享结构减少了权值的数量进而降低了网络模型的复杂度，这一点在输入特征图是高分辨率图像时表现得更为明显。同时，下采样阶段利用图像局部相关性的原理对特征图进行子抽样，在保留有用结构信息的同时有效减少数据处理量。

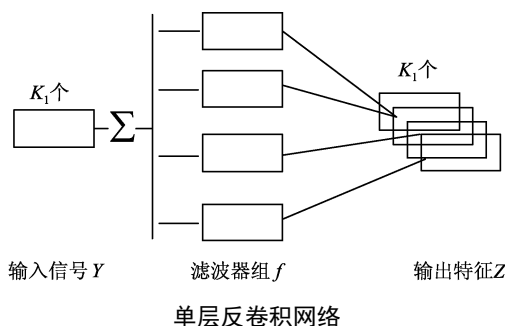
(2) 反馈深度网络。

反馈网络并不是对输入信号进行编码，而是通过解反卷积或学习数据集的积，对输入信号进行反解。前馈网络是对输入信号进行编码的过程，而反馈网络则是对输入信号进行解码的过程。

典型的反馈网络有反卷积网络、层次稀疏编码网络等，这里以反卷积网络为例进行讲解。M.D.Zeiler 等人提出的反卷积网络和 Yann LeCun 等人提出的卷积神经网络的思想类似，但在实际的结构构件和实现方法上有所不同。卷积神经网络是一种自底向上的方法，该方法每层输入的信号经过卷积、非线性变换和下采样 3 个阶段的处理，

进而得到多层信息。相比之下，反卷积网络的每层信息是自顶向下的，网络通过滤波器组学习得到卷积特征并重构输入信号。层次稀疏编码网络和反卷积网络非常相似，只是在反卷积网络中，分解图像采用矩阵卷积的方式，而在稀疏编码中，分解图像采用矩阵乘积的方式。

我们再来看单层反卷积网络。单层反卷积网络是通过先验学习，对信号进行稀疏分解和重构的正则化方法，其结构如下图所示。

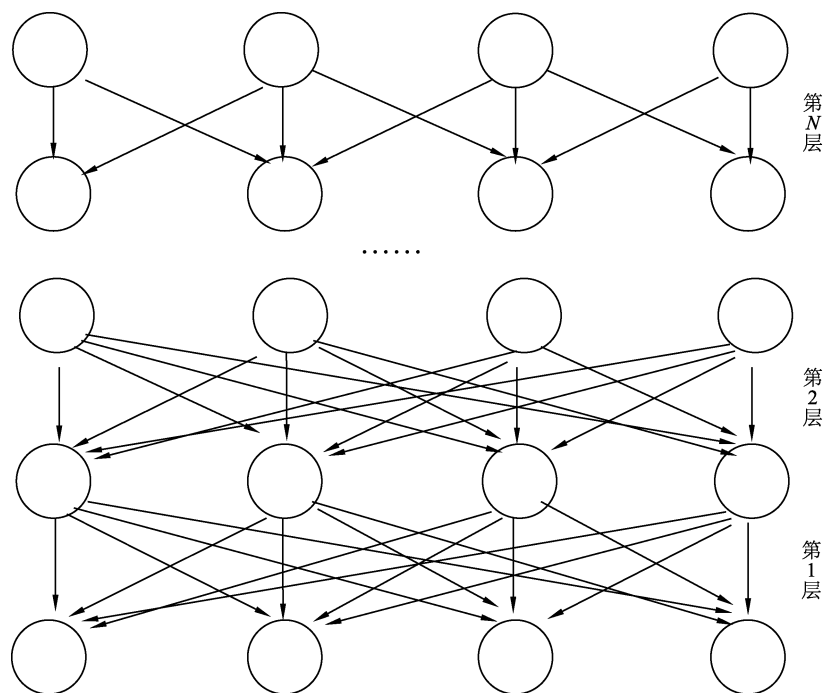


上图是一个单层反卷积网络，输入信号 Y 由 K_0 个特征通道 Y_1, Y_2, \dots, Y_{K_0} 组成，其中任意一个通道 Y_C 可看作 K_1 个隐藏层特征图 Z_K 与滤波器组 $f_{K,C}$ (个数为 $K_0 \times K_1$) 的卷积。

$\sum_{K=1}^{K_1} Z_K * f_{K,C} = Y_C$ 这个公式是一个欠定（未知数的个数多于方程个数）的函数，为了求得其唯一解，需要引入一个关于特征图 Z_K 的正则项，且该正则项使得特征图 Z_K 趋于稀疏。

反卷积网络通过上面所述单层反卷积网络进行多层叠加，可得到反卷积网络，如 129 页的图所示。

反卷积网络的特点在于通过求解最优化输入信号分解问题计算特征，而不是利用编码器进行近似，这样能使隐藏层的特征更加精准，更有利于信号的分类或重建。



反卷积网络的结构

(3) 双向深度网络。

双向深度网络由多个编码器层和解码器层叠加形成，每层可能是单独的编码过程或解码过程，也可能同时包含编码过程和解码过程。双向深度网络的结构结合了编码器和解码器2类单层网络结构，双向网络的学习则结合了前馈深度网络和反馈深度网络的训练方法，通常包括单层网络的预训练和逐层反向迭代误差2个部分，单层网络的预训练多采用贪心算法训练每层网络。网络结构中各层网络结构都经过预训练之后，再通过反向迭代误差对整个网络结构进行权值微调。其中单层网络的预训练是对输入信号进行编码和解码的重建过程，这与反馈深度网络的训练方法类似。而基于反向迭代误差的权值微调与前馈深度网络的训练方法类似。典型的双向深度网络有深度玻尔兹曼机、深度信念网络、栈式自编码器等。

2. 从产品经理的角度讲解深度学习技术

上面笔者对深度学习的各种主要神经网络技术进行了解析，下面笔者用综述的语言对深度学习技术进行深入浅出的讲解。我们已经知道深度学习作为机器学习领域的

一项新兴技术,其动机在于建立能够进行学习的神经网络。深度学习的本质是对观察到的数据进行分层特征表示,将低级特征表示抽象成高级特征表示。

深度学习主要借助神经网络技术,神经网络最基本的单元是神经元。对神经网络的研究开始得比较早,早期的感知器模型是神经网络的雏形,也被称为单层神经网络(无隐藏层)。然而,感知器只能做最简单的线性分类任务,甚至无法解决简单的异或问题。但是当网络增加一个计算层之后,不仅可以解决异或问题,而且具有非常好的非线性分类效果。1986年,Rumelhar和Hinton等人提出了反向传播(Back Propagation, BP)算法,解决了两层神经网络需要复杂计算的问题,从而带动了业界使用两层神经网络进行研究的热潮。

与单层神经网络不同。理论证明,两层神经网络可以无限逼近任意连续函数。也就是说,面对复杂的非线性分类任务,两层神经网络可以完成得很好。但是,神经网络仍然存在若干问题,尽管使用了BP算法,一次神经网络的训练仍然耗时太久,局部最优解仍是困扰训练优化的一个问题,这使得神经网络的优化较为困难。同时,隐藏层的节点数需要调参,使用起来不太方便,工程人员和研究人员对此多有抱怨,神经网络的研究再次进入低谷。

2006年,Hinton在*Science*和相关期刊上发表论文,首次提出“深度信念网络”概念。与传统的训练方式不同,深度信念网络有一个“预训练”过程,这可以让神经网络中的权值找到一个接近最优解的值,之后再使用微调技术对整个网络进行优化训练。这两项技术的运用大幅度减少了训练多层神经网络的时间。Hinton给多层神经网络相关的学习方法起了一个新名字——深度学习。

深度学习极大促进了机器学习的发展,受到了世界各国相关研究人员和互联网公司的重视,深度学习应用最广泛的三个领域是语音识别、图像识别和自然语言处理。

长期以来,语音识别系统大多采用高斯混合模型(GMM)来描述每个建模单元的概率。这种模型构造简单,方便使用大规模数据对其进行训练,且模型有较好的区分度训练算法,保证了该模型能够很好地被训练。但GMM实质上是一种浅层学习网络模型,特征的状态空间分布不能够被充分描述。而且,使用GMM建模,数据的特征维数通常只有几十维,这使得特征之间的相关性不能被充分描述。最后,GMM建模实质上是一种似然概率建模方式,一些模式之间的区别即使能够通过区分度训练得到,但是效果有限。

从 2009 年开始，微软亚洲研究院的语音识别专家开始和 Hinton 合作。2011 年，微软推出基于深度神经网络的语音识别系统，这一成果将语音识别领域已有的技术框架完全改变。采用深度神经网络以后，样本数据特征间的相关性信息得以充分表示，连续的特征信息结合构成高维特征，科研人员可以通过高维特征样本对深度神经网络模型进行训练。由于深度神经网络模拟了人脑的神经架构，因此，通过逐层进行数据特征提取，最终可以得到适合模式分类的理想特征。

图像识别是深度学习最早尝试应用的领域之一。在 2012 年的 ILSVRL 中，Krizhevsky 等人训练了一个大型的深度卷积神经网络 AlexNet，它包含 8 个学习层（5 个卷积层和 3 个全连接层）。Krizhevsky 等人借助 ImageNet 这个足够大的数据集、GPU 强大的计算能力、深层的 CNN 网络、随机梯度下降（SGD）和 Dropout 等优化技巧，最终赢得了 2012 年的 ILSVRC，实现了 Top 5 的错误率为 15.3% 的好成绩。这一成绩震惊了机器学习领域，促使大量的研究人员开始进入深度学习领域。2013 年，ZFNet 将 Top 5 的错误率降到了 11.2%。2014 年，Christian Szegedy 等人设计的 GoogLeNet 将 Top 5 的错误率降到了 6.7%。2015 年，微软亚洲研究院的何凯明等人设计了一个多达 152 层的 ResNet，它比之前所有的深度神经网络都深，但却降低了复杂度，从而更易于训练，ResNet 将 Top 5 的错误率降到了 3.6%。而人类识别的错误率是 5%~10%，深度学习在图像识别领域的分辨能力已经超过了人类。

自然语言处理是深度学习的另一个重要应用领域。数十年以来，自然语言处理的主流方法是基于统计的模型，人工神经网络虽然是基于统计方法的模型之一，但在自然语言处理领域却一直没有得到重视。美国 NEC 研究院最早将深度学习引入自然语言处理的研究中，其从 2008 年起采用将词汇映射到一维向量空间和多层一维卷积结构去解决词性标注、分词、命名实体识别和语义角色标注四个典型的自然语言处理问题。NEC 研究院构建了一个网络模型用于解决四个不同问题，都取得了相当精确的结果。总体而言，深度学习在自然语言处理上取得的成果还无法和其在图像识别、语音识别方面取得的成果相比。

深度学习主要利用了神经网络技术，因此，深度学习所面临的技术挑战主要来自神经网络所面临的问题。随着层数的不断增加，神经网络中的各种问题开始涌现，笔者在这里将问题总结为以下几点。

可用的、带标签的数据集较少。深度学习是一种监督学习，监督学习通过训

训练数据集学习得到一个分类模型,并将该分类模型应用到对未知数据(测试数据)的分类中去。因此,分类模型的好坏强烈依赖于训练数据集的数量和质量,一个好的训练数据集能够得到一个好的分类模型,从而能够对未知数据得到一个较好的分类结果。截至 2018 年,可用的、带标签的训练数据集仍比较少,这可能会限制深度学习的进一步发展。2012 年的 ILSVRL 提供了包含超过 22000 个类别、1500 万张高分辨率照片的数据集,这使得 AlexNet 的成功成为可能。但是随着深度神经网络的进一步加深(例如,2015 年的 ResNet 已经达到惊人的 152 层),训练这样的一个深层的神经网络需要更大的数据集。

网络复杂度的上升和可解释性的降低。深度学习的一个很大的局限性是神经网络的可解释性不强,有人形象地把神经网络比喻成一个黑盒,就是因为它的内部知识的表示很不直观。因此,提取隐藏在训练数据集背后的知识以及如何表示这些知识也是一个很重要的研究方向。

随着深度学习的进一步发展,神经网络的复杂度也在逐渐上升。2012 年, AlexNet 的层数达到了 8 层,而其参数和神经元的数量是巨大的, AlexNet 具有 6000 万个参数和 65 万个神经元,其复杂度可想而知。并且,越来越多的参数不仅需要更大的数据集,还需要更强的计算能力,否则神经网络的训练时间将是不可忍受的。

过拟合。训练神经网络的过程其实就是不断调整网络中神经元之间的连接权重,使得网络预测的输出值与实际值之间的误差最小化的过程,这种过程不断进行下去,直到权重收敛。在这个过程中,网络的训练误差不断减小,在未知数据上的测试误差也在不断减小,但当训练误差减小到某个值之后,预测误差反而增大,这说明模型的泛化能力降低,这就是神经网络的过拟合现象。最近几届 ILSVRC 中提出的神经网络都不同程度地对过拟合提出了解决方案,由此可见过拟合给神经网络带来的重大影响。

在神经网络中,数据扩增和 Dropout 是两种常用的正则化方法。数据扩增其实就是获取更多的数据,然而,数据通常是难以获取的、昂贵的。在很多时候,收集更多的数据意味着需要耗费更多的人力、物力。在原始数据上做些改动,以获得更多的数据似乎是个不错的选择。以图片数据集为例,可以将图片做各种变换,比如将原始图片旋转一个角度、添加随机噪声、截取原始图片的一部分等。Dropout 每次将一半的隐藏层随机删除,一直迭代,直到训练结束。可以这样解释,运用了 Dropout 的训练过程,相当于训练了很多个只有一半数量隐藏层的神经网络,每个这样的“半数网

络”都可以给出一个分类结果，这些结果有的是正确的，有的是错误的。随着训练的进行，大部分“半数网络”都可以给出正确的分类结果，少数的错误分类不会对最终结果造成大的影响。

局部最优解。神经网络的训练存在多个极小点，梯度下降不能保证求出全局最优解，为解决这个问题，许多专家提出了多种方案，如使用改进的BP算法、增加动量项、自适应学习率、和别的优化算法组合等。

梯度消失现象。我们常常使用sigmoid作为神经元的输入、输出函数，对于幅度为1的信号，在反向传播梯度时，每传递一层，梯度衰减为原来的0.25。随着层数的增多，梯度指数衰减后，低层基本接收不到有效的训练信号，这就是梯度消失现象。对此，我们引入ReLU函数代替sigmoid函数。采用sigmoid等函数计算激活函数（指数运算）时，计算量大；用反向传播算法求误差梯度时，求导涉及除法，计算量也很大；而采用ReLU函数时，整个过程的计算量小很多。对于深层网络，sigmoid函数反向传播时，很容易出现梯度消失的情况（在sigmoid接近饱和区时，变换太缓慢，导数趋于0，这种情况会造成信息丢失），从而无法完成深层网络的训练。而ReLU函数会使一部分神经元的输出为0，这就造成了网络的稀疏性，并且减少了参数的相互依存关系，避免了过拟合问题的发生。

无监督学习因不需要带标签的训练数据集而被广泛应用，但是在深度学习中，它已经被监督学习的成功而覆盖。但是，无监督学习还是应该引起我们的关注，很多时候，人类的意识是基于对周围事物的观察，而不是基于训练。

3.1.8 生物特征识别技术

所谓生物特征识别技术就是将计算机技术与光学、声学、生物传感器和生物统计学原理等高科技手段密切结合，利用人类固有的生理特性（如指纹、人脸、虹膜等）和行为特征（如笔迹、声音、步态等）来进行个人身份的鉴定。

身份认证在人类社会生活中早已有之。传统的身份认证方法（如使用身份证、护照、钥匙、智能卡、密码、口令等）存在携带不便、容易伪造、容易遗失、容易因使用过多或不当而损坏或不可读、密码易被破解等诸多问题，其安全性低、可靠性差。而每个人所固有的生物特征，具有与其他人不同的唯一性和在一定时期内不变的稳定性，而且不会丢失、不易伪造和假冒，所以被认为是终极的身份认证媒介。

生物特征识别在国家安全、公安、司法、金融等领域有着广泛的应用，如重要部门的合法进入、计算机网络的登录控制、信用卡持卡人的身份验证、犯罪嫌疑人的身份辨认等。早在 20 年前，IBM、微软等公司就成立了 BioAPI 联盟，其宗旨是制定生物特征识别应用程序接口的工业标准。生物特征识别技术备受世界各国的重视，“9·11”恐怖袭击事件发生之后，以美国为首的西方国家将生物特征识别技术作为关系国家安全的关键技术加以发展。美国连续发布了 3 个法案，强调在边境检查、执法、民用航空等领域应用生物特征识别技术。2003 年 6 月，在国际民用航空组织（ICAO）发布的规划中，也建议其 188 个成员国在护照中加入生物特征识别模块，大部分西方国家已经立法支持 ICAO 的这项规划。从 2005 年开始，多个国家开始发放“生物特征护照”（又被称作“电子护照”），它存储了一些可用于识别旅客身份的生物特征信息。

仅在 2015 年，美国就发放了超过 1550 万本护照和护照卡。

2016 年，荷兰的一位企业家在斯德哥尔摩使用植入式芯片成功登上了飞机。

截至 2018 年，护照标准仍由 ICAO 管控。2015 年，ICAO 在 9303 号文件的修订版中规定了可机读护照和生物特征护照的格式。未来，如果我们真的要把纸质护照换成植入式芯片之类的东西，主导这次转换的很可能还是 ICAO。

我国幅员辽阔、人口众多，未来必将是生物特征识别技术的应用大国，开展生物特征识别技术的研究对信息安全和国家安全具有重要意义。

可供选择的人体所固有的生物特征有许多，能够用来认证身份的生物特征应该具有以下特性。

普遍性：每个正常人都应该具有这种特征。

唯一性：不同的人应该具有各不相同的特征。

可采集性：所选择的特征可以定量测量。

稳定性：所选择的特征至少在一段较长的时间内是不变的，并且特征的采集不随条件、环境的变化而变化。

在实际的生物特征识别系统中，还应考虑如下因素。

性能：识别的准确性、速度、鲁棒性以及为达到所要求的准确性和速度所需要的资源。

可接受性：使用者在多大程度上愿意接受这种生物特征识别系统。

安全性：用欺诈的方法骗过系统的难易程度。

理论依据：是否具有相关的、可信的研究作为技术支持。

提取的特征容量、特征模板是否占用较小的存储空间。

造价：价格是否在用户可接受的范围之内。

隐私保护：是否具有非侵犯性。

用于身份认证的生物特征可以分为两类：生理特征和行为特征。生理特征与生俱来，多为先天的，包括指纹、掌纹、手形、脸型、虹膜、视网膜、耳郭、DNA（脱氧核糖核酸）等；行为特征则是习惯使然，多为后天的，包括笔迹、步态、击键动作等。

我们一起看看常见的生物特征识别技术。

1. 指纹识别

指纹是指人的手指末端正面皮肤上凸凹不平产生的纹线。纹线有规律的排列形成不同的纹型。纹线的起点、终点、结合点和分叉点被称为指纹的细节特征点。指纹识别就是通过比对指纹的细节特征点来进行鉴别。由于指纹具有终身不变性、唯一性和方便性，几乎已经成为生物特征识别的代名词。

其实，我国古代就利用指纹（手印）来证明个人身份（画押）。随着计算机技术的发展，美国联邦调查局（FBI）和法国巴黎警察局于 20 世纪 60 年代开始研究开发自动指纹辨认系统（Automatic Fingerprint Identification System, AFIS）用于刑事案件的侦破。20 世纪 90 年代，用于个人身份认证的自动指纹识别系统得到开发和应用。自动指纹识别系统通过特殊的光电转换设备和计算机图像处理技术，对指纹进行采集、分析和比对，可以自动、迅速、准确地鉴别出个人身份。指纹识别技术主要包括指纹图像增强、特征提取、指纹分类和指纹匹配几个部分。

指纹图像增强。一般包括规格化、方向图估计、频率图估计、生成模板、滤波等几个环节。Hong 等人采用同时具有频率选择和方向选择的 Gabor 滤波器来增强指纹图像，能在指纹图像质量很差的情况下取得很好的效果，并减少了计算局部区域方向图的开销。

特征提取。最常用的细节特征的定义是由 FBI 提出的，它将指纹图像的最显著特征分为脊终点和分叉点，每个清晰指纹一般有 40~100 个这样的细节点。指纹特

征的提取采用链码搜索法对指纹纹线进行搜索，自动指纹辨认系统依赖于这些局部脊的特征及关系来鉴别身份。

指纹分类。常见的有基于神经网络的分类方法、基于奇异点的分类方法、基于脊线几何形状的分类方法、基于指纹方向图分区和遗传算法的连续分类方法。

指纹匹配。指纹匹配是指纹识别系统的核心步骤，匹配算法包括图匹配、结构匹配等，但最常用的方法是用 FBI 提出的基于细节模型来做细节匹配（即点模式匹配）。

指纹识别是国际公认的应用广、造价低、易用性高的身份认证技术。市场上推出的产品有指纹门锁、指纹考勤机、指纹门禁机、指纹保险箱、指纹储物柜、指纹无线鼠标、指纹 U 盘，以及内置指纹识别装置的手机、平板电脑、笔记本电脑等。十多年前，把指纹识别技术同 IC 卡结合起来是最有前景的应用方向，到 2018 年，最有前景的应用方向是指纹支付、扫手开柜门取货等应用。

2. 虹膜识别

虹膜是眼球前部含色素的环形薄膜，由结缔组织细胞、肌纤维组成。虹膜图像中含有极其丰富的结构和纹理特征，作为生物识别特征，虹膜具有如下优势。

虹膜在妊娠 3 个月的时候开始形成，整体结构在妊娠 8 个月的时候开始创建，大概在 2~3 岁的时候稳定，以后不随年龄的增长而变化。

每个人的虹膜纹理在人群中的分布是随机的或者说是混沌的，虹膜的形成依赖胚胎中胚层的初始条件，不受遗传的影响，即使拥有同样的基因型（如同卵双胞胎，甚至是同一个人的两只眼睛），二者的虹膜也是不相关的，因此虹膜可以说是唯一的。

虹膜的内部组织被水样液和角膜所包围，与外界环境隔离开来，不易受损。

不易被假冒，一般的外科手术不能改变虹膜的纹理。

在识别的过程中，不需要和被识别者物理接触，具有非侵犯性。

虹膜具有活体组织的特点，由于虹膜肌肉间复杂的相互作用，瞳孔直径一直在小的范围内有规律地变化，且随光线强度变化而收缩，使得虹膜具有高度的防伪性。

1987 年，Flom 和 Safir 首先提出了自动虹膜识别的思想。1993 年，剑桥大学的

Daugman 提出了基于 2D Gabor 变换的虹膜识别方法，其主要思想是：首先构造二维 Gabor 滤波器，然后用它对图像进行滤波，获得相位信息之后，根据相位所在的象限编成 256 字节的相位码。Daugman 的开创性工作使得自动虹膜识别成为可能，商业虹膜识别产品的核心软件大多基于 Daugman 的算法。

近年来，越来越多的科研工作者正投身于虹膜识别产品国产化的研究中，虽然起步较晚，但是国内的部分研究成果已经达到了国际领先水平。中国科学院自动化研究所的模式识别国家重点实验室是国内最早从事虹膜识别研究的单位之一。2006 年 9 月，模式识别国家重点实验室作为中国虹膜识别技术的权威，参加了由国际生物识别组织举办的生物识别技术测评，其虹膜识别算法的速度和精度得到了国际同行的认可。此外，模式识别国家重点实验室的虹膜图像数据库已经成为国际上规模最大的虹膜信息共享库。我国的虹膜识别系统从硬件到软件都拥有完全自主知识产权，突破了西方国家的技术垄断与封锁。

张鹏飞等虹膜识别专家提出了一种基于多纹理特征融合的新的虹膜识别方法。该方法对虹膜图像进行 Gabor 小波变换后提取不同分辨率、不同方向下的纹理特征作为虹膜的全局特征，在滤波后的子窗口图像上运用灰度级共现矩阵提取虹膜的局部特征。通过加权欧几里得距离和最小距离分别对全局特征和局部特征进行分类识别，设计了基于模糊推理系统的全局特征和局部特征融合方法，以提高虹膜识别系统的鲁棒性。

3. 声纹识别

声纹是一项根据语音波形反映说话人的生理、心理和行为特征的语音参数。人类语言的产生是人体语言中枢与发音器官共同作用的一个复杂的生理物理过程，人在讲话时使用的发声器官（舌、牙齿、喉头、肺、鼻腔）在尺寸和形态方面各不相同，所以任何两个人的声纹图谱都有差异。每个人的语音声学特征既有相对稳定性，又有变异性。这种变异可能来自生理、病理、心理、模拟、伪装，也与环境干扰有关。尽管如此，由于每个人的发音器官都不相同，因此，在一般情况下，人们仍能区别不同的人的声音或判断某些声音是否属于同一个人。声纹识别也叫说话人识别，可以看作语音识别的一种。但它不同于语音识别，并不注重语音信号中的语义内容，而是希望从语音信号中提取人的特征。在处理方法上，语音识别力图对不同人说话的差别进行归一化，而说话人识别则力图强调不同人之间的区别。

声纹识别可分为与文本相关的声纹识别和与文本无关的声纹识别两种。与文本相关的声纹识别要求用户按照规定的内容发音，每个人的声纹模型被逐个精确地建立，而识别时也必须按规定的内容发音，因此可以达到较好的识别效果，但需要用户配合，如果用户的发音与规定的内容不符合，则无法正确识别用户。而与文本无关的声纹识别则不规定说话人的发音内容，建立模型比较困难，但用户使用方便，可应用范围较宽。根据特定的任务和应用，两种声纹识别方式具有不同的应用范围。比如，在银行交易时可以使用与文本相关的声纹识别，因为用户自己进行交易时是愿意配合的；而在刑侦或侦听中则无法使用与文本相关的声纹识别，因为你无法要求犯罪嫌疑人或被侦听的人配合。

声纹识别涉及两个关键问题：一个是特征提取，另一个是模式匹配。特征提取的任务是提取并选择说话人的可分性强、稳定性高的声学特征或语言特征。与语音识别不同，声纹识别的特征必须是“个性化”特征。虽然大部分声纹识别系统用的都是声学层面的特征，但是表征一个人特点的特征应该是多层面的，包括：与人类发音机制的解剖学结构有关的声学特征（如频谱、倒频谱、共振峰、基音、反射系数等）、鼻音、深呼吸音、沙哑音、笑声等；受社会经济状况、受教育水平、出生地等影响的语义、修辞、发音、言语习惯等；或受父母影响的韵律、节奏、速度、语调、音量等特征。

从建模的角度出发，声纹识别模型可以使用的特征包括：声学特征；词法特征；韵律特征；语种、方言和口音信息；声道信息等。

根据不同的任务需求，声纹识别还面临特征选用的问题。例如，在刑侦中，希望不使用声道信息，也就是说希望弱化声道对识别的影响，因为我们希望不管说话人使用什么声道，系统都可以辨认出来。而在银行交易中，希望使用声道信息，即希望声道对识别有较大影响，从而剔除录音、模仿等行为带来的影响。

总之，较好的特征应该能够有效区分不同的说话人，但又能在同一说话人语音发生变化时保持相对稳定；不易被他人模仿或能够较好地解决被他人模仿的问题；具有较好的抗噪性能等。当然，这些问题也可以通过模型方法去解决。对于模式匹配，有以下几大类方法。

(1) 模板匹配方法：利用动态时间规整（DTW）进行对准训练和特征序列测试，主要用于固定词组的匹配（通常为文本相关任务）。

(2) 最近邻方法：训练时保留所有特征矢量，识别时对每个矢量都找到训练矢量中最近的 K 个，据此进行识别。

(3) 神经网络方法：有很多种形式，如多层感知、径向基函数 (RBF) 等，可以显式训练以区分说话人和背景说话人，其训练量很大，且模型的可推广性不好。

(4) 隐马尔可夫模型方法：通常使用单状态的隐马尔可夫模型，或高斯混合模型，是比较流行的方法，效果比较好。

(5) VQ 聚类方法：效果比较好，算法也不复杂，和隐马尔可夫模型方法配合可以得到更好的效果。

(6) 多项式分类器方法：有较高的精度，但模型存储和计算量都比较大。

声纹识别需要解决的问题还有很多，如以下问题。

(1) 有限的训练及测试样本问题，即在声音不易获取的应用场合，能否用很短的语音进行模型训练，而且用很短的时间进行识别。

(2) 声音模仿（或播放录音）问题，即怎样有效区分开模仿的声音（录音）和真正的声音。

(3) 在有多个说话人说话的情况下，怎样有效提取目标说话人的声纹特征。

(4) 怎样消除或减弱声音变化（不同语言、内容、身体状况、时间、年龄、情绪等）带来的影响。

(5) 鲁棒性问题，即怎样消除声道差异和背景噪声带来的影响等。

4. 人脸识别

人脸识别具有主动性、非侵犯性和对用户友好等优点。

人脸识别系统主要包括四个组成部分，分别为：人脸图像采集及检测、人脸图像预处理、人脸图像特征提取以及人脸图像识别与匹配。

(1) 人脸图像采集及检测。

人脸图像采集：不同的人脸图像都能通过摄像机采集下来，比如静态图像、动态图像、不同位置的图像、不同表情的图像等都可以得到很好的采集。当用户在采集设备的拍摄范围内时，采集设备会自动搜索人脸并拍摄人脸图像。

人脸检测：人脸检测在实际中主要用于人脸图像的预处理，即在图像中准确标定

出人脸的位置和大小。人脸图像中包含的模式特征十分丰富,如直方图特征、颜色特征、模板特征、结构特征及 Haar 特征等。人脸检测就是把其中有用的信息“挑”出来,并利用这些特征实现人脸检测。

主流的人脸检测方法基于以上特征并采用 Adaboost 算法。Adaboost 算法是一种用来分类的方法,它把一些比较弱的分类方法组合在一起,形成新的很强的分类方法。

在人脸检测过程中使用 Adaboost 算法挑选出一些最能代表人脸的矩形特征(弱分类器),按照加权投票的方式将弱分类器构造为一个强分类器,再将训练得到的若干强分类器串联组成一个级联结构的层叠分类器,以有效提高分类器的检测速度。

(2) 人脸图像预处理。

对人脸图像的预处理基于人脸检测结果,对图像进行处理并最终服务于特征的提取过程。系统获取的原始图像由于受到各种条件的限制和随机干扰,往往不能直接使用,必须在图像处理的早期阶段对它进行灰度校正、噪声过滤等图像预处理工作。对于人脸图像而言,其预处理过程主要包括人脸图像的光线补偿、灰度变换、直方图均衡化、归一化、几何校正、滤波以及锐化等。

(3) 人脸图像特征提取。

人脸图像可使用的特征通常分为视觉特征、像素统计特征、人脸图像变换系数特征、人脸图像代数特征等。人脸图像特征提取就是针对人脸的某些特征进行的。人脸图像特征提取也被称作人脸表征,它是对人脸进行特征建模的过程。人脸图像特征提取的一种重要方法是基于知识的表征方法。基于知识的表征方法主要根据人脸器官的形状以及它们之间的距离特性来获得有助于人脸分类的特征数据,其特征分量通常包括特征点间的欧氏距离、曲率和角度等。

(4) 人脸图像匹配与识别。

人脸图像匹配是指将提取的人脸图像的特征数据与数据库中存储的特征数据进行搜索匹配,并设定一个阈值,当相似度超过这一阈值时,就把匹配得到的结果输出。人脸图像识别就是将待识别的人脸特征与已得到的人脸特征模板进行比较,根据相似程度对人脸的身份信息进行判断。这一过程又分为两类:一类是确认,即一对一进行图像比对的过程;另一类是辨认,即一对多进行图像比对的过程。

5. 步态识别

步态识别是一个新的发展方向，它指从相同的行走行为中寻找和提取个体的步态特征，以自动地实现身份识别。在智能视频监控场景中已经广泛应用的自动步态识别系统，融合了计算机视觉、模式识别、视频/图像序列处理等技术。

首先，由监控摄像头采集人的步态，通过检测与跟踪获得包含步态的视频帧，经过对采集的视频帧进行预处理，分析提取该人的步态特征。即对图像序列中的步态进行运动检测、运动分割、特征提取等步态识别前期的工作。

其次，经过进一步处理，使其成为与已存储在数据库中的步态相同的模式。

最后，将新采集的步态特征与数据库中的步态特征进行比对、识别，匹配成功即进行报警；匹配失败则监控摄像机继续进行步态的采集。

因此，一个智能视频监控的自动步态识别系统，实际上主要由监控摄像机、计算机、一套性能良好的步态视频序列处理与步态识别算法组成。其中，最关键的是步态识别算法。

自“9·11”事件以来，远距离身份识别技术的相关研究备受关注。与其他生物特征识别技术相比，步态识别的突出特点是能远距离识别。因此，步态识别的研究已经越来越引起国内外学者的关注。已研究出的步态识别算法有如下几种。

(1) 二维步态识别算法。

对于每个步态序列而言，一种改进的背景减除技术被用来提取人的空间轮廓。这些轮廓的边缘，按照逆时针方向展开为一系列相对于质心的距离模板特征。这些模板特征通过使用主元统计分析方法来训练，从而得出步态形状的变化模式在特征空间中的轨迹表达。识别时，采用时空相关匹配方法和基于归一化欧氏距离的最近邻规则，并引入个人的体形等生理特征，以用于必要的步态分类、校验。

(2) 行走步态识别算法。

该算法源于“从行走运动的时空模式中可得到人体的外观”的观点。对每个序列而言，背景减除过程用来提取行人的运动轮廓，这些轮廓在二维空间中被描述为一个序列的复数配置，并随时间变化。利用 Procrustes 形状分析方法，从该序列的复数配置中获取主轮廓模型作为人体的静态外观特征。实验结果表明，该算法获得了令人满意的识别性能。

(3) 时空轮廓分析算法。

该算法源于“行走运动导致人体轮廓随着时间的变化而变化”的直观想法。对于每个序列而言,背景减除与轮廓生成相关算法用于检测和跟踪行人的运动轮廓,这些随时间变化的二维轮廓形状被转换为对应的一维距离信号,可以通过特征空间变换来提取步态特征,并基于时空相关或归一化欧氏距离度量,以及标准的模式分类技术进行最终的识别。实验结果表明,该算法不仅获得了令人满意的识别性能,而且对计算能力的要求较低。

(4) 基于模型的算法。

该算法源于“行走运动的关节角度变化包含着丰富的个体识别信息”的思想。首先,结合人体模型、运动模型和运动约束等先验知识,利用 Condensation 算法对行人进行跟踪。然后,从跟踪结果中获取人体主要关节的角度变化轨迹。这些轨迹经过结构归一化和时间归一化后,作为动态特征用于身份识别。

(5) 基于 Hough 变换的算法。

这是一种基于新的特征提取方法的自动步态识别算法,该算法仅从腿部的运动进行身份识别。对于每个序列,用一种基于图像色度偏差的背景减除算法来检测运动对象。在经过后处理的二值图像序列中,利用边界跟踪算法获取对象边界,在对象边界图像上,局部应用 Hough 变换检测大腿和小腿的直线,从而得到大腿和小腿的倾斜角。用最小二乘法将一个周期内的倾斜角序列拟合成 5 阶多项式,把按傅里叶级数展开得到的相位与振幅的乘积定义为低维步态特征向量。在小样本的数据库上用 Fisher 线性分类器验证所研究算法的性能,正确分类率为 79.17%,在步态数据库不是很理想的情况下,算法也取得了较好的识别率。

(6) 基于三维小波矩理论的算法。

广义多尺度分析理论针对不同的应用图像或信号库,得到最优小波分解,并在人体步态识别中与二维小波矩结合进行应用。在三维物体的表示方面,三维小波矩理论作为三维物体的一种无冗余的描述和识别方法。与现存的方法相比,它不但具有平移、缩放和旋转不变的特性,在径向上还增加了多尺度分析的特性。可以根据不同的需要,提供多层次的特征描述,在引进球面调和函数加速算法和 Mallat 算法后,可以使小波矩的计算得到双重加速。有人计划搭建实用的三维物体检索平台,以便进

一步完善该算法。

此外，有人基于“人体生物特征不仅包含静态外观信息，也包含行走运动的动态信息”的思想，提出了一种在判决级上融合人体静态特征和动态特征的身份识别方法。此方法在不同融合规则下的实验结果表明，融合后的识别性能优于使用任何单一模态的识别性能。

6. 笔迹识别

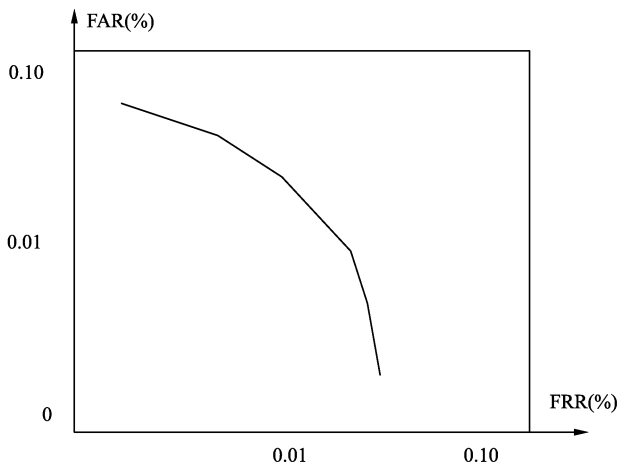
笔迹是书写者自身的生理特征和后天学习过程的综合反映。笔迹识别也是人们进行身份认证的重要手段之一。

计算机笔迹识别主要分为脱机识别和联机识别两类。脱机笔迹识别的对象是写在纸上的字符，通过扫描仪、摄像机等设备将字符输入计算机以后，进行分析与鉴定；而联机笔迹识别则通过专用的数字板或数字仪实时采集书写信号，除了可以采集签名位置等静态信息，还可以采集书写速度、运笔压力、握笔倾斜度等动态信息。显然，较脱机笔迹识别而言，联机笔迹识别可利用的信息更多，同时识别难度也更大。

笔迹识别主要研究两类问题，即文本相关的笔迹识别问题和文本无关的笔迹识别问题。文本相关的笔迹识别对于书写内容有限制，因此相对容易解决，然而应用范围较窄；而文本无关的笔迹识别，由于对书写内容不加限制，问题变得更加复杂。截至本书完稿时，对于文本无关的笔迹识别，还没有准确而高效的识别系统投入市场。

7. 生物特征识别系统的性能指标

衡量一个生物特征识别系统的两个重要性能指标是：误拒率（False Rejection Rate，简称 FRR；或 False Non-Match Rate，简称 FNMR）和误认率（False Acceptance Rate，简称 FAR；或 False Match Rate，简称 FMR）。FRR 是指系统拒绝真正的生物特征拥有者的错误率；FAR 是指系统将冒充者误认为生物特征拥有者的错误率。对于理想的系统来说，这两个错误率都应该是零。但在实际应用中，这两个指标是相关的，当 FRR 比较低时，FAR 就会比较高；反之亦然。系统往往需要在 FRR 和 FAR 之间取一个平衡点。用 ROC（Receiver Operating Characteristic）曲线能够很好地反映 FRR 和 FAR 之间的关系，ROC 曲线如下图所示。



ROC 曲线

曲线上的点表示在某个给定的匹配阈值下得到的 FRR 和 FAR。在现有的技术水平下，无法使 FRR 和 FAR 同时达到最小，需要调整匹配阈值来满足不同应用的要求。我们把调整匹配阈值称为调整“操作点”。好的系统应该允许对操作点进行自由调整。

身份认证涉及身份辨认和身份验证两方面的应用。以指纹识别为例，身份验证就是通过把一个现场采集到的指纹与一个已经登记的指纹进行一对一比对来鉴定身份的过程。作为验证的前提条件，指纹必须在指纹库中已经登记，且以一定的压缩格式存储，并与其姓名或其标识（如身份证号）联系起来，根据所提供的标识来鉴定身份的合法性。验证其实是回答“他是他自称的这个人吗？”这样的问题。在银行交易等安全性要求高的应用场合，用到的是身份验证。身份验证通常采取“宁漏勿错”的策略，因此应尽量减小 FAR，而放宽 FRR 指标。身份辨认则是把现场采集到的指纹同指纹模板数据库中的指纹进行逐一比对，从中找出与现场指纹相匹配的指纹，这是“多选一”问题。身份辨认其实回答了“他就是我们要找的这个人吗？”这样的问题。在刑侦应用中，通常将一个身份不明的人的指纹与指纹库中有犯罪记录的人的指纹进行比对，以确定此人是否有犯罪记录，采取“宁错勿漏”的策略，此时应尽量减小 FRR，而放宽 FAR 指标。FRR 实际上也是系统易用性的重要指标。由于 FRR 和 FAR 是相互矛盾的，这就使得在应用系统中，要权衡易用性和安全性。对于指纹识别系统来说，一个有效的办法是比对两个或更多个指纹，从而在不损失易用性的同时，极大地提高

系统的安全性。

8. 多模式生物特征识别技术

迄今为止，还没有一种单独的生物特征识别技术可以满足所有实际的需求。单个生物特征有其固有的局限性，例如，一部分人是没有指纹的，此时单靠指纹识别就有问题，戴隐形眼镜的人不能单靠虹膜识别来认证，用人脸识别技术无法区分双胞胎，等等。而且，基于单个生物特征的身份认证系统所能达到的准确率是有限的。表 1 给出了几种生物特征识别技术所能达到的 FRR 和 FAR，表 2 给出了实际应用所要求的 FRR 和 FAR。

表 1 几种生物特征识别技术所能达到的 FRR 和 FAR

| 生物特征 | FRR (%) | FAR (%) |
|------|---------|---------|
| 人脸 | 4 | 10 |
| 指纹 | 2.5 | 0.01 |
| 手形 | 1.5 | 1.5 |
| 虹膜 | 6 | 0.001 |
| 语言 | 15 | 3 |

表 2 实际应用所要求的 FRR 和 FAR

| 应用 | FRR (%) | FAR (%) |
|------|---------|---------|
| 身份验证 | 0.1 | 0.1 |
| 身份辨认 | 10.0 | 0.0001 |

许多学者致力于多模式生物特征识别技术的研究，通过融合多种生物特征，来提高识别的准确率、扩大应用范围，以满足实际的需求。Brunelli 和 Falavinga 于 1995 年提出了利用多种特征进行个人身份认证的方法，并在匹配层次上整合了人脸识别和说话人识别，取得了较好的识别效果。该方法实现的难点在于需要准确构建正负样本集，选择若干映射参数，这在小样本情况下有很多困难。Kittler 等人提出了融合理论框架，同时比较了加法准则和乘法准则等算法在融合中的优缺点。Jain 等人于 1998 年提出将指纹识别与人脸识别的结果相融合，于 2000 年提出确定每个用户的特定参数的方法。

2005 年，由清华大学电子系自主研发的“TH-ID 多模式生物特征身份认证识别系统”通过教育部组织的专家鉴定。该系统包括两大部分内容：利用人脸、笔迹、

签字和虹膜 4 种生物特征进行身份认证（识别和验证）的 4 个子系统；融合多种生物特征进行身份认证的系统。他们构建了基于统一数据库的人脸、笔迹、签字、虹膜 4 种生物特征的多模式生物特征融合的身份识别认证系统，可以对融合模式进行选择。例如，人脸与签字融合的身份识别/验证系统，人脸与虹膜融合的身份识别/验证系统等。这可以有效克服单一生物特征常有的缺陷，极大地提高身份认证的准确度，从而也为利用生物特征进行身份认证的实际应用打下了坚实的基础。这一技术已被香港伟清创新科技有限公司等数家公司支持并进入产品开发阶段，集成这一技术的人脸认证旅客自助通道系统已经在深圳罗湖口岸正式投入使用，通过率超过 97%。

小结：在信息时代，如何准确鉴定一个人的身份、保护信息安全，已成为一个必须解决的关键的社会问题。传统的身份认证媒介由于极易伪造和丢失，越来越难以满足社会的需求，最为便捷与安全的解决方案无疑就是生物特征识别技术。它不但简单快速，而且利用它进行身份认定，安全、可靠、准确，同时更易于和公共安全、公安监控等系统整合，实现自动化管理。由于其广阔的应用前景、巨大的社会效益和经济效益，已引起各国的广泛关注和高度重视。

3.1.9 机器人技术

机器人技术是 AI 技术的一种，机器人是自动执行工作的机器装置。它既可以接受人类指挥，又可以运行预先编写的程序，还可以根据 AI 设备的行动纲领行动。它的任务是协助或取代人类的工作，例如生产业、建筑业中的高强度工作、危险的工作。

1. 机器人的构成

机器人一般由执行机构、驱动装置、检测装置、控制系统、复杂机械结构等部分组成。

执行机构是机器人的本体，其臂部一般采用空间开链连杆机构，其中的运动副（转动副或移动副）常被称为关节，关节个数通常为机器人的自由度。根据关节配置形式和运动坐标形式的不同，机器人执行机构可分为直角坐标式、圆柱坐标式、极坐标式和关节坐标式等类型。出于拟人化的考虑，常将机器人本体的有关部位称为基座、腰部、臂部、腕部、手部（夹持器或末端执行器）和行走部（对于移动机器人而言）等。

驱动装置是驱使执行机构运动的机构，它根据控制系统发出的指令信号，借助动力元件使机器人运动。它的输入是电信号，输出是线位移量、角位移量。机器人使用的驱动装置主要是电力驱动装置，如步进电机、伺服电机等，也有采用液压、气动等装置驱动的机器人。

检测装置是实时检测机器人的运动及工作情况的机构，其根据需要将情况反馈给控制系统，与设定信息进行比较后，对执行机构进行调整，以保证机器人的动作符合预定的要求。检测装置的传感器大致可以分为两类：一类是内部信息传感器，用于检测机器人各部分的内部状况，如各关节的位置、速度、加速度等，并将所测得的信息作为反馈信号送至控制器，形成闭环控制。另一类是外部信息传感器，用于获取机器人的作业对象及外界环境等方面的信息，使机器人的动作能适应外界情况的变化，使之达到更高层次的自动化，甚至使机器人具有某种“感觉”，向智能化的方向发展。例如，利用外部传感器给出工作对象、工作环境的有关信息，利用这些信息构成一个大的反馈回路，从而大幅提高机器人的工作精度。

控制系统有两种：一种是集中式控制，即机器人的全部控制工作由一台计算机完成。另一种是分散（级）式控制，即采用多台计算机来负责机器人的控制工作，当采用上、下两级计算机共同完成机器人的控制时，作为上级的主机常用于负责系统的管理、通信、运动学和动力学计算，并向下级计算机发送指令信息；作为下级的从机，各CPU分别对应一个关节，进行插补运算和伺服机构控制，以实现给定的运动，并向主机反馈信息。根据作业任务的不同，机器人的控制方式又可分为点位控制、连续轨迹控制和力（力矩）控制。

2. 机器人技术的发展

随着智能控制、机械、通信、电子等技术的进步与发展，机器人技术及其产业在生活中的方方面面发挥着越来越重要的作用。人类对于机器人有着长期的刚性需求，这主要是由于以下原因：机器人是制造业发展的基石；在工业上应用机器人可以提升精度和效率；机器人可以渗透到医疗、教育、服务等生活的诸多方面；机器人产业具有重要的社会价值。在过去的几十年里，制造业得到了快速发展。但是，全球化经济模式已明显改变了制造业的格局。随着劳动力成本的提高，大量劳动密集型产业面临转型升级和区域性转移的挑战。工业机器人具有工作效率高、稳定可靠、重复精度高、能在高危环境下作业等优势，在传统制造业，特别是劳动密集型产业的转

型升级中可发挥重要作用。

机器人的发展历史大体可以概括为三个阶段：第一个阶段的机器人只有手，即以固定程序工作，对外界的信息没有反馈能力；第二个阶段的机器人具有了一定的反馈能力，即有了感觉，如力觉、触觉、视觉等；第三个阶段是“智能机器人”阶段，这一阶段的机器人已经具有自行学习、推理、决策、规划等自主处理事务的能力。

第一代机器人一般可以根据编程人员所固化的程序，完成一些简单的重复性工作。这一代机器人从 20 世纪 60 年代后期开始投入使用，由于其操作简单、可靠性高、成本低廉，已经在很多中小型工厂中得到了普遍应用。

第二代机器人是感知机器人，即自适应机器人。该类机器人在前一代机器人的机械结构、电气结构的基础上增加了“感知”能力。这类机器人在制造汽车、航空器、船舶等产品的自动化生产线中得到了广泛的应用。其工作效率、工作能力较第一代机器人有明显提升。

第三代机器人将具有识别、推理、规划和学习等能力，它可以把感知和行动智能化地结合起来，因此能在非特定环境下作业，故被称为“智能机器人”。这类机器人还处于试验阶段，正在向实用化的方向发展。在过去的几十年里，机器人学和机器人技术的发展引人注目，具体体现在如下方面。

- (1) 机器人产业在全世界范围内迅速发展。
- (2) 机器人的应用范围遍及工业、农业、国防等各个领域。
- (3) 形成了新的学科——机器人学。
- (4) 机器人向智能化的方向发展。
- (5) 服务机器人成为机器人界的“新秀”。

机器人是先进制造技术及自动化装备的典型代表，是人造机器的终极形式。其需要的理论及实践知识包罗万象，涉及机械、电子、液压、传感器及其应用、自动控制原理、通信技术、网络技术、计算机、人工智能等多个学科及领域，是多种高科技技术理论的集成成果。其应用领域主要有两个：制造业领域和非制造业领域。这里对制造业机器人（即工业机器人）和非制造业机器人分别进行简述。

3. 制造业机器人

“工业机器人”一词由美国《金属市场报》于 1960 年提出，美国机器人协会将其

定义为：用来进行搬运机械部件或工件的、可编程的多功能操作器，或通过改变程序可以完成各种工作的特殊机械装置。这一定义已经被国际标准化组织所采纳。随着工业机器人的深度和广度的延伸及其智能化水平的提高，工业机器人已在众多领域得到了应用。工业机器人已广泛应用于汽车及汽车零部件制造业、机械加工行业、电子电气行业、橡胶及塑料工业、食品工业、木材与家具制造业等行业。在工业生产中，弧焊机器人、点焊机器人、分配机器人、装配机器人、喷漆机器人及搬运机器人等工业机器人都已被大量采用，并从传统的制造业领域向非制造业领域延伸，如采矿机器人、建筑机器人以及水电系统中用于维护、维修的机器人等。在国防军事、医疗卫生、食品加工、生活服务等领域，工业机器人的应用也越来越多。

下面介绍几种常见的工业机器人。

搬运机器人。为了提高自动化程度和生产效率，制造企业通常需要快速、高效的物流线来贯穿整个产品的生产及包装过程，搬运机器人在物流线中发挥着举足轻重的作用。搬运机器人一方面具有人类难以达到的精度和效率，另一方面可以承担大重量和高频率的搬运作业，因此被广泛应用在搬运、码垛、装箱、包装和分拣作业中。

打磨抛光机器人。机械零件形状不断向复杂化、多样化的方向发展，导致打磨抛光机器人少有统一的方案。在打磨抛光加工中，机器人的工作方式有两种：一种是机器人夹持被加工工件贴近加工工具（如砂轮、砂带等）进行打磨抛光加工；另一种方法是机器人夹持打磨抛光加工工具贴近工件进行加工。

移动式工业机器人。对于制造大尺寸产品，如航空航天产品，传统的工业机器人无法胜任。首先，大尺寸产品由于重量和尺寸较大，不易移动；其次，工业机器人相对产品而言尺寸不足，如果单纯地按比例放大，机器人的制造和控制成本将变得十分高昂。因此，移动式工业机器人是一个很好的解决方案。

码垛机器人。最早将工业机器人用于码放物体的国家是日本和瑞典。在 20 世纪 70 年代末，日本第一次将机器人用于码垛作业。1974 年，瑞典 ABB 公司研发出全电控式工业机器人，主要应用于工件的取放和物料的搬运。此外，德国、意大利、韩国等国家的工业机器人的研发水平也相当高，如德国的 KUKA 等。

4. 非制造业机器人

由于技术进步、制造成本下降，机器人正悄无声息地进入平常人的生活、学习和

工作中。它们在家可以做家务，在战场上可以侦查，甚至还可以在医院自主进行精细的手术。在教育、医疗、军事和服务行业中，小型机器人正发挥着重要的作用。下面列举几个非制造业机器人的典型实例。

教育机器人。教育机器人是以激发学生学习兴趣、培养学生综合能力为目标的机器人成品套装或散件。它除机器人机体本身外，还有相应的控制软件和教学课本等内容。近年来，教育机器人逐步成为中小学技术课程和综合实践课程的良好载体。常见的教育机器人既有人形教育机器人，也有非人形教育机器人。青少年可以与机器人实现基本的互动交流，并可以对机器人的结构进行改装以扩展其功能。通过了解机器人的结构部件和所需知识等，操控者可以较为便捷地掌握机器人的基础功能，并开拓思维进行创新。

炒菜机器人。炒菜机器人不仅实现了煎、炒、炸等中式烹饪过程的智能化，还可以轻松做出意大利、希腊、法国等地的菜肴。只需轻轻一按，就可以享受到世界各地的地道美食，烹饪过程不粘、不糊、不溢，而且安全、节能、无油烟。

医疗机器人。随着机器人产业的快速发展，医疗机器人的发展已经得到了全世界的高度关注。随着我国人口老龄化程度的加剧和科技的发展，人们对生活各个方面的要求也在提高。以医疗机器人为代表的服务机器人蕴藏着极大的发展潜力，服务机器人未来的市场规模很可能会超过工业机器人。美国已经把治疗机器人、假肢机器人、康复机器人、心理康复辅助机器人、个人护理机器人、智能健康监控系统确定为未来的6大研究方向。

军事机器人。在立国之根本的国防业，机器人也已经在悄然发挥作用。军事机器人具有常规作战人员不具备的优点：灵敏度高、载重能力强、高续航能力、可在特殊场合进行特种作战等，最重要的是可以减少人员的伤亡。其在战场上可进行电子干扰、信息采集、侦查、突击、排雷、爆破、运输等任务。

小结：机器人产业发展迅猛，相关技术的发展也日新月异。因其适应能力不断增强，智能化程度不断提升，预计未来机器人会被应用在人类生活的方方面面，大到宇宙探索，小到衣食住行。届时，人类需要思考机器人大面积普及后的伦理、法律等社会问题。如人工智能会不会造成人类的大面积失业？智能化的最终结局会不会是人类的逐渐退化？机器人最终有无可能成为人类的敌人？这一系列的问题值得我们深思。

3.1.10 文本分析和自然语言处理

自然语言处理使用和支持文本分析，它借助统计学习方法和机器学习方法，为理解句子结构及其意义、情感和意图提供方便。

自然语言处理是文本挖掘的研究领域之一，此领域探讨如何处理及运用自然语言。

对于自然语言处理的发展历程，可以从哲学中的经验主义和理性主义说起。基于统计的自然语言处理是哲学中的经验主义，基于规则的自然语言处理是哲学中的理性主义。在哲学中，经验主义与理性主义的斗争源远流长，这种矛盾与斗争也反映在了自然语言处理上。

早期的自然语言处理具有鲜明的经验主义色彩。如在 1913 年，马尔科夫提出马尔科夫随机过程与马尔科夫模型的基础就是“手工查频”，具体工作是统计《欧根·奥涅金》长诗中元音与辅音的出现频率。1948 年，香农把离散马尔科夫概率模型应用于语言自动机，同时采用手工方法统计英语字母的出现频率。

然而，这种经验主义方法到了乔姆斯基时代出现了转变。

1956 年，乔姆斯基借鉴香农的工作，把有限状态机作为刻画语法的工具，建立了自然语言的有限状态模型，具体来说就是用“代数”和“集合”将语言转化为符号序列，建立了一大堆有关语法的数学模型。这些工作非常伟大，为自然语言和形式语言找到了一种统一的数学描述理论，催生了一个叫作“形式语言理论”的新领域。在这个时代，经验主义被全盘否定，理性主义取得“胜利”。

在 20 世纪 50 年代末至 20 世纪 60 年代中期，经验主义“东山再起”。学者普遍认为只有详尽的历史语料才能带来“靠谱”的结论，于是一些比较著名的理论与算法诞生，如贝叶斯方法、隐马尔可夫模型、最大熵算法、Viterbi 算法、支持向量机等。世界上第一个联机语料库也于这个时期在布朗大学诞生。

但是总体来说，那个时代依然是基于规则的理性主义的天下，经验主义虽然取得了不俗的成就，却依然没有受到太大的重视。

20 世纪 90 年代以来，基于统计的自然语言处理开始“大放异彩”。首先是在机器翻译领域取得了突破。1990 年，在芬兰赫尔辛基举办的第 13 届国际计算语言学会议的主题为“处理大规模真实文本的理论、方法与工具”，研究的重心开始转向大

规模真实文本,传统的仅仅基于规则的自然语言处理越来越不被重视。学者们认为,大规模语料至少是对基于规则的方法的有效补充。

1994年至1999年,经验主义开始空前繁荣。句法剖析、词类标注、参照消解、话语处理等算法几乎把“概率”与“数据”作为标准方法,成为自然语言处理的主流。

总之,理性主义在自然语言处理的发展史上是有重要地位的,也“辉煌”了几十年。但是,基于规则的理性主义在当下这个时代被提及得比较少,用得也比较少,主要是由于它自身存在以下几个缺陷。

(1) 鲁棒性差。过于严格的规则导致对非本质错误“零容忍”。

(2) 研究强度大,泛化能力差。一项研究需要语言学家、语音学家和各个领域的专家互相配合,且很难应用机器学习的方法,难以普及。

(3) 实践性差。基于统计的方法可以根据数据集不断对参数进行优化,而基于规则的方法就不可以,这在当前数据量巨大的情况下,影响是致命的,因为前者常常可以通过增大训练集来获得更好的效果,后者则死板许多,结果往往不尽如人意。

但理性主义还是有很多优点的,同样,经验主义也有很多缺陷,两种方法各有所长、各有所短。一个学科通常有多个研究角度,某个角度在某个特定的历史时期对提高生产力“更有用”,所以重视的人更多。但“有用”不代表胜利,暂时的“无用”更不能说是科学层面上的失败,尤其是在中文自然语言处理的发展还不成熟的时期。笔者认为基于统计的自然语言处理在很多方面并不完美,基于规则的自然语言处理的发展空间还很大,需要更多的人去关注。

在结合深度学习进行自然语言处理的时候,笔者一直在思考一些问题,其中有一个核心问题:神经网络究竟是怎样让各种 NLP 任务被如此完美地完成呢?数据在神经网络中到底发生了什么?词向量、分布式表示、word2vec、glove 等名词分别代表什么?它们的具体关系是什么?它们是否是平级关系?

1. DeepNLP 的核心: 语言表示

最近有一个新名词——DeepNLP (DeepLearning+NLP=DeepNLP) 开始流行。当常规的机器学习发展到了一定的阶段以后,风头慢慢地被后起的深度学习夺去。深度

学习开始引领新一波高潮，因为深度学习有机器学习难以比拟的长处，当深度学习进入自然语言处理领域，自然要带来翻天覆地的变化。

先看一下数据特征表示问题。数据表示是机器学习的核心问题，在过去的机器学习阶段，特征工程大行其道，人工设计大量的特征以解决数据的有效表示问题。而到了深度学习时期，完全不用考虑如何标识数据特征，end2end(端到端)“一步到位”，Hyperparameter(机器学习模型中的超级参数)自动帮你选择关键的特征参数。

深度学习如何能在自然语言处理中发挥应有的作用呢？很明显，先不提如何设计出很强势的网络结构，不提如何在自然语言处理中引入基于神经网络的解决方案，首先需要把语言表示这一“关”过了——如何让语言表示成为神经网络能够处理的数据类型。

我们先看如何用数据表示图像、语音。

在语音处理中，由音频频谱序列向量所构成的矩阵作为前端被输入神经网络进行处理。在图像处理中，由图片的像素构成的矩阵展平成动态数组后，组成动态数组序列被输入神经网络进行处理。在自然语言处理中，将每一个词用一个向量表示出来。

在图像和语音领域，最基本的数据是信号数据，我们可以通过一些距离度量，判断信号是否相似。在判断两幅图片是否相似时，只需通过观察图片本身就能给出回答。而语言作为人类进化了几百万年所产生的一种思维信息的表达工具，其具有高度抽象的特征。文本是符号数据，两个词只要字不同，就难以刻画它们之间的联系，即使是“麦克风”和“话筒”这样的同义词，从字面上也难以看出这两者意思相同(语义鸿沟现象)。词语并不是文字简单地相加就能表示出来的，在判断两个词是否相似时，需要更多的背景知识才能做出回答。

根据上述内容可知：有效表示出句子是神经网络发挥出强大拟合计算能力的关键前提。

2. 自然语言处理中词的表示方法

接下来将按照上面的思路，说明各种词的表示方法。截至本书完稿时，词的表示分为词的独热表示和词的分布式表示。

(1) 词的独热表示。

自然语言处理中最直观、最常用的词的表示方法是词的独热表示。这种方法把每个词表示为一个很长的向量,这个向量的维度是词表大小,其中绝大多数元素为 0,只有一个维度的值为 1,这个维度就代表了当前的词。关于词的独热表示的资料很多,这里简单举个例子进行说明。

“话筒”表示为 [0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 ...]。

“麦克风”表示为 [0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 ...]。

每个词都是“茫茫 0 海”中的一个 1。这种词的独热表示如果采用稀疏方式存储(也就是给每个词分配一个数字 ID)会非常简洁。比如在刚才的例子中,话筒记为 3,麦克风记为 8(假设从 0 开始标记)。如果要编程实现的话,用哈希表给每个词分配一个编号就可以了。这个简洁的表示方法配合最大熵、SVM、CRF 等算法已经能很好地完成自然语言处理领域的各种主流任务。

但是它也有不足之处:向量的维度会随着句子中词的数量类型的增加而增大;任意两个词之间都是孤立的,根本无法表示出语义层面上词与词之间的相关关系,而这一点是致命的。

(2) 词的分布式表示。

传统的词的独热表示仅仅将词符号化,不包含任何语义信息。如何将语义融入词的表示中?Harris 在 1954 年提出的分布式假说为这一设想提供了理论基础:上下文相似的词,其语义也相似。Firth 在 1957 年对分布式假说进行了进一步阐述和明确:词的语义由其上下文决定。

基于分布式假说的词的表示方法,根据建模方式的不同主要可以分为三类:基于矩阵的分布式表示、基于聚类的分布式表示和基于神经网络的分布式表示。尽管这些不同的分布式表示方法使用了不同的技术手段获取词的表示,但由于这些方法均基于分布式假说,因此,它们的核心思想也都由两部分组成:选择一种方式描述上下文;选择一种模型刻画某个词(下文称“目标词”)与其上下文之间的关系。

在详细介绍词的分布式表示之前,需要将自然语言处理中的一个关键概念描述清楚:语言模型。语言模型包括文法语言模型和统计语言模型,一般我们指的是统计语言模型。

统计语言模型是把语言（词的序列）看作一个随机事件，并赋予相应的概率来描述其属于某种语言集合的可能性。给定一个词汇集合 V ，对于一个由 V 中的词构成的序列 $S=(w_1, \dots, w_T) \in V_n$ ，统计语言模型赋予这个序列一个概率 P ，来衡量 S 符合自然语言的语法和语义规则的置信度。

用一句简单的话来说就是：语言模型就是计算一个句子的概率大小的模型。这有什么意义呢？一个句子的概率打分越高，说明它越像人说出来的自然句子。

常见的统计语言模型有 N 元文法模型等。形式化地讲，统计语言模型的作用是为一个长度为 m 的字符串确定一个概率分布 $P(w_1, w_2, \dots, w_m)$ ，以表示其存在的可能性，其中， w_1 到 w_m 依次表示这段文本中的各个词。

基于矩阵的分布式表示。基于矩阵的分布式表示通常又被称为分布式语义模型，在这种表示下，矩阵中的一行对应词的表示，这种表示描述了该词的上下文分布。由于分布式假说认为上下文相似的词，其语义也相似，因此，在这种表示下，两个词的语义相似度可以直接转化为两个向量的空间距离。

常见的 Global Vector 模型（GloVe 模型）是一种对“词—词”矩阵进行分解从而得到词的表示的方法，属于基于矩阵的分布式表示。

基于神经网络的分布式表示——词嵌入。基于神经网络的分布式表示一般被称为词向量、词嵌入。基于神经网络的分布式表示技术通过神经网络对上下文，以及上下文与目标词之间的关系进行建模。由于神经网络较为灵活，这类方法的最大优势在于可以表示复杂的上下文。在前面基于矩阵的分布式表示方法中，最常用的上下文是词。如果使用包含词序信息的 n -gram 作为上下文，当 n 增加时， n -gram 的总数会呈指数级增长，此时会遇到维数灾难问题。而神经网络在表示 n -gram 时，可以通过一些组合方式对 n 个词进行组合，使得参数个数仅以线性速度增长。有了这一优势，神经网络模型可以对更复杂的上下文进行建模，以包含更丰富的语义信息。

前面提过，为了选择一种模型来刻画目标词与其上下文之间的关系，我们需要在词向量中捕获任意一个词的上下文信息。同时，上面提到了统计语言模型正好具有捕捉上下文信息的能力，因此，构建上下文与目标词之间关系的最自然的一种思路就是使用语言模型。

2001 年, Bengio 等人正式提出神经网络语言模型 (Neural Network Language Model, NNLM), 该模型在学习语言模型的同时, 也得到了词向量。请注意: 词向量可以被认为是用神经网络训练语言模型的副产品。

(3) 神经网络语言模型与 word2vec。

我们已经对分布式表示以及词嵌入概念的层级关系有了理性的认识, 但这跟 word2vec 有什么联系呢? 下面梳理一下相关概念。

神经网络语言模型。上文讲过, 通过神经网络训练语言模型可以得到词向量, 那么, 究竟有哪些类型的神经网络语言模型呢? 据笔者所知, 大致有以下几种。

- Neural Network Language Model, NNLM。
- Log Bilinear Language Model, LBLM。
- Recurrent Neural Network based Language Model, RNNLM。
- C&W 模型。
- CBOW 模型和 Skip-gram 模型。

word2vec 与 CBOW 模型、Skip-gram 模型的关系。这里我们引出一个术语: word2vec。上面 5 种神经网络语言模型只是逻辑概念上的东西, 我们需要通过设计将其实现出来。其中, 实现 CBOW 模型和 Skip-gram 模型的工具是 word2vec, 实现 C&W 模型的工具是 SENNA。

统计语言模型就是给定几个词, 在这几个词出现的前提下来计算某个词出现的概率。CBOW 模型也是统计语言模型的一种, 即根据某个词前面的 C 个词或者前 C 个连续的词来计算某个词出现的概率。而 Skip-gram 模型则相反, 它是先确定某个词, 然后分别计算在它前后出现某几个词的概率。

以“我爱魔都”这句话为例。假设我们现在关注的词是“爱”, 那么, $C=2$ 时它的上下文分别是“我”“魔都”。CBOW 模型就是把“我”“魔都”的独热表示作为输入, 也就是把 C 个 $1 \times v$ 向量, 分别跟同一个 $v \times n$ 系数矩阵 W_1 相乘得到 C 个 $1 \times n$ 隐藏层, 然后 C 个隐藏层取平均, 只算一个隐藏层, 再跟另一个 $n \times v$ 系数矩阵 W_2 相乘得到 $1 \times v$ 的输出层, 这个输出层的每个元素代表的就是词库里每个词的事后概率。输出层需要跟“爱”的独热表示做比较计算。这里需要注意的是 v 通常是一个很大的数 (比如几百万), 计算起来相当费时间。除了“爱”那个位置的元素, 其他元素肯定要算在 loss 里面。word2vec 用基于赫夫曼编码的 Hierarchical softmax 筛选掉

了一部分不可能的词，然后又用 negative sampling 去掉了一些负样本的词，所以时间复杂度就从 $O(v)$ 变成了 $O(\log v)$ 。Skip-gram 的训练过程与此类似，只不过输入/输出刚好相反。

Word Embedding 的训练方法大致可以分为两类：一类是无监督的预训练或弱监督的预训练；另一类是端到端的监督训练。无监督的预训练或弱监督的预训练以 word2vec 和 Auto-Encoder 为代表。这一类模型的特点是，不需要大量的人工标记样本就可以得到质量还不错的 Embedding 向量。不过因为缺少了任务导向，可能还无法解决我们要解决的问题。因此，我们往往会在得到预训练的 Embedding 向量后，用少量人工标注的样本去调整整个模型。

端到端的监督模型在近几年越来越受人们的关注。与无监督模型相比，端到端的监督模型在结构上往往更加复杂。同时，也因为有着明确的任务导向，端到端的监督模型学习到的 Embedding 向量往往也更加准确。例如，通过一个 Embedding 层和若干个卷积层连接而成的深度神经网络可以实现对句子的情感分类，可以学习到语义更丰富的词向量。

词向量既能够降低维度，又能够捕捉到当前词在上下文中的信息（表现为前后距离关系）。我们将其用来表示词语、句子并作为神经网络的输入是合适的。

小结：当你做某一项具体的自然语言处理任务时，如果要用到词向量，笔者建议你：选择使用别人训练好的词向量（注意：需要使用相同语料内容的词向量）；训练自己的词向量。

3.2 从企业视角看 AI 技术

笔者在撰写这本书的过程中，已经启动了自己的创业项目（AI 相关方向）。笔者认为从创业的角度来看，可以将 AI 行业划分为三层：最底层是基础平台层，包括数据服务、人工智能芯片、服务器、云计算等，该层是通用且非常必要的；中间层是技术提供层，计算机视觉、机器学习等技术都属于这一层；最顶层是应用层，包括无人驾驶车、无人机、机器人等，该层承载着很多应用。

如此众多的应用给有识之士提供了广泛的创业机会，大家需要保持理性的态度去创业。

要评价一家人工智能（AI）企业是否成功，不妨从四个角度加以判断：

是否具有全球化的视野。

是否正在做最简单、最实际且可以落地的事。

是否寻找到了合适的合作伙伴。

对商业模式是否有清晰的判断，以及是否有随环境变化而灵活调整商业模式的心态。

全球技术融合的趋势越来越明显。人工智能已经不是一场区域化的战斗，而是一种全球化的竞争。

（1）从客户角度看。

观察企业服务市场会发现，中国的市场规模虽然非常大，但公司付费率，尤其是中小企业的付费率非常低。与此同时，虽然美国的公司总数比中国少，但其中中小企业的付费能力一般非常强，约为中国企业的 10~30 倍。如果人工智能企业只局限于一个市场，就将失去很多机会。

（2）从技术角度看。

中国、美国、以色列等国都拥有非常棒的技术人才，不仅有奇妙的想法，更有一些非常棒的系统设计方式。因此，全球技术交流的趋势将会愈加明显。

举个例子，如今非常火的无人驾驶，其阻力之一是激光雷达的成本高昂，这造成普通消费者购买无人驾驶车的成本非常高。试想，购买一辆车的总价也就几十万元，若为了安装无人驾驶系统再掏几十万元，有几个人会心甘情愿呢？很多公司在尝试打破这个壁垒，且处理方法不尽相同。中国创业者拥有较强的工程能力，能够把一些比较成熟的技术的性能和稳定性较之前提高很多倍；美国创业者则会投入庞大的科研经费去探索一条新的技术路线（也许这条技术路线并不成熟）；以色列创业者则会运用非常成熟的技术，但在系统设计、产品设计上却会采取一些非常不一样的方法，让更多成熟技术可以被应用到一些不成熟的产品上。

笔者认为，在 AI 这个领域，大家在技术层面上一定是在全球范围内互相交流、互相融合的。AI 其实有三个比较重要的环节，分别是感知、认知和决策，三者的难度不断攀升。根据不同难度和不同行业的构成元素来看，现在存在较大创业机会的行业主要是以下几个。

无人驾驶。很多人觉得无人驾驶在 AI 领域非常难实现，但其实难度主要存在于系统层面，因为它需要很多数据，需要多方面的技术，还需要决策、控制等能力。无人驾驶在 AI 层面的实现其实很简单，为什么呢？大家可以想象，开车本身并不困难，因为其本身存在诸多规则，司机是根据交通规则来完成驾驶行为的。因此，对比认知计算方面的技术，无人驾驶相对而言还是比较简单的。

工业自动化。工业自动化在风电等方面的基础架构已经非常完善，所以将 AI 技术应用到工业环境中会有非常大的前景，而且现在的技术也可以做到。

健康。笔者认为这一领域的主要问题是能够拿到数据，因为不管是做医学图像处理，还是做一些辅助诊断，数据本身的处理、清洗、标注等都非常重要。只要过了数据关，应用 AI 技术相对比较简单，商业化方面会有比较多的机会。

教育。将 AI 算法应用到教育这个特定场景的发展已经比较成熟。

(3) 从企业的角度看。

AI 公司的商业模式并不清晰。不管是 ToB 还是 ToC，不管是提供技术还是提供解决方案，不管是提供产品还是提供服务，在不同的场景、不同的领域，商业模式可能完全不一样。

AI 公司到底应该选择 ToB 还是 ToC？二者各有利弊。如果选择 ToB，发展可能会更快，但“天花板”也会更低；如果选择 ToC，那么，因为要跟巨头“掰手腕”，所以难度更大，但成功之后的回报也是巨大的。所以，创业者需要根据实际情况对整个商业模式的逻辑有一个明智的判断。

AI 公司到底应该提供技术还是解决方案？在美国，很多公司虽然只提供一项非常简单的技术，但也能创造较高的收入和利润；但在中国，输出技术需要较长时间才能获得利润。此外，中国的很多巨头在跟你做生意的时候，如果你的解决方案很好，它们通常希望把你“挖”过去。

AI 公司到底应该做产品还是服务？这方面要具体情况具体分析。比如，如果想在医疗行业做 AI 服务，笔者感觉会比较难。因为医疗行业的产业链非常长，就诊流程也非常复杂，如果想为医院提供一项好的服务，不仅需要很多人力，也需要积累很长时间，所以应该针对医疗行业提供非常棒的产品，比如自动分析 CT 图像的产品、辅助诊疗的产品等。

因此，在我们对人工智能行业有超高期望值的时期，更需要理性地选择一个技术可以实现、市场相对成熟且能够产生成熟商业模式的领域，这样的创业才可能产生更大的价值。

3.3 AI 产品经理应该知道的 AI 技术

AI 产品经理需要了解的 AI 技术实现包括深度学习、卷积神经网络、递归神经网络、深度神经网络、Q-Learning.....

AI 产品经理需要了解的技术分为互联网前端技术、终端设备技术、手机技术、家庭网关技术、后台程序技术、自助服务终端技术等。

AI 产品经理需要了解的行业应用包括图像识别、人脸识别、文字识别、语音识别、无人驾驶等。

AI 产品经理应该了解如何运用 AI 技术来构建 AI 产品，AI 产品经理用 AI 技术赋能细分领域时应该了解如下内容。

- (1) 相同的产品是否可以在其他领域进行拓展。
- (2) 组合产品是否有互补品，或进行模组化穿透的能力。
- (3) 产品价值链在设计、采购、生产、物流、渠道、促销、品牌等环节是否可以独立形成新的商业模式。
- (4) 产业系统是否可以形成“产业链”“产业生态圈”的商业模式。

笔者解释一下以上部分内容。当产品完全相同时，你需要做的是向产品之外的领域拓展。例如都做二手车金融，这个时候看的是谁分析大数据的效率更高，因为放款给用户的效率决定了产品的胜负。二手车的贷款相比于银行放款给企业，属于小额贷款。因为放款金额小，所以每笔交易的手续费比例较高。这个时候用 AI 进行赋能可以明显地降低放贷成本，形成竞争优势。

另外，笔者再强调一下对穿透能力的两方面认知：一方面是理性的认知，另外一方面是感性的认知。

理性的认知：

这种认知包括以下两个方面。

一是横向方面，即便语音识别完全正确，但是不理解语义，这样的系统对用户也没用。

二是纵向方面，例如，商汤科技做人脸识别的同时瞄准银行、保险领域“做深”……

感性的认知：

这种认知指的是，对我们的核心技术进行深挖以领先市场一小步，形成自己的门槛，持续给行业赋能，即感情不变，能力要变。

AI 能够在很多方面实现产品的升级，具体包括以下几点。

人机交互方式的升级。从键盘到触摸屏是从传统互联网到移动互联网的升级，而从触摸屏到人脸识别，再到语音识别、语义分析，AI 天生具有解决人机交互问题的能量。

实现精准运营。AI 可以帮助转换产品运营数据的考核维度，将考核指标从日活跃用户数量（Daily Active User，DAU）转换为月活跃用户数量（Monthly Active User，MAU），实现精准触达用户。

解决创新乏力的问题。传统产品通常是一个网站、一个 App，抑或一个小程序。而在 AI 时代，创新的空间更大，产品可以软硬结合，也可以跨行业创新。

实现个性化发展。AI 可以使为消费者提供个性化的服务和产品变成现实。例如，智能工厂的 C2M 模式，首先是消费者提出需求，企业收集庞大的需求数据；然后智能工厂在 AI 技术赋能下实现数据智能分流，从而实现个性化生产。另外一个例子是，在学会自然语言识别和分析后，AI 能够理解用户，这个时候 AI 可以实现千人千面、一对一的交流服务。因此，AI 产品经理可以打造个性化的 AI 产品。

非 AI 技术背景的人员转行从事 AI 产品经理工作时，除了要掌握 AI 的技术原理，在学有余力的情况下还要注意以下两点内容。

摸清自己行业的特点。例如，大数据分析（电商类网站的千人千面）、图形图像识别（直播平台运用图像识别技术发现违规行为）、知识库文本处理技术（智能客服）等。

补充短板知识。做软件出身的 AI 产品经理需要补充硬件知识，做硬件出身的 AI 产品经理需要补充软件知识，而做非 AI 层产品的产品经理需要补充第三方产品的相关知识。例如，如果想补充语音领域的知识，可以了解科大讯飞、思必驰、出门问问等公司的产品；如果想补充图像领域的知识，可以了解商汤科技、

旷视科技、云从、依图等公司的产品。

随着数据的累积、AI 的迅猛发展，希望了解 AI、期待从事 AI 产品经理工作的人士越来越多。摆在希望从事 AI 产品经理工作的人士面前的问题如下：“AI 技术的能力边界在哪里？”“时下有哪些可以落地的 AI 技术？”“各种 AI 技术各自的优点、缺点是什么？”“从产品经理的视角来看，应该知道的技术、相应的切入角度是什么？”

3.3.1 AI 产品经理对 AI 技术的理解

AI 的发展已有近 70 年的历史，AI 在技术实现上可归为 6 种途径：符号主义、连接主义、学习主义、行为主义、进化主义和群体主义。这 6 种途径并非泾渭分明，它们只是从不同的角度提出了解决方案，如学习主义就用到了人工神经网络来实现。产品经理所理解的 AI 应该是广义范畴的 AI。

截至本书完稿的 2018 年年底，今日头条估值高达数百亿美元，今日头条系产品的主要特征是对 AI 和机器学习的应用。

时下流行的机器学习以及深度学习实际上是对符号主义、连接主义以及行为主义理论的进一步拓展。

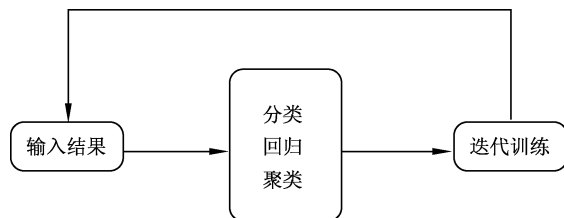
对于机器学习的理解，笔者认为产品经理可以从 3 个问题入手，即机器学什么、机器怎么学、机器做什么。

首先，机器学习需要学习的内容是能够表征此项任务的函数，即输入和输出的映射关系，从信息论的角度来看，其学习的目标是确定两个状态空间内所有可能取值之间的关系，使得熵尽可能最低。熵越低，信息越有序。

其次，机器怎么学。要实现学习目标，就要教给机器一套评判的方法。不同于告诉机器每个具体步骤该如何操作的传统方法，机器学习把对机器描述过程演进为对机器描述结果。从数学角度来看，就是为机器定义一个合适的损失函数，能够合理量化真实结果和训练结果之间的误差，并将之反馈给机器继续做迭代训练。

最后，机器学习究竟要做什么。其实，机器学习主要做三件事：分类、回归和聚类，其中分类和回归属于监督学习的范畴，而聚类则属于无监督学习的范畴。多数人工智能应用都是通过把现实问题抽象成相应的数学模型，分解为这三类基本任务的有机组合，并对其建模求解的过程。

机器学习的过程如下图所示。



机器学习的过程

3.3.2 热门的 AI 算法

首先讨论常见的机器学习任务及其常用算法。

(1) 回归。

回归是一种对连续型数值变量进行预测和建模的监督学习算法。

回归算法有很多种，其中常用的主要有两种。

线性回归（正则化），它最简单的形式是用一个连续的超平面来拟合数据集。

回归树（集成方法），该方法又被称为决策树，通过将数据集重复分割成不同的分支来最大化每次分离的信息增益，从而让回归树很自然地学到非线性关系。集成方法包括随机森林（RF）或梯度提升树（GBM），以及最近邻算法和深度学习。

(2) 分类。

分类算法是用于分类及预测的监督学习算法，许多回归算法都有对应的分类形式，分类算法往往适用于类别（或其可能性）的预测。其中常用的算法主要有五种。

逻辑回归（正则化）。逻辑回归通过逻辑函数将预测映射到 $0\sim 1$ 的区间，因此预测值可被视为某一类别的概率。

分类树（集成方法）。对应于回归树的分类算法是分类树。通常，它们都是指决策树，更为严谨的说法是“分类回归树”，这也就是非常有名的 CART 算法。

支持向量机。支持向量机使用一个名为核函数的技巧，将非线性问题转换为线性问题，其本质是计算两个观测数据的距离。支持向量机算法所寻找的是能够最大化样本间隔的决策边界，因此又被称为大间距分类器。

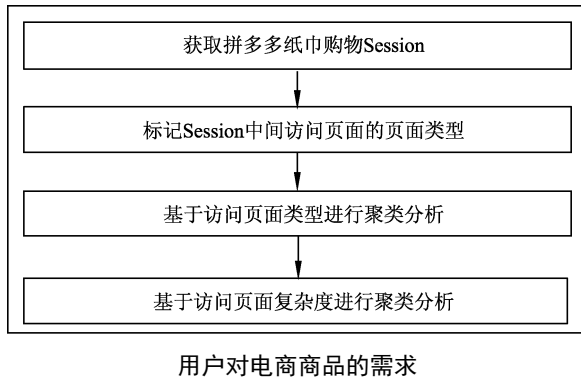
朴素贝叶斯。基于条件概率和计数的简单算法，其本质是一个概率表，通过训练数据来更新其中的概率。它预测新观察值的过程，就是根据样本的特征值在概率

表中寻找最有可能的类别。被称为“朴素”的原因是，其核心的特征条件独立性假设（例如，每一项输入特征都相互独立）在现实中几乎是不存在的。

深度学习方法。

（3）聚类。

聚类算法是基于数据的内部结构来寻找样本自然族群（集群）的无监督学习任务，使用案例包括用户画像、电商商品聚类、社交网络分析等。下图用聚类算法表示用户对电商商品的需求。



聚类常用的算法主要有四种。

k 均值。这是基于样本点间的几何距离来度量聚类的通用目的的算法。由于集群围绕在聚类中心，因此，结果会接近于球形并具有相似的大小。

仿射传播。基于两个样本点之间的距离来确定集群，其结果倾向于更小且大小不等的集群。

分层/层次。其主要完成层次聚类的目标。

DBSCAN。这是基于密度的聚类算法。它将样本点的密集区域组成集群；其最新进展是 HDBSCAN，它允许集群的密度可变。

在机器学习领域，两个最为重要的概念是“维”和“特征选取”。其中，“维”通常指的是数据集中的特征数量（即输入变量的个数）；而“特征选取”指的是，从你的数据集中过滤掉不相关或冗余的特征。特征选取主要包括四种方法。

方差阈值。摒弃观测样本中那些观测值改变较小（即方差小于某个设定的阈值）的特征，这样的特征价值极小。

相关性阈值。去掉那些高度相关的特征（这些特征的特征值变化与其他特征非常相似），它们提供的是冗余信息。

遗传算法。可用于不同任务的一大类算法的统称。它们受进化生物学与自然选择的启发，结合变异与交叉，在解空间内进行高效的遍历搜索。

逐步搜索。逐步搜索是一个基于序列式搜索的监督式特征选取算法。它有两种形式：前向搜索和反向搜索。

特征选取与特征提取不同，其关键区别在于：特征选取是从原特征集中选取一个子特征集，而特征提取则是在原特征集的基础上重新构造出一些（一个或多个）全新的特征。

特征提取主要用来创造一个新的、较小的特征集，但仍能保留绝大部分有用的信息，其主要包括三种方法。

主成分分析。主成分分析是无监督学习算法，它用来创造原始特征的线性组合。新创造出来的特征之间都是正交的，也就是没有关联性。具体来说，这些新特征是按它们本身变化程度的大小进行排列的。第一个主成分代表了数据集中变化最为剧烈的特征，第二个主成分代表了变化程度排在第二位的特征，以此类推。

线性判别分析。线性判别分析是监督学习算法，它必须使用有标记的数据集。

自编码器。它是用来重新构建原始输入的，关键是在隐藏层搭建比输入层和输出层数量更少的神经元。这样，隐藏层就会不断学习如何用更少的特征来表征原始图像。

总体来说，对机器学习算法进行归类非常棘手，因为其背后的实现原理、数学原理等存在差异，其分类标准存在多个维度。

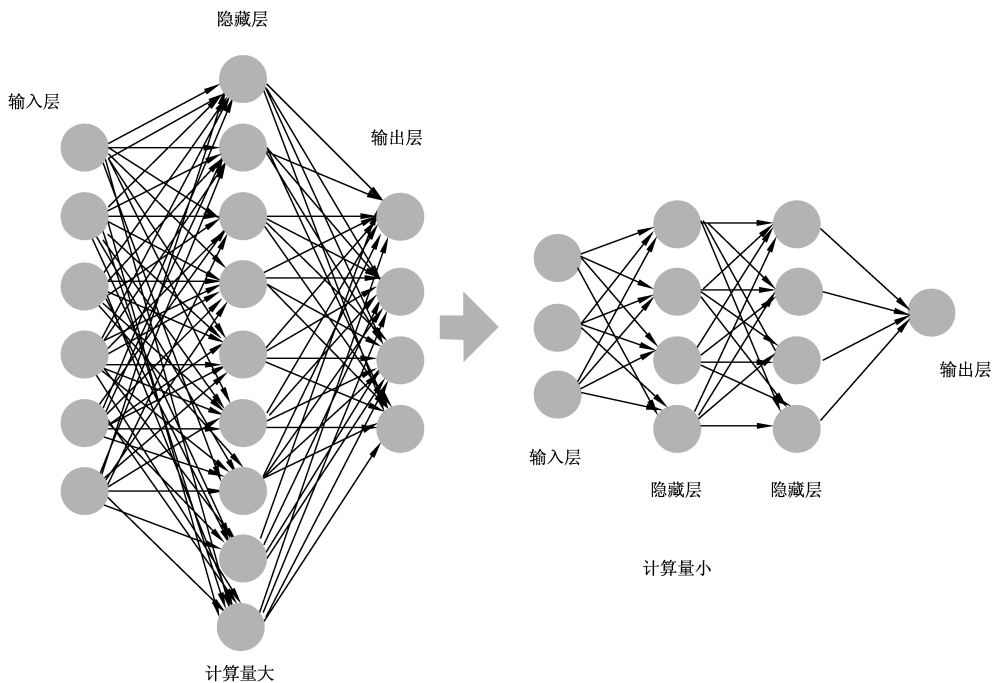
3.3.3 深度学习大热背后的力量

笔者认为，深度学习大热的原因来自以下三点。

（1）深度学习与传统算法之间的区别。

传统算法在语音识别、物体识别方面无法有效展开应用，重要原因之一就是这类算法无法使用语音及图像的高维度数据在高维空间学习复杂的函数，在这类高维

空间进行学习通常也意味着对算力的极大消耗,即使在算力极为丰富的阶段也很难有效满足其算力需求。因此,深度学习应运而生。深度学习采用多层调参、层层收敛的方式,将参数数量始终控制在一个较为合理的水平,使得原本不可计算的模型可计算了。传统算法与深度学习的区别如下图所示(左边是传统算法的结构,右边是深度学习的结构)。



(2) 深度神经网络具有完备性。

从理论上来说,深度神经网络可以表征任何函数,因此,深度神经网络可以通过不同的参数及网络结构来对任意函数进行拟合,这排除了无法学习到复杂函数的可能性。

(3) 深度学习的特征选取完备。

深度学习具有自动学习特征的能力,这是深度学习也被称作无监督学习的原因。从深度学习模型中选择某一神经层的特征后就可以用来进行目标模型的训练,而不需要人为参与特征选取。

3.3.4 AI产品经理应该了解的深度学习算法及其应用

1. 深度学习算法

AI算法从专家系统到特征工程,再到深度学习的这个过程,人工参与在逐渐减少,而机器的工作在逐渐增加,深度学习算法主要包括两种。

(1) 卷积神经网络。

卷积神经网络是近年才发展起来,并引起广泛重视的一种高效识别方法。卷积神经网络已经成为人工智能领域的研究热点之一,K. Fukushima在1980年提出的新识别机是卷积神经网络的第一个实现网络。随后,具有代表性的研究成果是Alexander和Taylor提出的改进认知机,该方法综合了各种改进方法的优点并避免了耗时的误差反向传播。

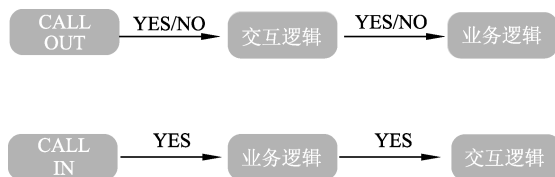
卷积神经网络的基本结构包括两层:特征提取层。每个神经元的输入与前一层的局部接受域相连,并提取该局部的特征。一旦该局部的特征被提取,它与其他特征间的位置关系也随之确定下来。特征映射层。网络的每个计算层由多个特征映射组成,每个特征映射是一个平面,平面上所有神经元的权值相等。特征映射结构采用sigmoid函数作为卷积网络的激活函数,这使得特征映射具有位移不变性。此外,由于一个映射面上的神经元共享权值,因而减少了网络自由参数的个数。卷积神经网络中的每一个卷积层都紧跟着一个用来求局部平均与二次提取的计算层,这种特有的两次特征提取结构减小了特征分辨率。

卷积神经网络主要用来识别位移、缩放及其他形式扭曲不变性的二维图形。由于卷积神经网络的特征检测层通过训练数据进行学习,因此在使用卷积神经网络时,可以避免显示的特征抽取,而从训练数据中进行学习。此外,由于同一特征映射面上的神经元权值相同,因此网络可以并行学习,这也是卷积网络相对于神经元彼此相连网络的一大优势。卷积神经网络的局部权值共享的特殊结构在语音识别和图像处理方面有着独特的优越性,其布局更接近于实际的生物神经网络,权值共享降低了网络的复杂性,特别是多维输入向量可以直接输入网络这一特点避免了特征提取和分类过程中数据重建的复杂度。

(2) 递归神经网络。

在实际应用中,我们会遇到很多序列数据,如自然语言处理问题中的单词、语音处理中每帧的声音信号、股票价格的时间序列数据等。

下面以某自然语言处理产品为例进行解释。



某自然语言处理产品的结构

为了建模序列问题，递归神经网络引入了“隐状态 (Hidden State)”概念。一个递归神经网络中可以存在一个或多个隐状态，计算时每一步使用的参数都是一样的，即每个步骤的参数都是共享的，这是递归神经网络的重要特点之一。同时需要注意的是，递归神经网络的输入和输出序列必须是等长的。由于这个限制的存在，因此经典的递归神经网络的使用范围比较小。但也有一些问题适合用经典的递归神经网络进行建模，如计算视频中每一帧的分类标签，因为要对每一帧进行机选，所以每一次的输入和输出序列等长。

当要处理的问题为“输入是一个序列，输出是一个单独的值而不是序列”时，需要对递归神经网络中的最后一个 h 进行输出变化，这种结构通常用来处理序列分类问题。如输入一段文字，判别它所属的类别；输入一个句子，判断其情感倾向；输入一段视频并判断它的类别等。对于输入不是序列而输出是序列的情况，只需要在序列开始时进行输入计算或把输入信息作为每个阶段的输入即可。这种网络结构可以处理的典型问题是从图像生成文字，此时输入的是图像的特征，而输出的序列是一个句子。

递归神经网络最重要的一个变种是输入为 N 而输出为 M ，这种结构又叫作 Encoder-Decoder 模型，也可称之为 Seq2Seq 模型。在实际的生产、生活中，我们遇到的大部分问题的序列都是不等长的，如机器翻译中源语言和目标语言中的句子往往没有相同的长度。为此，Encoder-Decoder 结构会先将输入数据编码成一个上下文向量 C 。得到 C 的方式有多种，最简单的方法就是把 Encoder 的最后一个隐状态赋值给 C ，也可以对最后的隐状态做一个变换得到 C ，还可以对所有的隐状态做变换。在得到 C 以后，就用另一个递归神经网络对其解码，这部分递归神经网络被称为 Decoder。具体做法就是将 C 当作之前的初始状态输入到 Decoder 网络中。

由于这种 Encoder-Decoder 结构不限制输入和输出的序列长度，因此应用范围广泛，常见的应用范围有如下几种。

机器翻译，Encoder-Decoder 的经典应用是在机器翻译领域被最先提出的。

文本摘要，输入是一段文本序列，输出是这段文本序列的摘要序列。

阅读理解，将输入的文章和问题分别编码，再对其进行解码，得到问题的答案。

语音识别，输入是语音信号序列，输出是文字序列。

在 Encoder-Decoder 结构中，Encoder 把所有的输入序列都编码成一个统一的语义特征 C ，然后再解码。因此， C 中必须包含原始序列的所有信息，它的长度就成了限制模型性能的瓶颈。如在机器翻译中，当需要翻译的句子较长时，一个 C 可能存不下那么多信息，这会造成翻译精度的下降。而 Attention 机制通过在每个时间输入不同的 C 来解决这个问题，每一个 C 会自动选取与当前所要输出的结果最合适的上下文信息。

2. 深度学习的不足

虽然深度学习在多种识别、感知任务中表现优异，但笔者在实践中也发现深度学习存在以下三方面的不足。

(1) 数据基础问题。我们在实际生产、生活中所收集到的数据往往都是小数据，而不是大数据。比如，手机上的个人数据，教育、医疗、基因检测与实验、学生测验、客服问答系统中的数据，都是小数据。只有得到可以从大数据迁移到小数据的通用模型，才能使得更多领域真正用上人工智能，但基于大数据迭代得到的深度学习模型无法胜任小数据场景的业务。

(2) 模型可靠性问题。深度学习模型非常脆弱，稍加移动、离开现有的场景数据，它的效果就会降低。因此，深度学习模型的可靠性是一个重要的问题。对机器学习来说，由于训练数据和实际数据存在区别，因此，训练出来的模型被用于处理它没有见过的数据时，效果也会大打折扣。解决这一问题的办法是迁移学习，它能把可靠性提升一个台阶。

(3) 应用问题，特别是机器学习模型在个性化推荐方面的问题。比如，在手机上推荐信息、服务的时候，它要适应个人的行为。因为任何个人的数据都是小数据，所以，个性化的问题就是如何把云端的通用模型适配到终端的小数据上，让模型在终端也能正常工作。这是迁移学习比较擅长的事情。

例如 ,Siri 是一款云端通用型机器人 ,但是 Siri 并不能完全基于笔者本人的工作、生活提供个性化建议。

3.3.5 AI 技术的落地

AI 领域的机器学习、深度学习技术是一个无限逼近极限的系统。产品经理了解 AI 技术以后 ,一个关键责任是落地。切入 AI 产品的重要角度是如何在算法日渐完善、数据逐步丰富、算力快速增长的时代交付符合用户心理预期的产品。

这不仅仅是 AI 技术问题 ,更是产品经理捕捉人性特点的问题。

AI 产品经理需要提升自己的两种能力 :一是对 AI 技术的认知能力 ,二是捕捉人性特点、甄别需求的能力。第一种能力决定了产品经理能不能充分挖掘技术的潜力 ,第二种能力决定了产品经理能否在当今的 AI 技术水平下打造出符合用户人性特点的产品。

这些能力需要我们持续学习 ,毕竟一款产品的成功是一家企业成功的标志。

3.4 AI 技术小白的学习资料

本节是笔者给想要进入 AI 这个领域 ,但不知道从哪里开始的初学者准备的学习资料和一些学习建议。

1. 机器学习

有关机器学习领域的最佳介绍 ,请观看 Coursera 上吴恩达的机器学习课程。它解释了机器学习的基本概念 ,能让你很好地理解最重要的算法。

有关机器学习算法的简要介绍 ,可以观看 envatotuts + 的 Machine Learning Distilled 课程。

Toby Segaran 的 *Programming Collective Intelligence* 一书是很好的资源 ,通过这本书可以学习如何用 Python 实现机器学习算法 ,它需要你动手做许多实践项目。

Tom Mitchell 在卡内基梅隆大学教授的机器学习课程 Another Course On ML 也非常有价值。

2. 深度学习

笔者遇到的比较好的深度学习介绍是 *Deep Learning With Python*。 *Deep Learning*

With Python 这本书不会深入到困难的数学，也没有一系列的先决条件，而是描述了如何用简单的方法开始学习深度学习，如何快速开始构建并实践深度学习中的一切。它介绍了最先进的工具（Keras、TensorFlow），并通过几个实际项目，介绍如何实现更好的效果。

Edx 中的 Great Introductory DL 课程也是非常好的资源。

为了更深入地了解，可以参考以下资源。

Geoffrey Hinton 在 Coursera 上讲授的 Neural Networks for Machine Learning 课程会带你了解人工神经网络的经典问题。

MIT 出版社的 *Deep Learning* 一书，三位作者（Ian Goodfellow、Yoshua Bengio、Aaron Courville）都是“大牛”，笔者建议读者阅读英文版原书。

斯坦福大学的 UFLDL Tutorial 是很不错的资源。

deeplearning.ai 的教程也非常有参考价值。

Michael Nielsen 的 *Neural Networks and Deep Learning* 一书写得深入浅出。

Simon Haykin 的 *Neural Networks and Learning Machines* 一书也非常棒。

3. 人工智能

《人工智能：一种现代的方法》（*Artificial Intelligence: A Modern Approach*）是关于“守旧派”AI 最好的一本书。这本书概述了人工智能的各个领域，并解释了你需要了解的所有基本概念。

来自加州大学伯克利分校的 Artificial Intelligence 课程是一系列优秀的视频讲座，其通过一种非常有趣的实践项目（训练 AI 玩 Pacman 游戏）来解释 AI 的基本知识。笔者推荐大家在观看视频的同时阅读《人工智能：一种现代的方法》，因为它是基于这本书来讲解的，并从不同的角度解释了很多概念，以使概念更容易理解。它的讲解清晰、透彻，对初学者来说是非常不错的资源。

如果你对人工智能感兴趣，则可能很想知道人的大脑是怎样工作的，下面这几本书会通过直观有趣的方式来解释大脑的工作原理。

Jeff Hawkins 的有声读物 *On Intelligence*，以及 Ray Kurzweil 的有声读物 *How to Create a Mind* 和 *Principles of Neural Science* 笔者都非常推荐。后两本书深入谈论了脑科学、神经解剖学等学科的内容，非常有趣，但内容很长。

4. 数学

以下这些内容可以帮助你开始学习 AI 之前掌握非常基础的数学知识。

(1) 微积分学。

可汗学院 (Khan Academy) 的 Calculus 视频课程。

麻省理工学院 (MIT) 的多变量微积分讲座, 即 Lectures On Multivariable Calculus。

(2) 线性代数。

可汗学院 (Khan Academy) 的 Linear Algebra 视频课程。

麻省理工学院 (MIT) Gilbert Strang 教授主讲的 Linear Algebra 视频课程。

布朗大学 (Brown University) 的 Coding the Matrix 课程。

(3) 概率和统计。

可汗学院 (Khan Academy) 的 Probability 课程、Statistics 课程。

Edx 的 Probability 课程。

5. 计算机科学

要掌握 AI, 你要熟悉计算机科学和编程。

如果你刚刚开始, 建议阅读《深入 Python 3》(Dive Into Python 3) 这本书, 你在 Python 编程中所需要的大部分知识在这本书中都会被提到。

要想更深入地了解计算机编程的本质, 可以观看麻省理工学院的经典课程“计算机科学及 Python 编程导论”(Introduction to Computer Science and Programming Using Python)。这是一门关于 Python 语言和计算机科学的基础课程, 是计算机程序方面最有影响力的课程之一。

6. 其他方面的资源

Metacademy 是一个学习算法的辅助工具, Metacademy 会把各个知识点联系起来, 就像游戏里的技能树一样。每个知识点都有简介, 而且会链接到那些优质的学习资源, 最重要的是, 它会画出通向这个知识点的知识图谱。

作为全世界最著名的数据挖掘、机器学习竞赛平台, Kaggle 早已成为数据玩家在学习机器学习相关知识之后“一试身手”的练兵场。

3.5 AI 技术小白的学习方法

除了上述资料，笔者对各位读者还有以下建议。

首先是“先易后难”学习法，即从本书列举的内容和资源中，先学习自己比较熟悉的、相对容易的、离自己已有知识边界比较近的内容，然后逐渐学习自己不熟悉的、相对难的内容。

在学习的时候很可能会遇到自学困境，产生自学困境的原因有如下几点。

- 直接原因：精力分散，不够专注，书读不下去，代码看不下去。
- 深层原因：驱动力不足以让自己专注于学习。
- 根本原因：没有足够渴望的目标，无法调动自身的能量形成驱动力。

归根到底，学不下去是因为不知道学习是为了什么，学习到底有什么用。漫无目标的学习注定是浪费时间。

笔者建议大家学习时要注意三个要点：目标明确、系统性强、足够深入，三者缺一不可。因此，笔者建议各位读者在学习前做好以下工作。

1. 明确学习目标。

首先，相信阅读本书的读者已经拥有了一个明确的目标：成为 AI 产品经理，跟着笔者一起用 AI 技术重新定义产品经理，用 AI 技术对产品进行赋能。在明确目标的指引下，比较容易找到什么“有用”。这虽然有点功利，但不得不承认，“有用”是大多数人产生驱动力的关键原因。

另外，我们可以以“在 AI 领域找到技术岗位工作”为基准进行学习。避免那种“东一榔头、西一棒槌”的“乱学”。

什么样的目标才算明确呢？“从事 AI 领域的工作”是不能被称为一个明确的目标的，“成为 AI 产品经理，用 AI 技术再定义产品经理，用 AI 技术对产品进行赋能”才是一个明确的目标。

利用 AI 技术进行赋能首先要搞清楚几件事：真正落地的 AI 领域都有哪些。每个领域都有什么样的代表企业。在这些企业中，都有什么样的技术岗位。每种岗位需要什么样的入门条件。

读者至少要能够定位你的目标岗位，明确这类岗位的技术要求，这样才有可能确

定需要学习的内容,从而选择正确的培训课程。

2. 制定明确的目标

如何制定明确的目标呢?笔者推荐大家采用如下步骤。

Step1:确定一个领域(图像、语音、NLP等)。如果你本来就有特别感兴趣的领域,比如人脸识别,那么可以直接由此入手。

如果你对 AI 了解得不多,只是想当前市场上最热门的领域,那也比较好办。多搜集一些招聘网站、社交媒体中的招聘信息,做一下数据分析,看看哪些领域招的人最多、薪水最高。

Step2:了解本领域当前的科研状况。最简单的方法是找 10 篇本领域核心期刊或会议的论文,认真读一遍。用谷歌找论文还是比较方便的,可以在谷歌页面的搜索框中直接输入相关领域的关键词,选取前 3~5 篇论文进行精读。读完再根据阅读中遇到的问题、产生的兴趣寻找其引用文献,或者重新搜索。

如果真的能认真读完 10 篇比较新的论文,相信你对该学术领域已经有最基本的了解了。

Step3:了解本领域的落地技术以及相关企业。AI 作为一个新兴方向,很多领域还处于研究阶段,真正能够应用到现实产品中的领域相当有限。截至本书完稿时,比较有代表性的领域也仅限于语音识别、语音合成、图片识别、人脸识别等。

当然,这里并不是说尚未投入使用的技术就不值得去了解或者投身其中了。这里只是想说,落地技术的范围并不算太广,了解起来也不难。

有了目标技术,再找企业就容易多了。虽然大公司掌控了当前 AI 领域的绝大多数人才和资源,但是也有越来越多的小企业在具体技术点上发力。

普遍来说,进大公司是为了平台,而进小公司则是为了“跟人”。AI 行业是一个强调学术背景的行业,一家公司也好,一个团队也罢,如果连一个具备最起码学术建树的博士都没有,那能走多远真的不好说。

从这一点来看,Step2 的调研过程也可以应用到此处。如果你有兴趣的小公司,尤其是刚刚成立不久的创业公司,不妨先评估一下技术合伙人的学术水平。

Step4:了解具体岗位的招聘需求。这里的具体岗位并不是指某公司的某个岗位,而是指同一类型公司的同一技术角色。

AI 行业的技术岗位，按角色可以简单地分为三类。

角色 1：科学家——研究理论，开发/改进算法。

角色 2：工程师——结合业务，训练模型。

角色 3：工程辅助人员——选择、清洗、标注数据等。

从实践来看，在一个团队中，如果工程辅助工作不是外包给第三方的话，则工程师本身也要肩负工程辅助的责任。虽然内部有分工，但工程师和工程辅助人员都属于一个团队，在职衔上也没有明显的区别。

一般来说，如果不是科班出身，没有在学校拿到相关专业博士学位，在入行的时候就不必指望成为 AI 科学家了。对于一般人而言，需要选择的是成为角色 2 还是成为角色 3。

当你选定了公司以后，注意先看看同类型的公司，至少先看看那些有代表性的公司的角色 2 和角色 3 是分开的还是合并在一起的。这一点，通过招聘启事的职位描述应该就可以找到。

从描述来看，角色 2 和角色 3 由不同成员来分担时，角色 2 显然比角色 3 “酷”多了。但正因为如此，两者的能力要求也存在区别。

通过领域、企业和角色共同定义了岗位之后，再根据岗位需求来反推需要学习的内容就有的放矢了。

3. 不要被“捷径”误导

笔者知道，很多人喜欢走“捷径”，比如去招聘网站用“AI”“人工智能”等关键词搜索一堆职位，看看某些职位要求的工具和语言是什么，直接去学，相当于从 Step1 直接跳到了最后。反正现在大多数职位都要求掌握 Python、TensorFlow，直接报个班学习怎么用 Python 调用现成的算法，或者怎么用 TensorFlow 处理数据不就好了？何必那么麻烦，还要看论文，学理论。

这种想法属于被“捷径”误导了。在 AI 行业从事技术工作，哪怕是做工程辅助工作，如果想要做得长久，有所发展，理论学习也是必不可少的。

如果要详细解释这一点，完全可以展开写一章内容。限于篇幅，此处仅举一个直观的例子。

工具就像武器，学会使用一种工具只是学会了使用这种武器最基本的招式和套路。

而理论学习则是学习策略，决定了未来在真实对战中，遇到对手攻击时，你选取哪些招式、套路，如何组合起来去迎敌。

当然，不排除现在有些公司跟风想做 AI，但自己没有人才，于是直接招聘，招聘要求上写着会用某工具就可以了。只学会使用工具做一些基本操作后，也许就可以获得这样的职位。但是这样的职位能长久存在吗？这样的职位能解决真正的问题，产生价值吗？做这样的工作能提高个人价值吗？

为了个人职业生涯的长远发展，我们仍需要打牢基础。

4. 制订学习计划

有了明确的目标，也就有了明确的范畴。同时，经过前文 Step2 到 Step4 的步骤，你也应该明白了学习到的知识以后能用来干什么，可以用到什么场景中去。

有了这些，就可以确定一个知识体系（至少是其中一个结构分明的分支），并进一步确定，对这个体系中的内容需要了解到什么程度。

有了知识体系和所需深度，就可以列举出所需要掌握的各个知识模块，并在此基础上划分学习内容。然后，根据需要学习的内容制订学习计划。

很多人在制订学习计划的时候会先规定学习时间。但是如果没有内容，光指定时间是没有意义的。制订学习计划的第一步应该是明确要学什么。具体步骤可以这样来走：明确知识范畴和应用目的，划定知识体系并确定深度目标，填充知识模块，列举针对具体模块的主要知识点。

这里可以借鉴写作文时列提纲的办法：在勾勒出轮廓之后，先把知识结构的骨干勾勒出来，分出篇章，然后列出大标题，再在其中填注小标题（知识点）。

这里很关键的一点是，计划中粒度最细的“小标题”到底有多大。

笔者建议：学习一个计划中的最小单位，如果是不脱产，即仅在业余时间学习，则掌握它的时间不应该超过一周；如果是脱产学习，则最好不要超过 2 天。

下面用一个例子来说明从确定目标到制订学习计划的过程。

从需求出发：笔者要学习基于机器学习的自然语言处理，学会开发聊天机器人的语言理解模块。

进行调研：通过向有类似经验的同事请教、到网上搜索相关文章和论文、研究开源项目等一系列手段，笔者了解到有两件事情必须要做：意图分类和实体提取。

要做这两件事情，有基于规则和基于机器学习/深度学习模型两类方法。基于规则的方法虽然直接、初始代价小，但是可扩展性差，所有规则都需要人工添加。基于模型的方法有较强的可扩展性，而且随着聊天机器人的用户和语料的增多，还可以通过反馈持续增强模型。在聊天机器人中应用机器学习/深度学习不仅符合产品发展的需求，也是当前业界的发展方向。

机器学习和深度学习的区别在于：前者适合数据量、运算资源相对较小，而开发者对业务理解较深的场景；后者则更加“自动化”，但对于数据量和运算量需求巨大，尤其是对人工标注的数据需求很大，先期投入巨大。

笔者决定在工作中应用机器学习模型，因此就先从机器学习学起。

(1) 确定目标：根据调研，确定了目标为机器学习。

接下来就是了解机器学习的学科框架。很多学校的计算机专业都开设了机器学习课程，找到相应的教学大纲和课本并不难。通过高等院校的教学大纲和课本目录来了解一个学科体系是最靠谱的方法。

机器学习界中有一位“大牛”——吴恩达，他在几年前就发布了一门机器学习的免费课程，因为其学术严谨，所以笔者将该课程作为了解机器学习的一个来源。

但是在一般情况下，不建议在最开始的时候就直接把公开课、培训课程作为构建学习体系的依据。

(2) 构建体系：通过对比多本机器学习著作，不难发现，机器学习理论是以模型为主要内容的。

之前在工作中要用到的意图识别和实体提取，正好对应分类模型和 seq2seq 预测模型。为了获取意图，可能还需要对原始语料做文本聚类。

因此，构建的体系可以以模型为主要节点。

(3) 填注内容：确定了主干节点以后，下一步就是填注课程的“血肉”。

归根到底，模型是用来解决问题的。比如，语言理解用到的分类、序列预测、聚类等，都是典型的机器学习问题，每一个问题都对应多个模型。每一个模型都有其适用的具体情形。

要获得模型，需要在算法和数据两个方面进行准备。模型还需要一个训练过程，训练过程一般迭代进行，期间要做多次验证，根据验证结果调优，最终通过测试来检

验模型质量。

通过进一步查询资料可知，要了解模型的运行原理，需要掌握求导、求微分、求积分、矩阵运算、概率统计等方面的知识。

另外，所有模型处理的内容都是数值，我们要把现实中的文本等人类可读信息转化成向量。这就需要掌握将文本转化为向量空间模型的能力，以及进行数据清洗、数据整理的能力，这些能力又都是以编码能力为支撑的。

(4) 制订提纲：当学习计划已经有了眉目以后，我们需要学习（复习）下列知识。

高等数学、线性代数和概率论的相关知识。

选择一批在实践中应用得较多的典型模型，学习其原理、数学推导过程和训练模型的算法。

训练模型、验证模型、测试模型、评判模型的指标。

将人类可读信息转化为数据的方法。

训练模型、测试模型的工具、框架和编程语言。

(5) 完善学习计划。

有了上面这些内容以后，再结合笔者之前介绍过的资料和自己当前的需求，就可以制订详细的计划了。

“机器学习”的学习计划如下。

A. 数学。

a. 数学分析：求导、求微分、求积分。

b. 线性代数：向量基本运算、矩阵基本运算、矩阵分解、多种矩阵的性质。

c. 概率统计：古典概率模型、贝叶斯公式、常见概率分布及其公式和特点。

B. 模型原理及公式。

a. 线性回归。

b. 逻辑斯谛回归。

c. 贝叶斯。

d. 决策树。

- e. 支持向量机。
 - f. CRF。
 - g. k 近邻。
 - h. 谱聚类。
 - i. LDA。
- C. 模型的构建和验证。
- a. 二元分类、多标签分类。
 - b. 归一化、正则化。
 - c. 验证、测试。
 - d. 交叉核实。
 - e. 精确率、召回率、得分。
 - f. ROC、AUC。
- D. 数据处理和向量模型空间的构建。
- a. 文本标注。
 - b. 自助投票、自适应提升。
 - c. 中文分词方法及原理。
 - d. n-gram 模型。
 - e. 文本特征提取。
 - f. 计算文档的 tf-idf 及其信息熵。
- E. 工具及语言。
- a. Python、Java、C#。
 - b. Python 库、Java 库、C#库。
 - c. 分词工具和词库。
 - d. word2vec。

F. 实践。

上述这些只是第一个版本的内容，大家可以先依据它制订一个为期两到三个月的

学习计划。在执行过程中，根据新的认识和具体需要随时调整计划。

5. 选择培训课程

选课容易，选对难。有的读者觉得选择培训课没有什么难度，现在互联网这么发达，把培训课当餐馆、电影院一样选就可以了。

此处笔者想说的是，选择培训课程容易，但选择有用的培训课程可就难了。

如果你选择了不适合的培训课程，浪费钱是小事，浪费自己的时间才可惜。

像 AI 这种热点、风口，如果因为总也学不到东西，错过了发展最迅速的阶段，错过的可能就是大好的职场机遇了。

(1) 培训课程的用处。

AI 行业的知识、技术，包括各种工程实践、最优方法都是纯粹的书面知识；其不像乐器、舞蹈这类文字不足以描述全部知识，而需要老师面对面指导和反复多角度演示才能掌握的技能。

对于这样的知识，培训课程的作用何在？省事。自己学习太累了，不如听老师讲。用耳朵听讲比用眼睛看书轻松，而且老师讲的内容比书上写的内容丰富、细致。

用自己的金钱和时间换取他人的知识和经验当然没问题，甚至可能是更高效的办法。但如果你觉得只要自己花了钱，就可以坐在那里，等着老师把知识“塞”进自己的脑子里，可就大错特错了。

在学习的过程中，学习者必须主动吸收、理解才能掌握知识。

而培训课程的作用，无外乎以下两点。

讲解细致。相对于力求清晰、简洁、逻辑性强，具有“攻击性”的书面资料，合格的课堂讲解应该丰富、细腻、包容性强，具有“防守性”。特别是和数学相关的内容，那些公式的推导过程，往往“一步千里”，如果有人能够给每个步骤提供一个解释，大家理解起来会比自己看书轻松很多。

和教师直接交流，获得即时反馈（答疑）。学习复杂的东西时，我们常有这样的体验：前面一大段都会，就是“绊”在一个点、一句话上，这一点不明白，后面的内容就全都弄不懂。自己的理解明明是这样的，可是习题答案却是那样的，是自己学错了，还是习题答案印错了？这种情况应该是这样的，那么那种情况呢？也是如此，还是有不同的解法？

在学习的过程中，疑问是难免的，如果疑问解决不了，则学习很难继续下去。无论是文字还是影音资料，都是只有传递信息的媒介，没有针对每个读者进行答疑解惑的过程，而培训课程恰恰可以弥补其不足。

（2）让培训课程更有效。

可能有些同学上培训班的目的之一是直接获得系统性的知识，而不必自己构建知识体系、制订学习计划。他们认为，自己就算确定了一个学习计划，也不太可能找到一个课程设置与自身计划完全一致的培训班。随便报个班，老师有教学大纲，自己听就是了，为什么要自己费力气来制订学习计划呢？此处需要提示这些同学一点：如果你自己都不知道想要学习什么东西，又怎么能够正确判断一门培训课程的质量呢？毕竟，培训（不只是培训，其实所有的教学都是如此）有效是教师和学生双方配合的结果。

教学双方能够配合好至少需要三个前提：教的人有足够的个人积累和授课技巧；学的人主动吸收课程内容并同步思考、理解；教师讲的内容与学生想学的内容相匹配，且张力适度。

学生在一无所知的情况下随便报个班，对于学习缺乏动力和认知，自己都不知道想从课程中获得什么；或者教的内容与学生的教育背景相差太远，学生一句都听不懂。在这种情况下，即使有幸碰到了好的老师、好的课程，学生能吸收的内容恐怕也非常有限。

通过制订学习计划，学生不仅能够明确自己的需求，而且能够了解授课内容之间的联系和优先级。

拿着自己的学习计划去找培训课程，能保证培训课程的有效性。原因如下：制订计划的过程是一个带着问题探索的过程，而“带着问题”很容易产生驱动力。当自己有疑问要解答的时候，就会有探索、学习下去的兴趣和好奇，不至于看见那些东西就想睡觉。对照自己制订的学习计划寻找满足需求的课程，有助于独立判断培训课程和培训教师的质量。制订计划也是一个了解自己的过程。这样能使得自己知道自身现在在哪方面有欠缺，缺到何种程度。同时对照培训课程的大纲和介绍，也可以辨识自己的知识背景和培训课程的内容是否匹配。如果要了解课程深度，则可以通过向授课教师或者培训学校进行咨询来完成；当然，了解自己则无人可以“代工”。

(3) 搜寻培训课程的途径。

当我们已经有了一份自己制订的详细学习计划以后,下一步就要寻找具体的培训课程了。

对此,笔者的建议是先从免费资源开始。

首先, AI 领域有非常经典的在线培训课程,这些内容都是完全免费的且非常容易找到。要选择有实战经验的老师讲的课,因为授课老师的这些经验网上一般没有。如果你和前面的例子一样,已经选定了机器学习作为学习目标,吴恩达的机器学习课程是一定要从头到尾学一遍的。每节课后面的习题都比较简单,要逼着自己做完。

其次,通过书籍、文章、论坛、微信/QQ 群、免费讲座和分享等渠道,去搜集大家的学习经验和各种培训班的信息。和有同样目标的人进行交流,经常能在不经意间得到优质资源。

再次,在尽量广泛地收集了大量培训信息后,对照课程大纲和自己的学习计划进行选择。教师的声誉、培训机构的声誉、网友的评价等都可以作为参考,关键是要考虑课程大纲与个人计划的匹配度。

最后,修改个人计划。要想找到和最初的计划 100%匹配的培训课程基本是不可能的(除非你的计划就是照着某个培训课程的大纲制订的)。这就不得不涉及对个人学习计划进行修改。修改的原则应该是什么?确保核心知识完备和大体结构完整,在细节上做出妥协。比如,你的自我判断是“打造 AI 产品是必须要补习的部分”,那么就不要找那些完全不讲 AI 产品实战而只讲模型的培训课。如果你确定了要学习打造 AI 产品的过程,就不要找只描述原理,把绝大多数时间都放在代码实现上的课程。但是如果课程的大致结构和自己的学习计划一致,只是有部分细节(如模型类型)和自己最初的学习计划不同,则可以再进一步考虑。如果你希望学习 CRF,而某个课程只讲 HMM,这也是可以选择的。因为两者都是 seq2seq 模型,只不过一个是判别模型,另一个是生成模型,二者有很多共性的地方。如果其他方面都合适,则不必为了一个模型而放弃一门课程。

笔者在“成为 AI 产品经理”课程中的一节课,主要内容就是按照上述方法一步步打造的。

(4) 实现有效学习的几个技巧。

在这里，笔者分享几个自认为行之有效的学习技巧。

关联。把新学到的东西和日常的工作、生活联系起来，进行思考。比如，将理论代入现实，对不同领域的事物、相似的内容进行对比等。以身边的实例来检测理论是否有效。这不仅能够加深自己对理论知识的理解，而且有助于自己改进处理日常事务的方法。

脑图。准备一个脑图软件。有什么发现、感想、疑问、经验等，全都由脑图把思路整理出来。如果对某个话题、某个题目有比较完整的想法，最好能够及时整理成文，并隔一段时间整理一下笔记，把分散的知识点整理成块，一点点填充自己的“知识地图”。

分享。把学习到的新知识、新理论、新方法分享给更多的人可以倒逼自己整理体系、记忆要点。把自己的感想、体会、经验分享出来能锻炼自己的逻辑思维能力和归纳总结能力。

(5) 制作知识速查手册。

这一点是针对学习 AI 知识的建议。AI 涉及的学科门类特别多，包括计算机、数学等学科的很多问题，AI 之所以让人“头大”，其实并不是真的有多难，而是系统比较复杂。

如果把常用的细小知识点都记录下来，按主题整理在一起，做成速查手册（小字典），需要用的时候迅速查找一下对应点，效果往往会非常好，这能让我们的学习 AI 技术之路顺畅不少。

第 4 章 AI 时代产品需求的特点和落地

4.1 AI 时代产品需求的新特点

AI 具有的前沿、综合、交叉学科的特点导致 AI 时代产品需求具有新的特点，AI 产品经理应该了解和掌握这些特点。

若想做好一款 AI 产品，除基础的技术设备、硬件设施以外，更核心的要素是人才。在技术层面需要非常优秀的数据科学家、数据建模专家、算法工程师等。在 AI 商业化的层面，需要综合素质十分强的产品经理。

对 AI 时代的产品经理而言，其对 AI 技术的理解可能永远无法在垂直领域达到算法工程师那样的高度，但是可以在应用领域找到自己的位置。因此，以下三点十分重要：懂行业中的深度痛点，能将问题描述清楚，同时不被传统思维束缚，提出新型的 AI 解决方案。理解算法，能够洞察技术的发展趋势，找到最适合当前行业应用的 AI 技术。能将解决方案和 AI 技术结合进行产品化，能将 AI 产品商业化。

非常资深的行业专家在思考问题时往往会给自己“设限”，比如公司老板说要做一个大型的购物中心，行业专家的工作流程可能就是：首先，给这个购物中心定位，它是一个高端的购物中心，或是一个低端的集贸市场；然后，再针对这个定位去选择进驻的品牌。但是，AI 产品经理的思维模式就不同了，他们不会直接定位购物中心是否高端，而是回归商业的本质，想办法使用 AI 技术让这个购物中心赚钱。AI 产品经理讲求生产效率、投资回报，而非购物中心是否高端。

对于算法工程师来说，由于只有某一类具体问题才能应用自己的数学模型，然后通过不断迭代去优化模型的参数，这在技术或者产品商业化层面是有先天劣势的，

所以需要产品经理帮助其拆解行业问题，分解出各个模块，并交给他们进行处理。

AI时代的产品经理要在自己所处的行业中扮演好“画龙点睛”的角色，这个“睛”就是商业化应用，即将技术和商业进行有机结合。

我们看看哪些行业已经开始应用AI技术了，大家可以从中找到感觉和受到启发。

(1) 交通运输业。

核心的生产效率问题：如何能更快速、更省油，人工成本更低，交通事故更少。

AI的解决方案：借助无人机、智能导航来减少时间成本和燃油成本；借助无人驾驶降低人工成本和事故成本。

应用案例：滴滴的派车算法、谷歌的无人驾驶系统、大疆的无人机、亚马逊的自伺服仓储。

(2) 传统制造业。

核心的生产效率问题：如何使得生产周期更短、人工成本更低、良品率更高。

AI的解决方案：使用能够精密作业且24小时不停歇的工业机器人。

应用案例：生产特斯拉的机械手臂；用于维修高空电缆的机械手臂。

(3) 高端制造业。

核心的生产效率问题：如何能更高效地设计出更优秀的产品方案。

AI的解决方案：通过学习历史数据，帮助系统更快地找到更好的设计方案。

应用案例：某芯片制造巨头通过AI算法将芯片的设计时间大幅缩短；某药企利用AI大幅提升药物晶体的选型速度。

(4) 教育。

核心的生产效率问题：如何能快速提高学生的学习成绩，如何能个性化地培养出更优秀的人才。

AI的解决方案：自适应考试、自适应学习、个性化职业生涯规划。

应用案例：Knewton通过自适应评测精确诊断学生对知识点的掌握情况，并制订有针对性的学习方案；iPIN的“完美志愿”根据个人特质和数千万人的职业发展经历，为学生制订个性化的职业生涯规划。

(5) 法律。

核心的生产效率问题：如何能更准确地预测案情走势，匹配法律资源。

AI 的解决方案：案情分析，案情推理，律师背景分析，法律文书写作。

应用案例：基于 IBM Watson 开发的 ROSS 和基于 iPIN 开发的“法律谷”可以帮法官分析案情、匹配律师。

(6) 财务。

核心的生产效率问题：如何提升对账、盘点、审计、合规的效率。

AI 的解决方案：智能数据录入，异常行为监测，自动文书写作。

应用案例：德勤率先将 AI 引入会计、税务、审计等工作中，代替人类阅读合同和文件。

(7) 金融。

核心的生产效率问题：如何在更短的资金周期内得到更高的投资回报。

AI 的解决方案：解读财报新闻，预测投资回报。

应用案例：高盛旗下的 Kensho 分析大量数据来进行智能投顾分析；某券商使用机器人代替数据录入员；Virtu Financial 用用户交易行为数据做投资，1500 天中仅有 1 天亏损。

(8) 医疗。

核心的生产效率问题：如何能更快速、更准确、更低成本地进行诊断，并给出治疗方案。

AI 的解决方案：疾病诊断，结合基因技术的精准医疗。

应用案例：IBM Watson 的肿瘤诊断模块；腾讯的肺癌诊断方案；国防科技大学的肺结核诊断方案。

(9) 文字工作。

核心的生产效率问题：如何更快地写作，如何让文字更优美，如何让情节更吸引人。

AI 的解决方案：标准文书写作。

应用案例：百度可以写古诗的写作机器人；今日头条可以报道体育赛事的写作机器人；iPIN 可以整理招聘需求的写作机器人。

(10) 艺术创作。

核心的生产效率问题：如何获得更多、更优质的灵感，如何构思更精妙的作品，如何降低制作成本。

AI的解决方案：灵感生成器。

应用案例：谷歌的绘画机器人；谷歌的谱曲机器人。

(11) 新零售。

核心的生产效率问题：如何能提升销售利润。

AI的解决方案：智能招商运营系统；智能定价系统；智能推荐系统；智能广告投放系统。

应用案例：万达用AI提升商铺坪效；沃尔玛用AI做动态定价；宝洁用AI辅助品牌广告投放。

(12) 编程。

核心的生产效率问题：如何提升编写代码的速度、提高测试的通过率。

AI的解决方案：自动代码生成器。

应用案例：ViV通过一系列的自然语言处理程序来实现语言识别，并且基于英语词汇库建立了复杂的自适应的计算机程序生成系统。

在笔者看来，每一次的技术革命都是效率的革命，蒸汽机的发明是人类体力劳动效率的提升；互联网的发明是人类信息传递效率的提升；AI的出现则是人类脑力劳动效率的提升，这次的提升能够从真正意义上解放人力，让人类有更多的时间思考新的问题，而把那重复的脑力劳动交给机器。

AI时代的产品经理要怎样抓住AI产品的核心特点呢？答案是回归商业本质。

找到当前阻碍公司发展或者提升行业效率的核心问题，按照业务流程进行逐级拆解，针对每级提出多重的解决方案，并对方案进行评级，以工程思维选取最优解。

构建自己的人工智能技术库，紧跟技术发展趋势。选取最合适的技术，应用到自己的解决方案中，看能否大幅提高效率，同时评估性价比，择机进行产品化。

笔者根据自己打造AI产品的实践，发现AI时代产品需求的新特点可以概括为如下四点。

(1) 长尾深度化。相比于互联网时代，AI 时代产品的需求来源不再是浮在浅处的简单的需求，而是更具有“长尾”特点、更有深度的需求。例如，电商平台除了主流的几大综合平台，又涌现出了垂直类目平台，然后又出现了掌握商品供应链上下游的电商平台。

(2) 数字化。新零售是马云首先提出来的概念，新零售最大的特点是人与货品的匹配，最核心的竞争力是：人和货品的匹配，线上、线下的融合，增加线下的人效和坪效。人和货的匹配离不开对数据的挖掘和利用，增加线下的人效和坪效离不开对数据的抓取、对数据制订规则并进行文本分析，这样才能提升服务用户的能力。

(3) 智能化。笔者对 AI 的定义和对 AI 的理解一直是：AI 不是为了超越人，而是为了在广义范畴上提升效率，优化某一事物。智能化的例子比比皆是，如无人扫地机器人、无人驾驶轮船、无人汽车等。

(4) 多元化。AI 产品的需求来自一切可以使得某一事物互联网化、多元化的场景。例如，老师讲课时用的智能黑板，家中的智能化妆镜等。

4.2 AI 时代寻找产品需求的方法

4.2.1 重新定义寻找产品需求的方法

随着消费者的个性化需求越来越受到重视，需要以恰当的方式提供恰当的产品给消费者，这给企业带来了前所未有的压力。依靠传统方法来规划和交付产品的企业将无法跟上不断变化的市场。

具有前瞻性思维的 AI 产品经理们正在设法利用所得到的数据进行需求预测，比如，用户需要什么？用户何时需要？用户希望通过何种渠道满足需求？通过 AI，产品可以弥合供应链和用户实际需求之间的鸿沟。

AI 为预测用户需求提供了后盾，并为产品需求提供竞争优势，帮助企业建立用户信任。企业运用这种创新方法来提高用户满意度，获得竞争优势，实现更高的品牌价值。

在 AI 时代，产品依然是围绕用户需求定义的，这个本质是不变的，不过需求的范围和需求的层次不同了，因为 AI 满足需求的效率更高，所面对的需求层次也更深。例如，以前电商产品经理只是对电商产品进行设计，而现在电商产品的框架大同小异，电商产品经理需要具备数据分析、多渠道获客、持续优化产品等能力。

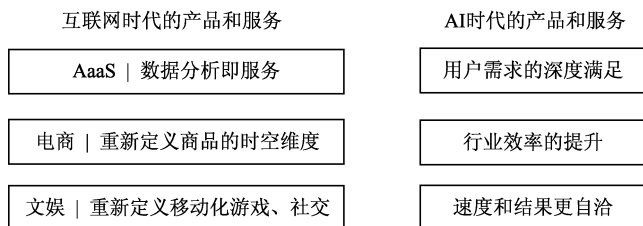
需求管理、需求定义、需求确认、需求跟踪等与需求相关的职责都是公司对产品经理最基本的要求。因为产品经理是公司产品的负责人，而产品是用于满足用户某种特定需求的。即使我们来到了 AI 时代，产品依然是围绕用户需求定义的，这个本质是不变的。

在互联网时代，电商平台的目标是让顾客能够快速、精准地购买到合适的产品，即构建人与商品之间的关系。移动社交平台或产品的出现是为了降低人与人沟通的成本，即构建人与人之间的关系。数据分析公司的 AaaS（数据分析即服务）产品的诞生是为了对传统企业分散、庞杂的数据进行汇集、分析、挖掘，从而重新构建企业与用户的关系。

AI 时代的产品本质上是全面优化和提升上述场景中现有技术的手段，实现用户体验的升级和解决方案效率的升级。

互联网时代的产品经理构建的是基础设施，即从人与人、人与物、人与数据的关系的角度构建了桥梁，实质上是优化了信息存储和互通的方式，因此产品经理关注的主要是入口和流量的走向。

而 AI 实际上给人类带来的是技术创新驱动下的产业升级，本质上更关注产品本身给用户的体验价值，如下图所示。



互联网时代与 AI 时代的差别

AI 技术的飞跃发展为产品设计和需求定义带来了新的思路和逻辑，具体可以总结为以下七个方面。

(1) 产品逻辑化繁为简，用户学习成本降低。AI 产品的目标就是降低用户的使用门槛，尽量减少用户的交互流程和难度，让产品的使用过程更接近用户的自然行为。

例如，语音交互产品与传统的点击鼠标、敲击键盘、触摸手机屏幕等交互方式不同，用户通过说话即可完成唤醒、查询、关闭和一系列复杂的人机交互操作；借助人脸识别身份验证，登录产品需要的输入账号、输入密码、输入验证码等内容，变成了在镜头前露个脸。

由于新技术的诞生，产品经理不能再用“线性思维”设计产品，需要了解更多的技术可能性，尝试用“颠覆式思维”设计产品。

(2) 从用户角度考虑投入产出比。AI 产品由于具有更复杂的系统架构和实现逻辑，因此，实现某一功能往往伴随着高昂的代价，而与此匹配的功能的价值却在很多情况下与代价不成正比。因此，对于产品经理来说，选择更容易展现商业价值的需求作为产品的切入点很重要。尤其当产品或功能还没有被用户认可或当产品属于一个新的市场时，最终的实现效果和价值都很难预估。产品经理应选择用户最“痛”的点或者直接和利益相关的点作为需求切入点。

这也是 AI 产品在制造业中落地相对较快的原因，尽管研发 AI 产品的成本很高，但产生的直接回报却更大。下图是采用 AI 代替普通工人识别地板的品质，系统能够迅速识别并发现地板是否有裂纹、结疤、虫眼等方面的问题，使得生产地板的效率大大提升。



采用 AI 代替普通工人识别地板的品质

另外，广告精准投放、电商平台中的搜索推荐都是效果比较直观、价值量化相对容易的场景，产品经理应在所处行业中找到这些场景。

(3) 算法可解释性差，产品需要逐渐获取用户的信任。使用复杂算法模型的 AI 产品对于用户来说大多属于“黑盒”产品，工程师或产品经理均无法很好地解释产品的实现原理。在很多领域中，如果不能证明算法的有效性和准确性，用户就不会对产品付费，甚至会对品牌产生强烈的抵触情绪。

例如，在某些基于数据挖掘的商品推荐系统中，推荐商品的广告如果让用户有一种被忽视的感觉，或有一种被侵犯隐私的感觉，是会遭到用户投诉甚至被弃用的。

AI 产品首先需要通过在某个具体场景的预测和推断能力证明自己的技术实力，进

而树立专业形象，步步为营地争取用户的信任。当公司和品牌都处于起步阶段时，更忌讳大步向前，因为这样更容易遭到用户的抛弃。

(4) 传感器技术的飞速发展带来了多元化的交互行为。AI与传感器技术的融合产生了良性循环，传感器采集的数据用来对模型进行训练，模型的完善也提升了传感器采集数据的效率。

例如，在无人驾驶产品中，关键系统和解决方案分别是传感器、高精度地图、高级辅助驾驶系统（Advanced Driver Assistant System, ADAS）和车联网。传感器作为无人驾驶车的感应系统，用来接收和感知行驶时环境的动态变化；高精度地图为无人驾驶车提供全局视野；ADAS负责对静态、动态物体进行辨识、侦测与追踪，从而预先让驾驶者察觉到可能发生的危险，有效增加驾驶的舒适性和安全性；车联网能够保证传感器数据的更新，保证无人驾驶车的状况与周围环境的信息处于实时更新的最新状态。

正是因为传感器技术的发展及其成本的降低，ADAS的部署成本大幅降低。ADAS起初只能安装在高端车上，随着时间的推移则逐渐地被使用在入门级乘用车上。传感器不仅在机器人、无人驾驶等领域有着广泛、成功的应用，日新月异的传感器技术也为产品设计和定义提供了更大的想象空间。产品经理应学会合理利用多种传感器创造更多的交互方式来满足用户的需求。

(5) 产品的需求并不一定来自确定的因果关系。在过去，产品经理根据用户明确的需求设计产品，产品研发出来的结果会和原型设计保持一致。但是AI产品经理需要完全不同的思维模式，不应该再花大量的时间和资源来寻找确定的因果关系，而是应该通过大量的数据挖掘手段探索出相关性，并用数据指导产品设计。产品经理输出的需求未必是确定的页面内容，而是一堆规则和策略。

例如，Google Adwords 关键词竞价广告是一种通过使用谷歌的关键词广告或者内容联盟网络来推广网站的付费推广方式。设计这个产品的产品经理一定不会告诉算法工程师给什么样的用户推送什么样的广告信息，因为产品是“千人千面”的。产品经理只需要给广告主提供后台的“推广喜好配置”功能及“推广效果管理”功能即可。至于最终用户打开的页面是怎样的，完全由算法模型计算得出结果，即搜索结果页面都是基于商家偏好配置和用户精准匹配算法实现的个性化页面。

产品经理在定义需求前应该充分了解技术水平和可利用资源的局限性，避免定义一些研发人员很难实现的需求。一个完整的 AI 产品体系的搭建通常需要考虑基础设施、数据采集、数据处理、推理和决策等若干环节，产品最终的实现效果取决于上面所有因素的协同。

例如，设计一个能在多种场景进行复杂交互的机器人产品时，由于对交互实时性要求较高，系统应该具备足够的硬件，包括计算硬件（GPU、CPU、FPGA、ASIC 等）、存储硬件以及各种智能模组（如视觉模组、语音模组）等，因此需要产品经理在提出需求的同时综合考虑配套硬件的要求。

（6）不同场景对算法模型的准确率、召回率的要求大相径庭，需要在需求设计时区别规定不同场景对算法模型的衡量标准。

（7）从不同的角度思考 AI 时代产品的需求。

综上所述，AI 是个前沿、综合的交叉性学科，应该在广义范畴上理解 AI，同时利用 AI 技术重新设计产品，重新定义寻找 AI 产品需求的方法。

4.2.2 外部方法

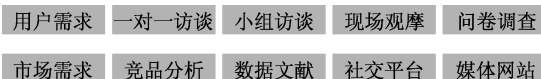
笔者根据多个 AI 产品的实战经历，总结了一套寻找 AI 产品需求的外部方法和内部方法。

向外部寻找 AI 产品需求的主要方向如下。

（1）用户需求：更快、更准、更新、更智能。

（2）市场需求：创新、垂直、弹性。

外部方法即在公司以外的资源中寻找需求场景，具体包含下图所示的方法。



向外部寻找 AI 产品需求的方法、渠道

例如，通过一对一访谈、小组访谈、现场观摩、问卷调查等方式寻找用户的需求。真正懂用户需求的产品经理并不多，即使是一些初创企业的创始人也未必真的懂，因为 AI 产品不仅要懂谁需要什么，用什么满足，还要懂得如何运用 AI 技术来提升产品的体验。

4.2.3 内部方法

在 AI 时代，AI 产品的需求除了向外部寻找，挖掘产品需求更重要的在于向内部寻找。

向内寻找需求是自己了解自己，自己给自己定位的过程。在市场竞争中，自己更容易获取自己的数据，也更了解自己公司的战略，所以向内寻找需求是一种更科学的方法。

向内寻找 AI 产品需求的具体方法如下。

(1) 企业战略。例如，3M 便签就来自公司胶水部门和纸张部门的战略融合。

(2) 头脑风暴。通过收集企业产品、运营、技术、市场等部门有价值的建议和意见，对其进行分析、挖掘，以形成新的产品需求。

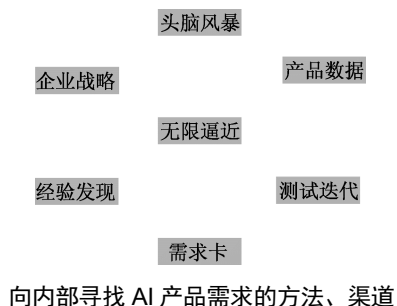
例如，在 2012 年年初，艺龙推出“摇一摇”住酒店，这个功能促使艺龙团购酒店的订单量达到携程团购酒店订单量的 70%。这个功能一方面源于“智能位子”，另外一方面源于将用户 ID 和酒店属性进行匹配。

(3) 产品数据。从产品数据中挖掘需求是一直存在的寻找需求的方法。在 AI 时代，笔者不建议仅仅从数据中寻找逻辑的推导关系，而应该从数据中多找关联关系。

例如，饿了么通过对平台的订单数据进行分析，发现订单中出现频率最高的品类是麻辣烫、沙县小吃、兰州拉面、黄焖鸡米饭，将这样的数据与消费升级联系起来，会发现消费升级的空间非常大。

AI 产品需求的最大特点和方法是“无限逼近”，这一特点是符合事物发展的基本规律的。寻找 AI 产品需求的重点是向内挖掘且无限逼近。

向内部寻找 AI 产品需求的方法、渠道如右图所示。



小结：在 AI 时代寻找产品需求首先需要重新定义 AI 产品的需求。在此基础上通过具体的向外、向内寻找来实现对需求的精准挖掘。AI 产品的核心是提升效率，使得重复无趣的工作由 AI 来完成。

4.2.4 AI 时代的产品创新源自产品经理的自我增强

在寻找 AI 产品需求的基础上，AI 产品的创新源自产品经理自身的持续增强。

1. AI 产品经理需要不断自我增强

机器学习里有一种算法叫自适应增强算法，实际上产品经理的成长经历也是一个不断自适应增强的过程。这样的过程被成功的产品经理证明是正确的。如滴滴投资人王刚是这样评价程维的。

程维一贯的做法是组织先行，向高人请教。
他和柳青把杨致远请来做滴滴的高级顾问。
我每次跟程维聊天，都能感觉到他的眼界、格局、野心又有了很大的变化，每次都不一样，每次都有意外的收获。其他投资人跟我的感觉也基本一样。
程维有极强的学习能力，今天他已经开始跟马云、马化腾、柳传志等中国顶级的商业领袖学习，也向孙正义请教。这是个垂直的电梯，每上升一层都有压力，但每上升一层以后看到的风景都不一样。有了商业领袖的帮助，他的解决大家出行困难的梦想正在逐步实现。
他是独立思考和极致执行的完美诠释。

滴滴投资人王刚对程维的评价

根据笔者多年来做产品的经历，笔者认为产品经理每年都要面对一个新的市场环境，每月都要完成一次迭代更新，这种迭代既是迭代自己的产品，更是迭代产品经理自己的思维方法。食物过期很容易辨识，但是思维过期了有时候不好辨识，影响却更大。

不同的思维方式会对同样的事物产生不同的理解和认知，不同的理解和认知会使自己对事物产生不同的概念。例如：当移动互联网时代到来以后，如果还是不停地劝用户在 PC 上生产、阅读内容，会导致产品错过移动互联网时代，这样的例子不胜枚举。

在大型互联网公司取得成功的产品经理离职后创业之所以没有产生杰出的战果，往往是因为沿用原有的成功方式。例如：在大型互联网公司里面，大多数产品经理只需要做好功能，分析好每日产生的大量数据即可，而新产品往往没有大量的数据。但正是因为新产品没有大量可供参考的数据，没有一个确定的商业模式供产品经理去添加功能，所以才更能锻炼产品经理的综合思考能力，迫使产品经理从 0 开始思考产品发展过程中每一步的增长点。

工作 10 年以上的产品经理大多接触过 PC 互联网产品、移动互联网产品和近两年的 AI 产品，这三种产品不是替代关系，而是一种更加深度融合的关系。打造 PC 互联

网产品是产品经理的基本功课，有了PC互联网产品经理以后，CRM/ERP/WMS等软件才开始讲究用户使用体验，而不仅仅是作为一个录入信息的工具。

移动互联网的信息在产生效率上是非常不同于PC互联网的。比如，在PC互联网时代，订机票、订酒店都是提早在PC上订好；在移动互联网时代，用户习惯变成了何时需要何时定。再比如，以前很多主播在PC上进行直播，移动互联网时代的很多主播都是走到哪播到哪。信息的产生效率呈指数级增长。

笔者个人认为，错过PC互联网时代，你可以花一年时间来补；错过移动互联网时代，你可以花三年时间来补；而错过AI时代，可能一生都无法弥补。

因为三个阶段的特点完全不同。

(1) 第一个阶段：PC互联网时代。

记得刚离开阿里巴巴时，经常有朋友邀请笔者帮忙规划产品，做LED照明灯的老板要做一个灯的商城，开经济型酒店的老板要做酒店用品的商城……一名没有产品经验的产品经理，可能要花上一年的时间才能打磨好一款不错的PC互联网产品。

(2) 第二个阶段：移动互联网时代。

在移动互联网时代有三类组织。

由传统PC互联网企业转型成为移动互联网企业的组织。典型代表是腾讯。马化腾在2013年的GMIC上说微信获得了一张移动互联网时代的“站票”。2014年，马云宣布阿里巴巴“All In”移动互联网，随后手机淘宝的流量超过了网页端淘宝。

移动互联网新贵。典型代表是今日头条、滴滴、美团等。

移动互联网深耕者、回滚者。典型代表是抓住90后、95后等较年轻用户碎片化获取信息、进行社交的特点的组织。

(3) 第三个阶段：AI时代。

有了PC互联网、移动互联网的发展，AI产品成为互联网发展的必然结果，原因有以下三点。

数据。数据越来越多会促使AI主动去表达自己、发展自己，AI开始对事物有主动认知。例如：Siri会以开玩笑的方式回答用户的刁难问题。

举一个例子，淘宝每日产生的数据量在 2015 年就达到了 7TB。2016 年 12 月，微信全球月活跃用户数已经达到 8.89 亿人。

算力。计算芯片的架构在 2018 年还都是按冯·诺依曼体系来做的，相信未来我们可以看到非冯·诺依曼架构的出现。随着量子计算的发展，笔者认为 AI 会出现新的变化。

AlphaGo 需要 1920 个 CPU 和 280 个 GPU 才能完成它的计算，这在以前是不可想象的。深度学习的崛起确实是它的原始的动力，但是真正能够让它稳定往前走的动力是算力的发展。

算法。算法的突破相比较算力和数据量的发展是一个更加缓慢的进程，算法的进展往往以 10 年为一个小的周期。但是算法一旦出现系统级的突破，往往会产生工业革命级的变革。

2. AI 产品与传统产品的区别

(1) 市场环境。AI 产品诞生的市场背景是每个垂直的细分领域均有 App 产品，这就需要 AI 产品的用户体验比原来的产品好上很多倍或者速度比原来的产品快很多倍才能赢得市场。

(2) 供应链产业。单纯做 App 时是不需要考虑供应链的，但是广义范畴上的 AI 产品涉及从数据获取到数据分析再到数据应用的全流程，因此少不了对硬件的使用。例如，用深度摄像头采集更多的数据，用物联网设备采集人和物体的行为数据，均需要软件、硬件的融合。AI 产品更加考验产品经理的综合素质，除设计、管理好传统软件的上下游以外，还要深挖供应链。例如，当你的摄像头等 AI 产品的主体硬件出现 bug 的时候，你需要联系摄像头厂商，这个时候需要调动摄像头的研发团队甚至一个工厂来配合你。

(3) 需求。AI 产品的需求的变化趋势是精准、个性化、成本降低等。例如，在新零售时代，需要货来匹配人，这需要 LBS (Location Based Service, 基于位置的服务) 和智能传感器的数据以实现精准的匹配。线下商铺原来是不知道哪个用户来逛街，哪个潜在消费者在哪个商品前停留得更久，节假日的什么时段用户结账是需要排队的……这些需求正好适合 AI 产品大展身手。AI 产品能够在提供个性化服务的同时降低运营成本。

3. 未来只有 AI 产品

在未来没有人会做一款非智能的产品，未来的产品只有 AI 产品。主要原因是 AI 产品能提升效率。例如，以前没有智能翻译时需要人工翻译，这个工作既累人又没有新的创作，而借助语音识别可以实现实时翻译，使得效率出现指数级的提升。此外，AI 产品能提升用户体验。比如，现在回到家以后，带有人脸识别的智能锁自动开门；然后用户通过智能音箱可以打开空调，拉上窗帘，享受音乐。这种体验比到处找遥控器不知道要好多少倍。

4.2.5 小结：成为 AI 产品经理的心路历程

笔者觉得做产品经理的历程是一个不断自我增强的过程，是把所有的“坑”都踩过以后，屡败屡战、越战越勇的过程。

AI 是时下和未来一段时间内先进的生产力。产品经理要观察和掌握先进的生产力，并把先进的生产力在企业产品中落地。依靠过去的经验来应付不断变化的世界是没有前途的。例如：在 PC 互联网时代，设计师通过个人电脑上传作品，装修公司通过个人电脑生产、宣传内容；在移动互联网时代，产品如果还用那套模式就会被时代所抛弃。在 AI 时代也是一样的，产品经理只有提早准备好 AI 时代的思维方法，才不会被时代所抛弃。再以家装为例，AI 时代的家装是 AI 的家装，装修用的材料变了，合作伙伴也变了。以前装修平台只需要跟建材生产企业合作，但在 AI 时代，装修平台必须跟数据公司合作，和被 AI 赋能的供应链企业合作，同时需要跟 AI 算法公司合作。

AI 时代的产品经理要掌握先进的生产力，时刻牢记，做产品经理是一个不断自我增强的过程。

4.3 AI 时代产品需求落地的方法

4.3.1 AI 产品的落地

本节先讲述 AI 产品落地的例子，然后总结 AI 产品落地的相关知识。

1. 如何快速落地 AI 应用

如果说前 50 年是 AI 的潜伏期，那么，2016 年以来就是 AI 的全面爆发期。我们

在各行各业都可以看到 AI 的身影，无论是各厂商推出的自研 AI 算法，还是满屏幕的基于 AI 的智能化应用。AI 技术已经渗透到生活的各个领域当中。那么，在 AI 技术实际落地的过程中，会遇到什么样的瓶颈？深度学习的框架这么多，企业该如何选择适合自己的那一个？

笔者在这部分聚焦 AI 产品、AI 技术的落地，讲述如何构建高可用的异构 AI 计算平台，如何流程化地生产和发布 AI 产品，如何使用深度学习优化工具等内容，总结了一套快速落地 AI 应用的路径和资料。

我们先看一下美团配送网络智能规划系统。截至本书完稿，美团已承担日均千万级订单量的配送，为上百万商家和上千万用户提供服务。在这个过程中，如何打造配送网络的规划工具、制订规划策略至关重要，仅仅依靠人工来对千万级订单量的网络进行调整，不但效率低、成本高，而且标准不一。根据调研，美团的技术专家就美团配送网络的总体设计和规划、配送范围的设定和算法演进等内容展开了深入研究。

美团配送网络的整体设计包括三个核心概念：取件范围、送件范围以及配送范围。取件范围表示加盟商的“势力范围”，每一个加盟商对应一个独立的配送团队，取件范围即各个加盟商的配送团队能够在哪些地方取件；送件范围表示加盟商的配送团队可以送达的范围；配送范围是指商家的配送范围，只有在这个配送范围内的用户才可以看到这个商家。

那么，美团怎样规划取件范围和配送范围呢？在取件范围的规划上，美团以效率、体验和公平为基准点，制订了生成划分方案。即基于聚类和多目标优化生成候选划分方案，将生成的候选方案输入仿真系统，通过真实的历史订单、骑手配送情况来模拟配送过程，得出效率和体验评估报告。在配送范围的规划上，美团首先将配送场景划分为常规的配送范围、恶劣天气时的配送范围、闲时的配送范围、新商家的配送范围等，通过 Block 决策、导航路线图、中间指标、商家推荐、机器学习预估多种技术手段，分析出不同场景下较优的配送范围。

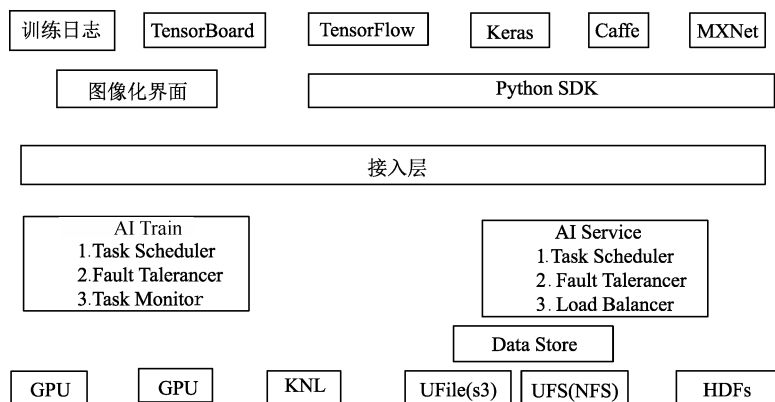
我们再来看看 UCloud 的例子。

笔者在前文提到，AI 产品在 ToB 端较为高效，如何构建商用 AI 平台是值得思考的。笔者经过与 UCloud 高级研发工程师宋翔交流，结合 UCloud 在研发 AI PaaS 类产品时遇到的问题，以及在构建 AI 平台时遇到的挑战与解决方案，总结了如何构建

一个高可用的异构 AI 计算平台。

业界主流的 AI 算法和框架非常多，算法包括 VCG、LSTM、RestNet、Fast-RCNN 等，而框架有 TensorFlow、Keras、Caffe、MXNet 等，其组合种类更加让人眼花缭乱。那么，如何在繁杂的组合里选择合适的算法和框架？UCloud 考虑的五大要素是：算法兼容性、平台扩展性、集群化、纵向拓展、易用性，用户可以以这五大要素为基准点，根据平台构建的目的来搭配合适的算法和框架。

基于以上五大要素，UCloud 构建了自有的 AI 基础平台，里面包含 AI Train 和 AI Service 两大核心功能，如下图所示。



AI 基础平台整体架构

图中，最上层包括训练日志和 TensorBoard 框架，其下面是图形化界面和 Python SDK。在图像化界面里面主要完成一些基本的部署操作。接入层下面为平台的核心——AI Train 和 AI Service，底层封装了所有的硬件和存储接入。这个平台看起来虽然简单，但基本上包含了 AI 商用平台的常用功能和接口，可以说是“麻雀虽小，五脏俱全”。

AI 框架的构建非常重要的两点是保障平台的兼容性和可靠性。在保障兼容性上，UCloud 采用了容器封装和数据接入两种方法实现环境的分离，将不同功能的 AI 框架、依赖环境、硬件设备、存储分隔开来，让不同的模块实现不同的功能。在可靠性的实现上，UCloud 主要采用负载均衡、请求调度算法优化、性能监控以及高可用的部署等方式，完成全局的弹性扩容。

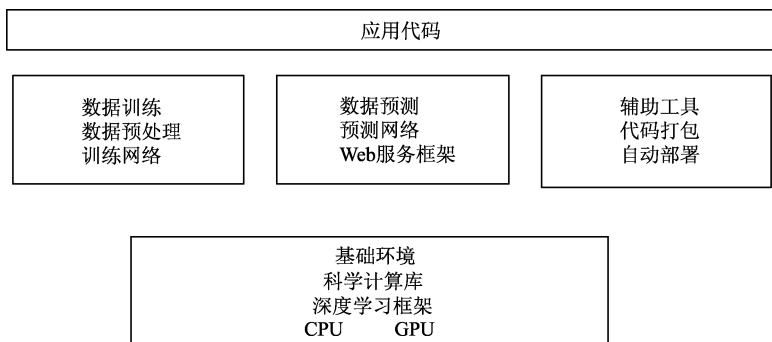
值得一提的是，UCloud 在 GitHub 上开源了自身基于 AI 基础平台的公共镜像库

和 SDK，用户可以通过开源的 SDK 将 UCloud 开源的镜像库打包成自己可以应用的镜像。开发者可以在本地使用镜像，也可以将镜像放在 UCloud 的在线服务或者训练平台，这些都是完全兼容的。

2. 如何轻装上阵玩转 AI

在产品化 AI 服务的过程中，普遍会遇到管理、使用、封装基础资源等问题。如何提升 AI 产品的研发效率，简化非核心业务的研发流程是各个公司非常关心的问题。通过与 UCloud 高级研发工程师范融交流，结合实践，笔者在这里详细分析如何流程化生产和发布 AI 产品，节省开发时间和成本。

在本地开发 AI 应用之前，开发者需要提前配置基础的环境，包括提供对 CPU、GPU 的支持，以及准备深度学习框架和加载科学计算库等，基础环境配置完成之后，即可进行开发工作。下图为 UCloud 本地开发 AI 应用的整个过程。



UCloud 本地开发 AI 应用的整个过程

下层是基础环境的配置，最上面的白框是应用代码部分，这部分主要解决两个问题：一个是数据的训练，另一个是数据的预测。一般来说，预测网络和训练网络是类似的。如果预测模型和预测程序不只在本地运行，则需要增加一个 Web 服务框架。

为了简化应用代码的部署和构建过程，UCloud 提供了一些开源的工具和框架，包括辅助工具里面的代码包和自动部署工具，开发者也可以在 UCloud 提供的 SDK 里面免费下载使用，以省去许多部署的烦恼。

本地训练 AI 算法虽然安全性比较高，但在资源申请和迭代开发方面具有一定的局限性，例如资源扩容较慢、多任务穿行时等待时间较长、调参困难等。范融认为，

将部分训练放在云端实现可以有效完成云端资源的按需申请、多任务并行处理以及任务参数记录等工作。在数据上云的实践方面，UCloud 提供了比较完善的工具和方法，用户只需通过代码改写、打包镜像、上传训练数据以及启动任务四个步骤即可轻松上云。

通过与来自英特尔的何普江交流，笔者总结了使用深度学习性能优化工具的实践经验，包括其在高性能计算库（MKL 和 MKL-DNN）和框架（英特尔发行版 Caffe）中的使用和注意事项，以及基于 CPU 的、高效的循环神经网络的实现过程。

英特尔的数学核心函数库（MKL）是一套高度优化、线程安全的数学例程、函数，能加快机器学习、科学研究、工程、财务和设计等领域的数学处理，并提供了对稠密及稀疏线性代数（BLAS、LAPACK、PARDISO）、FFTs、Vector Math、Summary Statistics 等内容的支持，具有标准的 API 以及高度优化的特点，能最大限度地发挥多核心和 SIMD 指令的优势。

与 MKL 不同的是，MKL-DNN 是一个开源的深度学习框架，通过两种方式可以使用它。一种是将训练好的模型直接导入模型转换器，然后再用 topo.txt（网络拓扑）生成 Inference.cpp（基于 MKL-DNN 的 Inference 代码，在初始化的时候加载 weights.bin），这套逻辑需要用到英特尔的网络拓扑模型的中间表示形式；另外一种方式是待 topo.txt 生成后，再用简单的代码转换器实现轻量级的 InferenceFramework.cpp（可以解析 topo.txt 并加载 weights.bin）。

Caffe 是一个清晰、可读性高的深度学习框架，英特尔发行版的 Caffe 同样也是开源框架，它基于 BAIR/BVLC Caffe 改进而来。英特尔在 BAIR/BVLC Caffe 的基础上添加了对流行检测网络（如 SSD）的支持，同时也支持多节点训练以及 Caffe、MKL 2017 和 MKL-DNN 三种引擎。

在对循环神经网络的优化上，英特尔主要在高效的 GEMM 实现、一些小操作的合并、并行化元素级操作、合理的数据排布以及低精度表示上进行了一系列改进，优化之后的 CPU 性能可以比 GPU 的性能高出不少。

以上是笔者了解的一些 AI 产品快速落地的“术”，下面我们再看相关的“道”。

3. AI 产品快速落地的“道”

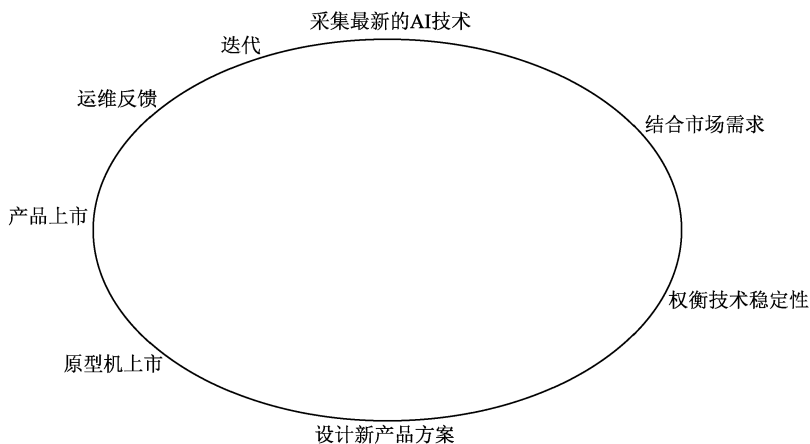
根据腾讯在 2017 年 11 月发布的《2017 互联网科技创新白皮书》显示，截至 2017

年 6 月,中国创业投资机构共有 767 项针对人工智能的投资案例。2017 上半年发生的融资额超过 150 亿元,累积融资额攀升至 635 亿元,占据全球融资总额的 33.18%。很多公司都在标榜自己有 AI 技术,更多公司则在尝试将 AI 融入自己的产品和服务之中。

当 AI 不再神秘之后,我们明白 AI 并不是一项只能掌握在少数人手中的“阳春白雪”。那么从落地的角度来看,这项技术如何才能带来价值,真正落地到行业中去呢?

举一个真实的应用场景:你在某视频网站上观看某个演员的电影,那么网站会显示这个演员的基本信息以及他所主演的其他电影,同时也会显示这部电影中的其他参演者以及他们的个人信息和参演电影,你还可以通过这些演员查到与相关的人员的信息,如朋友、家人等人的信息,形成一个由知识点相互连接而成的语义网络,而这就是一个简单的知识图谱(知识图谱是 AI 的内容之一)。

AI 产品落地的流程如下图所示。



AI 产品落地的流程图

上述流程图针对包含算法、框架、深度学习技术、硬件、操作系统等模块的 AI 产品。

4.3.2 裸眼 3D 智能试衣镜

本节以裸眼 3D 智能试衣镜的实战为例讲解打造 AI 产品的核心过程。首先,我们了解一下裸眼 3D 技术。3D 是 Three-Dimensional 的缩写,意思是三维图形。在计算

机中显示 3D 图形,是在平面里显示 3D 图形,其不像在现实世界(真实的三维空间)那样具备真实的空间距离。在计算机中显示 3D 图形,是让人眼“感觉”是真的图像。本例以柱状透镜技术(也被称为微柱透 3D 技术)为基础进行讲解。使液晶显示屏的像平面位于透镜的焦平面上,每个柱状透镜下面的图像的像素就会被分成几个子像素,透镜就能向不同的方向投影每个子像素。此时,双眼从不同的角度观看液晶显示屏,就可以看到不同的子像素。柱状透镜技术并不会像光屏障技术那样影响屏幕亮度,所以其比后者的显示效果要好。

介绍完裸眼 3D,接下来介绍智能试衣镜。智能试衣镜产品的定位在于替换传统的平面试衣镜,智能上传新购入的衣服,智能关联已经购买过的旧衣服,然后根据天气、场景、节日等为主人推荐服装搭配。

产品定位:代替传统试衣镜的高科技试衣镜。

技术选型: 软件部分:3D 人体建模(软件+模型)、3D 服装建模(软件+模型)服装贴合人体实现衣服上身效果软件、系统软件、体型生成算法、服装生成算法。 硬件部分:RFID 芯片、裸眼 3D 芯片板卡模组、ITO 裸眼 3D 大尺寸屏幕模组、深度体感摄像头模组。

AI 赋能点:人脸识别。

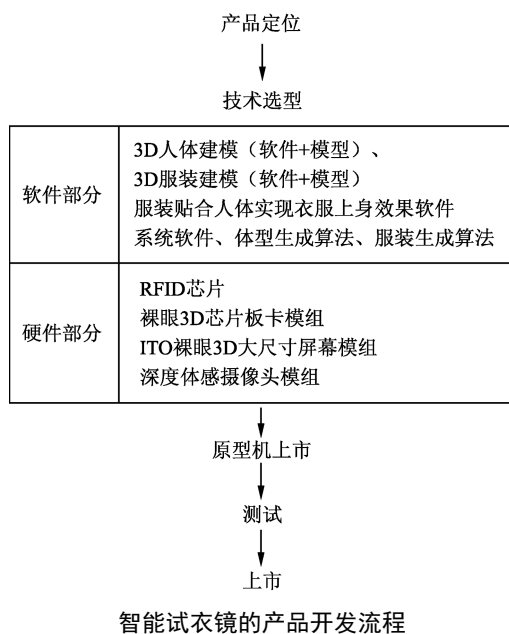
主要的产品开发流程如右图所示。

4.3.3 AI 衣柜

下面笔者将讲解打造一款新的 AI 衣柜的流程。

(1) 项目背景。

“明天早上穿什么衣服”是现在很多都市女性晚上睡觉前要思考的“人生大事”。爱美之心,人皆有之。很多女性每天在穿衣打扮方面花费再多的时间、金钱都觉得不



够。但是在生活节奏加快的时代，“如何能让自己更时尚，同时节省时间”可能是每位追求时尚的女性面临的重大难题，这该如何解决？随着 AI 的发展，借助计算机视觉、机器学习、语音识别等前沿科技，这一问题将被解决。

(2) 目标用户。

性别：女。

年龄：20~50 岁。

分布：一线城市、二线城市。

收入群体：中高层。

用户需求如下。

| 序号 | 需求 | 说明 |
|----|--------|---|
| 1 | 身体测量 | 测量用户身高、体重等参数 |
| 2 | 衣服搭配 | 识别用户的所有衣服，生成衣服搭配组合 用户可修改衣服搭配方式 |
| 3 | 穿衣计划 | 用户可创建一周的出行计划，重复使用该计划 根据天气、计划、活动场所生成一周穿衣计划 |
| 4 | 购买衣服建议 | 对用户身体、行为、天气、已有衣服等数据进行分析，生成健康、时尚的穿衣建议 根据大数据分析，融合电商平台推荐商品，让每次购买都是最好的选择 |
| 5 | 穿衣打扮社区 | 可“炫耀”自己的穿衣打扮，交流穿衣心得 |

(3) 产品理念。

私人定制的女性衣服搭配专家。

(4) 产品亮点。

推荐个性化的衣服搭配，让女性更时尚。

根据用户的出行计划，制订每周的穿衣计划。

通过大数据分析，生成个人穿衣建议，让自己更美、更健康。

为用户推荐最合适的衣服，让每次购买都是最好的选择。

和爱美的朋友一起交流，让自己变得更美。

(5) 产品形态。

AI 衣柜硬件。

AI 衣柜 App。

(6) 产品功能需求。

| AI 衣柜硬件的功能需求 | | |
|--------------|------|--|
| 序号 | 需求 | 说明 |
| 1 | 数据测量 | 测量用户的身高、体重、体脂率等数据 更新每天的天气情况 |
| 2 | 衣服识别 | 识别用户的所有衣服，建立个人衣服数据库 对所有衣服进行统计、分析、标签化 |
| 3 | 衣服组合 | 根据用户所有的衣服，生成衣服组合搭配 |
| 4 | 穿衣计划 | 由用户设定一周出行计划，包括目的地、场景等信息 根据用户出行计划，自动生成每周的穿衣计划 可设置衣服是否参与搭配计划 |
| 5 | 穿衣建议 | 根据用户身体信息，计算健康体重、生成健身建议 生成搭配组合建议 根据行为计划、天气、配饰、发型等，生成穿衣建议 |
| 6 | 商品推荐 | 根据大数据分析，个性化推荐商品给用户 购买商品、查看订单 |
| 7 | 照镜子 | 查看穿衣效果 |
| 8 | 交互 | 唤醒方式：语音、触摸屏 |

| AI 衣柜硬件的设计 | | |
|------------|------|-----------------------------------|
| 序号 | 需求 | 说明 |
| 1 | 产品外观 | 单组合、双组合、三组合、四组合衣柜 |
| 2 | 工业结构 | 传统木架衣柜 电线位于衣柜背后下部 显示屏、镜子一体化 |

| AI 衣柜 App 的功能 | | |
|---------------|-------|--|
| 序号 | 需求 | 说明 |
| 1 | 注册、登录 | 注册：手机号、微信、QQ 等 登录：密码、微信、QQ 等 |
| 2 | 个人信息 | 头像、姓名、收货地址、手机号、出生年月、性别 身体数据：身高、体重、体质等 |
| 3 | 绑定硬件 | 连接 Wi-Fi 硬件升级 设置硬件名称 |
| 4 | 衣服库 | 扫描衣服 衣服库预览 |
| 5 | 衣服组合 | 查阅衣服组合 |

续表

| AI 衣柜 App 的功能 | | |
|---------------|------|--|
| 序号 | 需求 | 说明 |
| 6 | 穿衣计划 | 用户设定一周出行计划，如目的地、场景等信息 根据用户出行计划，自动生成每周的穿衣计划 可设置衣服是否参与搭配计划 |
| 7 | 穿衣建议 | 根据用户身体信息生成健身建议 生成每件衣服的搭配组合，推荐建议 根据行为计划、天气生成穿衣建议 |
| 8 | 商品推荐 | 根据大数据分析，个性化推荐商品给用户 购买商品、查看订单 |
| 9 | 时尚搭 | 发布：穿衣搭配的照片、文字 点赞、评论、转发 |

| 项目计划 | 说明 |
|--------|--|
| 项目 1 期 | 计划实现的 App 功能：注册、登录、个人信息、衣服库、衣服组合 |
| 项目 2 期 | 计划实现的 App 功能：穿衣计划、穿衣建议 |
| 项目 3 期 | 计划实现的 App 功能：商品推荐、时尚搭 |
| 项目 4 期 | 经过前 3 期验证商业模式，具有议价权以后，与衣柜厂商、硬件厂商合作，开发衣柜硬件，上市 |

笔者对这个案例的点评是：应多调研一些成熟的技术方案，如 3D 人体建模、3D 服装建模，找到一条 AI 技术与市场需求平衡的发展策略。

第 5 章 AI 重新定义敏捷开发

5.1 AI 时代更需要敏捷开发

5.1.1 什么是敏捷开发

敏捷开发的英文简称是 Scrum。敏捷开发以用户需求为核心，采用迭代、循序渐进的方法进行软件开发。在敏捷开发中，软件项目在构建初期被切分成多个子项目，各个子项目的成果经过测试，都具备可视、可集成和可运行使用的特征。换言之，就是把一个大项目分为多个相互联系，但也可独立运行的小项目，并分别完成。在此过程中，软件一直处于可使用状态。

敏捷开发是跟项目管理（Project Management，PM）并存的，所以一提到敏捷开发，人们就会联想到项目管理。社会上很多项目管理培训机构会把敏捷开发作为项目管理的培训重点。

1. 敏捷开发的概念

那究竟什么是敏捷开发？我们站在产品经理或者创业者的角度来对其进行拆解。

有的公司有项目经理，而有的公司并没有明确的项目经理，而是由产品经理来兼任项目经理并担当其主要职责。

敏捷开发是相对不敏捷开发来说的，而这个不敏捷在软件工程里面有个专业的术语——瀑布式开发。

所谓的瀑布式开发，其实就是经典的软件工程方法。为了定义出一套完备的过程规范，使得软件开发的运作就像是机器设备正常运转一样而总结出来的项目管理方法论。

这套方法论分为 5 个阶段：需求分析、设计、编码、测试和维护。需求分析阶段通常定义系统的需求，明确系统的目标；设计阶段通常确定系统使用什么数据库，划分系统的模块，确定各个模块的功能；编码阶段用编程语言实现设计阶段的任务；测试阶段主要测试功能是否实现，以及是否正确实现、有没有 bug；维护阶段是根据用户新的需求重新修改系统，使系统正常运行，并更加稳定。

瀑布式开发的局限性非常明显，比如对市场变化和用户需求的响应太慢，更改成本太高等，有可能出现的情况是产品一推出市场就被宣告失败。

在瞬息万变的移动互联网时代，大家已经渐渐体会到了敏捷开发带来的优势，越来越多的互联网产品保持了一周发布一次的节奏。这么快的速度，就是为了迅速响应市场与用户的需求。

2. 敏捷开发的特点

(1) 小步快跑，快速迭代。

在移动互联网时代，随着时间的变化，市场环境、用户需求、竞争对手等因素都在时时发生着改变，需求方（比如用户、市场环境、运营人员、老板或者是产品本身）会不断赋予产品新的需求来应对这种变化。为了让需求方尽快看到结果，并给予反馈，产品经理应该以“小步快跑”的态度来做产品，尽早交付新版本的产品。对于敏捷开发来说，可用的软件胜过完备的文档。传统的瀑布式开发要求使用产品需求说明书来记录详细的需求。采用敏捷开发的方法之后，有时候可以只画一个原型加点备注来告知需求，或者直接通过口头沟通来告知需求，这大大缩短了项目交付时间，从而达到了版本快速迭代、产品尽早交付的目的。

小步快跑意味着产品交付的时间间隔越短越好，也就是产品有较短的迭代周期（通常是 2~4 周）。通过这种频繁的迭代，可以让产品与用户形成良好的合作关系。用户可以及时反馈，产品则不断地完善和提高产品的用户体验。对于不能给用户或者产品带来价值的功能需求，则坚决不做。

(2) 有项目计划，但也拥抱变化。

敏捷开发不代表不做项目计划，敏捷开发更加注重计划的制订。因为敏捷开发就是为了能够及时响应用户和市场需求，所以并不会“死守”着计划不进行调整。一旦市场发生变化，即使到了开发阶段的后期，敏捷开发也欢迎改变需求，不断修正自己原先

的计划，利用变化来为产品创造竞争优势。同时，参与敏捷开发项目的团队成员也不害怕变化，因为这些变化意味着自己更了解市场需求，团队本身能够与市场、用户的需求同步。

（3）在版本开发周期内尽量不加任务。

尽管敏捷开发的目的是尽量让产品能够适应需求的变化，但也并不意味着可以毫无节制地添加和修改项目任务。事实上，可以把每个版本迭代看作一次小的瀑布式开发，敏捷开发并不是全盘否定了瀑布式开发，而是借鉴了它优秀的部分。比如，瀑布式开发对于那种需求比较确定的项目来说还是不错的，工厂里面的生产环节就可以采用瀑布式开发的项目管理方法。

每个版本都有自己的开始时间和结束时间，并在项目刚开始的时候就配置了相关资源来实现产品的需求。如果临时插入新的需求或是修改需求，多少会对项目的进度产生影响。所以，我们还是尽量在版本开始前就思考清楚，除非碰到特殊情况，否则尽量做到“在版本开发周期内不加任务”。

（4）团队配置也要敏捷。

为了实现项目的敏捷，在团队组成上也需要进行敏捷处理。一般来说，一个项目团队的人数要少于20个人，太多的话可以进行团队分割（事实上，很多大公司已经在这么做了）。如果可以，项目成员在一个办公室进行办公将会大大提高沟通效率，因为有什么问题可以直接面对面地解决。

在互联网公司里，大家应该都听说过“站立晨会”这个概念，这种会议只有适应敏捷开发的团队才能更好地完成。想象一下，如果一个场地站了几十个人，每个人说一下自己的日常工作，那一上午的时间就过去了，这样做的效率非常低，根本谈不上敏捷二字。

（5）敏捷也需要反思。

项目团队成员需要定期对前一个或者前一段时间的迭代进行反省、总结，以便调整自己的行为，提高项目的开发效率。因为很多不确定的因素都会导致项目原来的计划失败，比如项目需求的变更、人员的流动、市场的变化等。我们需要对这些变化做出不同的反应。对失败进行反思，吸取经验教训，提升自身对敏捷开发的认识，团队成员只有通过不断地总结、反思和调整，才能更好地保持团队的敏捷性。

了解了敏捷开发的基本概念和特点后,我们来看看敏捷开发的本质是什么。

事实上,敏捷开发的背后是两个非常流行的概念,一个叫作 MVP,另一个叫作精益分析。

3. MVP

MVP 是最小化可行产品(Minimum Viable Product)的简称。这个概念是 Eric Ries 在《精益创业》里提出来的,简单来说,就是指产品开发团队通过提供最小化可行产品获取用户反馈,并在这个最小化可行产品上持续、快速迭代,直到产品到达一个相对稳定的阶段。

开发 MVP 也是《精益创业》最为崇尚的方法。在产品处于探索期时,产品价值、市场切入点、用户群、用户体验平衡点、商业目标等都模糊不清,这时候就需要用低成本、快捷的 MVP 去探索以上问题,让产品找到更好的发展方向。

传统的产品设计思路是一步接一步,从车轮、车轴、动力装置到内部装饰,一个流程一个流程地做,最后得到一个完善的产品。正确的敏捷迭代应该是每一步的产品都是最小化可行的,虽然滑板车和汽车相去甚远,但滑板车也是代步工具,在验证了用户对出行工具的认可程度后,我们就可以去生产更加高级、完善的摩托车甚至轿车。

MVP 的几个原则可以概括为如下内容。

(1) 抓住产品的核心流程。

MVP 要求我们抓住最核心的产品流程,剔除多余的功能或者高级功能。比如,电商产品的核心目标是让用户在产品上下单买东西,那核心流程就应该是:浏览商品→挑选商品→下单付款→查询物流信息。围绕这个流程,可以剥离多余的功能需求(如个性化推荐、活动推荐、秒杀、分享、积分等)做一款 MVP。

当然,一款产品的核心流程具体包含哪些内容,这是需要团队和产品经理一起去商量的,在不同的业务、不同的场景下,核心流程会有所差异。

(2) 不同阶段的 MVP 的目标不同。

在产品从 0 到 1 的阶段,产品刚刚上线,这个时候 MVP 1.0 的目标应该是验证需求,看看设想的需求是否真实存在,设想的需求是高频的还是低频的,是刚需的还是非刚需的。

而在后续的迭代中，如果产品设想的需求和市场痛点相匹配，那么 MVP 2.0 则可以开始优化产品核心流程的用户体验，然后增加一些新的功能点。在下一个版本继续观察新增加的功能点是否符合用户需求，对新功能应该是做改进还是做下线处理。简单来说，在不断迭代的过程中，我们对 MVP 的关注点发生了一些变化，但最需要关注的还是产品的核心流程。

(3) 可以尝试任何产品形态。

随着移动互联网红利期的结束，一个创业公司通过开发一款 App 而成为独角兽的可能性已经越来越小了，这个时候 App 的开发成本、推广成本也已经相当高。所以很多创业公司会纠结到底是做一款 App，还是做一个微信公众号，这其实就涉及 MVP 的产品形态。MVP 可以是一个只有基本功能的 App，也可以是一个微信公众号、一个微信群，甚至是一个纸面原型、一个视频。只要能让用户直观感知到产品的价值，能激发出用户想要使用产品的欲望即可。

越来越多的产品通过微信公众号的形式进行验证，比如知乎的付费问答产品——值乎，最开始的 MVP 的产品形态主要是知乎服务号的自定义菜单或者朋友圈的分享。如果一开始就放在 App 端，等到用户发现时，可能已经错失了市场机会（同一时间，分答已经在市场上“跑马圈地”了）。

4. 精益分析

所谓精益分析，首先是你有一个想法或者灵感，然后通过 MVP 策略让产品快速上线，产品上线以后，通过数据来衡量用户的表现，如果表现好就继续优化，表现不好就下线。

在创业公司，大部分的创意都来自创始人的“灵光一闪”，大部分情况下只有将产品投放到市场上进行验证时才可以发现用户有没有需求。创业团队往往精力有限，因此要让验证的周期尽可能短，而 MVP 加上精益分析，通过数据就可以快速验证市场对于产品的反应。如果用户反馈很好，就可以继续加大投入，如果用户反馈产品有问题，也可以及时调整，避免浪费更多的精力。

5. MVP 的实践——分答的版本迭代之旅

分答是 2016 年度最引人注目的付费语音问答产品，用户可以快速找到能够给自己提供帮助的那个人，行家可以通过语音用一分钟的时间为用户答疑解惑。这款产品

由在行团队孵化，创造了新的传播与分享知识的方式。不仅是科学家，很多名人和各领域的专家也都加入了分答。产品上线仅 42 天，就获得了超过 1000 万的用户，付费用户的数量超过了 100 万，33 万名用户开通了答主页面，产生了 50 万条语音问答，交易总金额超过了 1800 万元，复购率达到 43%。

在如此傲人的成绩下，分答的产品总监朱晓华透露，分答的创意来自姬十三，从确定要做到产品正式上线，中间只用了一周的时间。

我们看看在一周的时间里，分答团队都做了哪些事情。

一晚做出框架：1 名产品经理。

一天做出原型：3 名产品经理。

一个周末打造产品主体：1 名设计师、1 名后端工程师、1 名前端工程师、1 名测试工程师、1 名运营人员。

一周业余时间进行完善。

两天完成内测。

一日引爆。

通过一周紧张的设计和开发，分答的第一个版本终于上线了。分答的第一版非常简单，只有最基本的提问和收听功能，而且还有不少的 bug，但这并不影响产品上线并验证用户需求。

在上线之后，分答团队基本上保持了每天至少发布 3 个新版本的节奏，这是怎样做到的呢？

除了“精益”“迭代”的团队认知之外，产品形态的选择是产品能够快速迭代的基础。分答第一版的产品形态不是手机客户端，而是在微信环境内使用的一个 H5 页面，并通过服务号提供通知、支付等服务。

比起直接开发手机客户端，基于微信进行开发有几个非常明显的好处。

用户使用成本低：不需要下载客户端，打开网页就可以使用。

开发成本低：iOS、Android 客户端开发起来费时、费力，而且还有各种适配问题。

迭代成本低：不需要先去应用市场提交新版本，再让用户下载安装包进行更

新，直接在服务器端进行更新就完成了迭代，可以做到用户完全无感迭代。

使用 H5 技术进行开发，极大地降低了开发和迭代成本。在微信环境里，入口虽然深，但通知和交易闭环都是完整的。

上线并不意味着战斗结束，如何精准、高效地记录和分析产品数据并迭代产品成为下一个阶段的工作重点。分答团队通过使用第三方数据分析平台，跟踪不同模块的使用频次，监控搜索效率，查看付费收听流程的转化率，通过观察产品的数据表现并收集用户反馈来不断迭代产品。

就这样，分答在短短 6 周内通过敏捷开发方法创造了知识分享领域的“数据增长神话”。

其实，敏捷是一种思想，敏捷开发就是敏捷思想的一个实际应用，用思想指导实际工作，思想就具化成一些行之有效的手段和原则，比如极限编程等就是敏捷开发的应用框架。

6. 什么是极限编程

为什么叫极限编程？极限编程就是强调把整个工作流程都做到极致，做到最好，包括极限的工作环境、极限的需求、极限的设计、极限的编程、极限的设计。极限编程要求交付用户的版本一定是能够使用的，重复冗余的代码一定要被重构。而极限编程所不提倡的其他内容（如开发前期的整体设计等）则一概忽略。

下面来简单介绍一下用户故事。用户故事是描述对用户有价值的功能，即敏捷开发里强调的管理需求的方式。

（1）极限编程的角色。

产品经理：负责编写用户故事，对故事优先级进行排序，验证用户故事。

项目经理：负责管理项目的进度和流程。

程序员：负责开发任务，并对代码进行单元测试。

（2）极限编程的工作方式。

产品经理从用户那里获取需求，编写用户故事，并确定需求优先级。

开发人员和产品经理共同对需求清单进行评估，确定迭代周期，并按照优先级确定在该迭代周期内应该完成的需求。

进入开发阶段,开发团队每日更新进度,产品经理可以调整合理的需求,开发团队要做到快速响应。

迭代周期结束,产品经理验收迭代结果,团队总结经验,并发布新版本。

(3) 极限编程的要点。

重视沟通:沟通方式不拘泥于形式,可以是谈话、会议或者直接画图等形式。

简单:以最简单的方式满足需求。

重视反馈:无论是用户的反馈还是团队内部人员的反馈,都要一视同仁。

勇气:有勇气不断优化产品功能。

5.1.2 产品总监带领团队执行敏捷开发的方法、步骤

一般是产品总监带领团队进行产品开发,实现产品的商业价值,本部分通过对产品开发前、开发中、开发后的关键节点进行梳理和解析来阐明产品总监如何带领团队实现敏捷开发。

1. 开发前

(1) 阐述、调整需求。

产品总监要在开发之前考虑清楚各个需求的逻辑、页面交互形式,不然会引发很多讨论,延长开会时间,影响效果。笔者在拙作《产品经理进化论:AI+时代产品经理的思维方法》中讲过如何理清需求的内在逻辑。

优先级一定要明确,这是实现页面的基础。如果程序员对某个需求提出质疑,最好不要任由发散讨论,能立刻调整的,要立刻调整;不能立刻调整的,要先保留意见,在真正开发前再次确认。

如果产品总监在阐述需求的同时,开发人员已经开始沟通如何实现的问题,这时,应建议开发人员先停下来听完所有内容,在后面的需求分解中再进行详细讨论,不然在需求源头错过问题,容易对需求产生认知偏差。

(2) 将需求分解为任务。

小团队在没有项目经理的情况下,建议项目负责人由经验比较丰富的由技术转岗的产品总监担任,这样比较容易把控项目进度,同时对需求的分解以及后期的进度追踪也会有很大的帮助。

要强化需求追踪人追踪需求的意识，不要流于形式（有这样一个角色，但是并没有发挥作用）。需求追踪人的存在可以分担项目负责人的责任，同时让开发人员之间的关联性更强，因为一般某个需求会同时涉及前端、后端、移动端、API等模块的开发人员。

开发人员要把需求拆得“细”一些，这样在拆的过程中更容易吃透需求、理清逻辑，发现逻辑上的问题。针对需求，相关人员要进行技术讨论，有疑问的地方要及时和产品人员沟通，越早发现问题越好。最后，要把每个任务对应的人以及需要的完成时间都记录下来。

产品总监要在讨论过程中时刻参与其中，尽量保证大家都能正确理解需求。

（3）确认当期目标。

开发前要完整地“过”一遍每个需求的分解情况，看是否有遗漏或者误差，有疑问的话马上提。必须按照需求的优先级来确认目标，明确哪些工作一定要做完。

2. 开发中

（1）中期会议。

一定要在中期进行进度回顾和确认，如果进度偏差很大，要适当调整目标，以免最终结果跟预期相差太大。

在中期的时候，可以检查一下人力资源的分配情况，如果有相对空闲的人员，那么可以再安排一些任务。

（2）测试。

在开发过程中，建议开发完一个功能就开始对该功能进行测试，以免测试任务堆积起来。

提测可以在项目结束前的2至3天开始，这样可以保证有足够的时间进行修正。

如果功能和业务强关联，可以让业务人员在提测阶段使用测试版本，一起发现问题。

开发人员在开发新功能的过程中，可能会遇到之前已经发现的bug或者之前没发现的bug，这时候应该通知测试人员跟进，而不是自己直接修改。因为你的修改并不一定真正地解决了问题，一定要进行回测，同时留下记录，以便给以后的追溯带来帮助。

(3) 其他。

团队每天都要开“站立晨会”，以描述每个人的进度情况。注意：描述一定要明确，如果描述不明确，其实等于没有开“站立晨会”。

在适应新开发流程的初期，项目负责人最好能够每天或者每两天确认一下组员们的真实开发情况，包括进度和需求理解上的情况。

刚开始尝试敏捷开发时，容易出现预估（目标、时间）不准确的情况，导致2周迭代一个版本比较困难，发版可能会经常延迟。面对这种情况，建议先砍掉部分需求，以保证2周迭代一个版本，让团队先形成一个周期惯性，再去优化周期内的效率问题（也就是之后再提高要求）。

(4) 敏捷开发的要求。

目标、质量验收标准不能改变。

完成目标的要求不能过低。

3. 开发后

开发结束后，组员们一定要复盘一下这期开发中存在哪些问题，如何解决或者减弱这些问题的影响，不要流于形式。毕竟对团队来说，这是新的开发流程，不适应或者做得不好很正常，不要怕出错，大家在错误中磨合、改善、成长才是最重要的。

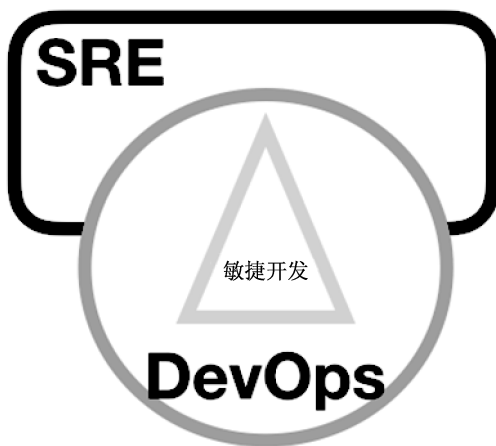
此外，每期也要对上期提出来的问题进行复盘，看相应问题在这期是否有改善。只记录不复盘相当于什么都没干。

4. 小结

PMP (Project Management Professional) 认证指项目管理资格认证。它是由美国项目管理协会 (Project Management Institute, PMI) 发起的，用来评估项目管理人员知识技能的资格认证考试，其目的在于给项目管理人员提供统一的行业标准。作为项目管理资格认证考试，PMP 已在国际上树立了其权威性。

PMP 的参与角色有：产品设计人员、开发人员、测试人员、项目负责人。

敏捷开发相关内容的关系如下图所示。



敏捷开发相关内容的关系

SRE 是指 Site Reliability Engineer (网站可靠性工程师), 它是软件工程师和系统管理员的结合。一名 SRE 需要掌握很多知识, 包括算法、数据结构、编程、网络编程、分布式系统、可扩展架构、故障排除等。

DevOps 是英文 Development 和 Operations 的组合, 是一组过程、方法与系统的统称, 用于促进开发 (应用程序/软件工程)、技术运营和质量保障部门之间的沟通、协作与整合。它的出现是由于软件行业日益清晰地认识到: 为了按时交付软件产品和服务, 开发和运营工作必须紧密合作。

敏捷开发是两者重叠的部分。了解 SRE 和 DevOps 的关系对了解敏捷开发的原理很有帮助。

5.1.3 认识 AI 时代的敏捷开发

本部分先讲为什么 AI 时代更需要敏捷开发, 然后再讲打造机器学习 MVP 的方法。

1. 为什么 AI 时代更需要敏捷开发

产品让用户“爽”的次数越密集, 用户越迷恋, 这是共识。

纯软件、纯硬件是很容易实现敏捷开发的, 但是 AI 产品是软硬一体的整体系统, 软件部分可以让工程师一晚上改很多次, 但是如果让擅长软件开发的团队开发 AI 产品, 则需要协调外部的资源, 甚至是一个工厂。所以, 如果想让产品快速被开发出来,

更加需要敏捷开发。

笔者总结的 AI 产品的公式如下。

敏捷开发方法论+软硬结合+AI=AI 产品

在 AI 时代,“不想做好测试的开发不是好的产品经理”,如果一个产品的测试工作做得不好,开发工作做得必然也不好。因为如果 AI 产品在生产过程中的测试工作做不好,产品上线后必然会积累很多的产品“黑洞”,到时候不知道哪个配件、哪个驱动、哪个算法、哪个程序有什么样的问题,解决问题的难度会大大提升。

综上所述, AI 时代更需要敏捷开发。

2. 如何让机器学习 MVP 结合敏捷开发

下面我们看一个机器学习 MVP 结合敏捷开发的例子。

MVP 是开发、设计团队用最小的成本,最大限度地验证产品的可行性。产品的可行性是指这个需求是否真实存在,这个产品满足需求的方式是不是对。

做机器学习产品的投入是长期的、持续的,带来的收入和回报也是巨大的,在开始之前,我们一定希望以较低的成本知道引入机器学习是否可以影响我们所面对的业务,机器学习产生价值的潜力有多大。

把一个业务真正用机器学习做之前,可以先通过两步做一个机器学习 MVP。

第一步:我们要选择正确的业务问题,并不是所有的业务问题都可以套在机器学习的框架里,有些业务问题适合用机器学习解决,有些业务问题不适合用机器学习解决。在项目管理中,用差的方法解决正确的问题,一定优于用好的方法解决错误的问题。

第二步:当找到一个机器学习可以解决的业务问题后,如何用最少的时间、最小的人力成本去证明机器学习可以解决它,带来令人满意的投入产出比呢?答案是做机器学习 MVP。

(1) 从分类器开始。

首先,我们看看机器学习擅长解决什么问题。这里笔者举一个周志华老师的《机器学习》一书中的例子,它很简单,却很深刻。这个例子是说要怎样判断一个西瓜是好的还是不好的。

这个问题的业务场景是什么呢，面对一个西瓜，我怎样在不交易、不打开的情况下，就知道它是好的还是不好的。如果我是消费者，我就可以用同样的价钱买到更好的西瓜；而如果我是瓜商，有了一套标准之后，我就可以更好地管理我的货品。

回到这个问题，一个西瓜是好的还是不好的，这是典型的机器学习的二分类问题。我们要知道，判断一个西瓜好不好有哪些可以用的数据。我们不能使用买卖西瓜之后的数据，比如买了西瓜之后，我打开就知道它好不好了，但这个数据没有价值。

我必须在破坏西瓜的前提下进行判断，这时候能用到的数据可能是西瓜的产地、纹路、重量、比重，敲击西瓜产生的声音、西瓜皮的质感等，这些是不打开西瓜就可以知道的数据。

刚刚我们的目标已经讲得很清楚了，判断西瓜是好的还是不好的。我们设定好的是 1，不好是 0，甚至还可以定义一个评分，评分可以是 0~1 之间的一个数。总体而言，可以设定一个机器学习的目标，我们称之为 Label。

(2) 构建真实世界的模型。

这看起来是一个很简单的场景，一旦我们具备了数据，就可以尝试建立机器学习模型。然而在现实中，当我们想用机器学习来解决实际问题时，也会这么简单吗？

第一，西瓜好不好是怎样定义的？是大，还是甜？是皮不厚，还是瓢脆？如果建立这个模型是为了西瓜的售卖，这些可能都是评价的因素，模型学习的样本也都需要基于这个标准来建立。如果我们仅仅是基于西瓜大不大来定义样本，而实际的应用场景是综合判断西瓜好不好，那么可能会得不到想要的结果。

第二，西瓜好不好是以什么为标准的？是用科学方法推测的、用仪器测量的？还是通过专家评测的？如果是后者，评测者是同一个人么？如果是不同的人，大家对好西瓜的判断标准一样吗？在现实情况中，判断标准很可能是不一样的，那就要想办法消除 Label 的偏差。

第三，在互联网的场景下，往往需要满足所有人的个性化的需求，有些人喜欢甜的西瓜，有些人喜欢脆的西瓜，那将问题定义为分辨西瓜的好坏是否合适？因为每个人对好西瓜的定义不一样，这时，这个问题可能就转化成了推荐一个西瓜给一个用户，他会不会喜欢。

第四，真实的应用环境是怎样的。假设我们需要一个在线、实时的西瓜分类器，拿到西瓜的那一刻马上能判断它好不好，那是不是某些当时不能马上拿到的特征就不能用了？如果好西瓜的判断标准在不断发生变化，或者西瓜本身的特性也在不断变化，那么，模型还需要能够跟得上这些变化，基于新的数据和反馈做自我更新、迭代。

可见，即便是简单的问题，也需要思考一下业务的方方面面，理清哪些因素、个性化要素、基础设施是需要考虑的。

（3）还原业务问题本来的面貌。

我们从判断西瓜的好坏回到具体业务，判断一个业务能不能用机器学习进行优化，要看三个要素。

第一，这个业务的目标值是什么。它不一定是唯一的，但一定有主次。这个目标是否可以量化、收集反馈、客观观测。什么叫客观观测，我说甜而你说不甜，这个事情无法判断，这就可能不客观，此时，含糖率这样一个客观的指标就是客观观测。

第二，样本应该如何构造。样本不应该违反因果关系，在 $y=f(x)$ 中， x 一定是在业务场景中所能知道的信息。在判断西瓜好坏的问题中，打开西瓜之前我们能知道的信息才可以作为 x 。同时，样本应该符合业务场景的真实情况，假设我们的业务场景是摸黑挑西瓜，我们看不见西瓜长什么样，只能敲，那西瓜的颜色就不能作为特征。

第三，样本的每一行代表什么意思。样本的每一行应该代表每次测量西瓜得到的数据，然后才是选择哪些数据作为 x ，这些我们已经讲得很清楚了。

3. 真实的机器学习的业务问题

当西瓜的问题说完后，我们再来看看真实的业务问题是怎样的。

（1）点击率预估。

一个推荐系统的目标是什么？它的终极目标一定是用户体验最优，但这个目标很“虚”，我们要把它量化，变成一系列可以测量的数据，比如说点击率、观看时长、购买率、好评比例等，这些就是 y 。

然后我们看哪些 x 是做出推荐排序的一瞬间（当客户请求时）系统知道的事情。系统知道客户的属性、特征，系统知道内容的特征、上下文的特征，但系统不知道这个内容最终是否会被展现和点击。系统可以知道内容在这个瞬间之前被点击了多少次，但一定不知道这个瞬间之后将被点击多少次，因为这样就穿越了。

有了 y 和 x ，就可以构造样本了。假设样本是给用户展现 10 条推荐的内容，这个样本的反馈可能是点击率和观看率，那么，每一次的展现就是一个样本。

(2) 简历匹配。

再举一个例子——简历匹配。简历匹配是系统根据某个人的简历向企业推荐这个人。

简历匹配有两个选择，第一个选择是企业通过面试、简历进行选择，即判断这个人是否适合企业。第二个选择是应聘者的选择，即要不要去推荐的企业面试？即便拿到了推荐企业的录用通知，要不要加入推荐的企业？

所以，这就变成了多点、双向的问题，在这种情况下，需要对问题进行拆解。我们可以不直接做个人被企业招聘的事情，而是分开来做，比如企业会不会邀请这个人去面试，以及这个人会不会接受企业的面试邀请，这样就能把问题做得更好。

总结一下，做机器学习 MVP 最优先的工作不是建立机器学习的模型，而是认真思考这件事情的业务场景到底是什么样的，要解决真正的问题。

4. 小结

一个机器学习能解决的业务问题，有四个关键点。

第一，能否转化成分类、回归问题。

第二，是否容易获取客观、无偏差的数据。

第三，问题的预测目标、因果关系是什么。因果关系越简单越好，如果是多因多果，或者说描述“因”的相关信息不方便获取，那是否可以把问题拆分成多个模型。特征往往是“因”的数据，或者是一些不是直接原因的数据，只要它不破坏这个因果关系即可。

第四，这个问题是否是一个真的业务需求。一个真的业务需求是指：在我们用机器学习做出预测后，业务是否会受这个预测结果的影响，这个影响点是否足够清晰、有效。因为业务人员会用对业务的影响来评估项目的效果，如果预测的结果并没有有效地影响业务，即使这个模型再好，也不会发挥作用。如果是推荐系统，预测了新的点击率后，可以按照点击率倒排来影响业务结果。但如果是游戏呢，如果预测这个人明天有 30% 的概率付费，系统该如何影响他？系统能不能影响他？

所以，你一定要思考，预测的结果在业务中会怎样被使用，这个使用会不会使业

务指标产生提升。如果你发现提升是很难的，那这本身就是个伪需求。然后你还需要思考，没有使用机器学习的业务，它使用了什么方法和数据，这些方法和数据有什么缺陷，哪些是通过机器学习可以改进的。

当以上的问题都有清晰的回答后，你就可以提出一个好的问题了，这时候你就成功 80%了，剩下的问题都相对简单了。

5.1.4 机器学习 MVP 的成本

打造机器学习 MVP 的第二步是在人力、金钱投入可控的前提下，构建一个有效的机器学习模型。

什么是可控的呢？建议 1~3 人/月的投入，因为更多投入会导致风险太高。我们会期望获得什么提升？不同业务得到的提升效果不一样，有些业务（比如说广告）的指标提升 1% 的收入就能增加几百万元的收入，而有些业务的指标可能要提升好几倍才有商业价值。

在机器学习成本分配中，如果最大的成本在机器学习本身，即调参、定义特征、模型评估、模型上线这些工作耗费了大量的时间，而问题的定义、数据的采集所用的时间非常少，笔者认为这是有问题的。一个机器学习项目，无论是通过合作的方式还是通过使用第三方平台的方式，都应该把更多的资源用在采集好的数据、定义好的问题上，这甚至要占用超过一半的时间。而其余的时间，才是真正做机器学习模型的时间。

1. 降低数据的成本

怎样降低数据的成本呢？这里笔者给大家一些思路。

第一，除非必要，否则只使用采集好的数据。因为数据采集是一件有成本的事情，一个公司的体系越复杂，采集数据的成本就越高，所以，只有某类数据采集起来很轻松，或者已经采集好了，你才应该去考虑使用它。

第二，如果你要开发新的数据源，首先要考虑的是成本。开发新的数据源是有风险的，机器学习最怕的是，说不清楚问题是算法导致的，还是数据导致的，抑或是问题定义导致的。所以，让可能出问题的环节越少越好。算法一般是不太容易出问题的，除非用错。而数据是很容易出问题的，所以我们尽量用简单、可靠、成熟的数据。

第三，在建模的过程中要尽量使用成熟的工具。在数据处理、特征计算和算法训

练的过程中，大量的工作是可标准化的，甚至是可以算法自动优化的，大量的问题其实是在产品引导中避免的。

总结一下，这一步总的思想是：能不制造新的风险点就尽量不要制造新的风险点，能降低不确定性就尽量降低不确定性。

2. 如何检验机器学习 MVP

经过前面介绍的步骤，我们已经有了机器学习 MVP，机器学习对业务的影响已经初见成效，如果业务有明显提升，那么祝贺你，你已经找到了新的价值增长点，优化后一定还会有更大的提升潜力；如果效果不明显，这里再给大家介绍如何检查机器学习 MVP。

第一步，检查问题的方向是不是对的，模型的效果是否符合预期，模型的优化目标是否有明显的变化，比如优化的目标是西瓜好不好，优化之后买到的好西瓜是不是变多了。如果不是，那就说明这个问题没有被解决。问题没有被解决的原因是什么？是指定了错误的目标，把模型用在了错误的环境，还是数据有问题。其实简单来说，要么是目标有错，要么是模型用错，要么是数据有问题，应该基于这三点来检查。

在真实的业务中，有时解决了一个问题会带来新的问题。比如，推荐的新闻的点击数增加了，是不是由于推荐的新闻变得更加娱乐化了，是不是新闻的点击变得更集中了，这可能并不是我们希望的，需要继续想办法来优化。

第二步，检查数据。哪些数据起了关键的作用，哪些数据是经验上认为会发挥作用，但实际上没有发挥作用的。重新检查这些数据，看是不是数据质量存在问题，导致模型没有发挥应该发挥的作用。还可以看一下下一步可以引入哪些新的数据。数据最好一批一批引入。

第三步，当检查完上面的事情以后，要制订下一步的方案，此时往往会有更新、更多的数据，因此需要调整目标，有可能是目标错了要改正，也可能是原来的一个目标不够了，需要加入几个新的目标，以使模型变得更平衡，还有就是要看模型在工程上能不能优化性能等。

下面笔者以问答的形式总结了在机器学习项目中执行敏捷开发的关键点。

Q1: 数据是否容易获取且客观无偏差？

笔者之前负责过一个新闻推荐系统，每到半夜点击率暴涨，刚开始笔者很纳闷，

最后知道了这是因为半夜有竞争对手的爬虫来“点击”新闻，我们知道以后就开始对这部分数据进行处理。

在大数据环境下，当数据量大到一定程度的时候，偏差可能会因为大量的样本以及超高的维度而不再重要。做广告系统的人可能知道，点击率是会被位置影响的，并有偏差的。把位置作为一个特征，并且利用一些特征工程的手法，可以把这个偏差吸收掉，得到一个无偏差的模型。所以有偏差不要紧，想明白原因是什么，大多数情况下总有方法解决它。

Q2: 针对数据不平衡问题，比较好的解决方法是什么？

对于机器学习问题，样本的数据分布跟现实的数据分布保持一致是最好的，如果极度不平衡，例如正样本极少，影响了建模，可以尝试以下策略：① 过采样，就是把 Label 中占比少的样本通过随机的方式多采样一些到训练样本中。② 欠采样，就是把 Label 中占比多的样本通过随机的方式采样一部分，这样可以平衡样本，同时还能加快训练速度，在一些较大的数据集上可以使用这种方法进行初步的探索。③ “通过多份欠采样训练样本，先分别训练模型再做集成学习”的方式能够充分利用全部数据信息，同时避免过于倾斜的样本对模型的影响，是一种比较好的方法。

Q3: 开发新的数据，首先要考虑成本问题，除了成本，还有哪些东西需要考虑呢？

答案是时间和风险。项目管理就是消除项目中的不确定性。互联网思维讲究的是快，如果速度太慢，会失去市场的机会。所以我们一再强调在可控的范围内，尽快做出一个有效果的机器学习 MVP。

同时，开发机器学习 MVP 的常见问题是对 y 的选择不合理，数据即便不够多，往往也能够验证。如果花费大量时间开发新的数据源，这对 MVP 是不合理的。

Q4: 在电商项目中，评价供应商好坏有个传统方法，就是根据经验设定几项指标并进行评分，根据评分来判断供应商的质量并按分数来分类管理。请问，这种问题用机器学习方法来解决与用传统方法来解决的根本区别是什么，价值是什么？

由指标生成评分的公式也是一种模型，但是是一种维度比较低的模型。因此，在数据量比较大的情况下，这种低维度的模型无法充分利用所有的信息。我们所说的机器学习可以通过算法和计算能力，使用大数据生成一个高维度模型，从而能够在指定的业务目标上更准确地判断“好坏”，“大数据 + 高维度模型”带来的准确率的提升

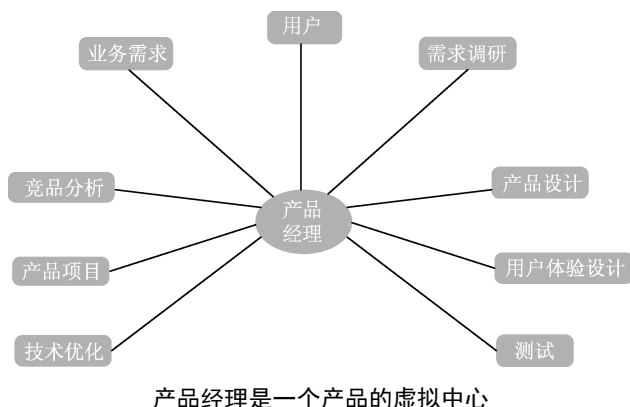
就是最终可以获得的价值。

Q5: 机器学习项目如何排期? 团队之间需要如何协作? 比如, 针对一个推荐模块的优化, 不同成员之间该怎样分工?

这件事情取决于产品总监、产品经理、创业团队的人员组成, 以及机器学习平台的建设进度, 是否已经有机器学习基础平台 (包括高维度机器学习模型和在线实时预估系统)。

5.2 创新的敏捷开发管理

在 AI 时代, 敏捷开发管理要求产品经理是一个产品的虚拟中心, 如下图所示。



敏捷开发管理要求产品经理在 AI 时代寻找前沿 AI 技术, 找到使 AI 技术与产品稳定结合的点, 形成有创意的 AI 产品, 且敏捷地获取市场反馈。

例如: 硬件设计、开模、加工、稳定性测试、匹配软件相互结合可以形成一个令用户兴奋的新产品。

AI 是产品的生产力, 敏捷开发管理是使生产力发生效力的催化剂。

在传统时代, 产品根据产品经理的活动而进化。而在 AI 时代, 产品自身根据数据而进化。因此, 没有明文禁止的, 都是可以敏捷的。

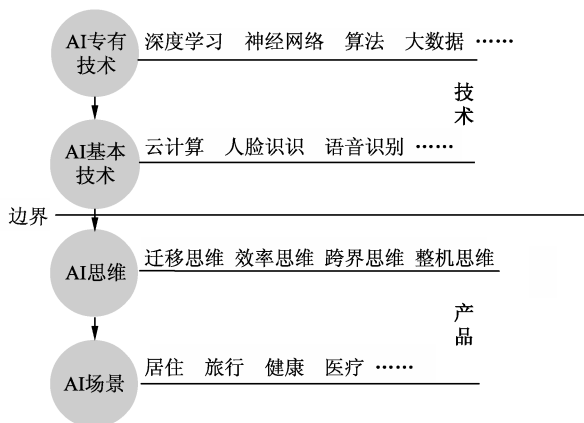
第6章 AI思维

6.1 AI产品经理需要AI思维

6.1.1 AI产品经理需要AI思维矩阵

AI产品经理最优先的工作是构建一套自己的AI思维矩阵，这是一个思维矩阵，不是技术架构也不是产品架构。

构建AI思维矩阵的第一步是忘记技术，把自己当成技术小白来看AI产品，以新手的心态来看AI技术，然后聚焦产品价值本身进行思考。笔者把这种思考称为AI思维，AI思维的矩阵如下图所示。



AI思维矩阵

AI产品经理偏重做管理和做方向性的决策，但如果懂技术会更好。

例如：在做智能客服产品时，产品经理澄清了需求，写出了产品需求文档。而技

术人员针对产品需求文档提出了种种质疑，比如某个数据从哪个系统获取，某个逻辑应该用系统的哪个字段来确定（当知识库里的 A 答案有歧义的时候，具体每一个回答的枚举、逻辑交互等）。产品人员和技术人员在这类问题上的思维是不一样的，技术人员主要考虑实现，而产品人员则考虑用户体验，这是两种思维，这时需要的是 AI 产品敏捷开发的新思维。只有在 AI 产品开发过程中加入敏捷开发才能解决上述问题。

6.1.2 AI 产品经理如何构建 AI 思维

互联网思维催生了小米、今日头条等一大波公司，随着移动互联网红利的消失，我们需要 AI 思维。

“AI 思维”是随着 AI 行业和 AI 企业的发展而诞生的。2017 年，百度创始人兼首席执行官李彦宏表示，以 AI 为核心的“新一代技术革命”正在兴起，成为人类社会又一次全新变革与发展的开端，创业者必须将自己的思维方式调转到“AI 频道”，才能拥抱时代的变革。

几年前，人们还在大谈特谈“互联网思维”，而在李彦宏看来，互联网只是“前菜”，AI 才是“主菜”。他认为，在互联网出现之前，人与人就可以交流，人与物却无法沟通；互联网时代到来后，人与人的沟通效率有所提升，但人与物的沟通还是非常少；在 AI 时代，人与万物交流的问题将得到解决，“唤醒万物”成为可能。

对 AI 持同样积极观点的，还有著名 AI 专家李飞飞。她认为，AI 是第四次工业革命的重要推动力，从医疗健康、交通出行、金融服务、生产制造，到能源、化工、农业……所有垂直产业都将因 AI 技术的发展而受益。

在 2017 年的中国国际大数据产业博览会上，AI 相关成果已成为各大企业竞相展示的当家产品：科大讯飞的智能陪护机器人“晓曼”，基于语音识别、大数据以及深度学习等技术，可以与独居老人聊天、陪儿童玩耍；蚂蚁金服推出的“扫脸付”，只需打开支付宝的“扫一扫”并进行授权操作即可进行支付……各种“黑科技”的涌现显示出 AI 赋能行业的力量。

1. AI 思维之软硬结合思考问题

在 AI 逐渐走上前台的过程中，如何用 AI 思维思考和解决问题？

AI 的特点是软件、硬件相结合,因此,要关注软件和硬件的结合处能够产生哪些创新,要从技术和体验的角度进行思考。

放眼全球, AI 已成为众多互联网企业“转型升级”的共同路径。从 2017 年开始,谷歌公司将未来的发展战略调整为“AI 先行”,在 2018 年的谷歌开发者大会上,谷歌公司展示了其在 AI 方面的诸多进展:用手机对准花朵,屏幕上就会显示花朵的名字;智能音箱可以免费拨打电话……

在中国,各大互联网企业也纷纷立足 AI 来谋划发展。百度早在 2013 年就建立了专注于深度学习的研究院,把 AI 提到核心技术的地位;小米创始人雷军也提出,要在国家层面进行 AI 发展的顶层设计,加强 AI 人才的培养与引进。

尽管 AI 看上去很美,但现状离要实现的目标还有距离。纽约大学心理学教授加里·马库斯认为,虽然人们现在已经可以利用 AI 来做一些很棒的工作,但是还没有办法构建可以治愈癌症、理解人脑工作的机器,还有很多认知层面的元素,如常识、推理、分析等,需要 AI 去学习。

2. AI 思维之大数据驱动发展

掌握 AI 思维的一项重要内容是要理解数据与算法的关系。在 AI 时代,数据的价值“秒杀”算法,“干净”的数据越多, AI 的学习能力就越强。但算法其实也在推动着社会的进步。

此前,数据量小、数据流通不畅一直是制约 AI 发展的关键问题。随着互联网技术的发展,海量大数据资源为 AI 的进化提供了条件。360 人工智能研究院副院长韩玉刚认为,数据量的大小决定了 AI 的“奔跑速度”。在大数据时代,人人都是 AI“数据燃料”的供应者。生活中的信息化设备已经无处不在,人的一切行为几乎都会产生数据,并流入互联网,汇聚成海量的数据信息。运营商、电商巨头、政府部门所掌握的海量大数据,绝大部分都是由单个人产生的数据汇聚而成的。

近年来,中国企业在大数据方面的发力为 AI 的发展提供了充足的资源。李飞飞表示,数据是推动 AI 发展的重要力量。中国拥有大量人口,产生了许多可供 AI 学者研究、应用的数据,催生了大量创新产品。

构建 AI 思维的最终目的与作用何在?笔者认为,如果用 AI 思维做互联网产品,能够实现“降维攻击”,比如与打字输入相比,语音输入就是降维攻击。这种思维还

可以延伸到家居、出行等各个行业。

最后，AI 产品经理构建 AI 思维时，一方面要注重实践，另外一方面要注重用 AI 赋能各行各业。

6.2 AI 思维的拓展

6.2.1 AI 思维的作用和重要性

AI 思维是指导打磨 AI 产品的指南针，AI 思维使你做出的产品具有 AI 属性，并采用创新的方法。

AI 产品是新产品，需要新的“DNA”。例如：梅西百货、西尔斯百货即使采用了 AI 技术，其依然不是 AI 电商平台，全新的组织结构和思维方式才使得亚马逊成为 AI 时代的电商巨头，亚马逊对 AI 技术的应用不仅体现在对 AI 算法和模型的应用，更体现在应用了 AI 思维。所以，对 AI 产品经理来说，AI 思维的重要性一定高于 AI 技术本身。

AI 是新时代的生产力，但是 AI 本身不是产品，AI 与产品结合（平台+数据+应用）才能产生价值，AI 思维可以改造流程中的任何一个节点，带来生产力的提升。

笔者一直强调“互联网已经成为过去时”，认识到这一点非常重要，因为认识到这一点代表着思维维度的提升，AI 需要时间进行数据积累，进而才能产生产业机会。

6.2.2 AI 生态思维

AI 是个生态。我们应该认识到 AI 不是单独的软件模块，也不是单独的硬件模块。AI 是应用在生活、工作的方方面面，通过硬件获取数据，结合算法建立模型的生态系统。

因此，产品经理在 AI 时代要用生态思维看 AI 产品的需求。

AI 时代互联网公司的需求=实体店（产品体验测试）+网络（快速更新及 A/B 测试）

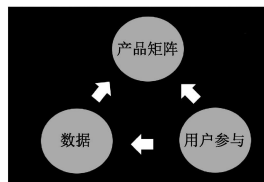
AI 时代 AI 公司的需求=互联网公司的需求+深度学习

深度学习公司的需求=原有数据+硬件收集到的数据

先从深度学习公司讲起。此类公司擅长算法、机器学习建模，但是一般没有大量数据，也没有较多的硬件资源。

AI 时代的互联网公司一般有一定的数据，但是缺少线下的流量，缺少感受线下体验的能力以及快速更新 AI 产品的能力。

基于 AI 生态思维构建的理想 AI 产品如右图所示。

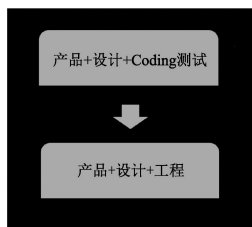


基于 AI 生态思维构建的理想 AI 产品

6.2.3 AI 工程思维

AI 时代的产品经理需要工程思维。每个工程都有个性，例如，工程的地理位置、参与人员、规模大小等属性各有不同，AI 要根据具体情况因地制宜，努力实现个性化。对 AI 产品里面的数据维度、需要的算法组合、算力情况也需要提前做好准备。

从“产品+设计+Coding 测试”的传统思维转换到“产品+设计+工程”的 AI 工程思维是打造 AI 产品需要注意的点。



AI 工程思维

AI 工程思维如右图所示。

6.3 AI 产品经理要懂得心理学知识

有很多创业者找笔者做产品顾问，笔者推掉不少，实在推不掉的，笔者会问他为什么选择笔者做产品顾问。创业者的回答有两个共同点：你懂人性，比一般的产品经理更懂，AI 产品一定是深谙人性的。你跨界且实干，你曾经长期以互联网人的身份调研企业需求，挖掘用户内心的需求，你曾经研究系统架构，你曾经带领研发、设计人员从 0 到 1 打造产品……你至今依然从事产品工作且在一线打磨产品。

讲这些内容，是想说笔者自己和身边的创业者在实践中的共识是：AI 产品经理尤其是从技术人员转岗成为 AI 产品经理的人需要补充心理学的知识，对人性的观察加上 AI 才能给产品赋能。

6.4 AI 产品经理的心力

产品经理是一款产品的灵魂，产品经理的心力决定产品的竞争力，而一名 AI 产品经理要想锤炼心力，需要对产品特点进行深入思考。

笔者最近和一位计算机专业毕业、从事编码工作 3 年后转行做产品经理的读者进行交流，他问笔者能否推荐一些心理学的书籍。由于担心耽误人家前程，笔者并没有推荐，因为锤炼心力可能比学一门计算机语言更难。AI 产品经理的心力决定了产品的质量，AI 产品经理的心力决定了产品获取用户的速度。

6.4.1 认知力：产品方向的致知力

认知论是一门有关知识的起源和性质的复杂而晦涩的学科。认知对产品经理来说是认识需求真谛的学问。

产品经理的认知并非“生而不变”，而是可以不断提升、不断学习的，市场环境、公司文化等都会对认知产生巨大的影响，拓宽认知范围，提高认知层次非常有必要。

打造产品的第一步是找到产品的方向，即找到正确的认知。不要把自己的“想象”当作普遍需求。例如，新零售一定是有市场和用户的，因为新零售重构货与人所处的场景，让货物更方便且高频地出现在人需要的地方。但是新零售不是无人货架。在 2017 年，有家无人货架公司经历了从呼声很高到裁员、撤架的过程，其原因是将无人货架放在办公室里是一种“想象”出来的需求，而不是客户愿意付费的需求。

把“想象出来的需求”当成需求的产品经理属于“自嗨型”产品经理，例如：某明星产品经理曾经令傅盛向其请教如何在海外做产品，但其后期做的自行车产品完全没有用户购买，公司连年终奖都发不出来，这就是把“想象出来的需求”当成真实需求所导致的。成功过的产品经理更容易犯把“想象出来的需求”当成真实需求的认知错误，因为他做出过成功的产品，所以觉得自己想象出来的需求就是真的需求。

明辨是非，做有益于用户、有益于市场的项目是产品长期发展的关键。

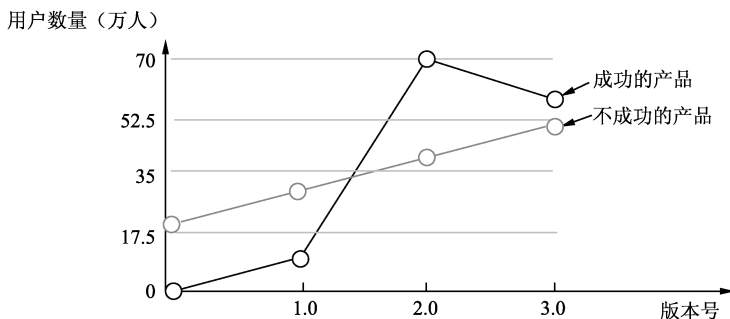
不要被伪需求所迷惑，要找到本心。伪需求是比“想象出来的需求”更加危险的产品，因为太夸张的“想象出来的需求”容易辨别，而伪需求不容易辨别。例如，智能灯泡、IoT 灯泡、用 App 开灯是伪需求，因为现阶段用 App 开灯的效率不如用传统开关开灯效率高。

辨别伪需求一方面需要考虑当下的技术发展水平,判断技术能否稳定支持这个需求,另外一方面要了解这个需求给用户体验带来多大的改善,只有用户体验超越原产品数倍以上,用户才有可能使用新产品。此外,新产品要足够有趣,用户才会对其产生兴趣,才愿意花时间体验。

只有走正确的产品道路,才能把产品做正确。

认知升级后会发现,找对产品方向,做出产品以后,剩下的就是在某个时间实现突破。

成功的产品和不成功的产品的用户数量增长曲线如下图所示。



成功的产品和不成功的产品的用户数量增长曲线

成功的产品的用户数量增长曲线不是呈 45° 的。作为一款科技产品,在每个版本节点,用户数量增长曲线会发生显著的变化。通过用户数量增长曲线的变化,可以了解到产品是会继续呈增长趋势,还是存在问题,需要调整。

知之者不如好知者,好知者不如乐知者,产品经理只有在研究产品时能乐在其中,才能快速提升自身对产品需求的认知。

6.4.2 方法论力: 事上练

产品经理在实际了解用户需求、迭代产品版本的过程中会遇到各种问题,如迭代版本导致用户不增长、团队对需求的认知不统一、项目增长率不理想等,产品经理在这些困难的磨炼下才能得到真正有价值的方法论。

产品经理如果只能在无需求、无用户数量增长困境、无竞品可打、无运营压力时做得很好,一遇到用户数量增长陷入困境、用户驱动力缺失、团队不齐心就手忙脚乱,什么事都做不成,这是不合格的。很多产品“大咖”只想通过“求神拜佛”使用户数

量增长，而并没有耐心地关注产品的一笔一画、一页一转，没有关注产品用户的一点一滴，以及用户数量长久不增长的真正原因。不经历难事便无法翻越产品的困境，而困境期间的琢磨、思考、调整就像鹰的重生，只有主动承担“拔掉指甲”的疼痛，才能打磨出伟大的产品。

产品经理需要在事上“磨”过以后方能成长。淘宝在2003年至2004年的时候如果不是给用户（C端）提供更加有趣、更加便宜的商品，为客户（B端）提供渠道边际成本趋近于零的商品营销管理平台，怎么会战胜实力强大的易趣。经历过淘宝从出生到长大过程的产品经理今日的魅力大部分是被“商家从少到多”“管理从混乱到规范”“用户从没有到被引进来”等问题磨炼出来的。

罗敏在2013年从好乐买离职，开始尝试不同的创业方向。先后“死”掉多个项目，从开始研发到投入市场到结束（甚至还没上线便结束），平均每两个月尝试一个方向。在经过一系列磨炼之后，罗敏于2014年3月决定做针对大学生的分期购物平台——趣分期，终于找到了用户数量增长模型——消费分期模型。

产品经理“事上练”就是去除幻想的需求，保存真实的需求，就是让自己的期望需求被打磨得恰到好处，这是实践出来的方法论。

产品需要小版本、大版本交叉发布以实现用户数量的增长，不能一直迭代却没有质变和跨越。曾经有位产品经理问笔者，产品经理应该吹毛求疵吗？按照“事上练”的产品驱动逻辑，产品经理既不能眼高手低，也不能用战术上的勤奋掩盖战略上的懒惰，既不能每天只关注产品的功能、外形、交互方式，也不能只关注产品的方向、战略，即产品经理既需要“吹”关键方向的“毛”，也需要“求”产品每个细节设计的“疵”。

产品经理要不停地迭代产品，直至找到理想的用户数量增长路径，找寻这条路径的方法是进行严谨的用户需求研讨、产品设计、开发等。

6.4.3 迭代力：初心不动、随机而动

产品经理的初心源自对产品的爱，当遇到用户数量不增长、产品需求处于“空想”的情况时，要保持初心不动，同时做到随机而动。

1. 时刻观察市场、用户的变化趋势

当淘宝网发现拼多多成立仅两年半，就斩获两亿用户，日订单量超过京东，成交总额达到 400 亿元的时候，迅速发现自身近几年一直关注大卖家、品牌卖家，导致中小卖家获取不到淘宝的流量。同时，近几年淘宝一直侧重运营江、浙地区，以及一线城市的核心城区，产品没有完全覆盖广大的三四线城市，忽略了低线城市用户的需求。因此，迅速做出了淘宝特价版来应对拼多多强势增长的情况，淘宝特价版的页面如右图所示。

2. 认知虽不从见闻中来，却要到见闻中去

产品经理要从见闻中观察产品的规模感。拼多多一出来就让用户觉得很多人在用，很多人在买，所以当一个新的消费者去购买的时候，他没有心理负担，他有安全感，因为大量的用户帮其做了选择。很多刚入行的产品经理求新、求异、求特、求变，但做久了就会发现热销的商品一定是中庸中带着一些不同，不同中又中规中矩的。例如：每代 iPhone 一上市，用户就在专卖店门前排队购买，而买到手发现苹果用的技术都是稳定的、市场之前就存在的技术，但是经过苹果的设计，又变得有那么一点不同。

除了规模感，一款产品要具备熟悉感，心理学说人类有怀旧等情感诉求，产品要让用户产生遇见故人的感觉，同时又要让用户找到一些陌生感，这样一款新的产品才有更大的机会去获取用户。例如，戴森吸尘器首先是一款吸尘器，但是在外观上又设计得十分简约，且重量轻，可以做到随用随取，这就是在熟悉中又有一种陌生。这种熟悉感容易让用户产生信任，不再怀疑产品的质量。

陈华创立唱吧 App 的初心是让用户 K 歌变得更简单有趣。当唱吧 App 获得超过 3 亿次下载时，陈华从线下 KTV 中发现了新的需求，唱吧 App 一方面开始做麦颂 KTV，另外一方面投资了咪哒 KTV，陈华做到了从见闻中来，到见闻中去。

3. 不要以昨天的眼光理解今日的用户和需求

2015 年，彭蕾来上海虹桥迎宾馆为“前橙会”的同学分享阿里巴巴的近况和阿里巴巴同学之间的往事和情谊。



淘宝特价版的页面

彭蕾提到，阿里巴巴早在16年前做的就是O2O，阿里巴巴国际产品的产品逻辑就是O2O，马云本人也拜访过客户。这种说法没错，但是笔者认为，这种想法是以昨天的眼光理解今日的用户和需求，是思考不出来支付宝红包这种产品的。

在日常的产品设计中，经常会听到一些“大佬”说：“这不就是我当年的某某思路吗？”首先，笔者承认这些“大佬”确实成功过，例如，当时笔者从0到1帮美特斯·邦威打磨“有范”的设计时，经常听到老板说：“这个就是19年前我开裁缝店时的思路”。笔者很佩服这些白手起家的老板的狼性和人格魅力，但是这种产品思路笔者非常不认同，19年前老板提出个体户概念时的市场环境和用户需求，跟2019年是有着本质的不同的。

同样的道理，在不同的市场里，不同的产品和服务才有机会赢得当时的用户。

4. 看准时机，适时出手，一击必中

抖音可以说是很火了，我们来看一下张一鸣的“打法”。一方面用火山小视频在正面战场牵制竞争对手，另外一方面差异化抖音的定位。竞争对手关注三四线城市或者乡镇用户群，抖音就关注一二线城市。内容和用户群定位十分精准，同时吸引头部KOL（Key Opinion Leader，关键意见领袖）参与，让这个目标用户群一进来就产生亲切感。

产品经理岗位是非常关键的岗位，有时候不是方向不好，而是对产品需求没做到初心不动、随机而动。

6.4.4 验证力：知行合一

产品思维、产品方法论和产品经理的心力都是在知行合一的过程中提升的。认知力和行动力是并行的。要反复从用户数量是否增长、增长速度是否适合、产品方向是否正确中验证知行是否合一。

在设计产品之前，产品认知力、方法论力、迭代力是奠定一版产品成功的关键。

只有当我们很清楚地看到认知力、方法论力、迭代力与产品方向、方法论之间的关系时，我们才能获取自己最应该获取的能力，做出最准确的判断。而决定产品一个版本的四个核心点，除了产品经理心力和产品力，还要加上代表了用户特点的熟悉感，以及使得产品增长的规模感。

第7章 AI产品美学

笔者在讲“成为AI产品经理”这门课时，许多学生对最后一讲“AI产品美学”有两个疑问：一方面好奇AI产品与美学的关系；另一方面想知道如何打造具有美学价值的AI产品。本章就来介绍一下AI产品美学的相关内容。

7.1 AI产品美学的科学性

AI产品美学有自己的研究对象和研究方法。具体的研究对象是一根、二脉、三通、四达。具体的研究方法由AI本身的特性决定，即以AI的跨学科、前沿、综合性的本质为依据，结合经济学、脑科学、生物学、神经科学等最新理论和实践进行研究。

7.1.1 一根：定义落地

AI是对人的模拟，人类具有符号化推理能力、逻辑思考能力，同时具有情感表达、创作、顿悟等能力。AI如果想要实现真正“智能”，就不能仅仅具有模拟人类抽象思维的能力，即推理能力和逻辑思考能力，还要具有模拟人类情感表达、创作等与形象、灵感有关的能力，而后者是美学所研究的对象。

美学家将审美、享受、感性、创造等能力视作人的思维认识的一部分。从智能的定义来看，AI就必然和美学相联系，因为智能除了包括应用知识、推理、思考的能力以外，还包括感知、体验等能力。

科学家将人类思维分成三种：抽象（逻辑）思维、形象思维和灵感思维。

7.1.2 二脉：行为落地

AI 可以通过认知人类的情感和行为，判断与人类交互应该采取的方式。

随着人类与计算机的关系越来越紧密，人们已经越来越意识到 AI 情感的重要性。例如，智能系统在与客户进行交流时，可以通过客户的服装打扮、面部表情、肢体动作来分析客户的情感。如果智能系统传递出相匹配的情感，那么就能使双方的交流更加平等、顺畅。

例如，当客户进入店铺时，为了使客户和商品得到准确的匹配，需要对进店客户进行分析，以定位目标客户群，从而进行营销策略的调整。通过对入店客流进行分析，可以对 VIP 客户、回头客，甚至潜在客户做到“入店即识”，将客户信息实时推送至店员手机，帮助店员抓住销售机会。

7.1.3 三通：认知落地

AI 可以通过模拟人的艺术风格发展出自己的创作风格。

从 AI 在艺术方面的应用来看，人类借助 AI 创作艺术作品甚至编写智能程序，让智能机器人在一定程度上独立进行艺术创作已经成为现实。

例如，美国艺术家科恩设计了艺术创作软件 Aaron，该软件所创作的绘画已具备了自己的风格，其作品甚至被美术馆收藏。天津大学孙济州的“中国水墨画效果计算机模拟与绘制系统”用计算机生成了水墨画。20 世纪 80 年代，美国音乐教授库柏设计了名为“音乐智能试验”的程序，通过分析和提取音乐大师作品中的旋律特征并加以重组，使得 AI 创作的音乐作品几乎能与大师的作品相媲美。而索尼公司设计的音乐程序则可以创作出酷似披头士乐队的作品。

这些 AI 所完成的艺术作品也可以成为美学的研究对象。艺术风格、计算机艺术的独特性、人与计算机在艺术活动中的关系（创作者、工具、欣赏者）都属于美学研究的范畴。

7.1.4 四达：交互落地

AI 产品应努力让人类情感和美学本质进行交互。

在 AI 模拟人类情感进行艺术创作的时候，也需要对人类情感和本质进行研

究。AI 在模拟人类的推理和逻辑思考方式时,探索出了非经典逻辑、不确定性推理等思维方式,针对人类的常识推理机制进行了有益的研究。智能体在模拟人类情感时,也需要对其进行系统的分析。例如,情感的产生机制,情感与人体机能、环境的关系,情感与性格、文化的关系,情感如何在艺术中表达等。

7.2 将美学融入 AI 产品

AI 美学是 AI 对人类情感和艺术美学的模拟。美学对 AI 产品的指导,不应该仅仅停留在总体感知和方向建议上,而应该应用于 AI 产品的设计中。

1. 情感量化

微软小冰、贤二机器僧等 AI 产品,能够针对客户谈话内容进行情感分析并根据分析结果进行对话。这种情感分析的基础是感情计算,即情感量化。

右图为笔者与贤二机器僧的对话。

情感量化的原理是利用计算机强大的存储、搜索和运算能力,进行计算、分析与情感相关的外在表现,如面部表情、心跳速率、皮肤温度等生理特征。

情感量化可以分析艺术家在作品中留下的能体现情感的痕迹,如色彩、形状、线条、文本等。在分析过程中,需要借助美学家、艺术理论家对艺术作品和艺术家的情感进行分析,这是美学参与 AI 研究的实例之一。

2. 程序创新

Facebook 与美国高校合作开发的绘画程序 CAN (创意对抗网络),宣称能够创造出新风格的绘画作品,并且能够通过图灵测试,让艺术爱好者分辨不出艺术作品是人还是机器创作的。CAN 与传统的 GAN (生成式对抗网络)的不同在于,CAN 在鉴别



笔者与贤二机器僧的对话

器中设置了两种标准：一种是“是否是艺术作品”，另一种是“所生成的艺术作品属于什么风格类型”。这两种标准被设置成对抗式的，即既要属于艺术的范畴，又要不同于既定的风格。这实际上是一种应用了学习偏差（最大限度地偏离既定风格）及鉴别艺术标准（最小限度地偏离艺术范围）的方法。也就是说，创新是在既定范围内对原有规则的突破和偏离。创新是建立在艺术风格和艺术标准基础之上的，而这正是美学研究的范围。而且，对 CAN 创作的作品进行风格鉴赏和艺术评判，也需要由美学家完成。因此，AI 产品的设计必须融合美学家、艺术家的能力。美学家、艺术家对 CAN 的创新性和艺术标准的研究和批评，正是 AI 美学的重要研究范围。

通过以上例子，我们可以看到 AI 美学并非仅仅是对艺术创作在宏观层面进行把握和指导，而是深入到具体的程序设计，对智能体创作出的美学作品进行分析和研究，这使得 AI 美学无论是在理论上还是在应用上都有其价值。

7.3 AI + 场景美学

AI 产品兼具感知性的美学和逻辑性的美学。好的 AI 产品是既让用户熟悉又超越常规场景的产品。

AI 产品在以下几个方面创造具体的场景美学。

7.3.1 懂用户之美

美是有属性和细节的，在不同的场景下，用户对智能产品有不同的渴望和需求。例如，尚未毕业的大学生需要的智能产品可能是智能手机，最好是兼具性价比和品牌调性的智能手机。工作了三、四年的人需要的智能产品可能是智能家庭影院、VR 一体机等。已经结婚、生子的人对智能产品的需求又会不同。所以美是流动的，对用户需求的把握要根据用户在不同阶段对美的不同需求来定。

7.3.2 懂痛点之美

因为美的流动性、用户需求的隐藏性，懂痛点之美显得尤为重要，互联网产品经理挖掘的不仅仅是用户的“痛点”，还有“痒点”甚至是“笑点”。那是因为很多产品能赢得流量即可。但是，AI 产品经理应该明白，AI 时代的各种应用多如牛毛，人们需要的是比原来的产品好 10 倍以上的产品，所以需要 AI 来增效。而用户的痛点往往

是用户表达不出来的，或者羞于表达的。笔者在《产品经理进化论：AI+时代产品经理的思维方法》一书中讲解过挖掘用户痛点及隐藏需求的方法。

7.3.3 爆款产品之美

用好的设计方法打造爆款产品是 AI 的核心目标之一。围绕用户去产出一个美的产品，打磨出爆款产品本身就是一种美。

笔者认为产品美学是一种节制，不只是功能的节制，更是产品思维的节制。产品美学靠的是内外兼修，向外修炼是指与用户互动、与粉丝互动，向内修炼是指围绕用户去产生一个美的产品，从不同的角度思考产品、思考生活。

7.4 AI 产品的美学目标

AI 产品要求产品经理把自己放在美学范畴中，把自己对生命的理解、对生活的热爱放在美学范畴中，把自己修炼成一个“美学产品”。让自己变得美，让身边的人变得美，让自己企业的产品变得美，让品牌变得美。

7.5 AI 产品美学综述

一般综述写在开头的居多，不过，因为 AI 产品美学是由笔者提出的，概念比较新，所以先讲 AI 产品美学的科学性，再进行综述。AI 产品美学包含广义的 AI 产品美学和狭义的 AI 产品美学。

广义的 AI 产品美学：产品经理的本质是产品管理者，产品经理需要考虑公司的战略布局、技术的潜力、团队的凝聚力、产品的具体设计（工业设计、外观设计、交互设计、UI 设计）等。

狭义的 AI 产品美学：产品和服务对用户交互行为的感知。

本书所讲的 AI 产品美学主要是广义的 AI 产品美学。产品经理具有美学意识在 AI 时代是非常重要的，原因如下。

(1) AI 产品已不再是单纯的软件产品，AI 产品需要立体的美学、结构的美学、工程的美学。

(2) AI 产品团队一般是由跨界的人员组成的，学会欣赏他人之美，AI 产品经理

才能驱动团队的成员发挥自身最大的创造力。

(3) AI 不仅模仿人的大脑，还通过模仿人的行为代替人完成一部分体力劳动。

7.6 如何培养产品美学观

1. 独立思考，不人云亦云

产品经理的产品美学观是一款产品区别于其他产品的根本，进行有意识的创造是 AI 时代对产品经理的本质要求。

在 AI 时代，我们会看到大量个人定制的产品。通过大数据分析，对用户的喜好做方方面面的调查，将一切都记录在案。通过分析用户的购物习惯、争吵方式、审美特征，给用户提供的贴合用户偏好的个性化服务。用户接收到的信息，会是依据用户的偏好发出的。

但这不是真正的个性化，这个被服务的用户只是机器人和 AI 工业流水线上中的一个服务对象，用户被标准化的作业“个性化”了。这个被服务的用户只是一个消费者，作为价值的需求端存在于经济系统中，这个被服务用户的“个性化”只是 AI 系统为自己积累的数据，是系统用来丰富算法的案例。

2. 设计产品时应该思考哪些是共性，哪些是个性

共性来自客观个体，来自用户对产品的综合感悟和观察。

3. 读“杂书”

AI 产品经理要多角度地感悟产品之美，发现产品之美，打造产品之美。

例如，读《史蒂夫·乔布斯传》可以了解简约设计之美；读唐诗可以了解对称之美；读阳明心学可以了解在逆境中泰然处之之美。

产品做得深刻的人，一定是对生命有深刻感悟的人。

第 8 章 AI 赋能产品的实例

8.1 AI+金融保险

1. 众安保险的存量机碎屏险——“赔你碎”的自动识别碎屏技术

(1) 解决了什么问题。

旧手机之所以缺少碎屏险，是因为有两个“障碍”：怎么证明提交信息的手机就是要投保的手机？怎么确认手机屏幕是好的？而“赔你碎”通过远程身份识别和智能碎屏识别等 AI 技术解决了上述问题。在保险有效期内，一旦用户的手机屏幕破裂，即可联系众安保险客服申请理赔，客服将根据用户所在位置安排工程师免费上门维修，或提供众安保险选择的第三方寄修中心的地址，以完成维修。

在用户体验方面，相较于同类型产品，“赔你碎”更加人性化，比如当用户为其他手机投保碎屏险时，“赔你碎”不需要消费者手动输入 15 位的 IMEI 数字，而是通过机器识别技术自动获取 IMEI 数字。

(2) 使用了哪些 AI 技术。

“赔你碎”利用远程身份识别技术确保用户上传的照片属于待投保设备，通过算法生成身份识别标识图片并自动刷新，采用智能碎屏识别技术确认手机屏幕是好的还是坏的，利用基于深度学习的图像识别算法，自动识别可疑碎屏，减少人工干预，降低人工审核成本。

2. 七灵科技的保险风险预测

(1) 解决了什么问题。

系统将运筹优化技术和机器学习算法应用于保险风险预测和自动化决策，解决

保险业务经营管理中数据处理和数据决策流程低效的问题，制定平衡风险、管控需求、把握市场导向的保险经营政策，并将数据可视化，实现业务分析、决策和监控，提升了保险公司的业务分析和业务决策能力。每年为保险公司节省综合成本约 360 万元，数据处理和分析的效率得到大幅度提升。

(2) 使用了哪些 AI 技术。

运筹优化技术、机器学习、数据可视化。

3. 法国巴黎人寿与联邦银行合作推出的“优利联年保险计划”

(1) 解决了什么问题。

“优利联年保险计划”提供寿险、年金双选择，还设计了内扣式附约功能，保户可添加“重大疾病”“意外伤害”附约，零附加费用率打造最优惠的保障平台。

(2) 使用了哪些 AI 技术。

“优利联年保险计划”利用 AI 智能平台结合大数据分析技术，有效控制投资波动，利用时间累积账户收益，协助保户实现“保障+投资”双赢。

“优利联年保险计划”主要采用了机器学习和数据挖掘技术，其进行数据挖掘的流程如下。

选取数据集。选取巴黎人寿与联邦银行的账户数据和相关数据。

数据预处理。对数据进行数据清理、数据集成、数据归约、数据变换和数据离散化预处理。

选择数据分析算法。利用预处理好的数据按照设计好的算法进行模型训练。

分析、总结、改进。分析、改进模型。

4. 市场波动智能分析系统

(1) 解决了什么问题。

自 2015 年股市出现异常波动以来，监测市场波动已逐渐成为监管机构的主要职责之一。为了提升分析市场波动的能力，实现对风险的预研、预判，需要对市场特殊时间节点的交易进行深入分析。使用机器学习等新技术，从股票账户信息、宏观信息、基本面等维度研究风险的传导路径，建立相关模型，分析市场波动，为进一步强化监管奠定基础。

市场波动智能分析系统具有以下功能。

辅助专家进行市场波动分析,提升工作效率。

帮助监管者发现市场中存在的、人类专家难以察觉的更深层次的问题。

独创的算法可以借鉴人类专家的经验 and 知识,使算法具有更强的可解释性。

(2) 使用了哪些 AI 技术。

市场波动智能分析系统利用机器学习技术对市场波动的特征和成因进行分析,以便跟踪、监控证券交易等市场的波动变化。在分析过程中,通过对宏观经济指标、交易行情等数据进行深度挖掘,快速完成市场波动传导路径的分析,有效防范风险,提升市场分析的工作效率。

5. 摩根大通证券交易执行软件

(1) 解决了什么问题。

摩根大通在欧洲股票算法业务中投入使用 AI 应用 LOXM,旨在以最佳价格和最快速度执行客户的交易指令。

(2) 使用了哪些 AI 技术。

LOXM 能够利用深度强化学习技术,从过往的数十亿条实盘和模拟盘的历史交易记录中学习、总结、归纳经验和教训,解决更加复杂的问题,比如怎样才能在不扰动市场价格的情况下大量抛售股票。

与一些私人银行提供的机器人投资顾问不同,LOXM 没有决策的能力,无法决定买卖标的,它的作用仅仅是买入、卖出。LOXM 未来的发展方向是识别特定的用户,从其言行举止中决定如何交易,并在客户同意的情况下进行交易。

6. 小结

众所周知,要想做一名称职的操盘手,必须要保持较高的效率,能够果断进行决策,但是人无完人,在瞬息万变的金融市场,没有人能够保证从不亏损,而 AI 凭借其高效、果断的特性,正在一步步地取代人类操盘手。高盛集团位于纽约的股票现金交易部门,其操盘手的人数已经从 2000 年的 600 人减少到了 2018 年的 2 人。在不到 20 年的时间里,AI 几乎取代了人类操盘手。

其实,一个普通操盘手的核心工作在于通过市场信息和自己对市场的判断,适时

进行投资交易决策，而这正是 AI 的优势所在。

AI 的数据收集能力和运算能力远超人类。无论是宏观经济指标还是市场价格数据，AI 搜集和整理相关数据的速度都要远远超过人类，再结合内嵌的算法模型，AI 能够在短时间内做出投资决策。

在深度学习的辅助下，AI 比人类更容易掌握市场的规律。人类操盘手做出的投资决策往往基于个人经验和传统的方法论，而 AI 则可以根据市场规律和市场大环境，迅速做出相应调整。

失手的可能性更低。人类操盘手的操作风险以及道德风险都是不确定的，这也让金融机构承担着随时可能出现重大损失的风险。如果使用 AI 操盘，首先，在“忠诚度”上就不必担心；其次，犯错的概率非常低。

所以，像高盛和摩根大通这样的知名金融机构自然愿意启用 AI 进行交易。

8.2 AI+医疗健康

AI 的优点是高效计算与精准分析，这将给医疗健康行业多个领域的发展带来机会。AI 在医疗健康行业主要的应用有：医学影像分析、医疗数据管理、症状分析及咨询、医疗机器人、可穿戴医疗设备、远程医疗、智能医疗决策、智能诊断、健康管理、脑机接口等。

AI 对医疗健康行业的颠覆将会是全方位的，主要表现在：颠覆传统的药企，帮助药企发现精准化、个性化的药品；颠覆传统医院，促使传统医院从固定到移动，从近程到远程；颠覆医生的诊断方式，将医生从烦琐的事务中解脱出来，成为诊疗规则的制定者和诊疗过程的监督者；颠覆病人的看病方式，逐步实现足不出户就能得到精准的、个性化的治疗方案，让患者拥有更好的治疗体验。

有专家认为 AI 在医疗健康行业的爆发顺序为：初级计算智能、中级感知智能、高级认知智能。专家预测在未来的 5~7 年，计算智能和部分感知智能将迎来发展机会，计算智能的基因检测和药物发现，感知智能的可穿戴医疗设备等将迎来爆发；在未来的 8~10 年，感知智能中的远程医疗和医疗机器人，认知智能中的智能决策和智能诊断将迎来爆发。

资本市场也非常看好 AI 在医疗健康行业中的应用。2017 年，医疗 AI 领域共发

生融资事件 32 起, 涉及金额约 11.48 亿元。

本部分对部分有代表性的医疗健康行业的 AI 企业进行简要介绍, 希望可以给读者带来启发。

1. 腾讯觅影 (<https://miying.qq.com/>)

以肺癌识别为例, 基于肺癌识别算法, 腾讯觅影对大于 3mm 的微小结节的敏感度已经超过 80%, 对于良性肺结核的检出率超过 84%。据悉, 普通医生的判断准确率一般为 60%~70%, 腾讯觅影已经超越普通医生的平均诊断水平, 与优秀医生的水平相当。

2. 依图 (<http://www.yitutech.com>)

依图从事 AI 创新性研究, 致力于将先进的 AI 技术与行业应用相结合, 建设更加安全、健康、便利的世界。从智能安防平台开始, 依图已经将 AI 应用于智慧医疗、城市大脑、智能硬件等领域。

在智慧医疗方面, 依图不仅为医院提供跨科室的多场景应用系统和数据分析平台, 还在科研课题上与数十家国内顶级三甲医院达成了深度合作, 参与多项国家“十三五”重大课题与国家自然科学基金项目, 通过医学与其他学科的跨学科融合, 共同推进医学的进步。

凭借在深度学习、计算机视觉、自然语言处理等领域的先进技术以及在医疗行业的深度积累, 依图面向放射科、超声科、儿科等多个科室提供临床诊疗辅助, 并结合前沿医学热点, 提供专业的科研大数据分析应用。

依图在医疗方面已经推出的产品有: care.ai 胸部 CT 智能辅助诊断、care.ai 儿童骨龄智能辅助诊断、care.ai 儿科门诊智能辅助诊断、care.ai 病历智能搜索引擎、care.ai 临床智能科研平台等。

3. 图玛深维 (<http://www.12sigma.cn/>)

图玛深维由高通公司深度学习和计算机视觉专家于 2015 年在美国创立。公司率先将深度学习技术引入智能医学诊断和医学数据分析领域, 图玛深维在北京、苏州、圣地亚哥设有研发团队。

图玛深维致力于研究开发基于深度学习技术的自动化医疗诊断系统与医学数据分析系统, 并计划推出针对多类肿瘤疾病、心血管疾病、脑血管疾病自动检测的诊断

产品。其 σ -Discover 系列深度学习诊断系统不仅大大提高了医生诊断效率与精确程度，还在一定程度上避免了因为经验差异造成的误诊，并且可以大大改善由于区域性医疗资源分布不均所导致的看病难、治病难的现状。图玛深维已经与北京、上海的多三甲医院达成科研合作协议，在产品的医学应用、技术研究及临床试验上进行全方位合作。

此外，图玛深维十分注重创新，同时制定了云端战略，将 AI、大数据分析和云计算相结合，大幅度提高医疗机构的服务能力和诊断分析水平。其云端 CAD 平台是一套基于云计算与云存储的辅助诊断平台，是智能医疗诊断平台发展的重要一环，它不仅可以实现远程智能诊断与信息共享，还能够与 CT、MR、CR、DR、ECT、超声、内镜等医疗设备完美结合。

4. 推想科技 (<http://www.infervision.com/>)

推想科技是一家 AI 高科技公司，致力于应用深度学习技术为医疗影像辅助筛查提供快捷、准确的解决方案，能够有效利用各类医疗数据创造出具有临床价值的产品，力求突破医疗影像辅助筛查现有的技术水平，全面推进医疗领域的发展。

推想科技提出的解决方案主要有深度学习科研平台、智能 X 射线辅助筛查产品、智能 CT 辅助筛查产品等。推想科技的深度学习科研平台集成深度学习核心算法与功能，拥有强大稳定的 GPU 运算能力，每秒可处理和运算超过 100 张高清医学影像的数据；智能 X 射线辅助筛查产品能够对心、胸部位的 20 多种不同病灶进行判断，在肺结节检测上的表现尤为突出；智能 CT 辅助筛查产品在早期肺癌筛查中，凭借其高性能的并行运算能力，能有效挖掘肺癌的核心特征点，高效判断不同序列影像是否存在疑似癌症的特征，实现肺癌早诊、早治的技术性突破。

5. 深睿医疗 (<http://www.deepwise.com/>)

深睿医疗坐落于北京中关村及杭州人工智能小镇。公司致力于开发医疗影像识别技术和产品，应用于各种恶性疾病的早期筛查和精确诊断。

2017 年年初，深睿医疗的研究成果开始临床验证后，正式成立公司并获得 3500 万元的 A 轮融资。深睿医疗已经完成第一款用于肺癌早期筛查的医疗影像诊断系统，并在多家三甲医院进行试用。围绕“疾病筛查”和“基础医疗”，深睿医疗也在研发针对乳腺癌、脑卒中、骨折等疾病的 AI 辅助诊断的相关产品。

深睿医疗的产品定位为,一是帮助基层医院做疾病的早期诊断和筛查;二是帮助三甲医院对复杂病例进行科学决策和辅助治疗。相对而言,三甲医院对 AI 技术的期望和要求更高,不仅仅是筛查,还需要结合其他检查信息对病变部位进行进一步分析,从而帮助医生进行科学决策,达到精准治疗的临床目的。

6. 森亿智能 (<https://www.syyni.com/>)

森亿智能是一家专注于医疗数据治理与挖掘的医疗 AI 公司。通过 AI 技术,解决当前医疗机构所面临的大量原始医疗数据质量不高、结构化和标准化程度不足、分散在“孤岛”中无法直接利用的问题。通过进行数据治理将低质量的数据自动转换为高质量的数据,推动数据驱动的医疗科研、医院管理、患者服务的发展,加速智慧型医疗系统建设。

森亿智能已拥有多项成熟的医学 AI 核心技术,包括:医学自然语言处理(中文)、医学知识图谱(中文)、基于医学 AI 技术进行数据治理等。依靠其所拥有的行业领先技术,森亿智能成功研发出了临床科研一体化平台,该平台以数据治理为理念,全自动化解决研究型医院的海量临床数据与科研数据的结构化及标准化问题,释放、挖掘数据的价值,推动数据驱动的医学科研与患者服务的融合发展。

7. GoogLeNet 系统的深度学习算法

(1) 解决了什么问题。

深度学习算法对制作肿瘤概率预测热点图很有帮助,但其结果常有噪声。研究者将算法额外定制化,模仿病理学家的诊断方式完成神经网络的训练,希望能训练出一个相当于或超越病理学家表现的模型,算法经过优化可确定是扩散到了淋巴结的乳腺癌还是扩散到了临近乳房的淋巴瘤。即使是针对不同医院、不同扫描仪得到的图片,模型也能较好地泛化。将深度学习算法应用到数字化病理学中,可以通过创造自动检测算法来补充病理学家的工作。

(2) 使用了哪些 AI 技术。

使用了深度学习、图像识别等 AI 技术。

8. 小结

AI 对医疗的赋能提升了医疗效率,降低了医疗成本。AI 对医疗领域的影响非常大,从辅助诊疗到药物研发,从基因测序到病历信息化等,相关领域借助 AI 都已经

取得了巨大的进步。同时，AI 也推动着医学检测、影像诊断等业务的发展。AI 赋能医疗已经成为未来医疗领域发展的趋势。

8.3 AI+工具软件

虽然我们已经处于 AI 应用大爆发的时代，但是具体到每个人的体验，恐怕很多人都会觉得“我每天的工作内容跟 AI 并没有什么关系啊。”

其实，很多都市办公族忽视了每天都在用的办公系统——微软 Office 365。没错，Office 365 早已经从昔日的桌面办公软件进化为全面 AI 化的生产力平台。

体现 AI 价值的并不是玄妙的科学幻想，而是在每个人触手可及的地方给人们的工作和生活带来改变。

作为全球科技巨头，微软对 AI 的价值观让人耳目一新。微软要让 AI 给每个人的工作都带来改变，而最好的平台就是微软自己的核心产品——Office 365。在这样的理念驱动下，全面 AI 化成为 Office 365 的首要目标。

每天使用 Office 365 工作的读者可能已经注意到，在 Word、Excel、PowerPoint 界面最顶部菜单栏的中间部位，多了一个有放大镜标志的搜索框，搜索框里有一行字“告诉我我想要做什么”，这个搜索框就是 Office 365 已经 AI 化最直观的展现。

但是，Office 365 的 AI 功能可不止这一个，微软最新的 AI 技术，包括微软认知服务、微软翻译（Microsoft Translator）、微软小娜（Cortana）、自然语言问答、Power BI、数据防泄漏保护等，已经全都融入 Office 365 中了。

通过“告诉我我想要做什么”，用户可以获得所有关于 Office 365 的功能介绍和操作指导。例如，智能个人助理微软小娜能够连接 Office 365 与 Outlook 日历，使 Office 成为具有语音对话功能的智能私人秘书；Outlook 的重点收件箱则可以通过智能算法过滤并自动分类邮件，用户可以快速完成回复会议请求、管理待办事项等工作。

除了上述通用办公功能，对于专业领域的人士而言，研究工具（Researcher）还可以检索 Bing 学术搜索中的公开资源并自动生成引用信息；智能查找（Smart Lookup）让你不用离开文档就能从网上找到更多相关信息；由微软翻译技术驱动 Presentation Translator 功能可以帮助用户将 PPT 的文本在 60 多种语言之间快速翻译，还能为中文、

英文等 9 种语言的语音演讲添加实时翻译字幕。

截至 2018 年，Office 365 已经整合微软的 53 项 AI 技术和特性，服务超过 150 万的企业付费用户。

除了被 AI 赋能的微软 Office 365，被 AI 赋能的工具软件还有以下类别：模型及技术开发、虚拟助理开发、身份识别、图像处理、人机交互、图像分析、视频分析、嵌入式 AI、工作效率管理、语音识别、图像识别、情绪识别、系统管理、声音模仿、运动追踪、电子身份构建、眼球跟踪、定位、质量检测、代码编写、个性化设置、仿真全景构建等。

下面笔者继续介绍几个实例。

1. 英伟达的可编程推理平台 TensorRT

(1) 解决了什么问题。

TensorRT 是一种高性能神经网络推理引擎，用于在生产环境中部署深度学习应用程序，可实现最大的推理吞吐量并提高效率。TensorRT 是一款可编程推理加速器，能提升现有和未来的网络架构的运算速度。通过 TensorRT 的大幅度加速，服务提供商能够以经济实惠的成本部署计算密集型 AI 负载。

(2) 使用了哪些 AI 技术。

TensorRT 包含一个可以优化在生产环境中部署的深度学习模型的库，可获取经过训练的神经网络（通常使用 32 位或 16 位的数据），并针对降低精度的 INT8 运算来优化这些网络。借助 CUDA 的可编程性，TensorRT 能够适应深度神经网络日益多样化、复杂化的发展趋势。TensorRT 的基本处理过程为：将 Caffe 的模型转化为 GIE 的模型，或者从磁盘、网络加载 GIE 可用的模型；运行 GIE 引擎；提取结果。

其知名客户有科大讯飞、京东、商汤科技、海康威视等。此外，TensorRT 已经支持所有流行的神经网络框架，并覆盖更多种类的 GPU。

2. 搜索和推荐服务技术方案 Hugefly

Hugefly 为在线零售商店和电子商务公司开发了实时搜索和个性化推荐工具。其工具能够改善搜索功能，例如现场检测、多面搜索、个性化搜索，以及集成和分析；该工具同样提供推荐功能，包括个性化推荐、协同过滤、基于内容的过滤、基于知识的过滤，以及集成和分析。

3. 谷歌：个人移动设备锁屏推荐技术

个人移动设备锁屏推荐技术为锁屏的个性化设置引入新的算法层，使得锁屏可以展示用户最喜爱的应用、视频、音乐等内容。这种“富”数据将被用于“喂养”机器学习系统，能够使锁屏的个性化推荐更加智能。锁屏界面会通过持续学习大量交互动作及语境数据完成升级，即它会根据用户的反馈和个人模式进行迭代。

应用该技术的设备具有如下特点。

这种新型的“富通知”交互方式展示了一种新的用户界面范式。锁屏不会受到“按照时间推送通知”的束缚，通知消息会在没有操作的前提下自动出现在用户的锁屏上。

在根据用户行为进行直接学习的过程中，个性化代理在利用云计算能力的同时，还利用存在于各种数字环境中的内部代理和外部代理。此外，这些个性化代理还可以处理并提供特定领域的的数据、信息及建议，这些领域包括股票市场的提示、经过优化的旅游路线等内容。最终推动个性化代理进行协作、竞争并向更优版本进化。

个人控制算法可以确保个体在未来的算法世界中保持能动性。随着决策由机器学习系统得到强化，用户需要可解释的算法、界面与方法，可以用明确和可理解的方式指导并控制这些个性化代理。

4. 阿里巴巴：工作效率管理、统一变更管理平台 Changefire

Changefire 平时对接了阿里巴巴集团多个业务单元的 100 多个变更系统，每天为各种业务线提供几万次变更信息查询服务来辅助故障定位。

Changefire 使用了故障定位技术，用机器学习技术自动统计并分类各条业务线的故障规则库，在故障发生时能利用规则库提升定位故障的效率，故障恢复时间同比可以减少 65%。

5. 寒武纪科技

2018 年 6 月，全球智能芯片独角兽企业寒武纪科技（以下简称寒武纪）正式宣布完成数亿美元的 B 轮融资。寒武纪方面称，B 轮融资后寒武纪整体估值达 25 亿美元。

作为 AI 芯片公司，寒武纪在终端领域以处理器 IP 授权的形式与全世界同行共享技术成果，帮助全球客户快速设计和生产具备 AI 处理能力的芯片产品。寒武纪研发

的寒武纪 1A 处理器 (Cambricon-1A) 是较早的商用深度学习专用处理器, 已经被数千万台智能手机应用。寒武纪智能终端处理器已衍生出 1A、1H、1M 等多个型号, 将为全球数以亿计的各类终端提供本地处理能力。

在云端, 寒武纪致力于为全球客户提供高性能、低功耗、高性价比的智能云服务器芯片。2018 年 5 月发布的寒武纪 MLU100 智能云服务器芯片 (Cambricon-MLU100), 适用于视觉识别、语音识别、自然语言处理等多种类型的云端 AI 应用场景, 其本身不仅可以高效完成多任务、多模态、低延时、高通量的复杂智能处理任务, 还可以与寒武纪 1A、1H、1M 系列智能终端处理器适配, 以端、云协作的方式提供前所未有的体验。

据中国半导体行业协会统计, 2017 年中国集成电路产业销售额达到 5411.3 亿元, 同比增长 24.8%。但这一全球最大的集成电路市场, 主要的产品却严重依赖进口。在赛迪研究院统计的数据里, 2017 年世界排名前 20 的半导体企业中, 美国企业占了 13 家, 这 13 家企业在中国市场的销售额合计是 667 亿美元。其中, 高通、博通、美光有一半以上的销售额是在中国实现的。

在巨大的市场机遇和压力下, 发展“中国芯”已经不仅仅是企业的动力, 更是企业的崇高使命。

(1) 寒武纪智能终端处理器解决了什么问题。

寒武纪智能终端处理器是专门针对深度学习而设计的, 在图像识别、语音识别领域比传统处理器的性能提高了至少两个数量级, 集成度也是传统处理器的数倍, 这让手机等移动设备搭载 AI 处理器成为可能, 其具备以下特点。

低功耗、本地智能处理。

理论峰值性能高。

卓越的完备性和前瞻性。

A. 支持视觉识别、语音识别、自然语言处理等多种任务;

B. 支持各类 CNN、DNN、RNN、LSTM、SOM 等模型;

C. 支持稀疏化网络, 支持 FP16、INT8 运算。

其面向智能手机、安防监控设备、可穿戴设备、无人机和智能驾驶汽车等各类终端设备, 在运行主流智能算法时性能功耗比全面超越 CPU 和 GPU。

(2) 寒武纪智能云服务器芯片解决了什么问题。

寒武纪智能云服务器芯片将其产品的神经网络加速能力拓展到机器学习领域，以适应更多的运算任务。

寒武纪 MLU100 与 MLU200 支持多种机器学习推理任务和训练任务的加速，前者偏重推理，后者偏重训练。此外，MLU100 面向中小型服务器，而 MLU200 面向企业级 AI 研发中心。

(3) 寒武纪 1H8、寒武纪 1H16、寒武纪 1M 解决了什么问题。

寒武纪新一代智能终端处理器包括面向低功耗场景视觉应用的寒武纪 1H8、拥有更广泛通用性和更高性能的寒武纪 1H16、面向智能驾驶领域的寒武纪 1M。与寒武纪 1A 相比，新一代智能终端处理器在性能功耗比、成本开销等方面进行了优化，性能功耗比再次实现较大提升，适用范围覆盖了图像识别、语音识别、自然语言处理、安防监控、智能驾驶等多个重点应用领域。

8.4 AI+商业服务

AI 赋能商业服务是笔者在撰写本书的过程中研究的课题之一。

企业数字化转型已经迫在眉睫，数字化转型给 AI 提供了一个非常大的机会，AI 是数字化转型的结果。

8.4.1 数字化转型下的 AI

什么是数字化转型？麻省理工学院斯隆管理学院数字化经济计划的首席研究科学家 George Westerman 说：“数字化转型是企业利用技术来彻底改变企业的业绩或方向。”这里以零售巨头 Target 和沃尔玛为例进行讲解。

Target 为了推动数字化转型，聘请了 Tesco 的 CIO——Mike McNamara，McNamara 在这家英国连锁店主导过多个数字化项目。2016 年，McNamara 的大部分时间都花在了将软件开发模式从外包转向内部定制（包括新的供应链应用），以更好地将库存情况与消费者需求进行匹配。

沃尔玛试图将其软件堆栈重新设计为定制平台，包括一个新的搜索引擎和多个云应用。沃尔玛的新电子商务平台是基于 OpenStack 的，旨在帮助沃尔玛更好地利用云

平台参与竞争。事实上，开源是沃尔玛数字化转型的重要组成部分。沃尔玛的团队在 2018 年推出了 OneOps，这是一个让程序员在不同云服务商之间进行测试和切换的平台，是沃尔玛的一个关键优势。

AI 是一个非常强大的、能够帮助企业进行数字化转型的方式，AI 能够帮助企业更好地服务客户，帮助企业更好地优化产业布局，促进企业实现产业、产品和服务渠道的创新。

2017 年 10 月，普华永道做了一项报告，报告认为：到 2030 年，AI 将为全球经济贡献高达 15.7 万亿美元的 GDP，AI 对经济贡献最大的两个国家将是中国（到 2030 年，AI 给中国 GDP 带来的增长将超过 26%）和美国（到 2030 年，AI 给美国 GDP 带来的增长将超过 14.5%）。在报告里面，普华永道对中国进行了分析，因为中国的 GDP 实际上以制造业为主，在这个情况下，普华永道预测中国在 AI 转型过程中会有更大的战略机会。

此外，麦肯锡全球研究院用三个月的时间做了一套分析，它的分析方法是把企业分成三类：第一类是积极采用 AI 的企业，第二类是选择性地采用 AI 的企业，第三类就是被动选择 AI 的企业。麦肯锡的研究人员发现，不管是哪个行业的企业，积极采用 AI 的企业的利润率明显高于其他企业。在某些特定的行业，比如金融、教育、医疗行业，积极采用 AI 的企业的利润率会高于行业平均水平 15%~20%。

8.4.2 AI 一定会颠覆所有商业模式

笔者从初次接触 AI 到做 AI 产品，经历了大概 12 年的时间。这次 AI 爆发非常不一样，因为大数据、算法合在一起带来了非常多的机会。

微软创始人比尔·盖茨讲过：人一般都会高估自己一年干的事，低估自己十年干的事。

从笔者的经验来看，笔者坚信两件事：第一件事是 AI 一定会颠覆所有的商业模式；第二件事是颠覆速度可能会比我们想象的要快得多。

AI 对商业模式的颠覆将会从两方面实现：一方面是对金融、零售等行业进行颠覆；另一方面是对市场、销售等岗位进行颠覆。AI 的发展速度在一定程度上很难控制，但是对 AI 的发展方向，我们应该做好选择，要用一定的方法去把控。

例如，微软很早就提出过一个非常清晰的思路，AI 要增强人的能力，让人类更有

成就，而不是取代人类。笔者的观点和微软一致。

8.4.3 AI 在商业中的具体应用

笔者在调研很多资料后提出了“AI 赋能”的概念，希望大家不要把 AI 仅仅用在极个别的产品当中，服务极个别的客户，而是要将 AI 强大的能力开放出来，让每一位 AI 开发人员、每一位数据科学家、每一家商业公司都可以应用 AI 进行赋能。

例如，微软在 2017 年成立了商业 AI 部门，组建了非常出色的团队。这个部门发现 AI 与传统的软件和服务不一样，传统的软件和服务是先确认是 A 还是 B，发布以后产品性能就不再变化了，有时候随着时间的变化，其性能甚至会变差，但是 AI 产品第一次发布出去的性能可能只有 20%~30%，需要以数据和产品来“饲养”产品，AI 产品需要跟业界、场景深度结合，利用数据训练算法，随着数据量越来越大，产品性能最终会达到 80%、90%甚至 100%，从而更懂商家，也更懂消费者，最终真正服务商业。

这里并不需要每一家零售企业从底层算法设计、数据分析、系统搭建做起，而是学会用 AI 赋能商业的思维。通过将 AI 算法、算力与自己的业务需求和数据结合搭建 AI 体系，通过提供解决方案的方式，帮助自己的企业更好地提升效率。

8.4.4 预测未来最好的方法就是创造未来

除了上文提到的内容，笔者还调研了 AI 对以下商业进行赋能的情况。

AI 能够赋能数据分析、广告营销、工作流程自动化、系统管理、客户管理、文件管理、云计算服务、供应链管理、市场调研、商业管理咨询等内容。

下面笔者挑几个有代表性的案例介绍一下。

1. 王府井百货的智能营销系统 DeepCreator

(1) 解决了什么问题。

王府井百货有大量的数据，既包括传统的 CRM 数据、ERP 数据、POS 系统中的数据，也包括线上商城、微信公众号以及有赞商城等平台的数据，这些数据在之前并未打通，而智能营销系统 DeepCreator 将这些线上、线下、内部、外部的数据全部汇总到基础数据平台。基于以消费者为中心的大量数据，智能营销系统 DeepCreator 结

合业务人员的经验与机器学习算法搭建模型,对会员或访客的行为轨迹进行分析,从而建立消费者画像,商家可根据消费者画像进行精准营销,实际应用有推荐顾客感兴趣的产品、确定呈现给不同顾客的优惠券信息等。

零售行业之前的营销非常粗放,比如店庆大促时给所有会员都发送同样的信息。而王府井百货的智能营销系统 DeepCreator 基于大数据分析,搭建精细化营销系统,同样是店庆大促,每一位顾客收到的优惠券信息却有所不同,如果是对价格敏感的顾客,就为他们推送打折信息;如果是喜欢尝鲜的顾客,则为他们推送新品的信息。此外,还可以根据顾客的喜好,推荐特定的品牌和品类。

(2) 使用了哪些 AI 技术。

搭建这套智能营销系统分三步。

A. 搭建大数据系统。

B. 归拢、整理数据,打通品牌、品类之间的数据壁垒,建立针对用户画像的标签体系。在归拢数据的过程中,很多数据的格式都不相同,需要进行数据的降噪和清理。在整理数据的过程中,采用图像解析技术提取纸质文档的内容,以提高效率。由于线下零售商关心场地周围的客户,故补充一些第三方的数据,例如商业 Wi-Fi 运营商、腾讯或者京东等平台的数据。

C. 搭建智能营销系统 DeepCreator。基于大数据分析,筛选不同的客户群体,找到对他们更有效的刺激点,改变零售企业与消费者之间的沟通方式。

2. 阿里巴巴结构化数据搜索托管服务 OpenSearch

(1) 解决了什么问题。

OpenSearch 是满足用户结构化数据搜索需求的托管服务,支持数据结构、搜索排序、数据处理的自由定制。OpenSearch 为客户的网站或应用程序提供简单、低成本、稳定、高效的搜索解决方案。

(2) 使用了哪些 AI 技术。

OpenSearch 的团队提供一套基础算法,并在基础算法上构建了自适应的模型。团队会根据用户网站不同的需求,为用户挑选最合适的模型,同时提供产品可视化配置,用户可以独立配置自己的词表。这套系统主要的应用领域是商业数据分析。

3. 阿里云 ET 工业大脑

(1) 解决了什么问题。

ET 工业大脑通过对企业“供”“研”“产”“销”全链路的数据进行分析，提供数据智能型工具，优化从企业信息化到装备制造、生产物料管理、物流运输管理、人力资源管理、数字化设计、模拟仿真、数字化控制的各个环节，利用 AI 帮助企业降低成本、提升效率。

(2) 使用了哪些 AI 技术。

A. 智能供应链：通过对历史销量数据、订单数据、车辆数据、高德地图数据的智能挖掘，对销量进行精准预测，对库存进行分析和优化，把订单推送给最合适的物流车辆，避免无效推送，最大限度地降低库存，提升物流配送效率。

B. 智能研发：通过整合从产品研发到设计、生产以及售后所有产生的数据，建立覆盖产品全生命周期的数字档案，通过 AI 算法定位关键因素，运用大数据分析技术从多维度进行动态仿真并优化，提供产品全价值链预测服务与优化策略。

C. 智能生产：在对制造过程进行数字化描述的基础上，运用 AI 算法研究参数变化对设备状态及整个生产过程的影响，并根据实时数据与现场工况动态调优，提供设备健康动态管理、工艺参数实时推荐等功能，为管理生产现场、提升生产效率提供强大的工具。

D. 智能营销：提供“攻守结合”的智能营销产品组合，品牌卫士实时智能评估品牌形象为“守”；客户洞察提供从产业洞察到挖掘潜在客户的多层次分析为“攻”。

4. 美团的图像质量排序模型

(1) 解决了什么问题。

通过图像质量排序模型，系统自动为商家挑选最合适的优质首图并进行展示，以吸引用户点击，提高业务指标。

(2) 使用了哪些 AI 技术。

图像质量排序模型采用 SiameseNetwork 结构，通过美团云进行算法训练，用“根据图像相似度进行排序”的思路解决图像质量评价问题。

8.5 AI+新零售

世界各主要国家都已经开始主动布局 AI，运用 AI 进行赋能，零售作为一国经济增长的有力支撑，应当率先运用 AI 进行赋能。

AI 作为数字技术发展的前沿，有望为数字经济发展带来新的技术红利，成为全球经济增长的新引擎。事实上，AI 领域吸引的投资额近年来不断增加，而零售可能是对技术嗅觉最为灵敏的行业之一，随着消费升级不断加快，AI 将创造者与消费者连接在一起，为新零售破局。

8.5.1 零售行业更需要个性化

在“2018 口碑新店商峰会”上，口碑业务部总经理陈盛这样说道：以前，用户消费时经常会问哪家更便宜、哪家有团购、哪家打折，但今天很少会听到“哪家便宜去哪家”的说法。以餐饮业为例，消费者更关心的是哪家更近、哪家可以外带，消费需求已从更优惠转变到更便捷。

这样的变化也出现在旅游消费领域。笔者调研同程 CEO 吴志祥时听到了如下说法：“截至 2018 年 1 月，根据同程的数据，越来越多的消费者选择旅游专线产品，旅游专线产品的销售额占比已经从以往的 3.9%增加到现在的 26.7%。其实，旅游专线产品比普通旅游产品要贵 10%~20%，却仍有很多消费者选择。这是因为消费者变‘懒’了，也变‘聪明’了，他们不愿意做很多消耗时间的事，而更愿意多花钱得到更好的体验。”

中欧商学院原战略学教授、菜鸟网络首席战略官陈威表示：“以往零售行业都是连锁化和标准化才能赚钱，因为不知道消费者需要什么，所以需要先把产品生产出来，然后再卖出去，但这样很容易形成库存积压。未来 10 年到 20 年，做零售行业需要更多的个性化。”

8.5.2 新零售化解供需矛盾

如何满足消费者的个性化需求和商家零库存之间的矛盾？究竟有没有解决方案呢？

口碑 CEO 范驰提出，这一矛盾的解决方案就是以智能算法为基础的新零售，“基于 AI 技术，商家可以给不同消费者匹配不同的优惠，一键实现‘千人千面’的精准

营销，精准提高每位客户的到店消费金额。”

事实上，长期以来，传统零售业受到电子商务的冲击，迎来关店倒闭潮；而线下体验严重不足的电子商务又到达了发展的“天花板”。于是从2017年起，各大科技巨头相继展开了对线下商业的布局。

腾讯研究院在2018年1月发布的《从移动互联时代向AI时代演进——2018数字前沿趋势预测》显示：互联网经过纯线上时代和O2O时代，如今正进入OMO（线上线下整合）新阶段。OMO出现的原因可以归纳为几种核心能力的出现：移动支付的高渗透率将互联网的商品和服务延伸到无数场景；传感器、计算机视觉等技术的广泛应用改变了企业认知消费者的方式，提供了精确的用户画像；小程序等工具打通支付、会员、公众号、社交分享服务，实现零售业商家和客户的智慧连接，对会员、商品、供应链、库存这些内容的聚焦将更加明显；无人客服、AR试妆等AI科技让消费者获得了更好的线下体验。

新零售首先是数字化，然后是互联网化，最后是智能化。数字化是把事物量化，使得事物可以被分析。互联网化是把上下游的信息打通，让商家更加了解消费者。进入AI时代以后，商家可以比消费者本人更了解他们的需求，在这种情况下，消费者能快速做出决定，商家也能依照自己的优势，为消费者提供产品。

新零售的本质就是不抛弃实体，兼容线下但又融合线上。相较于过去连锁店强调效率和成本，新零售特别强调体验。以盒马鲜生为例，它既非超市，又非餐厅，也不是社交场所，而是把几种业态融合起来并将商品的信息、服务的信息、会员的信息等所有数据打通，在此基础上实现自动化，帮助每个环节的数据相连，实现实时调拨和匹配。

零售在经历过电商热潮、O2O模式以及“互联网+”后，“新零售”这一概念引起了广泛的关注。根据马云对新零售的定义可以发现，线上、线下和物流必须结合在一起才能诞生真正的新零售。

当马云提出“新零售”这一概念之后，各行业也掀起了对“新零售”的热议。虽然对新零售的解读不尽相同，但对于新零售的本质和目的的认识却非常统一，即“更高效、更优质地服务消费者”。只有经过AI赋能后，零售才能更高效、更优质地服务消费者。

8.5.3 AI 赋能新零售的案例

笔者总结了 AI 对新零售进行赋能的具体应用，包括消费推荐、销售助理和导购、销售预测与管理、自然语言搜索、电商网站运营、可视化搜索、信息系统管理、库存管理、尺寸和造型设计、店内视觉设计、基于位置的营销和分析、实时定价与激励政策、产品定位等。接下来笔者详细介绍 AI 为新零售解决了什么问题以及分别运用了哪些 AI 技术。

1. 阿里巴巴无人零售解决方案

(1) 解决了什么问题。

阿里巴巴无人零售解决方案采取了多技术融合的方式，主要涉及三大核心系统：生物特征自主感知和学习系统、结算意图识别和交易系统、目标检测与追踪系统。

阿里巴巴无人零售解决方案将线下门店智能化，让阿里巴巴的核心能力延展至线下消费场景，使线下零售门店能够获得与线上旗舰店同样高效的运营能力。

(2) 使用了哪些 AI 技术。

A. 在识别人和商品上，注重多模态技术环境，混合使用了计算机视觉、传感器等技术，并叠加了非配合生物特征识别技术来降低误判率。

B. 对于货架区的消费者行为、动线轨迹、停留时间等信息，由多角度布控的摄像头来完成收集。

C. 将支付环节放在门的位置，鲁棒性比较强。结算区域有两道门，相互配合以完成支付环节。采用多模块融合的方案提高识别精度。由于现实场景下多人购物的复杂性，单独使用计算机视觉或传感器都无法达到实际应用需要的精度。

而在支付门里，系统完成对商品的识别依赖于结算意图识别和交易系统，其使用的是计算机视觉技术和传感器技术。如果消费者故意破坏商品或在结算时试图作弊，那么都可以追溯到并被视为不良记录，影响其下一次入场购物。系统还会接入芝麻信用的信用评分体系，一些破坏行为会被自动分类。

2. 7-11 的 X-STORE 无人零售商店

(1) 解决了什么问题。

根据笔者调研，7-11 的 X-STORE 无人零售商店有两个闸机：入口闸机和出口闸机。消费者从入店、结账到离开，全程无需使用手机、卡片或钱包，只需“刷脸”

就能完成消费。店内设备均具备智能功能，比如冷饮区的自动感应冰箱，当感应到有顾客走近就会自动开启。货架上设有小型显示屏，用来播放产品介绍视频、实现互动功能等。

据官方介绍称，X-STORE 无人零售商店初期将以内部测试为主，并配备驻店员工帮助用户得到更好的体验。希望借此找出不足和弊端，并总结出最适合的无人零售商店运营模式，最终推出完善的方案。

(2) 使用了哪些技术。

X-STORE 无人零售商店使用了商品识别、感应式自动化、智能语音互动、人流热点分析、自动结账 POS 系统、电子标签等技术。X-STORE 无人零售商店的研发历时 4 个多月，整合了 20 家厂商，测试店面积为 22m²，以探索、体验、超越三大概念为主，陈列了 800 多种商品，开放时间为早上七点到晚上七点，测试期间仅允许内部员工进入，且进入店内必须绑定 OPENPOINT，并进行脸部生物特征识别。

3. eBay 和 Rebecca Minkoff 合作的智能试衣间

(1) 解决了什么问题。

智能试衣间通过带触摸屏的镜子及可调节的灯光，帮助顾客找到满足自己尺码、颜色和穿着场景的服装。主要希望通过智能化的手段和亲身体验来提升线下实体店的购物体验，创建线下店铺的独特竞争力。

(2) 使用了哪些技术。

消费者进入商店，通过镜子浏览店铺中的所有商品，提交试穿申请，衣服就会被导购员摆放在试衣间。顾客可以通过调整灯光亮度和颜色来模拟使用场景，镜子能感应衣服上的 RFID 标签并显示在屏幕上，然后镜子给出搭配建议。如果需要试穿其他颜色或尺码的衣服，也能通过屏幕下达指令，让导购员送来。当顾客试穿满意后，可以直接在镜子上通过 PayPal 付款。试衣间里还安装了 Kinect，可以记录、追踪试衣者的动作。

这套系统在 Nordstrom 的西雅图分店和圣何塞分店已经投入使用。

4. 美团 App 的客服页面、公众号自动回复解决方案

(1) 解决了什么问题。

通过使用机器学习、深度学习和自然语言处理技术，计算用户问题与知识库中已

存在问题的匹配度,提高直接返回唯一答案的正确率,让用户可以更直观地看到所提问题的答案,并给用户推荐相似问题,以提高智能客服的服务能力,降低人工客服的参与率。

通过配套的智能客服运营平台,可以让运营人员实时更新和发布知识库,降低业务线扩展和知识库维护时技术方的参与度。

(2) 使用了哪些 AI 技术。

计算用户问题与知识库中问题的匹配度时,使用了以下方法。

A. 用美团云深度学习框架对短文本进行分类,计算短文本分在各个问题类型上的得分。

B. 用自定义的需求解析模板,计算用户问题可能映射到某一个需求上的得分。

C. 使用多种深度学习和自然语言处理算法,计算用户问题与语料库中问题的相似度。

应用此技术的产品有美团 App、美团外卖 App、开店宝 App。此外,美团为了给商业赋能,还采用了基于自然语言处理的文本模型。例如,通过美团云的深度学习框架,对美团的海量评论数据进行计算和分析,对于评论进行多维度理解,针对不同维度分别建立分类器,以判断向用户提供的商品介绍、评论是否为商家以消费者身份“刷”出来的。

8.6 AI+机器人

之所以称为 AI 机器人,是因为它有相当发达的“大脑”。在脑中起作用的是中央处理器,这种机器人跟操作它的人产生直接的联系。最主要的是,这样的机器人可以执行按目的安排的动作。

AI 机器人给人最深刻的印象是它是一个独特的能进行自我控制的“活物”。其实,这个“活物”的主要器官并不像真正的人那样微妙而复杂。

AI 机器人具备形形色色的内部信息传感器和外部信息传感器。除具有传感器外,AI 机器人还有效应器,能使机器人的手、脚等器官动起来。由此可知,AI 机器人至少要具备三个要素:思考要素、感觉要素、反应要素。

我们称 AI 机器人为“自控机器人”,以便使它同普通的机器人区分开来。它是

控制论产生的结果，控制论主张这样的事实：生命和非生命有目的的行为在很多方面是一致的。机器人是一种系统，这种系统过去只能从细胞生长的结果中得到，但在 21 世纪已经成为人类能够制造的东西。

AI 机器人能够理解人类语言，用人类语言同操作者进行对话，在它自身的“意识”中单独形成了一种使它得以“生存”的外界环境。它能分析出现的情况，能调整自己的动作以满足操作者所提出的全部要求，能拟定所希望的动作，并在信息不充分的情况下和环境迅速变化的条件下完成这些动作。当然，要它和人类一模一样，暂时是不可能办到的。不过，AI 机器人正在迅猛发展中，有人正在建立计算机能够理解的某种“微观世界”。

笔者将 AI 机器人分为如下几类：聊天机器人、服务机器人、工业机器人、特种机器人、医疗机器人、玩具机器人等。

下面详细看一下笔者调研的几个实际案例。

1. 微软亚洲研究院的智能聊天机器人——小冰

(1) 解决了什么问题。

微软将智能聊天机器人小冰作为用对话型 AI 解决情感交流问题的一种工具，希望通过小冰的情感计算去探索更多的未知领域和需求，让整个 AI 产业拥有更大的想象空间。

小冰可以作诗、写文章、使用全双工语音打电话，并提供智能客服、智能媒体平台等商用服务。

微软曾经和日本的 LAWSON 合作，对小冰进行了一次商业化试水。小冰在和两千万 LAWSON 用户进行对话交流的过程中，能找到合适的机会给用户打提供折扣券。在传统营销模式的转化率不足 10%的情况下，用户使用小冰派送的折扣券在 4 日内到店完成消费的比率高达 57%。

(2) 使用了哪些 AI 技术。

小冰的技术核心是情感计算框架，不同于普通聊天机器人使用 RuleBase 方法来处理对话情景和上下文关系，小冰的情感计算框架中的上下文是覆盖了长尾的。其情感计算框架从大量数据中还原人世间的情感反应，再通过一套专门的排序方法，去拟合近似于人的对话，让机器人在某种程度上拥有对话的主动权。

2. 丰田集团的助残及养老护理机器人

(1) 解决了什么问题。

室内单臂机器人能照顾残障老人的起居、生活。这个机器人可以帮残障老人完成一些基本活动，如开门、把食品从食品柜拿出来等，特别适用于家庭陪伴。

(2) 使用了哪些 AI 技术。

机器人的手臂装有智能系统，向外伸缩时可识别物体；其手臂末端的夹子可以夹取铅笔、水杯等物品。机器人可以在家里自主移动，人类也可以通过智能手机控制其移动。

它有三种主要的模式：抓取、声控拾取、手动控制。

3. 波士顿动力的机械狗 Spot

(1) 解决了什么问题。

Spot 具有出色的平衡能力，且采用纯电驱动，能极为安静地行走、小跑、上楼梯，其被踢之后甚至能恢复到初始姿势。

用户使用一个连接在笔记本电脑上的游戏手柄即可控制它，操控半径可达 500 米。

(2) 使用了哪些 AI 技术。

Spot 安装了大量的传感器，包括深度摄像机、姿态传感器，以及肢体的本体感受传感器，从而帮助机器狗完成各项复杂的动作。

4. 拣货机器人 Fetch Robotics

Fetch Robotics 可利用摄像头、图像处理和导航技术等，根据订单内容，使用机械臂将位于货架某一位置上的特定商品取下来，然后交给名为 Freight 的自动货架车运回打包。在不工作时，Fetch Robotics 会自动返回充电桩进行充电。

Fetch Robotics 适用于制造业、零售业、供应链管理三个领域。

Fetch Robotics 还有大型搬运机器——Freight 500/Freight 1500。Freight 500 能够搬运 500 千克的有效载荷。Freight 1500 则设计用于搬运托盘类的标准货箱，其尺寸更大，能够搬运 1500 千克的有效载荷。Freight 1500 的前端和后端装配了激光雷达传感器，以及一个 RGBD 摄像头。

5. 哈佛大学的外骨骼装置 Soft Exosuit

Soft Exosuit 由柔软的纺织品和传感器制成,并由安装在背部的电池和发动机驱动。其绑带系统和动力系统经过精心设计,能够实时监测使用者各部位的张力、穿戴者的状态(跑步、行走、蹲伏等),并代替人体的肌肉(如大腿或者小腿的肌肉),从而帮助中风患者、肌无力患者及其他人士提升运动能力。

该装置也可使士兵具备更强的运动能力。

在研发过程中,研究者使用 Vicon T 的摄像机系统进行运动捕捉,用一个完全仪器化的分离式跑步机来测量地面反作用力。

通过比较在装置激活、装置无动力、无装置三种情况下每个关节的轮廓和运动范围,研究者可以确定该装置本身如何影响步态,以及该装置如何改变用户的运动行为。通过分析步态动力学指标和运动学动作来研究该装置在多大程度上辅助人类。通过比较关节力矩和辅助力,研究者能够监测用户和机器人之间的同步程度。

6. 哈尔滨工业大学开发的微型纳米机器人

该微型纳米机器人具备位置检测传感器,可实现自动编程运动,并具备多种功能强大的附加模块。

与传统机器人相比,微型纳米机器人具有超级灵敏、超高精确度等特点,可以在极微小的尺度下完成传统机器人无法实现的各种观测和操控制业。该机器人由外部磁场控制,能在血管内以每秒 10 微米的速率“游”动,将原子级别的药物输入细胞中。

通过这种微操控的形式对药物进行测试,能够使测试的效率大大提升,且能够让老药有新用法。

7. 小结: AI 机器人的发展方向

尽管 AI 机器人取得了显著的成绩,但控制论专家们认为它尚未达到智能水平的极限。问题不光在于计算机的运算速度不够和传感器种类偏少,还在于其他方面,如缺乏编制机器人理智行为程序的设计思想。在 2018 年,科学家连人解决最普通的问题时的思维过程都没有破译,又怎能让计算机的思维速度变得更快、更有规律呢?因此,没有认识人类自己这个问题成了机器人发展道路上的绊脚石。制造“能够生活在不固定环境中的智能机器人”这一课题使人们对发生在人类大脑中的认识和自我认识过程进行了深入研究,导致出现了“等级自适应系统”学说,这种学说正在

快速发展。作为组织智能机器人进行符合目的的行为的理论基础，我们的大脑是怎样控制我们的身体的呢？纯粹从机械学观点来粗略估算，我们的身体具有两百多个自由度。当我们进行写字、走路、跑步、游泳、弹钢琴这些复杂动作的时候，大脑究竟是怎样对每一块肌肉发号施令的呢？大脑怎样在最短的时间内处理完信息呢？研究表明，我们的大脑根本没有参与这些活动，大脑——我们的中心信息处理机“不屑于”去参与这些活动，它根本不去监督我们身体的各个运动部位，动作的详细设计是在能力比大脑皮层差得多的组织上进行的。这很像用高级语言进行程序设计，只要指出“间隔为 1 的从 1 到 20 的一组数字”，机器自己会将这组指令输入操作系统。“一接触到烫的物体就把手缩回来”这类最明显的指令甚至在大脑还没有意识到的时候就已经发出了。

把一个大任务在几个皮层之间进行分配，这比给构成系统的每个要素规定必要的动作要合算、经济、有效。在完成这样或那样一些复杂的动作时，我们通常将其分解成一系列小动作（如起来、坐下、迈右脚、迈左脚），教给小孩各种各样的动作可归结为在小孩的“存储器”中形成并巩固相应的小动作。同理，感知过程也是如此组织起来的。感性形象是听觉、视觉或触觉脉冲的固定序列或组合，或者是序列和组合二者兼而有之。学习能力是对先前并不知道且在相当广的范围内发生变化的生活环境的适应能力，这种适应能力不仅是整个机体所固有的，而且是机体的单个器官，甚至是某个功能所固有的，这种能力在同一个问题应该解决多次的情况下是不可替代的。可见，这种适应能力在整个生物界的合乎目的的行为中起着极其重要的作用。

控制机器人的关键在于模拟动物的运动能力和人的适应能力。AI 机器人具有处理大规模信息的能力，等级自适应结构的出现提高了控制机器人的质量，降低了不确定性，加快了动作的速度。为了发挥各个等级和子系统的作用，必须使信息量大大减少。因此，算法的各司其职使人们可以在不确定性大大减少的情况下完成任务。

人们根据机器人的智力水平决定其所属的机器人类别，有的人甚至依此将机器人分为以下几类。

受控机器人：第零代机器人，不具备任何智力性能，如由人来掌握、操纵的机械手。

② 可以被训练的机器人：第一代机器人，拥有存储器，由人操作，动作的计划和程序由人指定，它只是记住和再现出来。

③ 感觉机器人：机器人记住人安排的计划后，再依据外界这样或那样的数据（反馈）算出动作的具体程序。

④ 智能机器人：人指定目标后，机器人独自编制操作计划，依据实际情况确定动作程序，然后把动作变为操作机构的运动。它有感觉系统和智能模拟装置，能识别周围及自身情况，具备意识和自我意识。

8.7 AI+安防

在安防领域，AI 正在全面接管以视频业务为主要应用的社会治安防控体系和行业智能可视化管理两大领域。从后端延伸出智能，从被动逐渐转为主动，从人工发展到自主判断，AI 对安防行业的渗透和影响正在促进传统安防业务模式的转型，加速技术的融合和创新发展。

8.7.1 安防市场的发展特点

在安防产业智能化的演进过程中，首当其冲的是集成领域，很多没有 AI 技术储备和交付能力的集成商最终倒在了安防 AI 化的路上。安防市场的发展具有以下特点。

（1）政策扶植叠加市场需求，双轮驱动安防市场的快速发展。

自从国家安全委员会成立以来，国家安全的重要性被提升到一个新的高度，九部委全力推进公共安全视频监控的全域覆盖和全网共享，“十三五”相关发展规划也将安防作为重点领域，充分显示出国家对安防的高度重视。我国智慧城市建设如火如荼，平安城市、智能交通等领域的需求成为安防行业的重要推动力。在政策扶植和旺盛的市场需求的共同驱动下，我国安防市场有望持续扩大。

（2）AI 技术日渐成熟，智能安防或成第一“着陆场”。

随着高清监控的普及，安防监控领域的数据量呈爆炸式增长，催生了市场对智能化技术的需求。芯片、算法和数据支撑 AI 技术在安防行业的应用，GPU、FPGA、TPU 等智能芯片大大提高了运算效率，深度学习算法增强了图像和视频分析的准确率，使得系统在解决视频结构化和人脸识别等问题方面更智能，安防市场天然的优质数据源为 AI 落地提供了沃土，智能安防或成 AI 产业化的第一“着陆场”，即将开启新一轮成长周期。

（3）智能化帮助安防产品量价齐升，智能安防的市场规模远超千亿元。

智能化对提升安防设备市场规模的作用表现在两个方面：从量的角度来看，随着平安城市、智能交通的建设，安防设备的部署密度将会增加，而且智能化设备对已部署设备的替换也将带来新的增量；从价的角度来看，智能化前端和后端设备的逐步建设将带来设备单体价值的提升。假设到 2020 年，智能化前端和后端设备的渗透率为 6%，那么，预计我国安防硬件的市场空间可达 1650 亿元。此外，深度学习等 AI 技术有望使得安防产业链中的软件、集成及运营等环节取得突破性进展，带动安防产业规模地扩张，安防智能化的市场空间广阔。

国内安防企业众多，智能安防赛道上群雄逐鹿，市场集中度不断提升。全球安全形势恶化，海外市场迎来发展良机，海康威视、大华股份等业界巨头进军国际市场，并逐步被海外市场认可。可以预见，随着全球行业洗牌加速，国内外市场份额将进一步向龙头企业集中。

运用 AI 赋能安防的企业有海康威视、大华股份、苏州科达、千方科技、东方网力、熙菱信息、汉邦高科、佳都科技等。

8.7.2 安防企业进行规划建设的典型实例——佳都科技

（1）战略一：投资云从科技，布局 AI。

从 2016 年起，当整个产业都在聚焦安防 IT 化战略时，佳都科技选择战略投资云从科技。从安防的智能化之路来看，不得不说当初这个决定实在高明。云从科技是一家专注于计算机视觉与 AI 技术的高科技企业，核心技术源于黄煦涛教授。作为起草与制定人脸识别国家标准的企业之一，云从科技拥有视频认知相关的人脸识别、行为分析等核心技术。

从 2017 年起，佳都科技开始提出“视频云+大数据应用”整体解决方案。战略投资千视通，佳都科技的战略意图很明确，就是重点布局机器视觉和智能大数据技术。佳都科技已经形成了以视频云+、智能交通管理平台、视频联网平台、视频图像侦查平台、社会网格化治理平台等一系列安防产品和解决方案，形成一套完整的产品体系，并以“感知互联、数据驱动、智能应用”为技术理念，促进多种信息技术的融合和深度集成。

佳都科技在人脸识别、车辆识别、特征识别等机器视觉相关技术、产品和解决方

案上持续进行战略投入，并形成了完整的可视化技术体系，支撑其各条产品线和解决方案在 AI 方面的迭代升级。

（2）战略二：聚焦高成长行业。

中国已经建成世界上最大的视频监控网——中国天网，摄像机总数超过 2000 万台，并利用先进的 AI 和大数据技术进行社会治安防控体系和行业智能可视化管理的建设。在安防众多的应用领域，智能轨道交通和智慧城市是佳都科技 AI 战略聚焦的两大高成长行业。

对 AI 技术而言，其生命力不在于技术有多么重大的突破，而在于在实际场景中解决用户的需求，并得到社会和市场的广泛认可。依托提前布局 AI 所掌握的技术，并结合垂直行业的应用特征，佳都科技已经成功将 AI 应用于轨道交通智能化领域。佳都科技拥有自动售检票系统、站台门系统、综合监控系统、通信系统（含视频监控子系统）四大核心智能化系统解决方案，产品体系已经全面覆盖地铁、有轨电车、快速公交等细分市场，并在广州、青岛、天津、厦门、武汉、宁波等全国 17 座重点城市建立了标杆示范项目。

在智慧城市方面，佳都科技将六大体系、三大阶段作为建设理念，以视频监控业务为阵地，以社会治安防控体系和行业智能可视化建设为重点，向智慧城市这个战略高地聚焦“火力”，全面覆盖雪亮工程、天网工程、平安城市、群众性治安防控工程等业务领域，并取得骄人战绩。

（3）战略三：突破界限，构建 AI 生态。

佳都科技凭借技术积淀和 AI 实战经验，以及在实战中对技术和业务界限的清晰把握，围绕 AI 生态的建设，拓展传统安防业务的范畴。佳都科技在 2018 年明显加快了实现这个战略规划的步伐。

首先，在合作方面，佳都科技与云从科技、千视通、方纬科技等企业展开深度合作，继续完善人脸识别、视频结构化、交通大数据分析等核心技术，并推动这些智能技术在交通、金融、安防等领域的落地。

其次，在 AI 产业基金设立方面，佳都科技动作频频。先是联合粤科集团设立投资基金，聚焦 AI 核心技术和产品，以智慧城市项目、先进集成电路制造和设计项目等为投资方向。接着又携手广州基金和创钰投资成立 AI 产业并购基金，通过技术整合及产业并购的方式进一步加强对 AI 和大数据等智能化技术的探索与创新，实现内

生增长及外延式扩张的双层布局,加快 AI 产业生态圈的建设。

跨媒体感知技术是 AI 在传媒行业的创新应用。佳都科技已经与羊城晚报报业集团达成战略合作,双方将充分发挥各自的资源优势,开创“AI+媒体”的新模式,加强双方的产品在智能新闻、智慧发布、智能广告、跨媒体感知等方面的深度融合,引领智慧生活。

2018 年,一台后端服务器就可以并发处理包含 2 亿张人脸的数据库,城市大脑可以利用一个城市各个十字路口的摄像机,为救护车规划用时最短的通行路线……

8.7.3 AI 赋能安防的实战化应用

2018 年,AI 赋能安防的实战化应用在网络安全、身份认证、视频监控、犯罪预测、刑侦排查、工业安全、警用治安、家庭安全、空域安全、案件侦破、灾难预测、隐私保护等领域实现了落地。笔者从中挑选了几个案例进行具体介绍。

1. 阿里云的智能护航方案

(1) 解决了什么问题。

智能护航方案为金砖国家领导人会晤期间的网络安全提供护航,涵盖事前检测、事中防护、事后复盘整个护航生命周期。通过这套方案,安全能力可以像水、电一样让政府、事业单位的网站快速远程接入,并具有弹性扩容能力,极大地提升了保障工作的便利性和响应速度。会晤期间,阿里云启用 AI 云博士,24 小时守职,在线解答客户安全问题,时刻向有安全隐患的客户通报隐患,确保通知客户环节更简单、更快捷、更高效,从而让更多的安全专家可以集中精力解决难题。云博士能主动帮助客户发现服务欠费、数据被清洗、CPU 跑满、实例被释放等高危情况,向客户提前预警。

(2) 使用了哪些 AI 技术。

阿里云具备国内最丰富的攻防对抗数据和样本,从而令阿里云能充分利用自身保障网络安全的经验,凭借强大的机器学习能力和计算能力,使得智能护航方案能够及时从海量的安全数据中抓取到攻击线索、漏洞信息和威胁情报等高价值信息并进行分析预警,遇到问题第一时间联合相关部门进行应急响应。会晤启动前两个月,经相关部门授权,阿里云专家利用智能护航方案对云上政府类网站做了一次全方位“体检”,以找出网站可能存在的漏洞,并提出优化、容灾建议,做到不留安全死角。

2. 苏州科达的交通管控平台

(1) 解决了什么问题。

烟台高速交警支队成立之初，信息化基础比较薄弱，而烟台高速公路的路况相比其他区域更为复杂，且由于编制原因，警力资源与辖区实际管控范围不成正比。因此，在 2015 年，烟台高速交警支队通过招投标方式，请苏州科达做了高速交警支队未来三年的建设规划，欲以智能化为主轴解决由于人员配置不足与信息化水平偏低而带来的繁重的交通治理任务。该项目包含烟台市内的所有高速公路，以及荣乌高速烟台段、沈海高速烟台段、烟海高速烟台段、威青高速烟台段等 7 条主要高速公路。在这些高速公路上，一期建设了 108 个视频球机、70 个全景枪机、40 个主线卡口、43 个出入卡口、105 个违停球机、34 个应急车道抓拍机、17 个执法站前端智能设备等装置。

另外，该系统融合了交通、公安、社保、卫生等多个应用场景，使用前端系统作为数据支撑，融合高清视频监控系统、卡口信息管理系统、路况信息管理系统、违法信息管理系统、车辆速度诱导控制系统、指挥调度集成系统、重点车辆动态监管系统和社会化服务系统等应用和服务支撑系统。在前端、后端智能化系统结合部署的模式下，项目建成后可以确保在辖区范围内，无论昼夜，都可以对道路交通进行实时监控，实现信息采集、卡口抓拍、信息诱导、区间测速、缉查布控等一系列智能化应用。深度学习技术能够更精确地帮助识别车辆的各种特征，包括车牌、车型、车体等。此外，该系统中配备的相关模块能对进出烟台高速公路的车辆进行分析，实时了解高速公路几百公里范围内的车辆从哪里出发、去往哪里，为相关管理部门根据实况有效地调度警力提供准确依据。整套产品可以有效降低交通事故率，从烟台高速交警支队返回的数据来看，整套系统投入使用后，每个月事故率相比前期平均水平下降了 40% 左右。

其优势包括：A. 采用深度学习等技术，系统适用性强，在识别车辆特征及违法行为时准确率有很大提升。B. 前端产品采用了一体化设计，在一台设备中集成了抓拍、识别、处理、存储、上传等功能，不需要再配主机设备，安装部署非常方便，在低成本的前提下，稳定性却有所提升。通过定制护罩实现了更高等级的防护，能够适应高盐碱环境。

(2) 使用了哪些 AI 技术。

该项目的技术架构采用视频专网及公安网“双网、双平台”架构。据悉，在此之

前,烟台高速交警支队已经部署了基础版的交通管控平台,可实现部分基础应用。苏州科达基于该版本,定制研发了“双网、双平台”架构,实现了烟台高速交警支队视频专网与公安网内各业务支撑系统的互联。智能交通安全管理系统由应用支撑系统、指挥调度与服务系统两大部分组成,应用支撑系统采集前端数据,为指挥调度与应急事件指挥、重点车辆监管等服务提供数据支撑。当发生重大交通事故时,智能交通安全管理平台可以与接警和处警系统对接,实现高速公路交通事故类接警信息与数据共享。与此同时,指挥中心也可通过平台获取的信息对突发事件进行决策指挥与调度处理。

此外,苏州科达还为烟台高速交警支队建设了视频会议系统,支队可以与烟台市公安局进行可视会议、协同办公,与莱山大队、福山大队进行可视会议、远程办公及可视指挥调度。在日常工作中,视频会议系统可作为多方工作交流的媒介;在遇到突发事件时,配合智能交通管控平台,支队可在第一时间获取现场的相关信息,及时与各大队召开视频会议,分享录像资料,沟通、协调多方资源,开展扁平化指挥调度。

在管控违法行为方面,系统可实现如下功能。

A. 监控重要路段,比如对容易超车路段的车辆速度做严格控制。

B. 烟台盛产水果,高速公路上有时会出现果农兜售水果,这类活动危险系数大,系统部署后可以做到及时发现、及时处置。

C. 货运车辆超载往往会导致重大交通事故。系统可以发现超载车辆的规律,及时查处、打击超载行为。

D. 对烟台高速境内的重点车辆(如危化品车、大卡车、大客车等)进行自动识别并监控。系统可以自动识别车辆并分类,提供查处危化品车辆违规通行、在恶劣天气对车辆进行分流等管理服务。

3. 智能门铃 Nest Hello

(1) 解决了什么问题。

应用了 AI 的摄像头,在访客还未按门铃之前,门铃就可以识别走过来的人。如果识别出访客是邮递员,它可以通过应用程序通知用户。用户可以通过门铃上的扩音器和传声筒说话。如果此时用户正在忙其他的事情,不方便说话,它可以播放提前录好的音频,如“请将包裹放在门外”。

(2) 使用了哪些 AI 技术。

Nest Hello 装备了扩音器、传声筒及高质量的深度摄像头。

4. 照片反向识别技术

D-ID 是一家创业公司，2017 年加入了著名孵化器 Y-Combinator。

D-ID 的研究者试图通过其算法减少人脸识别技术带来的隐私泄露危险。其算法能对用户发出的照片进行反向识别，使得照片在能用肉眼看清的情况下无法被机器识别，从而保护用户隐私。

其主要服务对象分为三类：需要保护员工、用户图片隐私的公司；需要保护病人隐私的健康管理公司；政府及相关部门。

8.7.4 AI 赋能安防的四大方向

前端化：随着芯片的集成度越来越高，处理能力越来越强，许多厂商推出了智能 IPC、智能 DVR 和智能 NVR，将一些简单、通用的智能应用移植到前端设备中。未来将有更多、更复杂、专用的智能应用在前端设备中出现，在前端设备上实现的优势在于组网灵活、低延时、低成本，能减轻后端的压力，为大规模部署提供了可能。

云端化：已有的智能化产品大多是将多种智能应用固化在某一类硬件中，每台硬件设备提供一种或有限的几种智能化服务。未来，硬件资源的概念将逐步淡化，智能化将以服务模块的方式提供给客户。云端会根据客户的需要（功能、路数等）提供服务，实现资源按需分配，最大程度地满足客户需求，提高资源利用率。

平台化：每个安防厂商在推进自己的智能化解决方案时，需要对软件平台及其配套的硬件设备进行整合，这个整合方案的兼容性、稳定性、安全性等标准越来越趋于统一。尽管国内主流的监控厂商基本都具备自己的软件平台，但是这些平台的定位和规模大小都不太一样。未来几年，安防监控的应用类型将越来越清晰，其技术标准、开发接口等将越来越趋于统一。大厂商制定标准，小厂商兼容标准的合理产业模式将逐渐形成。所以，推出自己有主导力的解决方案是安防企业发展过程中必须要考虑的问题。

行业化：智能化解决的是行业客户在业务应用中存在的问题，因此，智能化需要向行业化方向进一步深化。首先，智能化厂家要从行业出发，定位目标行业和细

分市场,确定自己的发展方向。其次,在具体行业中深入理解业务应用、业务流程等,剖析行业问题,寻找解决之道。最后,结合自身的技术积累,为行业客户提供优质的智能化解决方案,未来的智能化将是行业的智能化。

8.8 AI+农业

完整记录一头生猪一生的运动轨迹,没跑完 200 千米不许出栏;实时识别小猪的进食情况和叫声情况,以判断其健康程度……听起来很“魔幻”的 AI 养猪方法,已经在我国部分养殖公司试用,这些都来源于 AI 在农业领域的广泛应用。

2018 年 6 月,阿里云在上海发布了“ET 农业大脑”,希望将 AI 与农业深入结合。除了生猪养殖,这一系统还应用于苹果及甜瓜的种植,整套系统具备了数字档案生成、全生命周期管理、智能农事分析、全链路溯源等功能,未来还将在石榴、生菜种植中落地。

当“绿色革命之父”——诺曼·博洛格在 1970 年获得诺贝尔奖时,诺贝尔奖委员会表示:“20 世纪 70 年代出生的人比其他任何时间段出生的人都要多,而他给这个饥饿的世界提供面包。”博洛格帮助引入抗病、高产的作物品种和先进的农业技术,并成为游戏规则的改变者,促使农业产量大幅增加,帮助数百万人免于饥饿。

在博洛格获得诺贝尔奖半个世纪以后,我们生活在农产品产量稳定增长,但耕地总量却在持续减少的世界。气候模式的变化和水资源的缺乏正在改变某些地区的农业生产能力。与此同时,世界人口继续增长,预计到 2050 年,全球人口将达到 90 亿人,其中大部分增长集中在发展中国家。在这些国家,快速的经济扩张使人们对卡路里和蛋白质的需求不断增加。随着人口总量的增长和个人对粮食需求的增长,我们正在接近所谓的“马尔萨斯陷阱”,在这个假设中,人口增长的速度超过了粮食供应增长的速度,从而导致大规模饥荒。避免出现这个“末日场景”很可能是 21 世纪最重要的挑战之一。

消费增加推动了对农业生产增长的需求。全球各地的种植者都在应对这样的挑战,但他们必须以一种不会对地球资源造成无法弥补的伤害的方式来应对这一挑战。为了平衡增加产量和节约资源的目标,研究人员和企业家正在研究如何在现有的碳足迹上发展可持续性增强型农业。就像博洛格那样,这些研究人员和企业家们可以借助植物

遗传学、化学、农学等学科的知识，以及机械等工具。

今天，他们还有全新的工具——AI 技术。

8.8.1 农业正在迅速实现数字化

尽管农业在行业数字化程度的调查中排名接近垫底，但其正在迅速实现数字化。高速种植设备能提供准确的种植信息，产量监测设备能提供关于收获物的各类信息，这类基础数据（投入与产出）是构建预测算法的关键。农民正在使用传感器和取样技术收集土壤水分和养分水平的数据。各种各样的农业信息管理系统使人们获取和操作数据变得更容易。农民可以使用智能工具来辅助实地考察，从移动应用到无人机，这些工具收集到的数据可以用来评估作物的健康状况，并监测不同季节的病虫情况。

随着这种转变，农业数据变得更加丰富和实用。这些数据正在为开发和部署农业 AI 铺平道路。AI 的应用主要是由科技行业推动的，农业也不例外。谷歌和 IEEE 正在举办名为 iNaturalist Competition 的竞赛，希望训练 AI 算法以识别 5000 多种不同种类的植物和动物。

8.8.2 农业 AI 技术迅猛发展

虽然 AI 已经成为技术社区的重要话题，但许多大公司、设备制造商及服务提供商还没有大力发展农业领域的 AI 应用。部分原因可能是缺乏对 AI 技术的认识，对 AI 技术的潜在应用不够熟悉。此外，AI 算法在农业领域的发展也具有挑战性。AI 应用需要大量数据来对算法进行恰当的训练，而在农业中，虽然有大量的空间数据，但大部分数据只能在每年的生长季节使用一次。因此，可能需要数年时间才能收集到对某个特定品种或某个特定农场具有统计意义的数据。通常，在农业领域收集的数据需要大量的预处理，然后才能被放心地用于训练 AI 算法。

截至 2018 年，与数据相关的挑战仍然存在。《华尔街日报》曾经撰文指出，在农场里，手机信号不够稳定或根本不存在，因此，很难将数据转移到可以被分析的地方。缺乏数据所有权和使用权的标准，以及收集和共享数据的困难，都导致农业领域的 AI 算法开发人员无法找到合适的数据库。幸运的是，像 Climate Corporation 公司的 Field View Drive、John Deere 公司的 JDLink，以及 Farmobile 公司的 PUC 这样的产品，正在使数据的收集和传输变得更加容易和无缝。

许多初创公司正在构建决策自动化工具,但在数据收集、数据准备和基准测试等方面仍然存在很大问题。在历史上,农场始终缺乏信息技术基础设施和数据仓库系统,而科技公司一直以来正是依赖这些技术来开发和部署 AI 应用的。在农业 AI 成功大规模部署之前,农场中的数据基础设施需要变得更加强大。

此外,有些新兴公司倾向于使用经过科学验证的、受统计规律控制的实地试验来量化其产品的收益,采用精益方法,即先在少量客户中进行推广,然后逐步改进。此举遵循了创办科技企业的策略,但在农业领域,种植者不会冒险在整个农场采用全新技术,这种方法很可能行不通。

在推出产品之前,主要的农业公司需要通过多年的实地测验来确保产品的性能和明显的效益,许多种植者希望看到新产品在自己的土地上能够有上佳的表现。因此,快速进入市场和快速扩张的心态可能不合适农业领域的创业,企业需要采用更加稳重的产品发布策略。

8.8.3 AI 技术在农业领域应用的前景

尽管在农业领域应用 AI 技术面临诸多挑战,但我们有理由相信,AI 技术在农业领域的成功和大规模推广将会成为现实。笔者整理了部分有望改变农业的 AI 企业。

Abundant Robotics 是从斯坦福国际研究院剥离出来的初创企业,它已经开发出自主采摘水果的技术。Abundant Robotics 利用机器视觉技术来探测棚架上生长的水果的位置,然后利用真空系统将其从树枝上摘下来,其已经获得 Google Ventures 的投资。

Resson 是孟山都公司扶持的初创企业,在加拿大和美国圣何塞都有办事处。Resson 已经开发出一种图像识别算法,与训练有素的人类相比,它能够更准确地检测和分类植物害虫与疾病。Resson 与麦凯恩食品公司合作,帮助其减少土豆在供应链上的损失。

AgVoice 是总部位于美国佐治亚州的创业公司,AgVoice 正在为作物观察专家和农学家开发自然语言处理工具包。这套系统可以发现导致大豆突然死亡的真菌疾病,并提示观察的位置和疾病的严重程度。

除此之外,Orbital Insights、Descartes Labs、Gro Intelligence 及 Tellus Labs 等创业公司也正在基于卫星图像、天气信息和历史产量等数据开发产量预测算法。Tellus Labs 声称,它们的数据比美国农业部的更准确,而且可提前美国农业部 1 个月给出预测数据。

有些人批评应用 AI 对农业环境的要求过于严苛，在某种程度上，这是正确的，但是随着计算能力的进步，这一情况正在改善。SlantRange 的总部位于美国旧金山，正在开发机器视觉系统来测量作物的数量，并检测杂草。该公司的植物计数算法最初是针对美国中西部种植区开发的，因此，在南非的试验中表现不太好，因为南非的种植密度较低，土壤反射性更强。然而一夜之间，SlantRange 团队就用新的数据重新训练了他们的算法。升级版软件在南非部署了仅仅两天，就报告了其发现的问题。

在农业中成功使用 AI 的最佳例子应该是孟山都公司扶持的 Blue River Technology (BRT)。BRT 是一家总部位于美国加利福尼亚州的创业公司，由斯坦福大学的两名研究生乔治·赫劳德和李·赖登于 2011 年成立。其中乔治·赫劳德是已经取得许多成就的农业企业高管，而李·赖登则拥有深厚的 AI 和计算机视觉技术背景。最初，BRT 专注于用机器人来帮助减少莠苣密度，这个过程此前主要通过手工完成。现在，这家公司正在应用 See & Spray 系统来消除棉花田中的杂草。通过高度精确和有针对性的喷雾，See & Spray 系统可以减少 90% 的除草剂量，因为 See & Spray 系统利用 AI 来分析高分辨率的图像，能检测出杂草的存在及其准确位置。

8.8.4 AI 技术不仅仅是为了精确农业

AI 技术在农业领域的应用固然很重要，但利用 AI 技术去发现和开发新的、更高效的农业项目同样重要。然而，直到 2018 年，AI 技术还没有对化学和生物系统的数据进行充分利用。因此，在植物育种、生物技术、农业化学及供应链管理方面还存在很多利用 AI 技术的机会。事实上，与精确农业相比，AI 技术在育种、化肥及作物保护中的应用可能更为迅速，这主要有两个原因：第一，在过去十年里，开发人员对收集和存储数据非常谨慎，包括大豆品种的序列信息、合成物的结构活性、环境的生物降解能力等数据。第二，提高效率或加速农业产品研发的成本相当高。

根据 Phillips McDougall 的分析，将一种新的作物保护产品推向市场需要 11 年的发现和发展时间，分析 16 万种化合物，将每个产品商业化所需的支出超过 2.8 亿美元。在新的农业化学产品的开发过程中，整个行业每年的花费超过 26 亿美元。而 AI 的采用可以提高这一过程的效率。例如，致力于利用 AI 开发治疗疾病新疗法的初创企业 Atomwise 正与孟山都展开研究合作，以提高发现新作物的速度和可能性。这种合作基于模式识别技术，可以减小在早期化学发现过程中出错的概率。

AI 的优点也适用于植物育种。玉米育种通常被比作“从草堆中捞针”，因为玉米

是一个拥有 32000 个基因的大“草堆”，几代育种者都面临着搜索难题。从历史上来看，一个育种项目每年可以从成千上万个可用选项中选出大约 500 种组合进行试验。这种选择受到后勤和成本的限制。为了减少这些限制，孟山都的 AI 研究人员开发了一种算法，能够评估育种决策，并预测哪一个杂交品种将在试验的第一年表现出最佳的性能。这个算法正在被分子标记信息和现场试验信息所训练。孟山都全球育种主管迈克·格雷厄姆表示，这种算法可以优化育种过程，使育种者能够更快地把他们最好的想法投入到大规模的实地试验中。这一算法不仅加快了育种过程，而且与传统方法相比，还能将玉米育种的规模提高 5 倍。育种者可以利用 AI 工具完成更多的工作。

与此类似，先正达宣布与 AI for Good 基金会合作，以 AI 为基础工具进行育种，并提高现有作物生产方法的效率。先正达为 AI 研究人员提供包括种子遗传信息、土壤信息、天气信息及气候信息在内的数据集，其目标是开发出一种算法，来确定在哪些地区种植哪种作物。

总部位于美国圣路易斯的初创企业 Benson Hill Biosystems 也在将 AI 技术应用于植物育种和生物领域。他们利用自己独有的 CropOS 平台，寻找候选基因来提高作物光合作用的效率。CropOS 平台利用不同来源的数据，如 DNA 和 RNA 序列信息、现场试验观察的数据等，来预测获得特定表型反应所需的基因表达模式。每当获得新的数据集时，CropOS 平台就能重新校准并提高其预测能力。

8.8.5 AI 对农民的影响

在过去 60 年里，美国农民的数量急剧减少，剩余的农民则在育种、农作物保护、自动化，以及提高农业生产率等方面使用更先进的技术。无论是从近期还是从长远来看，农业领域的 AI 需要农民的积极参与，而农民也可以从这些新兴技术中获益。AI 将成为非常强大的工具，帮助组织应对现代农业中日益增长的复杂性。农民不仅从 AI 直接应用于农场中受益，而且还将从应用 AI 改善品种、保护作物以及培育新产品中受益。

8.8.6 AI 赋能农业的具体应用

1. 精准耕种：微眼科技的无人直升机飞控系统

(1) 解决了什么问题。

农药的最佳喷洒时间是晚上，但因无人机操控人员（简称“飞手”）看不清飞机，

飞机存在“飞丢”的风险，造成农业植保飞机无法在晚上作业。微眼科技自主研发的无人机飞控系统，使无人机具备了自主飞行能力，摆脱了对飞手的依赖，并具备抗5~6级大风进行作业的能力。据调研资料显示，该无人机飞控系统不但能使无人机保持一定高度的自主飞行，还可让无人机在空中与互联网自主交互。而被控无人机的自主飞行时速可达80千米，最远飞行里程约5千米。通过该无人机飞控系统，用户能充分发挥微眼无人机集群控制的优势，降低人工成本。同时，通过发挥无人机互联功能，用户能用无人机搭载多种传感器，组建空中物联网，从而进行土地调查与其他数据采集活动。除此之外，该无人机飞控系统在协助工业无人机进行地理测绘、抢险救灾、治安反恐、野生动物保护等诸多方面均有较大潜力。

（2）使用了哪些 AI 技术。

微眼科技攻克了单旋翼无人机 AI 控制系统和智能地面控制站等核心技术，研发出无人机飞控系统，能让无人机实现自主飞行。

2. 农场管理 SaaS 平台——佳格耘境

由于我国耕地分散，佳格耘境使用卫星遥感、计算机视觉技术来收集、处理、分析、存储数据，通过深度学习算法来识别耕地地块的地垄方向、植物间距、植物稀疏程度等要素。结合植动态、过去数十年的气象信息，计算每一块地的光合速率、蒸腾速率（土壤水分含量）等，最终推算出植物生长走势，从而指导种植。

用户可通过个人电脑、智能手机等设备，随时随地了解和预估天气变化、农作物的生长情况，从而及时进行农事安排、农机调配、农药喷洒等活动，在降低生产成本的同时，大幅增产增收。例如，一些大农场拥有几百亩地，但很少是连成一片的，用户可以利用佳格耘境进行资产管理，远程查看自家土地，进而对土地进行分级。

此外，佳格耘境通过气象信息预测天气以帮助农民判断是否需要灌溉，避免水资源浪费；根据一块地过去数十年的产量、灾害情况来判断它适合种什么品类的作物，进而评价土地资产的价值（已经有土地流转服务商利用佳格耘境做土地资产评估）；监测到虫害后，通过软件远程遥控植保无人机去撒药。

3. 牛的智能尾环 Moocall

该尾环可佩戴在任何种类、任何尺寸的牛的尾巴上，当牛要产崽时，农民能接收到短信提醒。尾环能很好地贴合牛尾，内置的齿状物能够在不干扰牛尾正常行为的情

况下保持尾环不跌落。

该设备能够在 190 个国家的移动网络中运行。养殖人员可以通过设备上的 LED 灯发出的信号对设备的状态进行甄别,蓝光表示设备正在正常运作,红光表示牛正在生小牛,琥珀色的光代表尾环电力不足。该设备的充电电池可以持续工作 30 天。

4. 阿里云的 AI 养猪技术

(1) 解决了什么问题。

养猪乍听起来是一个跟高科技毫不沾边的行业,甚至有些低端,但事实并非如此。养好一头猪或许并不难,但大规模养猪是一件非常复杂的事情,光是统计新生小猪的数量就已经非常困难了,更何况还要时刻监测猪场的环境和猪群的健康状况等。而且在猪场工作非常挑战人的忍耐力,工作人员需要能耐得住寂寞。因为考虑到防疫,从事繁育管理工作的工作人员平时不能离开猪场,重新进入猪场需要在隔离区待上 3 天。

由于普通的 RFID 耳标扫描距离短,感应效率低,不能快速完成区域性数据的读取,同时成本较高,比如每头猪要佩戴两个耳标,以保证数据读取的连续性。因此,每出栏一百万头猪,需要耗用一千万的 RFID 耳标,增加了养殖成本,减少了利润。如果采用有源 RFID 耳标,成本还将大幅增加,同时由于不便于挂在猪的耳朵上,推广的难度较大,失去了普遍应用的价值。

要建立生猪的可追溯系统,需要对猪的喂养过程、疾病治疗过程等进行全程跟踪记录,这些基础数据以往都是通过人工填写纸制表单,或是让工作人员手持相应设备手工输入到 App 中。所有的数据整理、数据分析工作,都只有在生产过程操作完成之后才能进行。如何快速、准确地收集生猪养殖过程中的日常数据,建立完整的生猪资产管理体系和业务可追溯体系,是养殖场需要解决的问题。

引入 AI 等先进技术对于提升养猪效率具有非常重大的意义。阿里云和特驱集团合作的首期落地业务包含猪群数量识别、猪群行为特征分析、疾病识别和预警、无人过磅等内容。这些以往需要依靠人力完成的工作,如今都可以由视频图像分析、人脸识别、语音识别、物流算法等技术来完成。比如,视频图像分析技术可以记录猪的体重、进食情况、运动强度、运动轨迹。如果一头猪长卧不起,AI 系统就会及时发出警报提醒人工介入,判断它是生病或怀孕了,还是在“偷懒”?

AI 系统还可以通过红外测温技术和语音识别技术,监测猪的体温和咳嗽的声音,

随时关注猪的身体健康。AI 系统还可以让每头母猪每年平均多产 3 头小猪仔，且将猪仔淘汰率降低 3%。

（2）使用了哪些 AI 技术。

AI 养猪基于机器视觉等技术，对每一头猪都建立一套对应的数字档案，里面记载猪的品种、日龄、体重、进食情况、运动情况等信息。通过记录母猪的生产数据，ET 农业大脑能预测母猪年生产力，即每头母猪每年提供活的断奶仔猪的头数。同时对生产力下降的母猪提前给出淘汰意见，一头母猪以前一年能产 20 头左右的健康幼仔，阿里云的目标是将这一数值提升到 32 头，由此保障养猪场的效益。

相比传统的 RFID 耳标及配套的设备，摄像头更容易安装，实际上，养猪场本身也会安装监控系统，在这个基础上进行数据采集和分析，不仅成本会下降，同时大量数据采集业务也无须再让工作人员动手，解决了数据采集必须依赖一线工作人员的难题。通过视频的实时传输，数据采集的及时性和准确性也会得到保证。

沿用传统的身份识别方法，在猪身上进行标记以建立唯一识别码，再结合视频分析等技术，可以提升识别的准确率。

从猪的配种开始，就可以通过 ET 农业大脑对母猪的行为进行分析，获知母猪的发情状况及交配时间，从而在恰当的时间安排配种，正确的交配时间和交配方法能够提升母猪的受胎率及产仔数。

除了配种和生产，在生猪养殖过程中，还需要随时掌握猪的体重情况，进而调整喂养计划。通过视频分析结合猪的品种、喂养标准和增重标准，可以对猪的体重进行测算，而无须使用传统的称重方法。

对养猪场来说，最关键的还是疾病防治和疫情监测。ET 农业大脑通过对猪的行为进行分析，结合自然语言处理技术和红外测温技术，能判断猪的健康状态，有针对性地进行疫情预警。

ET 农业大脑还能用视频图像分析技术检测猪的体态、分析猪的健康，并在这个过程中建立标准，进而确立猪每天睡多久，运动强度和频次怎样，以保证料肉比。

在后期配送阶段，可以按照将猪运送到客户那里的时间、路程，选择合理的物流方式。

8.9 AI+交通

2017年,国务院发布《新一代人工智能发展规划》(后文简称“规划”),提出推动人工智能与各行业的融合创新,智能运载工具、智能物流等交通运输领域的多项内容位列其中。

在智能运载工具方面,“规划”明确提出发展自动驾驶汽车和轨道交通系统,加强车载感知、自动驾驶、车联网、物联网等技术的集成和配套,开发智能感知系统,形成我国自主的自动驾驶平台技术体系和产品总成能力,探索自动驾驶汽车共享模式。发展消费类和商用类无人机、无人船,建立试验鉴定、测试、竞技等专业化服务体系,完善空域、水域管理措施。

在智能物流方面,“规划”明确提出加强智能化装卸搬运、分拣包装、加工配送等智能物流装备研发和推广应用,建设深度感知智能仓储系统,提升仓储运营管理水平 and 效率。完善智能物流公共信息平台 and 指挥系统、产品质量认证及追溯系统、智能配货调度体系等。

根据“规划”,国家将促进智能交通的发展,研究建立运营车辆自动驾驶与车路协同技术体系。研发复杂场景下的多维交通信息的综合大数据应用平台,实现智能化交通疏导和综合运行协调指挥,建成覆盖地面、低空和海上的智能交通监控、管理和服务系统。

下面笔者具体讲解在交通领域的具体应用中,AI解决了什么问题 and 使用了哪些AI技术。

1. 英伟达公司的道路数据采集技术和卷积神经网络架构

(1) 解决了什么问题。

该端到端学习系统希望为自动驾驶汽车提供策略。英伟达公司首先在收集数据车辆的挡风玻璃后安装三个摄像机,捕获道路的图像及驾驶员的转向角度等信息,然后利用收集到的数据训练卷积神经网络,使得卷积神经网络能识别路面特征。在卷积神经网络训练完成后,该系统能够分析实时图像并向无人车发送转向指令。当仿真效果较为理想时,公司使用测试车携带该系统进行道路测试,为自动驾驶提供技术积累。

(2) 使用了哪些AI技术。

该系统的数据收集过程与神经网络的训练细节具体如下。

A. 数据收集。英伟达公司使用人工驾驶并携带摄像机的车辆收集了美国伊利诺伊州、密歇根州、宾夕法尼亚州和纽约州高速公路的数据以及新泽西州中部地区的街道数据。道路类型包括双车道公路、带有停泊车辆的小区道路、隧道和未铺好的道路。天气类型包括晴朗、多云、有雾、下雪和下雨。道路类型、天气类型和司机的活动（保持车道、转换车道、转弯等）等特征用于标注数据。数据以每秒 10 帧的速度从视频中取样得到。

B. 卷积神经网络架构。使用 9 层卷积神经网络架构，其中包括一个归一化层、五个卷积层和三个全连接层。输入图像被分割成 YUV 平面并被传递到网络中。归一化器是硬编码的，在学习过程中不变。在网络中执行归一化任务时允许归一化方案被网络架构更改，并通过 GPU 处理。卷积层被设计用于进行特征提取，并通过一系列多样化层配置实验被经验性地选择。然后，在前三个卷积层中使用一个 2×2 的卷积核和一个 5×5 的卷积核，在最后两个卷积层中使用一个 3×3 的卷积核。在这五个卷积层后面加三个全连接层，以得到一个最终的输出控制值，也就是逆转弯半径。

C. 模型增强。为了减少系统转向命令输出之间的均方误差，研究者通过添加人工偏移和转动来增加数据，以教会网络如何从一个不利的位置或方向中恢复。

D. 仿真。将模拟器连接到一辆人类驾驶的数据采集车的前向车载摄像机，或者预先录制的视频。视频被输入模拟器的神经网络，判断应该给出的转向指令。模拟器记录偏心距、偏航信息和虚拟汽车的行程。当偏心距超过一米时，就会触发一次虚拟人类干预，而虚拟车辆的位置和方向会被重置。

2. 无人驾驶船舶

(1) 解决了什么问题。

劳斯莱斯和阿尔托大学及芬兰国家技术研究中心的研究人员合作开发了船舶驾驶台——oX 驾驶台。oX 驾驶台是一个智能的工作站，可以识别每一个走进来的人，并根据每个人的喜好做出调整。驾驶舱的窗户将展现船舶周围的实景，除了能够看到潜在的危险，系统还可以标注出船员看不到的海冰和拖船等信息。

无人驾驶船舶将由中央全息甲板统一控制，船长和船员能通过强大的 VR、AR 技术遥控全世界的船舶。

由于无人驾驶船舶无须设置船员及船员生活区，能节约船舶空间，减少运营成本、

主机负荷及二氧化碳的排放量。

该项目技术方面的工作由劳斯莱斯所属的 BlueOcean 团队担当,工作重点放在如何更经济、更方便地对无人驾驶船舶的推进系统和甲板设备进行远程遥控及自动化控制。

与此同时,欧盟开展了海上无人驾驶智能网络导航项目,由位于汉堡的海运物流和服务中心领导。该项目正从技术、经济和法律三个方面,对在公海内航行的一艘无人驾驶商用船舶的性能进行评估。

此外,国际船舶认证组织的研究人员也正在探索使用无人船舶沿挪威海岸线运输货物的可行性。

中国海事局和武汉理工大学正在合作进行“无人多功能海事船舶研究与开发项目”,主要研究自动驾驶船舶在商业和军事领域的用途。

3. 智能驾驶硬件

这一部分主要介绍地平线机器人开发的面向智能驾驶的高性能低功耗芯片——征途。

(1) 解决了什么问题。

征途专门针对智能驾驶场景及摄像头相关特性进行优化,具备对行人、机动车、非机动车、车道线、交通标志牌、红绿灯等多类目标进行精准、实时检测与识别的能力,可支持 L2 级别的高级驾驶辅助系统。

(2) 使用了哪些 AI 技术。

征途基于高斯架构研发,提供了“算法+芯片+云”的完整解决方案。

4. 智能车载系统

智能车载系统的代表是出门问问。

(1) 解决了什么问题。

基于 AI 技术打造的一整套语音交互方案,可应用于智能车载设备等。

(2) 使用了哪些 AI 技术。

A. 语音识别。出门问问的中文语音识别技术在移动搜索领域的识别率最高,并拥有整套离线语音识别技术。

B. 语义分析。出门问问的语义分析技术支持近 70 个垂直领域的语义理解，能够理解多达 11 个维度的自然语言查询。

C. TTS 语音合成。赋予机器一定的对话能力，让产品具有为用户播报信息的能力。

D. 垂直搜索。在整合移动数据的基础上，自主搭建搜索引擎，整合大众点评、高德地图等近百家合作方的优质数据，支持 LBS 搜索、跨领域垂直搜索，并自主研发基于网页搜索框架的垂直搜索工具。

E. 智能主动推荐。基于大数据进行个性化分析，在合适的时间推送合适的信息。支持类似 GoogleNow 的实时智能推送，如导航、天气、实时新闻等信息。

5. 智慧交通

AI 驱动的智慧交通解决方案的代表是 Active Scaler 的 Fleet Factors 平台。

(1) 解决了什么问题。

通过云和车辆、司机、ERP 系统、CRM 系统和其他商业系统的数据，使用深度学习引擎和大数据分析技术对车队的状态、车队的业务流程和数据进行实时观察，辨识出每辆车特定的行为模式，发现影响效率、安全的相关事件。在安全、运营效率、保险方面提供创新的车联网产品及服务，将传统汽车变成智能汽车。

A. 警告、报告及分析。进行实时分析，提供个性化的报告和警告，预测风险，提出建议。

B. 自动驾驶辅助。提供前方碰撞警告、车道偏离警告、物体和行人检测、道路危险检测等功能。

C. 多摄像头监控。利用 360° 外部摄像头、面向驾驶室的摄像头、黑盒记录行驶数据。

D. 分析驾驶行为。通过司机评分、车内行为反馈、驾驶员培训计划、车辆维修报告等信息帮助保险公司明确相应的法律责任。在安全方面，可以通过高级驾驶辅助系统及智能监控系统保障驾驶安全。

E. 车辆健康和维护。实施车辆状态分析，预测车辆运维花费。通过优化燃料使用效率和路线，监测车辆健康状况。通过车辆跟踪，以及事件和位置警报，降低运营成本和停机时间，提高运营效率。

F. 移动应用。地图及导航、车辆监测、评估车辆表现、支付费用等。

Fleet Factors 平台由 IMBOT、In-Motion Cloud 和 SDK/API 组成，其中 SDK/API 支持第三方软件或定制插件以提供个性化服务。

(2) 使用了哪些 AI 技术。

Fleet Factors 平台的核心集成部件 IMBOT 包括以下组件。

A. 处理器：Octa-core ARM 1.8GHz、Dedicated video DSP、3GB RAM、64GB Flash。

B. 照相机：面向道路的 1080P、面向驾驶员的 720P 摄像头。

C. 传感器：9 轴加速度计、环境/距离传感器、高精度 GPS、OBDII/CAN。

D. 通信：4G LTE 和 3G HSPA、802.11b/g/n、蓝牙、NFC、FM。

E. 显示：5 英寸多点触摸屏。

F. 安全：具备电池、Android 6 系统。

8.10 AI+文化产业

100 多年前，世界上首次出现了电影；50 多年前，世界上首次出现了电视；10 多年前，世界上首次出现了电商网络。2018 年，以 AI 为代表的现代科技正在向文化产业渗透，使文化产业的发展出现了新的重大机遇。

《文化部关于推动数字文化产业创新发展的指导意见》指出，要优化数字文化产业供给结构，“以供给侧结构性改革为主线，提升数字文化产业文化内涵、技术水平和产品质量。加强数字文化产业原创能力建设，鼓励全民创意、创作联动等新方式。发挥高新技术对内容创作、产品开发、模式创新的支撑作用，提高产品品质、丰富表现形式。深化‘互联网+’，深度应用大数据、云计算、人工智能等科技创新成果，促进创新链和产业链有效对接。提高不同内容形式之间的融合程度和转换效率，适应互联网和各种智能终端传播特点，创作生产优质、多样、个性的数字文化内容产品。”

2017 年，文化部发布了《文化部“十三五”时期文化产业发展规划》，在“推进供给侧结构性改革，推动转型升级提质增效”部分提及：推进“文化+”和“互联网+”战略，培育新型业态。加快发展以文化创意内容为核心，依托数字技术进行创作、生产、传播和服务的数字文化产业，培育形成文化产业发展新亮点。提升动漫、游戏、创意设计、网络文化等新兴文化产业发展水平，大力培育基于大数据、云计算、物联

网、人工智能等新技术的新型文化业态，形成文化产业新的增长点。在“加强科技创新与转化，提供发展支撑”部分提及：围绕文化产业发展重大需求，运用数字、互联网、移动互联网、新材料、人工智能、虚拟现实、增强现实等技术，提升文化科技自主创新能力和技术研发水平。支持数字文化资源开发关键技术研究与应用，加快文化产品数字化、协同化步伐，加强文化领域重要装备、工艺、系统、技术平台等相关研究。

国家频频出台的一系列规划和政策，为文化产业与 AI 的深度融合营造了良好的环境。其实，在文化产业领域内，AI 已有大量应用，但主要集中于音乐、新闻出版、视频、旅游等数字内容。随着文化产业数字化程度的提高，企业内部运营数据和第三方平台数据来源的多元化及算法的不断成熟，AI 可以解决的问题将越来越多，AI 在文化产业的应用范围也将越来越广泛。

AI 赋能文化产业的具体领域有：电影与动画制作、音乐创作、文学创作、体育运动及赛事转播、游戏电竞、移动社交、音乐流通、绘画创作、模拟训练、舞台及现场表演控制、健身等。

下面笔者具体介绍几个亲身调研过的领域。

1. 电影与动画制作

迪士尼和卡内基梅隆大学合作研发的与输入语音同步的动画生成技术。

(1) 解决了什么问题。

语音动画的制作对电影、电视和游戏制作来说非常重要，也非常耗时。动画人物在说话时，其言词与面部特征应相互匹配，并且表情需要表现自然，否则会影响观看体验。

研究者提出了一种深度学习方案，用以自动生成与输入语音同步且面部表情自然的动画。与之前程序化地生成语音动画相比，该算法能直接从数据中学会自然的协同发音效果。

此外，该方案中的系统需要调整的参数不多，可以实时运行，并与现有的动画重定向方法兼容。该方案还可让 VR 社交中用户的 3D 化身更加生动。

(2) 使用了哪些 AI 技术。

研究者首先对人类语音进行采样，根据演讲者的语言，快速建立用于训练的数据集，并以文本为音素标签，作为机器学习模型的输入。研究者使用滑动窗口的方法将

视觉帧单元重叠,使之与音素标签的组合单元相对应,以捕获局部情境分析和协同发音效应。

其中,选择音素标签的组合单元需要考虑语音的上下文和平滑度,以最小化成本函数。

此外,根据其算法的统计分析模型,模型识别了音素标签的组合单元所对应的嘴型和面部表情,即习得输出序列(动画序列)。该模型完成训练后,即可通过语音自动生成高质量的动画。

2. 音乐创作

AI 作曲系统 DeepMusic。

(1) 解决了什么问题。

给定时长、风格、乐器、情绪等基本要素,DeepMusic 可以在数秒内提供个性化的原创音乐。

(2) 使用了哪些 AI 技术。

使用反馈神经网络进行序列生成是一种普遍的做法,但是音乐相对于其他序列有其自身的特点:首先,单个音符所蕴含的信息量要比自然语言的单个单词所蕴含的信息量少很多,必须要多个音符组合在一起,才能给人带来“音乐”的感觉。其次,音乐这种序列要求前后有所呼应,满足乐理规范。研发者设计了一种自顶向下的层次化反馈神经网络,通过在不同时间尺度上对音乐进行建模来实现从粗到细地逐渐生成音乐,提高了音乐生成的质量。

3. 文学创作

麻省理工学院媒体实验室的恐怖故事智能写作程序 Shelley。

(1) 解决了什么问题。

Shelley 通过自然语言处理技术学会了“写”鬼故事。每小时它都利用一个关键词作为随机种子开始自己的写作,任何人都能用回复的形式把鬼故事接下去。

(2) 使用了哪些 AI 技术。

Shelley 结合了多层次循环神经网络和在线学习算法,可以随着时间的推移从人群反馈中学习。

4. 体育运动及赛事转播

语音交互看球助理 App。

(1) 解决了什么问题。

该 App 能提供足球赛事直播解说、介绍球队球员。App 不仅可以回答用户提出的关于足球的简单的事实型问题,还能够在一定程度上进行逻辑推理,并以语音、视频、图文等形式向用户传达球员基本信息、赛事数据、媒体资讯等与足球比赛相关的内容。

App 还会记录用户行为数据,从而构建独立完整的用户画像。在此基础上通过 AI 和大数据分析,推荐算法能为用户推送个性化内容,帮助用户直接锁定自己心仪的球队、球星,以及感兴趣的比赛资讯。此外,系统还能根据要求在赛后盘点每场比赛的高光时刻,从阵型统计、数据统计等多角度展开多维度的分析。

(2) 使用了哪些 AI 技术。

使用的三大核心技术分别为知识图谱、深度学习和自然语言处理。

为了建立足球行业的知识图谱,团队花费了大量的时间、精力去抓取散布在互联网各处的结构化及非结构化的数据,包括历年比赛中各球员、各球队的数据等,然后经过清洗、排歧、表达等处理过程,完成信息的聚合。

基于深度学习的自然语言处理能力被用于理解用户对话的语境与语义,从而实现人机交互。

垂行业语的料缺失限制了自然语言处理的能力,团队计划通过强化学习进行优化。

8.11 AI+教育

每年高考过后,填报志愿就成了考生和家长们关注的焦点,各种志愿咨询服务机构也应运而生。

2018 年 4 月,教育部发布了关于印发《高等学校人工智能创新行动计划》(后文简称“行动计划”)的通知。“行动计划”对中小学、高校等多层次的教育机构提出要求,未来将在中小学阶段引入人工智能普及教育。

2018 年 6 月,不少高考志愿咨询机构把 AI、大数据作为首推专业。

随着 AI 持续升温，AI 在教育领域的实际应用也逐渐崭露头角。AI 自适应教育作为 AI 与教育互融共生的产物，开始为更多的人所认识。伴随着技术的更迭，AI 自适应学习被寄予了让教育从“千人一面”转变为“千人千面”的希望。

自适应学习是一种强调学生个性化和自主性的教育方式。自适应学习系统通过提供相应的学习环境、实例或场景，让学习者自主发现规律，进行总结，进而形成理论，并能自主解决问题。自适应学习早在 20 世纪 80 年代便被提出，经历了人工、计算机编程与 AI 三个阶段。AI 阶段的自适应学习是对传统自适应学习的升级，也是对新型学习方式的探索。

能提高学习效率、可规模化的个性化教育是 AI 赋能教育的核心目标，“收集大数据→构建学习模型→输出学习建议”则是实现 AI 自适应学习的基本步骤。

AI 自适应学习的核心在于知识图谱的构建。而 AI 所做的工作，便是把所有的信息收集在一起，建立个性化的知识图谱，最终提升学生的学习效果，提出相关学习建议。

下面，笔者从 AI 赋能教育的实例里面挑选几个进行详细讲解。

1. 个性化学习——Hitalk

(1) 解决了什么问题。

作为面向成人口语的 AI 产品，沪江网校的 Hitalk 运用 AI 算法，根据学员在测试和学习中产生的数据评测学员水平，并分析学员的不足之处，有针对性地推送相应的学习任务，动态制订学习方案。

(2) 使用了哪些 AI 技术。

AI 技术在 Hitalk 中的主要作用是优化学习路径，提升学习效率。Hitalk 提供的服务有四大特色。

- A. 导航式学习，智能调整学习方案。
- B. 实景角色演练，与外教 1 对 1 地在线互动。
- C. 翻转课堂，学、练、习、测。
- D. 诊断式学习，智能匹配学习包。

在正式开始学习前，Hitalk 通过“机测+人测”的方式为学员定级，从听、说、读、

写四个方面将学员分为 12 个等级，根据学员的测试结果，以及其在学习产生的数据，学习方案将会进行智能、动态调整，Hitalk 称之为“导航式学习”。在学习期间，Hitalk 将会围绕“学、练、习、测”，从互动课件、实景演练、智能推送学习任务、人工测评等环节入手，先让学员自学课件中的知识点，然后在直播课上与外教进行实景演练。经过外教的点评指导后，再进行相同场景的二次对话。课堂结束后，诊断式学习针对课堂学习暴露的单词量不足、语法基础差等问题，推送专属学习包查漏补缺，强化薄弱环节。最后一环是“测”，通过定期的真实、人工评价介入来反映学员的进步，最终形成“学习、练习、批改和修正、再次巩固”的闭环。

2. 教育机器人——UniToy 机器人

教育机器人以云知声面向儿童智能机器人领域的 UniToy 机器人为标杆。

(1) 解决了什么问题。

- A. 先进的童声识别技术，可以更加准确地识别儿童用户。
- B. 儿童语义理解能够从儿童的角度出发满足儿童用户的需求。
- C. 高集成度的标准化产品，快速满足市场需求。

其产品可以实现以下功能。

- A. 童声识别。
- B. 海量内容供应。
- C. 极速响应。
- D. 童声合成。
- E. 儿童语义理解。
- F. 多轮对话。
- G. 设备互联互通。
- H. 差异化定制。
- I. 快速集成。

(2) 使用了哪些 AI 技术。

UniToy 机器人通过集成语音唤醒、语音识别、知识图谱、深度学习、推荐引擎、自然语言处理、语义理解、对话管理等技术，提供智能硬件模组、云端服务、手机

App 等整套专为儿童定制的系统。

3. 智能教学设备及系统——畅言智慧课堂

智能教学设备及系统以科大讯飞的畅言智慧课堂为标杆。

畅言智慧课堂是以建构主义学习理论为依据，基于“互联网+”的思维方式和大数据、云计算等新一代信息技术打造的智能、高效的课堂。其实质是基于动态学习数据分析 and “云、网、端”的运用，实现教学决策数据化、评价反馈即时化、交流互动立体化、资源推送智能化，创设有利于协作交流和意义建构的学习环境，通过智慧的教与学，促进全体学生实现符合个性化成长规律的智慧发展。

系统具体能实现以下功能。

一键开放投屏：采用自主专利的投屏技术，支持任意应用的投屏与讲解。

无网环境授课：采用自主专利的微云服务器，实现跨平台、不依赖互联网、满足高并发、稳定的师生互动授课。

海量优质资源：提供全学科、同步到课的优质资源与海量题库，支持本校资源、第三方资源引入，助力老师轻松备课、高效上课。

标准语言环境：电子课本提供带读功能，支持中文、英文语音评测，帮助学校构建标准语言的教学环境。

实用的教学工具：支持 PPT、Word 的原生态播放，提供白板、实物展台、课堂互动、微课录制等工具，满足全学段、全学科、多课型的常态授课。

多彩互动课堂：支持抢答、随机、分组等多种做题方式，提供投票、分享屏幕、拍照对比讲解等互动功能，助力老师打造多彩的互动课堂。

全程动态评价：从日常作业、随堂检测到周考、月考、大考，提供学习全过程的精准诊断与评价，推进智慧的教与学，实现个性化学习。

立体互动交流：创设有利于协作交流的学习环境，实现课堂的延伸，促进老师、学生与学生之间的互动交流，培养学生的自主学习习惯，提升学生的自主学习能力。

绿色学习环境：提供定制的学生终端，预置优质学习资源与应用，支持设备全方位安全管控。

8.12 AI+物流

随着无人仓、无人车、无人机的陆续出现，物流行业加速进化。

顺丰集团首席技术官兼顺丰科技首席执行官田民在 2018 年 2 月说：“今天，物流公司进入了一个全新的转型阶段，引领物流行业这场爆发式裂变的不仅仅是模式和资本，更是 AI、云计算等新兴技术。”

在这个物流行业发生剧变的时代，京东启用无人分拣中心、菜鸟网络使用无人机跨海送货、苏宁力推物流云仓……AI 对物流行业的改造成果如何？笔者经过调研得出如下结论。

(1) 装卸、搬运、分拣、包装、加工、配送等工作正在实现自动化、智能化。

走进京东昆山无人分拣中心，庞大的六轴搬运机器人负责用吸盘将货箱重新码放；自动引导运输车沿着地面的二维码搬运货架；分拣小件时，货架穿梭车从两排货架上将装有商品的货箱取下，放上传送带供分拣机器人分拣；拣选机器人利用 3D 视觉系统，从周转箱中识别出客户需要的货物，并通过工作端的吸盘把货物转移到订单周转箱中，拣选完成后，通过输送线将订单周转箱传输至打包区。整个分拣中心实现了真正的无人化，且效率大幅提升。据介绍，京东昆山无人分拣中心的作业效率为 9000 件/小时，在同等场地规模和同等分拣量的前提下，可以节省 180 人。

在分拣中心，AI 游刃有余；走出仓库，AI 借助无人机等设备参与最后一千米的派送。相比于封闭的仓库环境，户外环境要复杂得多，应用 AI 技术的门槛更高。2017 年 10 月，菜鸟网络公开进行无人机群组试验，满载 6 箱货品的无人机耗时 9 分钟，飞越近 5 千米的海峡，为农村淘宝提供物流服务。这是国内首次无人机群组跨海运送快递。

京东在无人机飞控调度中心、飞行服务中心、研发中心、制造中心等一系列配套技术与设施落地后，于 2017 年 6 月 28 日宣布其无人机送货正式进入常态化运营。

顺丰则对无人机寄予了更高的期望，其与合作伙伴研发出载重 300~2000 千克、续航能力 500~2000 千米、巡航速度 200~300 千米/小时的大型物流无人机，以期在未来能补充其支线物流的运营能力。

AI 对物流行业的改造逐步爆发。在物流公司方面，顺丰在 2016 年的研发投入达 5.6 亿元，到 2017 年 5 月，顺丰在无人机领域的专利数量达 64 件。“通达”系（笔者

注:对主流物流企业的统称,代指申通、圆通、中通、百世汇通、韵达)也开始注重科技和人才的投入,转运中心信息化、自动化程度逐步提高。

(2) 降低成本,提高效率, AI 助力物流行业突破瓶颈。

硅谷 AI 专家王亚莉说:“传统物流有较保守的生产线,较正规的运输线,需要工人值守的仓库,彼此之间相对独立而封闭,耗费了大量不必要的人力、物力、财力、时间,成本巨大,效率低下。”相比传统物流, AI 将带来人力成本的节省、周转效率的提高。

装卸、搬运是物流系统中最基本的功能要素之一,存在于货物运输、存储、包装、流通加工和配送等过程中,贯穿于物流作业的始终,物流机器人的应用直接提高了物流系统的效率和效益。

苏宁物流研究院副院长孟雷平说:“以苏宁南京云仓的 AS/RS 自动托盘堆垛设备为例,其效率是传统高位叉车的 4~5 倍,而且不需要人工操作,就能实现整托商品的自动上下架。机械化设备的投入大大减少了对人力的需求,降低了人力成本和管理难度。”未来 10 年,机器人解决方案有望节约 20%~40%的装卸、搬运成本。

再如,顺丰与腾讯云合作,实现对手写运单的认识。而在这之前,顺丰雇用了 8000 名输单员来输入手写的运单信息。

在效率方面, AI 一方面可依据历史数据对现有作业流程和库位规划进行优化,另一方面能带来运输路径的优化,提高配送效率。得益于 AI 的助力, 2017 年的“双 11”, 菜鸟联盟仅用时 2.8 天就将一亿个包裹送到消费者手中, 相比 2016 年减少了 0.7 天; 在 2017 年的“双 12”, 菜鸟联盟仅用 15 小时就送完 2016 年“双 11”全天的包裹量。

此外, AI 还将带来服务质量的提升, 信息化手段的使用可做到作业全程可追踪、过程可控和结果可预测, 将运作中许多不可控因素排除在外, 确保物流运营的质量。

(3) 重组生产要素, 重构运作流程, 形成新的物流生态。

下面, 笔者从快递送货、仓库仓储、交通信息、高精度地图、路径规划、贸易管理等 AI 赋能物流的案例中挑选出几个进行详细讲解。

1. 快递送货, 美团点评的实时智能配送系统

(1) 解决了什么问题。

美团外卖经过几年的高速发展, 凭借每天配送超过千万份订单的规模, 成为世界

上最大的配送平台。在此基础上，美团推出智能配送调度系统。

这套大规模、高复杂度的多人、多点、实时的智能配送调度系统，能够基于海量数据和 AI 算法，给最合适的配送员发送最优配送指令，确保平均配送时长不超过 28 分钟。

（2）使用了哪些 AI 技术。

在技术上，智能配送调度系统可以分为基础平台建设、垂直业务优化两个方面。

基础平台建设。团队开发了面向即时配送业务的机器学习平台、仿真平台、算法支撑平台。其中，机器学习平台提供若干数据预处理、机器学习算法工具包等内容，并实现底层大数据平台与线上系统的对接，提升了机器学习模型的迭代效率。仿真平台实现了对配送全过程的精准模拟，能够有效地进行各类配送策略的离线评估和分析。算法支撑平台主要包括离线/在线特征管理平台、线上策略管理平台和 A/B 测试工具等。

垂直业务优化。开发了配送网络优化系统、实时调度系统和定价系统。

A. 通过配送网络优化系统，为划分商家配送范围提供智能化的工具，辅助建立合理的底层配送网络结构。

B. 通过实时调度系统，实现订单和骑手的动态最优匹配，大大提升了配送效率。

C. 通过定价系统，实现对用户端、骑手端、商户端的合理定价，为即时配送的各参与方建立了良好的生态系统，实现了多方共赢。在配送过程中，有些数据无法直接采集，于是，美团通过骑手 App，采集骑手手机的蓝牙、Wi-Fi、GPS 等各类传感器的数据，将这些数据与配送场景结合，以挖掘更多信息，如上下楼时间、到店时间、出餐时间等。同时还可以精确还原骑手行为，包括骑行、步行、停留、走楼梯、坐直梯等。这些精细化数据对于提升调度和定价的精度都有很大帮助。此外，为了获取更精确和实时的数据，美团还向商家、骑手部署智能硬件。

在机器学习问题上，美团提出了很多解决方案。

首先，在特征工程上，构建多种泛化特征，比如区域的个性化特征，让模型能够感知区域特征。

其次，在模型架构上，针对各类问题单独建立补充策略模型，既保证整体效率的提升，又保证区域的个性化，不仅在区域这个维度上，在各种可能引起配送体验问题

的长尾维度上,模型均有较好的表现。

最后,在机器学习领域有一个跟这个问题有关的概念——小样本学习问题,美团正在探索利用迁移学习来解决这类问题。美团智能配送系统为了保证用户体验,有一套完备的供需平衡系统,通过实时监测供需关系,实时调整预计送达时间及配送费,将用户订单引导到运力更充足的商圈,或者引导用户错峰下单,保证订单能够快速、准时送达。

2. 仓库仓储,菜鸟网络的 3D 装箱技术

(1) 解决了什么问题。

菜鸟网络的 3D 装箱技术解决了运营优化问题,即如何使用体积最小的容器,依照最优的次序放置和打包物品,以降低运营成本。

利用该技术可以节省 5% 的包装消耗。在 2016 年的“双 11”,阿里巴巴产生了 6.57 亿笔物流订单,使用该技术能节省 3285 万笔物流订单的打包成本。

(2) 使用了哪些 AI 技术。

该技术将待打包物品的尺寸输入 Deep-Q 模型,计算出所需的最小容器,再用强化学习技术找到最佳的打包方式,利用组合优化技术使物品依照最优的次序放置,从而降低运营成本。

3. 交通信息,摩拜的大数据 AI 平台——“魔方”

(1) 解决了什么问题。

借助基于出行大数据的 AI 平台——“魔方”,摩拜单车不仅能够对已经发生的数据进行归纳和汇总,还可对未来任意时间节点、任意地点的共享单车的状态进行精准预测。通过将地点、时间、天气、运力、车型、人群及其他数百个变量因子纳入体系,“魔方”可以对特定地点未来某一时段的共享单车供给、用户需求、车辆使用频次、停放状况等做出科学预测,为精准运营提供有效指导。

随着共享单车扩张到几百万辆,提高规模化运营效率显得更为重要,这就需要硬件、固件、软件、设计、大数据、云平台之间的紧密配合,否则就会出现运营效率低、车辆损毁率高、用户体验下降等诸多问题。“魔方”是 AI 技术在共享单车领域的首次大规模应用,“魔方”已经在骑行模拟、供需预测、停放预测和地理围栏四大 AI 领域发挥巨大作用。

(2) 使用了哪些 AI 技术。

“魔方”的核心应用是对共享单车全天候的需求做出精准预测，为车辆投放、车辆调度和运营提供智能指引。

“魔方”利用 AI 及大数据技术，可以根据历史数据对每个地铁口的车辆需求进行预测，优化车辆投放方向和车辆投放策略，还可以根据天气情况预测第二天的骑行量，甚至可以实时监测车辆的车况，辅助运营人员进行车辆管理和回收。

另外，“魔方”还可以查询不同街区、不同时间段的车辆信息，智能识别车辆违章停放。

4. 高精度地图，百度地图的自动化图像采集技术

(1) 解决了什么问题。

百度地图同时部署了全景采集车、自行采集车、全景采集背包、室内图采集背包，以及正在测试的无人机采集设备，结合“1×0”单人采集、LiDAR 点云系统全景采集等核心技术，形成了业内首创的“多栖组合模式”，能满足多种地形、地貌及不同场景的地图数据采集需求。

百度地图的数据自动化处理程度已经达到 90%，能自动识别交通标志、地面标志、车道线、信号灯等上百种目标，相对精度达 0.1~0.2 米，准确率超过 95%，数据可以实现分钟级的更新。

(2) 使用了哪些 AI 技术。

自动化图像采集技术。

5. 小结：AI 与物流结合的走向

智能设备将重组物流的生产要素。南京邮电大学现代邮政学院、现代邮政研究院院长孙知信表示，智能硬件设备将使物流行业从人工向自动化、智能化方向快速发展，智能感知技术、信息传输技术，以及机械臂、机器人、自动化分拣带、无人机等智能硬件将在物流运作的各个环节广泛应用。

智能计算（尤其是智能物流云平台的建设）将重构物流运作流程。孙知信分析，智能计算将实现对供应链、实体物流的数字化、智能化、标准化和一体化的综合管理。以综合物流为出发点，应用现代 AI 技术及物流技术，使得供应链各环节的信息流与实体物流同步，产生最优化的流程并协同作业，实现货物就近入仓、就近配送，

提高产业链效率。

将形成全新的物流生态系统。在 AI、云计算、大数据、物联网等技术的协助下,依托铁路网络、公路网络、航空网络、水运网络及实体物流园区,可以形成覆盖线上和线下的物流生态系统,积极服务经济社会发展。

对传统物流企业而言, AI 的改造是机遇,更是挑战。数据、技术、场景是 AI 赋能运输物流的核心要素,三者缺一不可。

此外,构建物流标准化体系和诚信体系将是 AI 赋能物流的重点。不同企业的数据库编码、传输单据、承载单元等还存在着巨大的差异,需要建立统一的标准协议来支持 AI 赋能物流行业。同时, AI 技术的运用打破了物流行业传统的熟人模式,要想维系众多陌生关系,形成常态化交易,亟待建立社会化诚信体系。

8.13 AI+工业

8.13.1 AI 早已进入工业领域

2017 年,阿里云“ET 工业大脑”成为第一个进入车间的 AI 产品。光伏材料制造商协鑫较早采用了阿里云提供的云计算服务,其良品率提高了 1%,每年可节省上亿元成本。协鑫和阿里云的合作是中国工业制造领域的创新示范,阿里云在协鑫车间做的第一件事,是把生产线上所有端口的数据都上传到了“云”,然后调集上千台服务器的算力,在短时间内从数千个变量里找到了影响良品率的 60 个变量。接下来交由 AI 实时监测和控制这些变量,生产线只要“奉命行事”即可。

实际上,在工业领域布局 AI 落地的企业早已出现,作为引领自动化行业发展的先锋,深谋远虑的企业已经看到了数字化转型的优势,纷纷着手建立工业领域的 AI 平台。

(1) 西门子。

西门子中央研究院于 2018 年年初在慕尼黑演示了双臂机器人,借助 AI 的高度自动化,该机器人无须编程即可自主分工、协作。西门子中央研究院的研究人员表示:“我们只需要告诉机器人把哪个部件安装到导轨上,它就会执行这个操作。”这项任务描述了“单件定制化生产”的内涵,机器人从关联的软件模型中获取产品的相关信息,传统的机器人无法理解 CAD/CAM(计算机辅助设计和制造)模型,但新的机器人可

以理解。从某种意义上说，这就好像机器人能够理解不同的语言，使得人类不必再对其运动流程进行编程。

(2) 通用电气。

2018年，通用电气宣布与众多电气公司达成多年协议，并与纽约电力管理局达成广泛的协议，致力于成为全球首个全数字化的电力公司。如今，通用电气的综合智能运营中心已经开放，该运营中心是一家尖端的资产监控和诊断中心。其中，通用电气将部署和优化自己的资产绩效管理软件，该软件在通用电气的工业物联网平台上运行，用于监测、预测和提高 13 个燃气电厂和 1 个燃煤电厂的可靠性，这 14 个发电厂都使用通用电气或阿尔斯通的涡轮机和发电机。通用电气首席执行官 Russell Stokes 表示：“未来，通用电气将通过数字化转型，使工厂大幅度提高运营效率，创造更多收入。”

(3) 罗克韦尔自动化。

2017 年年底，罗克韦尔自动化宣布投资一家名为 The Hive 的硅谷创新与共创工作室，该工作室用来打造由创新群体与高科技初创公司组成的生态环境，致力于将 AI 应用于工业自动化领域。罗克韦尔自动化的企业发展高级副总裁 Elik Fooks 表示：“我们不断与领先的创新群体建立合作伙伴关系，比如我们与 The Hive 合作，期望通过工厂和企业一体化运营将工业生产力推向前所未有的高度。”其此举的目的包括通过共创解决客户问题，加速创新和发掘新兴技术，帮助制造业客户消除工厂基层和更高层信息系统之间的屏障，从而提高业务绩效。

8.13.2 从 AI 到工业智能

从计算机、互联网行业衍生出的智能化技术，正在以锐不可当的势头横扫全球各个领域，智能化与工业的结合更是引得全球瞩目。从德国的工业 4.0 到美国的工业互联网，从通用电气的 Predix 到 IBM 的 PMQ，可以看出，工业与 AI 的结合必将是下一个风口。

智能的核心在于决策和执行，而决策的核心在于感知和判断。在工业系统中，物联网技术，以及传感器、数据传输、数据管理等技术不断发展，为智能化的实施提供了可靠的感知基础。以前，工业界的大部分决策和反馈都以人为核心，这就导致系统中的很大一部分价值并没有被释放出来。系统越复杂，人的学习速度就会越缓慢，当人的学习速度比技术进步的速度更慢时，人就会成为制约技术进步和应用的瓶颈。而

AI 为工业带来的第一个革命性的改变就是摆脱人类认知和知识边界的限制,为决策支持和协同优化提供了可量化的依据。

8.13.3 AI 赋能工业的具体应用

借助 AI 全面实现工厂运营智能化,运营商能最大限度地减少人力,持续提升工业制造品质。笔者下面挑选几个案例详细讲述 AI 在工业领域的应用。

1. 库柏特机器人的香菇智能无序分拣流水线

(1) 解决了什么问题。

传统香菇加工企业依靠人工分拣不同种类的香菇,招工难一直困扰着相关企业。人工分拣速度慢,分拣一个香菇大约需要 5 秒,无法满足庞大的市场需求。此外,由于长期低头工作,工人容易疲劳,分拣质量也没有保障。

在技术上,香菇种类繁多、产地不同,这会给分拣造成困难。

香菇智能无序分拣流水线使用机器学习技术,同时使用传统方法进行分拣,可以识别 26 种香菇,分拣速度可以达到 0.8 秒/个,效率是人工分拣的 6 倍,精准率可以达到 99.7%。传统流水线需要 8 名工人,智能流水线则只需要 4 台机器与 2 名工人协同工作,经计算,在产量相同的情况下,一条智能流水线每年可为企业节约 20 万元以上的成本。

(2) 采用了哪些 AI 技术。

通过高分辨率工业相机配合图像感知算法,能够获取目标物体的类别和位置信息。利用机器臂、末端执行器和动态抓取控制算法,能够实现对物体的精准分拣。

2. 埃夫特的喷涂机器人轨迹自动规划技术

(1) 解决了什么问题。

用 AI 技术自动生成喷涂机器人的控制程序,解决了以往类似机器人编程困难的问题。

在柔性工业制造中,固定化编程难以满足多种产品混线生产,以及高频率更换产品的需求,埃夫特的喷涂机器人轨迹自动规划技术在很大程度上解决了这一问题,大幅减少了使用机器人过程中的编程工作。

(2) 使用了哪些 AI 技术。

机器人轨迹自动规划过程包含了三维重建、模型分解、模型部位识别、轨迹片段

规划及系统轨迹优化等技术。实现过程涉及模式识别、运动控制优化、轨迹规划与强化学习结合等问题。

在轨迹规划模块的训练中，埃夫特将已有的大量喷涂机器人的应用程序和对应的加工工件分解，生成了包含 14 万种模型碎片及轨迹片段的数据集，并在此基础上训练了一个深度神经网络，用于预测新的加工工件对应的喷涂机器人的加工轨迹。模型能够根据当前加工工件和前 5 个控制周期中的加工轨迹推理出当前控制周期需要选用的加工轨迹。经过调整后，轨迹规划模块能够较好地生成稳定、连续的喷涂机器人加工轨迹，并且能够根据不同的喷涂工艺对轨迹进行调整。

3. 日立的智能燃烧控制系统

(1) 解决了什么问题。

日立的智能燃烧控制系统能够通过分析发电机气缸内的压力、空气量、燃料的混合状态、点火时间等数据，自动学习燃烧的调整方法。基于训练，该服务可以根据燃料的种类和混合的状态来计算出最合适的燃烧控制值，并用于提高发动机的工作效率。部署了该智能燃烧控制系统的发动机可以将混合燃料的燃烧波动率控制在 3% 以下，使 15 千瓦的发动机的热效率达到 34%~41%。

(2) 使用了哪些 AI 技术。

机器学习。

4. 可口可乐通过 AR 技术辅助设备维修

(1) 解决了什么问题。

可口可乐通过 AR 技术使技术人员可以接收所维修设备的信息，并从远程专家那里得到帮助，这些专家可以看到技术人员所能看到的東西，并帮助诊断问题，辅助设备维修。

(2) 使用了哪些 AI 技术。

AR 技术。

5. 日立的汽车部件外观检查解决方案

(1) 解决了什么问题。

该解决方案无须手工选择需要的各种参数，可以直接利用深度学习来获取各种判定参数，从而提高传统判定方案的精度。同时，该方案避免了因为检测人员的个人差

异导致的品质不一致问题,保证了判定合格的产品品质相同。利用该方案,企业不需要太多的检测人员,避免了因为熟练工人退休导致的人员不足问题。

(2) 使用了哪些 AI 技术。

计算机视觉。

8.14 AI+家居

未来,我们的家居生活将以住宅为平台,利用综合布线技术、网络通信技术、安全防范技术、自动控制技术、音/视频技术将家居生活有关的设施集成,构建高效的住宅设施与处理家庭生活日程事务的管理系统。这些改变将提升家居生活的安全性、便利性、舒适性、艺术性,并打造出环保、节能的居住环境,宛如有一位管家在随时为你服务。

笔者下面从智能家电、智能家装、家庭自动化、家庭网络等方面分别介绍智能家居。

1. 智能家电:智能喂食器 Petnet

(1) 解决了什么问题。

该智能喂食器能实时监测宠物的状态,计算宠物摄入的卡路里,给出喂食建议,防止宠物营养摄入不足,也避免因过量喂食而增加的日常开销。

(2) 使用了哪些 AI 技术。

机器视觉、大数据。

2. 智能家装:Brain of Things 的智能公寓

(1) 解决了什么问题。

由于少子化与高龄化,老人独居现象越来越普遍,为了适应未来社会的特点,各国政府绞尽脑汁,希望能够解决房屋空置高、房价高、老人独居 3 大问题,智能家装可在一定程度上解决上述问题。

(2) 使用了哪些 AI 技术。

创业公司 Brain of Things 正在美国创建“机器人家园”,这些公寓配有大量传感器、自动化的灯具和电器,并透过传感器收集数据,使用机器学习算法构建行为模

型,使得系统能够学习、适应居民的习惯和偏好。居住者早上起床后,公寓里的百叶窗会自动打开,晚上百叶窗会自动关上;如果公寓感测到居住者在晚上起床喝水,会自动点起夜灯。

Brain of Things 的“机器人家园”配有约 20 个运动传感器,连管道都是联网和自动化的,还有宠物专用的自动化监测和喂食系统。居住者可以使用普通开关,也可以使用语音命令或手机应用程序来控制空调、电视、热水器等家电,随着时间的累积,公寓会学习居住者的偏好,并尝试提前提供服务。

3. 家庭自动化:海尔 U-home 智慧家庭解决方案

(1) 解决了什么问题。

海尔 U-home 智慧家庭解决方案以开放的态度整合各方资源,统筹所有智能家居技术,为用户提供全新的场景式生活体验,可提供数据采集、用户习惯分析等个性化服务,并主动向用户提供服务,打造真正的智慧家庭。

(2) 使用了哪些 AI 技术。

海尔 U-home 智慧家庭解决方案是海尔面对智慧家庭行业推出的一站式解决方案,它是集物联网、大数据、云平台和 AI 于一体的综合解决方案。旨在以家庭设备为核心来实现家庭云化,通过数据的采集及计算,感知、分析用户的生活习惯,以自然交互的方式向用户主动提供所需要的服务,为用户提供美好的生活体验。同时通过开源,向行业开发者、AI 提供商、大数据服务商、合作伙伴进行开放,搭建引领行业的 AI 开放平台,赋能智慧家庭行业,实现智慧家庭的规模化应用。

4. 家庭网络:天猫路由器

(1) 解决了什么问题。

天猫路由器可以作为整套智能家居系统落地的基础,不仅提供网络接入服务,更承担防火墙、家庭电子设备控制管理器、家庭多媒体中心等角色,除此之外,还涉及阿里巴巴的一项重要安全技术——安全加密通道。

(2) 使用了哪些 AI 技术。

天猫路由器采用了企业级配置,内置高通主芯片和蓝牙芯片,具备 2.4G 和 5G 双频天线,能够满足 Wi-Fi 无缝覆盖的需求。这款路由器本身可以自组网,可以连接多达 256 个连接点,支持蓝牙 5.0,具备蓝牙网关能力,可扩展支持市面上全部的连接

协议。它也具有开放工具集成协议，能够让设备迅速相互发现并联网。天猫路由器还能够借助边缘计算能力实时识别使用场景，自动匹配网速优化方案。

8.15 AI+法律服务

AI 赋能法律服务的历史源远流长。1987 年，关于法律与科技的第一次国际会议在波士顿举行，此后国际上法律与科技相结合的组织、公司陆续成立。但相对于 AI 赋能自动驾驶、金融、安防、医疗等领域的速度，AI 赋能法律服务的进展要慢得多。

8.15.1 政策方向

2017 年 4 月，最高人民法院印发《最高人民法院关于加快建设智慧法院的意见》。智慧法院是人民法院充分利用先进信息化系统，实现公正司法、司法为民的组织、建设和运行形态。2017 年 7 月，国务院在发布的《新一代人工智能发展规划》中提及建设智慧法庭，提出促进 AI 在证据收集、案例分析、法律文件阅读与分析中的应用，实现法院审判体系和审判能力的智能化。

8.15.2 技术变革

2017 年以来，深度学习技术在法律文本分类、法律文本自动生成、自然语言案例检索上有十分亮眼的表现。深度学习的发展离不开数据量的支持，事实上，司法公开力度的加大从根本上促进了 AI 赋能法律的发展。截至 2017 年 2 月，中国裁判文书网已公开超过 2600 万篇裁判文书。

8.15.3 行业态度

AI 技术的进步、司法公开给法律行业带来了深远的影响。在 2017 年，“法律+科技”的峰会不下百场，参与者不仅包括法律界人士，更延伸至社会各界专家学者、投资人、技术工程师、媒体与大众，而与之相关的话题讨论，从理论创新到未来趋势，从技术审判到职业变化，这些讨论正在一步步具体到当下的法律实务中，并且成果斐然，面向法检机构工作人员、律师、大众等不同群体的检索系统、管理软件、办案系统、辅助工具、法律咨询、智能分析报告等产品接连推出，AI 赋能法律迎来发展的黄金时期。

8.15.4 市场规模

2015 年全国共有超过 1800 万件诉讼案件，2016 年全国共有超过 2300 万件诉讼案件，但只有 20%左右的案件聘请了律师。造成这种情况的原因一方面是全国从业律师不足 40 万人，服务能力远不能满足每年多达 6 亿人次法律服务的需求。另一方面，法律服务专业性强，咨询、诉讼服务的费用高。

总体来说，AI 在法律领域的应用分为两大方向：一是为 B 端提供服务；二是直接服务于 C 端。

ToB 类：企业法律服务，为企业提供知识产权保护、合同分析、财税咨询、资产管理等相关法律服务，在整体上集中体现为一种法务外包服务。此外，还有其他模式的服务，如法律风险预警、智能合同和文书管理等。律所/律师服务智能化，包括律所管理软件、律师协同工具、案件流程化管理，以及智能化案件文书数据库、律师职业技能提升平台等。

ToC 类：直接面向个人，基于 App 或者微信平台，为用户提供案情咨询、律师匹配与对接等服务。尽管需求强烈，但整体依然属于低频需求，付费转化率较低，用户分散且难以与平台维系关系，如果通过推广来大量获取用户，则面临服务成本增加的问题，因此，基于 AI 的智能法律助手至关重要。面向法律援助中心，事实上，这是另一个服务个人的切入口，但付费者发生了改变，有人认为这种模式除了为公共事业做贡献，还看不到商业利益。

在 ToB 和 ToC 两个方向上，产品形态和技术壁垒有着质的不同。ToC 产品的关键点主要是从自然语言输入到法律要素的映射问题，常见的产品形式是自动法律咨询。而较为专业的案例检索，则涉及案件的深度研究，单一的裁判文书并不能反映案件的所有特征。

8.15.5 案例分析

下面笔者从 AI 赋能法律里挑选了几个有代表性的案例进行详细讲解。

1. 华宇元典的律师工作管理平台

(1) 解决了什么问题。

基于公司“元典睿核”这一底层法律知识服务分析平台，华宇元典开发了集

大数据智能辅助、云端协作、流程管理、知识管理等功能于一体的 SaaS 平台，方便律师管理案件项目、跟进流程进度、整理文档资料、高效完成任务。

(2) 使用了哪些 AI 技术。

云计算、数据挖掘。

2. 达观数据的法律信息智能化垂直搜索引擎

(1) 解决了什么问题。

通过对案件信息的解构和标签化处理，实现精准的关键词匹配，并判断裁判文书之间的相似程度，自动推送相关案件和法律法规，减少对相似法律文书的检索时间，提高案件分析效率。

(2) 使用了哪些 AI 技术。

达观搜索系统综合考虑了相关法律文书结果的权威性、时效性、多样性等指标，并利用自有专利的点击调序算法，自动根据用户搜索意图来提高排序精度。深度挖掘法律行业语言模型和语料资源，确保能够准确把握内容含义，并采用高效的全文索引技术，确保在 0.1 秒内完成亿级数量内容的搜索响应。

3. 法狗狗智能法律平台

(1) 解决了什么需求。

法狗狗智能法律平台利用区块链技术和比较成熟的图像识别技术来解决版权问题，降低了维权成本，从而解决了采用自然语言处理技术和知识图谱技术未解决的问题。

(2) 使用了哪些 AI 技术。

图片权利人可将图片上传至法狗狗平台，平台会利用卷积神经网络抽取图片特征，并将这些特征保存到区块链中。在发现疑似侵权图片时，会向整个区块链广播，向图片权利人告知侵权信息。图片权利人确认后，即可通过流程化维权服务进行维权。

4. “法理”智能问答系统

(1) 解决了什么问题。

通过大数据和 AI 高效地连接法律服务，同时用 AI 解决基础的法律问题，从而高效地为律师筛选案件。

(2) 使用了哪些 AI 技术。

数据挖掘、推荐系统。

5. 贵州省人民检察院的犯罪构成知识图谱

(1) 解决了什么问题。

贵州省人民检察院将 34 万个案件、908 万条记录、126 万份法律文书、920 万页卷宗重新组合、分析、计算，最终汇入贵州省人民检察院大数据司法办案辅助、大数据分析服务等系统。检察官通过这些案件和法律文书，总结出故意杀人罪背后的 24 个定罪要素和 22 个量刑要素。经过 3 次迭代升级，检察官还找到了故意伤害罪的 40 个定罪要素和 24 个量刑要素，而抢劫罪的定罪要素为 30 个，盗窃罪的定罪要素有 36 个。这些定罪要素和量刑要素组成了一个犯罪构成知识图谱。

(2) 使用了哪些 AI 技术。

知识图谱。

8.16 AI+社交

8.16.1 AI 如何赋能社交

AI 是对人类智力行为的研究。截至 2018 年，AI 并没有对社会科学和社会行为进行非常大规模的研究，除了意识，人类还有阅读、交流、集会等社会行为，这些行为是 AI 赋能社交的研究对象。

比如，SARA（有社交意识的机器人助理）能自动理解人的感受，感受到人际关系纽带，而且能够随着人的社交圈的不断成长和巩固，做出相应的反应。社交 AI 的核心是具有社交意识，它和情感 AI 是完全不同的。AI 的目标是建造能够自主决策的机器，社交 AI 的目标是建造能和人们建立联系的，并能够探知真实人类社会行为的系统，同时用自己的行为对交互行为做出合理的回应。

8.16.2 AI 赋能社交对生活的改善

很多人害怕 AI 会夺走我们的工作，这种看法是片面的，AI 让一些工作消失的同时也会创造一些新的工作，而且是更好的工作。AI 可以帮助我们生活在一个饱含关爱的世界。社交 AI 能够帮助我们去关心别人，能够帮助那些患有自闭症的人建立社交

关系, AI 会让我们的世界更具有社交性。

有人说社交 AI 可能会减少人们对真实社交生活的需求, AI 专家贾斯丁·卡塞尔则认为恰恰相反, 人类关注的是真实的人, 而社交 AI 只是训练我们的心灵并帮助我们与其他人建立关系。社交 AI 会提醒我们和其他人协作有多么重要, 和世界上其他人互相依赖有多么重要。

8.16.3 社交 AI 的现实应用

讨论社交 AI 不得不提美国的 ObEN 公司。ObEN 是一家区块链与 AI 创业公司, 专门开发支持个性化 AI (PAI) 的公有链协议和技术, 简单来说, 基于 PAI 项目, 任何人都可以生成自己的 AI 形象, 不仅长得像、说话像, 甚至连性格也像, 而且懂得如何思考。用户仅仅通过上传一段自拍视频, 便可创建属于自己的 PAI, 用户可以借助虚拟分身, 在社交网站上进行虚拟互动, 未来 PAI 还可以帮助用户管理日常工作, 这对明星而言尤其具有实际意义, 明星可以让虚拟的自己和粉丝进行互动。

8.16.4 AI 赋能社交的案例

1. Facebook 的文本理解引擎 DeepText

(1) 解决了什么问题。

Facebook 基于深度学习技术的文字理解引擎 DeepText, 可以精确地理解文字的上下文。DeepText 支持 20 多种语言, 每秒钟能理解数千篇文本, 准确率与人类相当。DeepText 被用在 Messenger 上, 用于推荐合适的对话回复。例如, 有人发布希望售卖物件的消息, 该引擎能提取出出售的具体物品、物品的售价等有价值信息, 并推荐 Facebook 提供的相应服务。

(2) 使用了哪些 AI 技术。

DeepText 使用了自然语言理解、深度神经网络等技术。DeepText 包含多种深度神经网络体系结构, 可实现文字级别和字符级别的学习。在模型的训练方面, 该技术使用了 FBLearner Flow 和 PyTorch, 训练好的模型可以通过 FBLearnerPredictor 平台运行。

2. WibbitzPC 浏览器插件

(1) 解决了什么问题。

使用基于文本生成的实用算法来分析文字, 并且提取有趣的信息。使用智能语言

处理技术将文本转换成吸引人的视频文件，视频包含图片、图表和利用文字转语音技术合成的声音。

(2) 使用了哪些 AI 技术。

WibbitzPC 浏览器插件使用了文本摘要、自然语言处理技术。

8.17 AI+人力资源管理

预计到 2020 年，国内人力资源行业的市场规模将达到三千亿元。从 2017 年到 2018 年，AI 赋能人力资源行业的融资层出不穷。下面笔者从职位匹配、背景调查、职位搜索、员工管理等被 AI 赋能的领域挑选出具体的应用进行深度讲解。

1. 职位匹配：趣校园互联网招聘平台

(1) 解决了什么问题。

虚假招聘信息引起学生财产、人身安全等受到损害的情况时有发生。对此，趣校园要求合作的招聘网站对发布的岗位信息必须进行严格审核，并引入蚂蚁开放平台的信用管控能力，帮助求职者识别潜在风险。

未来，求职者可以查看应聘企业的芝麻信用评分，从而有效规避失信和经营有问题的企业，还可以通过投诉举报系统直接在线举报虚假信息。

学生在支付宝内搜索“校园工作”即可进入该平台。平台几乎汇聚了行业内所有主流的提供兼职、实习、就业的招聘平台，包括智联招聘、实习僧、斗米兼职等。用户授权后，即可跨平台同步简历。该平台通过岗位搜索雷达功能将用人单位和大学生们的岗位需求精准匹配，不仅能显示出匹配职位的数量，还能筛选出最精准的 3 个职位，经用户授权后，自动投递给企业。

(2) 使用了哪些 AI 技术。

使用自然语言处理技术和卷积神经网络，综合企业接收简历、邀请面试、招聘入职、绩效考核和评价等数据，以及求职者的档案信息、简历文本、简历投递历史等数据，实现企业与求职者双方需求的高效率匹配。

2. 背景审查公司 Checkr

(1) 解决了什么问题。

使用身份证号码、住址、驾驶记录等数据对应聘者快速进行背景审查，提供包括

历史住址、犯罪记录等内容的背景调查资料,并用 AI 技术将记录快速分类并生成相应报告。

与传统的背景审查不同,Checkr 提供了一个客户只需要输入名字就能得到数据反馈的 API,简化了调查报告的手动操作过程,使得背景调查更加自动化,从而更容易链接到公司现行的入职流程之中。通过 Checkr 平台,公司能够通过在线审查的方式或者直接将 API 加入雇佣系统的方式来审查新员工。这能够大大简化公司对员工进行背景调查的流程,提高人员流动的效率。

(2) 使用了哪些 AI 技术。

机器学习、数据挖掘。

3. 职位搜索,谷歌的职位搜索引擎

(1) 解决了什么问题。

谷歌的职位搜索引擎囊括了几乎所有主要的在线招聘平台(LinkedIn、Monster、WayUp、DirectEmployers、CareerBuilder 等),还将列出在公司主页上抓取到的待招聘职位,并对职位进行筛选和分类,发挥搜索优势。

(2) 使用了哪些 AI 技术。

机器学习、自然语言处理。

4. 员工管理:微软的员工产出管理工具 Workplace Analytics

(1) 解决了什么问题。

该工具可以被安装在 Office 365 企业版上。它通过调取企业在 Office 365 上的邮件及日志元数据,帮助管理者掌握多位员工完成某一具体任务的时间,用来优化管理架构及任务分配。除此之外,该工具还有其他用途,例如,可分析顶级销售人员的性格特征与行为方式,并提供给其他销售人员借鉴。

(2) 采用哪些 AI 技术。

数据挖掘、机器学习。

8.18 AI+房地产

联合国发布《2018 年版世界城镇化展望》显示,全球大型和超大型城市的数量已经是十年前的两倍,到 2050 年,占据全球人口五分之一的 600 座大城市预计贡献

全球 60% 的 GDP，其中将出现更多来自中国、印度等国家的城市。可以看出，在世界范围内，亚洲和非洲国家的城市化进程和经济增长将有巨大的发展空间。

作为推动 GDP 增长和城市化进程的重要力量，房地产行业中的房地产企业近年来也不断调整策略以谋求更稳定、更快速的发展。

高周转策略已存在于房地产企业多年，房地产企业“将房地产行业干成快消费品行业”的现象说明了这些企业渴望降低风险。

在资源不断往头部企业集中、房地产企业融资压力巨大的环境下，压缩开发周期、快速回笼资金、高速周转成了房地产企业的共同选择。

根据大型房地产企业的周转信息，平均拿地以后半年内就需要开盘。传统的设计环节可能需要一个设计小组耗费 2 周以上的时间完成概念方案设计，这已经大大影响了拿地的决策效率。也因为如此，后续的方案设计难以在一开始就跟进，导致整体设计和反复修改占据了动工前的大部分时间。大量规划指标、动态的用户需求、周边地区的价值等也都需要人工进行搜索和评估。

不少房地产企业开始寻求“升级”，一方面在全国范围内通过品牌和产业结合积极拿地，另一方面大胆利用新技术提高拿地的效率和准确率。AI 在城市规划和建筑设计领域的介入，必定最大限度地提升房地产企业开发前期产品定位、评估、沟通和设计等环节的效率和准确率。

AI 赋能房地产行业能给行业带来什么？笔者建议对 AI 赋能房地产行业感兴趣的读者必须了解以下三大趋势。

(1) 定位分析。建立城市价值地图，提升设计定位、价值评估的准确性。

移动互联网时代为大数据的爆发式增长提供了可能。借助 GPS、手机、移动支付终端等获取的数据能够精准地反映城市中人的行为路径、模式、变化和规律，同时，由互联网企业产生的城市地图、房地产信息、商业价值等数据能更详尽地描绘城市复杂的“背景”系统。通过对各种线上、线下数据进行收集、整理、分析和标注，可以建立更加客观、科学的价值地图，进而作为定位设计和评估价值的依据。

AI 建立在大数据这个基础之上，可以追踪和掌握城市动态、复杂的环境，并通过大数据分析项目地块的价值，进而建立对项目地块和周边复杂环境进行科学判断的体系，让项目决策不再是凭经验、拍脑袋、赌运气，而是建立在更为客观、理性的分析

上,在项目立项之前就知道有多少“胜算”。

(2) 快速生成符合需求的方案,综合提升空间设计水平和施工组织能力。

一般而言,只有设计总监以上级别的建筑师能够快速而准确地在满足容积率要求和基地条件的前提下,找出不同产品的最优组合。绝大多数建筑师需要花费数天,依靠产品库的楼型在基地上反复计算、尝试,并验算排布可能性和极限。布局设计在没有多维数据支撑的情况下,很难既满足不同的需求又兼顾指标限制。

而 AI 可以对产品类型、布局合理性、视野、间距和日照等内容进行从整体到细节的评估和判断,快速生成上万种组合和布局方案,并结合用户需求和周边地块的数据,智能推荐和锁定方案。设计不再是单纯的经验组合、感性的表达,或者简单的套用,而是基于评估动态数据和空间形体,以及市场和产品需求,进行专业而科学的思考和判断。

(3) 引入 AI 实现高效、高质量增长,让每家房地产企业都有同等的机会。

房地产行业已经到了发展的拐点,随着 AI 技术在行业的落地,每家房地产企业都有机会在 AI 的介入下,更准确、快速地解决速度、时间和成本三者之间的矛盾,提升资源利用效率,在不断扩大规模的同时提升开发品质。

在 AI 的分析和推荐下,房地产企业将可以更精准地与设计机构、咨询机构和房产中介等不同阶段的参与机构进行更为有效的深度协作,共同实现产业纵深发力,提升整体效益。

在 AI 赋能房地产行业的众多案例里,笔者挑选了几个案例进行具体详解。

1. 建筑工程:麻省理工学院媒体实验室的自动建造机器人

(1) 解决了什么问题。

房地产行业传统的建造方法危险且建造过程缓慢,属于劳动力、能源密集型产业,而现有的多种建造机器人仅仅能完成建筑的某个部分,局限性比较大。该自动建造机器人作为一个数字化建造平台(Digital Construction Platform, DCP),可以自由移动,结合增材制造技术,理论上可以打印出任意尺寸的物品,建造房屋的所有部分。

(2) 采用了哪些 AI 技术。

该数字化建造平台以机器人手臂的形式呈现,作为一种自动化建造系统,可以根据需求现场搭建出建筑物。

其“手臂”的一端有一个喷头，可以将两种化学物质混合成一种液态的聚氨酯泡沫体，这种泡沫体会迅速膨胀并硬化。机器人不是直接用泡沫体打造整个建筑，而是打造两个泡沫体墙壁，一个泡沫体嵌套在另一个里，在布好管道及电线等内容后，可以在中间的空间注入水泥，保留泡沫体仅仅是为了增加建筑的绝缘性。

2. 建筑设计：小库科技的建筑设计智能化平台

(1) 解决了什么问题。

该平台将建筑领域 80%~90%的产品，比如住宅小区、产业园区的设计智能化，在自动化的基础上能够智能地判断、推荐、设计，让原本需要花费一到两周的设计工作在数秒之内就能完成，省去绘图过程后，建筑师只需要进行判断和选择。

其罗塞塔计划通过智能设计算法，分析、理解知名建筑师的设计方案和设计风格，提取出建筑师的设计逻辑、审美判断，生成特有的模型，在丰富平台模型库的同时，使平台更具创造力与想象力。

除了公有模型，小库科技还能为设计机构定制具有完整知识产权的算法模型，让其设计思想从只能在单一项目上呈现、每次面对新项目都要从头再来，跨越到普遍适用多个项目，能够最大化实现设计师的价值。

截至 2018 年年底，小库科技只给企业提供定制化算法，使设计变得更智能，未来则希望开放 API 及部分底层代码，让所有设计师都能在其平台上开发自己的设计算法，将设计交易变成算法交易。

(2) 使用了哪些 AI 技术。

采用了深度神经网络和机器学习技术。对设计院等机构提供的图纸、从渠道获取的海量数据进行机器学习，并在积累的建筑设计和城市规划算法的基础上，结合建筑师的人工标注和需求输入，实时生成大量符合建筑规范的方案模型，通过 AI 评估和判断哪些是更优的建筑方案，最后推荐至少十个优质方案给建筑师。

3. 建筑数据管理：智能地产数据录入平台 Earthworks Software

(1) 解决了什么问题。

传统工程项目往往会花费很多时间在建筑数据的录入上，建筑从业者一般要面临 300~1000 份不同作用的建筑图纸，而 Earthworks Software 智能地产数据录入平台可以解决上述问题。该平台通过 AI 技术可以准确识别近 99%的待录入图纸，并帮

助从业者进行分类。同时,该平台还可实现图纸的格式转换与编辑等功能。

(2) 采用了哪些 AI 技术。

计算机视觉。

4. 商业地产分析: 商业地产数据可视化及智能分析平台 REscour

(1) 解决了什么问题。

REscour 为经纪人、贷款人、开发商、投资者和分析师等商业房地产专业人士提供了一个平台,将数十种孤立的数据源整理归入一个统一的系统,可以为买卖房产做出更好的决策支持,并通过数据挖掘和行业趋势预测来提高市场洞察力。

(2) 使用了哪些 AI 技术。

数据挖掘,计算机视觉。

5. 智慧建筑。Building Robotics 的自动化控制软件 COMFY

(1) 解决了什么问题。

Building Robotics 公司提供了一款名叫 COMFY 的楼宇自动化控制软件,通过机器学习来自动控制办公室的灯光、温度等以帮助公司节省办公室的电费,同时通过收集员工工作空间的使用情况等数据,解决办公空间利用率不足的问题。

(2) 使用了哪些 AI 的技术。

物联网和机器学习。

8.19 AI+旅游

2018 年 6 月,“人工智能在旅游行业的应用前景与旅游创新”论坛在北京国际会议中心举行,来自科技、旅游、会展等多个领域的行业精英围绕 AI 赋能旅游展开深入探讨。

IBM(中国)全球企业咨询服务部航旅业解决方案总监陈宏甫在会上做了主题为“AI 时代下的航旅数字化重塑”的演讲。

百度度秘 ToB 业务高级经理李阳在名为“DuerOS 赋能酒店行业”的主题演讲中介绍的一体化解决方案也引起了酒店行业从业者的广泛关注。

携程集团在主题演讲环节做了“携程智慧旅游解决方案”的分享。据携程旅行网

AI 高级经理吕俊涛介绍，借助大数据中控系统，携程实现了 AI 在旅游行业的全流程应用。从旅行前的兴趣发掘到商品售卖，从目的地实时查询到售后和客户投诉的处理，AI 在客户旅行的各处发挥着作用。

下面笔者就旅游活动营销、产品价格预测及对比、旅游方案管理、服务虚拟助理等方面分别介绍 AI 赋能旅游行业的具体应用。

1. 腾讯全域旅游智慧平台

(1) 解决了什么问题。

云南省与腾讯公司联合打造了全域旅游智慧平台“一部手机游云南”，游客在出发前可通过“VR 看云南”“全景直播”等功能提前熟悉景点情况，提前规划好旅游线路。同时，智能客服将围绕精品线路定制，为游客提供一站式咨询服务。游客在旅游中可以通过扫码、人脸识别直接入住酒店。

在游览时，平台可为游客推送线路、提供导航服务，游客不仅可以通过扫码乘坐公交车和景区直通车，进入景区时还可以通过扫描二维码实现极速购票，并通过人脸识别系统进入景区，实现“刷脸”入园，避免出现排队、等待时间长等状况。

同时，游客可以通过小程序查找停车位、厕所，享受智能语音讲解服务。让游客通过一部手机就可以享受全方位的智能服务，增强游客旅游的舒适性、便捷性和自主性。

(2) 使用了哪些 AI 技术。

该平台集合了物联网、云计算、大数据、AI 等多项先进技术，并且整合了包括微信公众平台、小程序、腾讯云、微信支付、人脸识别、AI、智慧零售在内的多项核心技术能力。“一部手机游云南”项目通过腾讯的创新技术，构建了平台的多个体系。其中，数字身份体系用于用户行为理解、识别、追溯，精准提升用户体验和服务水平，如基于个人兴趣的旅游产品推荐等；数字消费体系用于支付、营销、服务行为与消费数据的耦合，多维合一；数字诚信体系用于建立游客、商家、政府三方共赢的诚信市场生态；地理信息开放体系用于实现开放、共享、精准的地图应用；基于 LBS 的数字化资源服务体系用于查找厕所、停车位、自驾营地资源等服务。通过实时动态的商业聚合服务，可以让用户随时随地获取所需的各种商业服务，大幅提升用户体验和用户转化率。

2. 服务咨询虚拟助理: 去哪儿网的智能客服——小驼

(1) 解决了什么问题。

消费者可通过去哪儿旅行 App 直接联系小驼, 并得到实时响应。与人相比, 该智能客服能提供全天候服务, 并且不会产生人工误差, 储备的知识无差错。

(2) 使用了哪些 AI 技术。

工程师通过对常见问题进行大数据分析、梳理, 为该智能客服建立了完善的问答模型, 使之能分析用户语义, 并通过自然语言与用户交互。例如, 消费者打出“订”“小驼”时, 系统便会根据订单信息判断消费者意图, 并基于用户行为特征, 准确给出智能联想选项, 引导用户进入标准化客服流程, 高效解决问题。该智能客服具有学习能力, 能通过分析用户的语言, 完善自身的知识库结构, 提升服务水平。

8.20 AI+信息通信

国际电信联盟将 2018 年世界电信和信息社会日的主题确定为“推动 AI 的正当使用, 造福全人类”, 旨在呼吁各成员国重视 AI 在应对贫穷、饥饿、教育不平等和环境破坏等重大挑战方面所具有的巨大潜力, 希望各方积极开展全球对话, 引导 AI 领域的创新发展。

AI 是全球新一轮产业竞争的核心驱动力。

通信智能服务在我们身边解决问题的例子非常多。

比如, 当宽带出现故障时, 技术人员可以通过用户家中光猫设备的指示灯判断故障的类型, 找到排障方案。但技术人员上门服务耗时较长, 教会所有用户看懂指示灯又不现实, 有没有什么方便又快捷的办法?

再比如, 办理手机补卡等高敏感度业务时, 营业厅服务人员会请用户配合做身份验证。但由于人脸的变化, 单凭人眼有时并不能确认业务办理人是否为证件持有人, 这个难题该如何解决?

下面从虚拟助理、通信设备、网络通信软件、用户业务等方面分别介绍 AI 赋能通信服务的具体案例。

1. 华为的 AI 解决方案 SoftCOM AI

(1) 解决了什么问题。

在过去 10 年，网络在提升用户体验的同时，网络部件的分工越来越细，网络的管理运营越来越复杂，网络的复杂度使得网络管理的难度超过了单人的能力，这是当前运营商网络面临的结构性问题。这些结构性问题仅仅依靠压缩费用或者提升人员技能已经不能解决，要用架构性创新解决结构性问题。

SoftCOM AI 解决方案对运营商的价值可以总结为三个倍增和一个提升：三个倍增是运营效率倍增、能源效率倍增、资源效率倍增；一个提升是用户体验提升。运营效率倍增是指将需要大量人工的故障分析定位转为自优自愈；能源效率倍增是指根据话务量关闭空闲载波，或在 DC 中调整 CPU 主频从而降低能耗；资源效率倍增指通过智能选路，均衡网络负荷，提升资源利用率；提升用户体验是指自动生成网络设计、规划、配置，将业务上线的时间缩短为原来的十分之一。

(2) 采用了哪些 AI 技术。

华为 SoftCOM AI 解决方案在原有全云化网络构架的基础上，引入已经成熟的以机器学习为核心的 AI 技术，使网络拥有了智能。整个构架有两个核心：一个核心是 AI 训练平台，用于训练数据，输出模型或算法；另一个核心是智能管控中心，用于收集数据，根据训练平台输出的模型和算法，推理出网络动作指令并执行。整个构架使得网络从自动化业务部署和动作执行，走向智能化的故障自愈、自我优化和自我管理，最终实现永不故障的自治网络。

2. 云脑科技的 AI PaaS 平台

(1) 解决了什么问题。

通信行业的用户行为数据呈现多源、高维的特点，针对用户的业务服务场景也非常多样，针对这样的行业特点，可以借助深度学习，充分挖掘高维数据的潜力，同时灵活应对精准推荐、精准营销等不同的业务目标。

(2) 采用了哪些 AI 技术。

利用无监督学习抽取用户多源、高维的行为特征作为中间特征向量。利用监督学习实现对各类业务的预测或分类。

3. 大象声科单通道语音增强方案

(1) 解决了什么问题。

大象声科单通道语音增强方案基于听觉场景分析及深度学习技术，实现了单通道语音增强的实用化。其芯片级单音轨实时语音提取方案，基于人类听觉系统感知和处

理环境声音的机理，采用深度学习算法进行降噪，有效实现了噪声与人声的分离，以及通话过程的噪声抑制和语音提取、回声抑制、语音增强。该技术可用于 iOS、Android 等操作系统。

(2) 采用哪些 AI 技术。

语音识别、深度神经网络。

8.21 AI+城市公共服务

城市公共服务资源是有限的，而保障公民基本权利所需的公共服务一定是多层次、多方面的，因此，总会有些细节和需求顾及不到或不能得到很好地满足。因此，需要引入社会资源进行补充或提高已有资源的使用效率。随着互联网和移动互联网的快速发展，公共服务的数字化建设和应用也在快速推进。

AI 赋能城市公共服务从大处来看，主要体现在政府对智慧城市建设的投入，治理模式从结果调整型管理向预见型管理转变；从小处来看，主要应用为提高城市公共服务的效率、增加人文关怀、监测与优化公共环境等方面。

下面笔者将从公共事务治理、城市规划、基础设施建设、公共科学研究、环境保护等方面详细介绍 AI 赋能城市公共管理的应用案例。

1. 阿里巴巴联合蚂蚁金服、申通地铁集团推出的售票机

(1) 解决了什么问题。

新一代售票机可以让乘客以语音交互的形式买票，其具有指定站名或票价购票、目的地模糊搜索购票等功能。乘客只需要站在机器旁，告诉机器“我要去中山公园”，机器就会自动调用云端的服务，确定距离目的地最近的地铁站，然后直接出票。

(2) 使用了哪些 AI 技术。

新一代售票机将计算机视觉技术融合到语音识别技术中，解决嘈杂环境下的远程降噪问题。为系统添加了光学摄像头模块，通过识别机器视角内人员的嘴部动作来识别说话人，并判断说话人与机器的距离，进而确定拾音角度。这些视觉信号将会与音频信号相融合，通过一个大的麦克风阵列及配套的软件信号处理器进行抑制干扰与信号增强工作，最终做到准确拾音。

2. 广州南站停车场智慧生活圈服务平台

(1) 解决了什么问题。

利用室内导航系统，旅客在广州南站可以实现反向寻车。在接客时，只需输入接站人员的车牌号码或车位号码，被接者即可通过室内导航系统获得车辆的精准位置及推荐线路。

送客完毕或回到南站自行取车时，可以很方便地查询车辆的停放位置，并可以快速获取取车路线，该系统还有正向引导功能。

当旅客需要停车时，系统可以根据旅客的目的地，以及车场空位热力图，推荐最优停车区域，并提供推荐路线。

旅客停车完毕后，在车库内可快速确定自己的位置，同时可获得前往目的地的推荐路线。

除了停车，在站内到达层，旅客同样可以通过室内导航系统获知自己当前的位置，并获得前往目的地区域的推荐路线。

在该停车场进行支付也非常便捷。旅客在取车前，可以扫码支付停车费，旅客也可以在收费岗通过微信或支付宝扫码支付。停车场还提供无感支付服务，绑定车牌、进行认证并绑定支付宝账户后，即可不停车、免操作支付。

(2) 使用了哪些 AI 技术。

计算机视觉。

3. Facebook 的全球人口分布地图

(1) 解决了什么问题。

该系统利用深度学习技术分析卫星图像数据，目标是制作全球人口分布地图，以此指导公司的无人机项目，帮助更多的人接入互联网。以相对较少的照片（约 8000 张在印度上空拍摄的图片）训练其模型，训练得到的模型能鉴别拍摄自其他 20 个国家的照片，分析照片中人类生活的信息。

该系统已经分析了代表 2160 万平方千米土地的 1560 万张照片。识别人造物的错误率低于 10%。Facebook 之后打算结合其他数据，生成准确的人口分布地图。

（2）使用了哪些 AI 技术。

研究者仅需简单判断样本照片中是否有人类居住的痕迹并进行标注，使用标注后的照片训练神经网络即可。标注时只需要在“是”与“否”之间进行简单的二选一，而不用标注房屋、汽车等人造物。训练结束后，能得到一个简单的分类器，此时部署该系统，它就自动分析卫星图像，判断图像中是否有人造物，再结合普查等数据，最终能生成较为准确的人口密度图。

4. 南加州大学的偷猎预防系统、野生动物安全保护助手（PAWS）

（1）解决了什么问题。

在野生动物保护区等受保护的区域内，大象和其他濒危动物的活动范围可能很大。但是，巡逻员在各个时间段内能巡逻到的范围却十分有限，因此预测偷猎地点与规划巡逻路线以发现偷猎者成为难题。该系统使用机器学习算法对巡逻员以往的巡逻数据进行分析，预测未来可能发生偷猎的地点，PAWS 的预测能力能使巡逻尽可能快速且有效。

（2）使用了哪些 AI 技术。

PAWS 通过机器学习算法学习以往巡逻所获得的数据，预测未来可能发生偷猎的地点。并利用博弈论模型帮助产生随机的、不可预测的巡逻路线，让巡逻轨迹变得更有价值。

5. ET 环境大脑

（1）解决了什么问题。

ET 环境大脑能够对气温、风力、气压、湿度、降水、太阳辐射等信息进行交叉分析，辅助政府、公益机构实现对生态环境的智能监管。还能基于地理位置系统，集成生态环境质量、污染源、污染物、环境承载力等多方数据，融合经济社会、气象水、互联网等数据，综合判断环境质量。

（2）使用了哪些 AI 技术。

ET 环境大脑基于阿里云的云计算、物联网与 AI 能力，实现保护生态环境相关决策的科学化、生态环境监管的精准化、生态环境相关公共服务的便民化。具体实现以下四个功能。

A. 分析研判。ET 环境大脑提供全景式生态环境形势研判，加强对生态环境质

量、污染源、污染物、环境承载力等数据的关联分析和综合研判，强化对经济社会、基础地理、气象水文和互联网等数据资源的融合利用。

B. 智能调度。ET 环境大脑结合中国气象局数据、环保部门环境监测数据，借助海量数据和智能算法，实现对空气质量、水质污染、自然灾害与极端天气的监测和预警。

C. 环境监管。ET 环境大脑综合阿里云 AI 视频解析、图像识别、语音识别、实时计算等技术能力，结合阿里云物联网技术，更快、更准确地监测和预警自然灾害、极端天气及环境风险源。

D. 云上生态。生态环境大数据服务平台通过 API 提供便捷的环境数据服务，让所有人都可以随时获取大气、地面、水体等环境数据。

海洋渔情预测是 ET 环境大脑的一个典型应用，ET 环境大脑将海洋环境大数据、数据智能和 AI 技术应用于金枪鱼的渔情发现中，通过对海面高度、分层水温、叶绿素含量、气象、渔船行动轨迹等数据进行分析，从可能影响金枪鱼产量的众多因素中提取最关键的影响因素，据此搭建出了金枪鱼渔情预测模型，为金枪鱼的发现提供可靠信息，避免盲目捕捞，减少出海成本。

后记：AI 重新定义产品经理

一、站在巨人的肩膀上

要利用 AI 进行赋能，首先要站在巨人的肩膀上。大公司有数据和算法开发人员、算力资源，并且有传统产品累积的数据基础。中小企业可以在大公司设计好的算法和已经研发好的 AI 模型上重构自己的业务以满足自身业务对 AI 的需求。

无论是大公司还是小公司，提出更有深度的问题是 AI 赋能的重中之重，提出问题以后就问题设计出更优的解决方案是 AI 时代对产品经理提出的新要求。

互联网时代到处是需求，只要提出一般深度的问题就可以做出产品，然后就问题进行跟踪即可。

而 AI 赋能需要产品经理站在更高的地方发现更有深度的问题，设计更好的解决方案。这里需要重新思考问题，重新思考设计方案，即重新定义产品经理。

二、跨界运营

AI 时代不再是独守一个功能就可以打败市场的时代，一方面，AI 技术越来越复杂，未来 AI 甚至可以自我学习、自我迭代；另一方面，金融、保险、教育、医疗、工业、娱乐、出行、酒店等多个行业需要用 AI 进行赋能，AI 融入了各行各业，这对产品经理提出了新的要求，产品经理的职责不再仅仅是设计方案并跟踪实现，AI 时代需要产品经理跨界运营 AI 产品，具体步骤如下。

第一步，了解产品所在行业的行业知识、用户特征。

第二步，建设适合行业的推广渠道，制订策划方案。

第三步，制订符合 AI 技术发展趋势的技术方案，自动生产优质内容。

第四步，增加跨界内容，把握跨界的边界，做到 AI 赋能产品的全生命周期。

第五步，适应螺旋式上升的运营特点。

三、软硬一体

AI 产品不再是纯软件的和纯硬件的。

AI 时代的产品经理首先应该明确，只有硬件和软件成为一个整体，系统才能具有智慧。纯软件产品无法端到端地面对用户，纯硬件产品没有一个聪明的“灵魂”，软硬一体才是 AI 赋能的根本方法。

四、自律

自律不是因为懒惰，更多的是因为时间有限，我们不得不做出选择。产品经理要学会隔离自己擅长点和精力点之外的需求事务，相关的具体内容，建议阅读拙作《产品经理进化论：AI+时代产品经理的思维方法》。

五、5G

在 5G 时代，设备进行信息传输时不必再通过有线连接，设备与设备之间的联系更为紧密，在高速网络与 AI 的帮助下，身边的设备也更“懂”用户。怎么懂？懂哪些？有方案吗？这是留给 AI 产品经理的问题。

六、量子计算

随着数据的积累、AI 的崛起，需要 AI 运算处理的数据差不多每 24 个月就增长 1 倍，模型的复杂度是原来的 5 倍，这就需要计算能力强大 10 倍的芯片，摩尔定律已经无法满足需求，就像台积电的张忠谋所讲：“AI 芯片的计算能力在 6 年间增长了 30 万倍，远超摩尔定律”。

AI 对计算能力的需求只有量子计算可以支撑。

