



如何运用 AI 创造 业务价值

12 个真实成功案例

IBM 如何提供帮助

企业可以利用 IBM 深厚的行业、职能和技术专业能力, 丰富的企业级技术解决方案以及基于科学的研究创新来释放 AI、分析和数据的潜能。如需了解关于 IBM Consulting 提供的 AI 服务的更多信息, 请访问 ibm.com/services/artificial-intelligence。

如需了解关于 IBM Software 提供的 AI 服务的更多信息, 请访问 ibm.com/Watson。

如需了解关于 IBM Research® 提供的 AI 服务的更多信息, 请访问 research.ibm.com/artificial-intelligence。

如需了解关于 MIT-IBM Watson 人工智能实验室的更多信息, 请访问 mitibmwatsonailab.mit.edu。



许多关于人工智能的常见概念其实是误导性的神话，只是在众多新兴技术的炒作周期中成为暂时性热点。

摘要

■ 深入 AI 最前线

IBM 商业价值研究院 (IBV) 与 MIT-IBM Watson 人工智能实验室携手合作，针对全球范围内超过 35 个人工智能 (AI) 实施中深度学习项目的参与者开展了一项调研。我们与来自十多个行业的业务和技术专家展开对话，深入了解了他们的 AI 目标、挑战和经验。

■ 止步于锥刀之利还是追求大规模转型

根据调研结果，AI 采用率呈持续增长之势，但大多数组织尚未充分运用 AI 来推动广泛转型。许多组织只是运用 AI 来应对一些孤立的业务挑战。到 2022 年底，我们估计只有四分之一的大型企业将从 AI 试点转为 AI 运营。¹

■ 揭开迷雾，破除 AI 神话

随着许多企业纷纷开始采用人工智能，企业最高管理层和其他领导者务必时刻保持警醒，不要轻易相信 AI 的一些神话，比如“AI 无捷径”、“只有深度学习才是真正的 AI”。他们应当根据 AI 的现状来做出明智的决策。

■ 多个不同行业真实企业的成功经验

本文揭穿了 AI 领域的五个神话，通过数据和真实案例揭示了当前企业运用 AI 的真实现状，让企业领导者及其团队能够学习和借鉴其他企业的实践经验。



引言

感知与现实的差距。

如今,各类媒体经常将人工智能 (AI) 视为缓解经济萎靡不振的灵丹妙药,但企业高管们仍然想了解当前企业运用 AI 的真实现状。这些 AI 采用者运用了哪些方法?取得了哪些成果?

IBM 商业价值研究院与 MIT-IBM Watson 人工智能实验室携手合作,与超过 35 家组织开展了对话,希望从中探索这些问题和其他一些问题的答案。通过这些对话,我们掌握了参与深度学习项目的业务专家和技术专家运用人工智能在现实商业领域创造切实价值的真实图景。

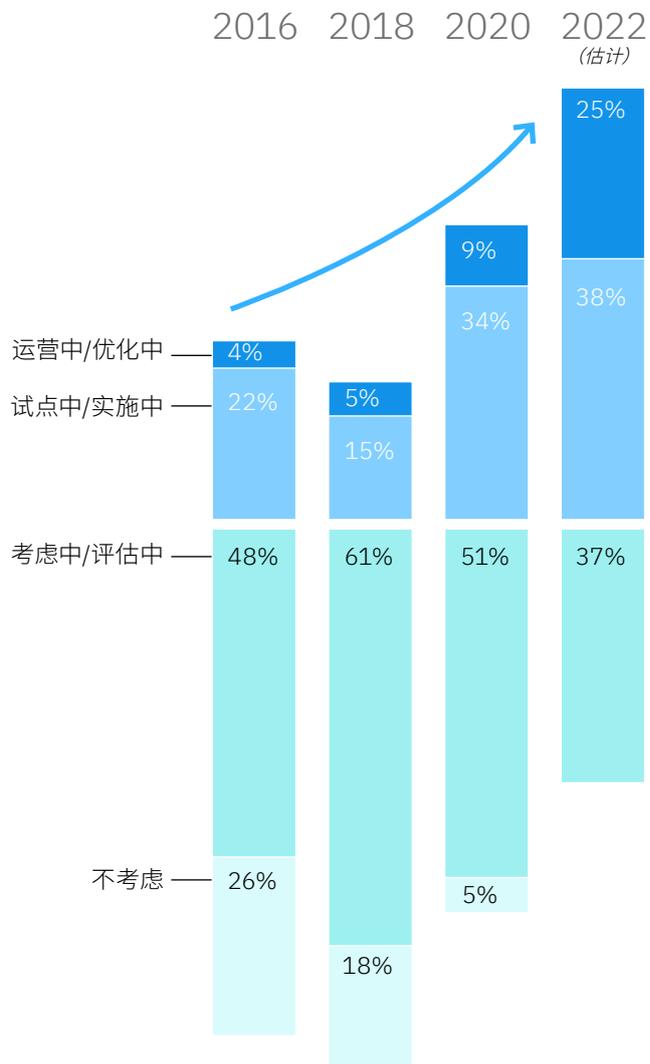
AI: 从概念到实际应用

人工智能在其技术采用曲线或成熟度周期中持续稳步上升 (见图 1)。

图 1

AI 采用* 2016–2022

到 2022 年底,我们认为四分之一的大型企业将从 AI 试点转为 AI 运营。¹



* 注意: AI 采用包括试点中、实施中、运营中或优化中。详见尾注 1。

受新冠疫情驱动，企业找到了 AI 采用的精准方向，并加速了 AI 采用。在新冠疫情期间启动 AI 试点项目的企业数量相比 2018 年增长了一倍多 — 最近的数据表明，此类企业继续保持增长之势。²

尽管调研数据反映 AI 采用呈上升趋势，但并未呈现 AI 采用的全貌，无法让许多业务和技术领导者通过对标分析了解其企业采用 AI 的真实情况。

为了探索这一全貌以及 AI 正在帮助企业解决哪些挑战，我们针对全球范围内参与深度学习项目的个人开展了一项调研。从 2021 年 4 月到 2021 年 8 月，我们与来自十多个行业的业务和技术专家展开对话，深入了解了他们的 AI 目标、挑战和经验（见图 2）。

图 2

我们调研的范围和规模

我们的调研侧重于如何应用定制化 AI 方案来解决不同的业务问题。

机器学习领域



受访者



我们从 AI 的现状中看到了什么？

AI 能否成为营收增长的驱动力？这当然是毫无疑问的。对于一部分创新型 AI 采用者，例如 NVIDIA 和 NavTech 等企业，AI 有助于创造全新的产品，甚至开创全新的商业模式。

但只有极少数企业能够应用 AI 来实现如此大规模的转型。大多数企业都是应用 AI 来解决孤立的实际业务问题。全球范围的组织都在广泛采用 AI 来帮助降低成本、改善客户和员工体验、提高赢单率、优化供应链绩效等等。

我们还发现，许多关于 AI 的常见概念其实是具有误导性的神话，它们在众多新兴技术的炒作周期中成为暂时性热点。不幸的是，这些误解往往会阻碍组织开展更务实的 AI 工作，并分散组织的注意力。

在接下来的报告中，我们将通过从调研中获取的相关见解和实际案例来揭穿五个最普遍的 AI 神话。我们与 35 家 AI 采用者企业的超过 55 位专业人员展开了对话，从中获取了宝贵的观察结果和真实案例，旨在帮助企业厘清事实和假象。这些信息让企业能够深入本质，并与其他同类企业进行对比，从而更加有效地发挥 AI 的影响力和价值。（希望了解详细信息的读者可以从附录中找到 12 个详细的案例研究）。

观点

神话与现实



神话 1

AI 是一把万能钥匙



神话 2

只有深度学习才是真正的 AI



神话 3

降低成本是 AI 的最佳发力点



神话 4

AI 无捷径



神话 5

AI 仅为眼前的问题提供价值

神话 1
AI 是一把万能钥匙

神话 2
只有深度学习才是真正的 AI

神话 3
降低成本是 AI 的最佳发力点

神话 4
AI 无捷径

神话 5
AI 仅为眼前的问题提供价值

附录

神话 1

AI 是一把万能钥匙

现实

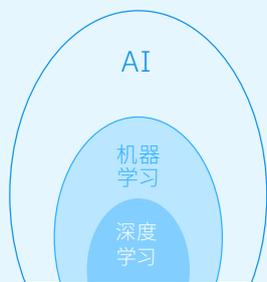
适用性至关重要。AI 驱动的业务改进与许多技术密不可分。

例如，在众多 AI 技术中，深度学习通常最适合用于解决与视觉、语言和其他预测模型中的底层（通常是大型）数据集相关的问题。从虚拟助手到欺诈检测，深度学习正在改变我们的工作和娱乐方式。在这些场景中，传统机器学习技术可能不太有效。

但 AI 并不一定适用于应对每一项业务挑战或实现每一项期望成效，尽管在媒体宣传时经常会过度神化 AI 的适用性。组织首先需要确定 AI 是适用于更广泛的战略计划，还是仅适用于解决特定的业务问题。本文在“[重新思考您的 AI 方法](#)”部分中更全面地探讨了这一主题。企业可以从评估其整体“数据财富”入手，并分析具体的业务问题。



AI、机器学习和深度学习的定义



深度学习是机器学习的一个子集，而机器学习又是人工智能的一个子集。这有点类似于俄罗斯套娃。这些技术通常与其他前沿技术形成互补，包括基于物联网的机器人、传感器和执行器以及虚拟接口等。

AI的概念和起源

AI旨在通过计算机来处理过去只能依靠人力执行的任务。但是，随着人类在准确性、速度和处理能力方面逐渐达到瓶颈，AI才开始真正获得发展动力。

AI诞生于20世纪中叶，但直到21世纪才开始进入技术成熟期。1955年，达特茅斯学院和哈佛大学的两位数学教授以及贝尔实验室和IBM的两位研究科学家建议“在（那一年）夏天开展为期2个月的……人工智能研究”。这项提议在摘要中指出，“将尝试探索如何让机器使用语言，形成抽象和概念，帮助人类解决各种难题，并实现自我改进。”³

这就是人工智能的第一个正式定义。此后，学术界和企业开始竭力构建越来越强大的人工智能。

什么是机器学习？

在麻省理工学院出版社的《深度学习》一书中，作者给出了机器学习的定义：“人工智能系统需要从原始数据中提取模式从而获取自有知识的能力。这种能力被称为机器学习。”⁴换句话说，计算机能够从复杂的数据集中进行学习，并通过不断学习变得更加智能。

如今，机器学习系统已广泛应用于各种领域，包括选择最相关的结果、响应关键字搜索以及解析视觉图像等。在此类人工智能应用中，深度学习是一项日益广泛采用的技术。⁵

深度学习的定义

深度学习是机器学习的一个子集，其灵感来源于人类大脑中的神经网络的运作方式。在当今投入使用的机器学习技术中，最重要的一项技术非深度学习莫属。深度学习具备以下能力：

- 处理图像和自由文本等非结构化数据
- 通过非线性关系建模来解决复杂问题
- 无需在目标任务上预先编程即可学习关系
- 利用持续产生的新数据提高预测能力

对于有足够数据可用的复杂问题，深度学习通常可以提供比其他机器学习方法更加出色的性能。

欧洲服装零售商运用 AI 提高效率和可持续发展能力

需求预测和销售效率始终是消费品和零售行业的核心；即使是些微改进也会对业务产生巨大影响。

服装零售商 BESTSELLER 希望提高其需求预测准确性，并借此充分提高服装产品的销量。当时，这家服装零售商售出了其成品的 78%。这个成绩在充满变数的时尚行业已经算相当不错了。但是，如果 BESTSELLER 可以增加其预测算法中的产品属性粒度，就可以进一步提高效率。

在确定传统分析技术已经达到极限之后，BESTSELLER 团队利用其服装图像作为输入训练了一个卷积神经网络 (CNN)。(CNN 是一种人工神经网络，通常用于分析视觉图像。) 这样一来，BESTSELLER 可以根据未包含在其结构化数据集中的其他特征来对产品进行分类。

通过这些数据输出至核心预测引擎，BESTSELLER 将销售效率提高到了 82%，并将所需的样本设计数量减少了 15%。在全球销售因疫情而持续低迷之际，实现这一改进已难能可贵。此外，该企业还减少了打折、捐赠或丢弃的服装商品，从而积极改善了可持续发展能力。

美国食品调味品公司 McCormick (味好美) 运用 AI 来弥补初级食品科学家在经验上的不足，让他们能够达到拥有 20 年专业经验的高级科学家的水平。

营销平台利用机器学习技术提高响应率

一家营销和广告代理商运用机器学习模型来预测消费者对客户营销活动的接受程度。这家代理商将此功能融入到了其面向所有客户的数据和分析平台中。准确性方面改进一小步，只要规模够大，就可以创造数百万美元的额外销售额，因此回报率相当有吸引力。

代理商发现这项改进可以将响应率提高 20%-30%，但同时也会增加存储、训练和处理额外数据及模型参数的计算成本。所幸的是，在迁移至云端之后，该团队可以更加清晰全面地了解成本状况，从而更有效地管理这些成本。最终，他们保持了较高的响应率，提高了计算效率，并将处理成本降低了约三分之二。

AI 为呼叫中心、食品科学等领域提供助力

法国合作银行集团 Crédit Mutuel 广泛采用深度学习来协助人工呼叫中心客服，每月节省数万小时的工作量。

同样，在另一个完全不同的行业中，美国食品调味品公司 McCormick 运用 AI 来弥补初级食品科学家在经验上的不足，让他们能够达到拥有 20 年专业经验的高级科学家的水平。

我们调研中的其他一些案例重点展示了如何运用定制化 AI 方案来解决不同地理区域、行业甚至职能领域的不同业务问题。通常，在选择正确的数据集来解决问题之后，正确的方法也将变得更加清晰可见。BESTSELLER 和营销代理商的成功经验就充分说明了这一点。

神话 1
AI 是一把万能钥匙

神话 2
只有深度学习才是真正的 AI

神话 3
降低成本是 AI 的最佳发力点

神话 4
AI 无捷径

神话 5
AI 仅为眼前的问题提供价值

附录

神话 2

只有深度学习才是真正的 AI

现实

大型企业正在运用混合数据科学、传统机器学习、深度学习和预处理技术来解决各种不同的业务问题并实现重大业务价值。

在过去十年中，AI 研究中的许多进展都发生在深度学习领域。随着社交媒体、搜索、零售和流媒体的兴起，以及在商业模式中嵌入深度学习的其他 B2C 平台的爆炸式增长，坊间开始流传关于 AI 的这样一个神话 — 只有深度学习才是人工智能。

实际上，深度学习只是企业分析工具箱中用于实现 AI 的众多工具之一（参见第 10 页的图 3）。

从概念上来说，对深度学习成本的担忧可能会导致 AI 研究的未来方向和性质面临重大挑战（请参阅第 11 页的“成本会导致深度学习消亡吗？”）。在实际应用中，企业通常可以通过比较概念验证或试点中的实际结果来确定深度学习有哪些不足。

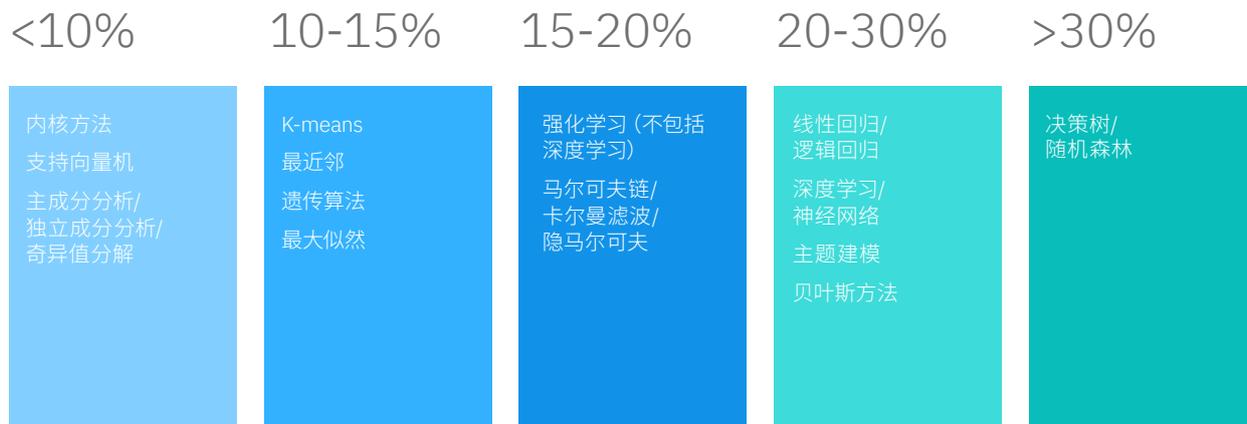


图3

深度学习并非放之四海而皆准

组织采用的机器学习技术因具体业务问题而异。

使用每种机器学习技术的组织百分比



信息来源: 2021 IBVAI 能力调研, 未公布的数据。问 16A. 您的组织采用了哪些机器学习 (ML) 技术?
请选择所有适用项。

毕马威运用深度学习和其他分析功能帮助客户节省数百万美元

全球税务、审计和咨询公司毕马威发起了一场内部黑客马拉松, 旨在确定如何采用最优方式来减少客户研发项目、投资和税款抵免等归档流程中所涉及的手动工作量。这一做法可以帮助客户减少每年的纳税额, 从而创造切实的业务价值。该公司发现, 使用即开即用的文档发现软件 (其效果与手动关键字搜索相当) 可以实现 55% 的准确率, 而使用深度学习可以实现超过 70% 的准确率。

最优方式则是采用基于规则的机器学习, 可实现超过 85% 的准确率。通过自动处理这些流程, 毕马威可以帮助特定客户提高成本效益, 每年节省数百万美元的税费支出。采用这种方法之后, 毕马威的一位客户在其研发支出上获得了额外 40% 的税收抵免。

成本会导致深度学习消亡吗？

神经网络自 1950 年代就已经存在。深度学习经历了 AI 寒冬的投资饥荒后幸存下来⁶，从 2000 年代后期至今，它一直沐浴在夏日的阳光下。

近段时间以来，随着数据处理能力和计算能力的大幅提升，再加上结构化和非结构化数据的快速增长，深度学习的应用日益广泛。

随着数据持续呈指数级增长，再加上业界预期摩尔定律即将走到尽头（如果还没有的话），一些 AI 研究人员担心维持这一趋势所需的财务和环境成本将大幅增长。在 IEEE Spectrum 于 2021 年发布的一篇文章中，Neil C. Thompson 等人就指出：“显然，如果使用更加强大的计算能力来构建更复杂的模型，并用更丰富的数据进行训练，您就可以利用深度学习实现更加出色的性能。但是，需要多高的成本才能维持这种计算能力呢？成本是否会上升至阻碍技术发展的程度？”⁷

例如，据报道 OpenAI 的 GPT-3 在开发和训练上花费了 300 万美元，而 Alphabet 子公司 DeepMind 的 AlphaGo 仅在训练上就花费了 3500 万美元。

成本如此巨大且呈迅速增长之势，企业也将面临更加两难的处境——一方面要满足对于更复杂的模型、更多的数据和训练以及更强大的计算能力的需求，另一方面又要接受预算和效率的固有业务现实，应当如何在这两者之间找到合理的平衡点？研究人员必须要解决这个难题，否则技术发展可能会陷入停滞。⁸

各种研究组织正在探索应对之策，包括采用不同的硬件解决方案和全新的 AI 学习方法，以及寻找创新方式来融合数据和参数丰富的强大深度学习与基于推理和规则的经典符号技术。

这些研究工作前途未卜，正如 IEEE Spectrum 的一篇文章做出的总结：“虽然深度学习如同飞火流星般急速兴起，但其未来可能会充满坎坷。”⁹

与此同时，组织在使用深度学习时需要密切关注成本与性能之间的平衡点——尤其是要参照其他 AI 工具。

Zzapp Malaria: 使用 AI 造福世界，而不仅仅是惠及企业

2020，疟疾造成约 62.7 万人死亡，其中 96% 发生在非洲。¹⁰ 2021 年度 XPRIZE AI 优胜者 Zzapp Malaria 致力于开发 AI 驱动的方法来抗击疟疾，并通过专有移动应用直接将方案传达至实地。

鉴于现有的卫星图像识别技术无法有效识别小型水体，Zzapp Malaria 开展了一个试点项目，旨在利用卷积神经网络来分析视觉图像以检测小型水体（这有可能是携带疟疾病毒的蚊子的滋生地）。该方法实现了大约 75% 的准确率，但无法足够清晰地确定哪些因素可以改善预测。虽然这种方法的效果不错，但还不够好，无法扩展至其他地理位置。

作为改进方案，该团队利用 CNN 从图像中提取 50 个地形和其他特征，并将这些特征用于传统的线性回归方法，以确定出现积水的可能性。其效果与之前的方法相当，但能够更加清晰透明地确定哪些因素可以改善预测。这样一来，团队就可以更加轻松地解释数据，也就更易于扩展至地形差异较大的地域。该团队运用这种 AI 驱动的成功经验来调整其方法，不断扩大适用范围，以帮助降低其他地区的疟疾发病率。





神话 1
AI 是一把万能钥匙

神话 2
只有深度学习才是真正的 AI

神话 3
降低成本是 AI 的最佳发力点

神话 4
AI 无捷径

神话 5
AI 仅为眼前的问题提供价值

附录

神话 3

降低成本是 AI 的最佳发力点

现实

运用 AI 来解决业务问题确实有助于降低成本,但这并不是 AI 的唯一作用。领先的组织正在从战略层面积极运用 AI 来建立差异化优势,并通过提升流程效率、推动业务增长和开展商业模式创新来释放营收潜力。

成本固然重要,但增长、创新和社会公益更加重要。IBV 研究表明,许多组织一直将以客户为中心的增长列为 AI 对业务影响最大的首要领域(参见第 15 页的图 4)。

我们想知道这种对营收的关注是否更多出于愿望,而非现实。不过,在调研期间与高管对话的过程中,我们发现了几家成功将愿望转化为现实的企业。



图 4

AI 价值驱动力 2016–2020

企业专注于以客户为中心的营收收入增长。



信息来源：见尾注 11。

AI 推动保险合资企业实现营收增长

总部位于印度的通用保险合资企业 IFFCO-Tokio 决定在客户提交获批索赔后直接向客户支付维修费用，从而改善客户体验。

第一步是更好地捕获碰撞受损汽车的图像。然后，团队运用深度学习对汽车型号、受损的零部件和损坏类型进行分类。AI 系统可以确定零部件是需要维修还是更换，

并提供成本预估，同时让人工评估员参与其中以降低欺诈风险。

最终，项目取得了巨大成功。IFFCO-Tokio 只用了不到一年时间就收回了成本，不仅将结算成本降低了 40%，将客户接受率从 30% 提高至 65%，而且还提高了客户满意度、客户保留率和客户获得率。AI 不仅是提高效率的利器，更是收入增长的强力助推器。

神话 1

AI 是一把万能钥匙

神话 2

只有深度学习才是真正的 AI

神话 3

降低成本是 AI 的最佳发力点

神话 4

AI 无捷径

神话 5

AI 仅为眼前的问题提供价值

附录

神话 4

AI 无捷径

现实

尽管 AI 模型的应用场景因行业和部门而异,但越来越多的“现成”基础模型和预训练模型可以为企业数据科学家提供更具成本效益的切入点。

自 2016 年以来,IBM 商业价值研究院 (IBV) 持续采用系统化方法来量化企业 AI 的各种趋势。2020 年 IBV 的一项意外发现就是,“技术可用性”再次成为 AI 采用的头号障碍(2018 年“技术可用性”曾让位于“技能和其他因素”,参见第 17 页的图 5)。我们不禁要追问:技术可用性为何再次成为 AI 采用的头号障碍?

我们对此的解释仍然是,组织最终意识到掌握适当数据技能的员工是让 AI 技术发挥效用的必要条件,但只有这一项条件是远远不够的。许多企业都聘请了数据科学家来训练各种数据集,这些数据科学家们也在尽职尽责地完成这项任务。但是,针对每个业务问题,企业通常都需要采用与之前不同的 AI 模型,似乎都要从零开始。企业无法通过简单有效的方法来重复利用过去的成果。



图 5

AI 障碍

2016–2020

技术可用性再次成为 2020 年 AI 实施的头号障碍。

2016 2018 2020

AI 障碍

技术可用性

技术型人才或技术专业能力的可用性

监管限制

高管支持程度

数据和信息使用方面的法律/安全/隐私问题

跨企业边界以及与外部合作伙伴共享的数据治理和政策

应用于决策和了解决策情景的数据量/数据可用性

客户就绪程度

合作伙伴或利益相关者的就绪程度

组织认同/就绪/文化契合度

信息来源：“Shifting toward Enterprise-grade AI: Confronting skills and data challenges to realize value.” IBM 商业价值研究院。2018 年 9 月。 <https://www.ibm.com/thought-leadership/institute-business-value/report/enterpriseai>。图 1，AI 实施面临的障碍：2016 年与 2018 年对比，AI 的业务价值，未发布的数据。问题 9. 您的组织在实施人工智能方面面临的障碍是什么？选择前 5 项。

最近，一些 AI 捷径方案开始帮助组织发挥其 AI 驱动的解决方案的优势。AI 领域的预训练模型和基础模型就相当于软件领域中的“现成”方案。这些模型可以为启动新的 AI 项目提供更加行之有效的切入点。

那么这是如何实现的呢？具体来说，组织可以利用从解决一个问题中收集的 AI 知识来解决其他相关问题，因此不必生成全新的数据集就可以建立更加成熟的模型。这种方法的关键是迁移学习，也就是将最初专门针对一项任务而训练的模型改用于处理不同的任务。例如，组织可以将用于识别汽车的模型应用于识别卡车。

业界已经设计了许多不同类型的预训练模型来解决一个或一些特定的业务问题，而且越来越多的大型通用模型（例如，Alphabet 的 BERT、OpenAI 的 GPT-3）现已可

用于应对各种挑战。目前，中国“悟道 2.0”模型的参数规模已经率先达到万亿级别，位居全球第一。

基础模型可以通过以下重要方式创造价值：

- 改善经济性：跨多个用例摊销成本
- 改进成果：利用更加庞大和强大的数据集来提高准确性
- 新功能：能够更有效地整合多模式数据。

但实际情况远非如此简单。调整预训练模型有时会导致新数据训练性能大幅下降。美国医疗设备制造商 Boston Scientific（波士顿科学）就遇到并成功解决了这一问题。

Boston Scientific 仅投入 5 万美元就实现了 500 万美元的成本节省

Boston Scientific 希望实现自动化的支架检查流程，从而提高查找缺陷（例如连接断裂或表面缺损）的准确性。准确的检查对于临床治疗的成败关系重大。美国食品和药物管理局根据对患者的风险来严格监管“逃逸率”（也就是漏检的缺陷部件的比例）。

Boston Scientific 的数据科学经理 Eric Wespi 评论说：“人员目测检查通常速度缓慢，成本高昂，还可能带来不必要的质量风险。”该公司配备了大约 3000 名专家负责目测检查工作，每年要花费数百万美元的成本。

Boston Scientific 实施了一个基于规则的自动化系统，可以使用尺寸测量和其他方法来捕捉常见问题。该团队将系统调整为相对保守，即设定可忽略不计的误报率。然而，5%-10% 的误报率仍然过高。许多符合质量要求的零件被标记为缺陷品。

卷积神经网络 (CNN) 支持分析视觉图像，因此特别适合解决此问题，但这样的模型需要大量的数据。该团队没有足够的数据从零开始训练这些模型。该团队还认识到，收集或生成这些数据不切实际，而且成本过高。

解决方案？首先，该团队专注于粒度更小、范围更窄的任务，从而缩小问题的范围。然后，他们利用已有的现成开源 AI 模型来应对重新定义的挑战。最后，他们使用较小的数据集来进一步优化此系统。

成效如何？该公司只投入了大约 5 万美元，就直接节省了 500 万美元的成本，同时还提高了准确性。

企业领导者应当谨记，基础模型和预训练模型在节省时间和资金方面成效斐然，但并不一定适用于帮助企业建立差异化优势。这些模型（其中一些是开源模型）对于所有组织都公开可用，因此只适用于解决对差异化优势影响不大的业务问题。或者，组织可以专注于定制工作，通过引入更多数据（通常是专有数据或专门集成的数据）来提升竞争优势。

组织最终意识到掌握适当数据技能的员工是让 AI 技术发挥效用的必要条件，但只有这一项条件是远远不够的。

神话 1

AI 是一把万能钥匙

神话 2

只有深度学习才是真正的 AI

神话 3

降低成本是 AI 的最佳发力点

神话 4

AI 无捷径

神话 5

AI 仅为眼前的问题提供价值

附录

神话 5

AI 仅为眼前的问题提供价值

现实

企业内部和企业间出现的 AI 网络效应正在推动整个企业创造真正的业务价值。

得益于不断增长的数据源和日益增强的数据利用能力，组织积累了越来越多的数据财富。通过从战略层面推动深思熟虑、合乎伦理的 AI 项目，企业不仅将获得丰厚的经济回报，而且也将积极拥抱开放式创新。而这又进一步提升了规模经济效益，一些领先的 AI 采用者更是成效显著。正如我们在 2020 年的 AI 业务价值报告中所指出：

“网络效应（即使只是发生在企业内部）似乎进一步放大了 AI 投资的效益。初步分析表明，组织在其业务运营的一个领域投资发展 AI 往往会增强其他领域的适应性和弹性，从而实现相应的财务收益。例如，针对一个职能中的数据治理和访问策略的改进可以扩展至相邻的其他职能，从而融合为跨工作流程合作和协作的一部分。组织中的核心或骨干职能（例如财务、IT 或人力资源）具有特别强大的跨组织影响力，这些职能中的 AI 投资尤其能够发挥这种网络效应。”¹²



例如，将 AI 人才从一个部门或项目轮换到另一个部门可以实现专业知识的交叉赋能，并为整个企业中的人员提供持续组织学习的机会。这种方法有助于提高整体 AI 敏锐度，避免因人才问题而导致项目停滞不前。

AI 与其他数字化转型技术（例如云、物联网、安全和数据管理）之间的网络效应和协同作用进一步提升了 AI 的价值创造能力。¹³

与许多新兴技术一样，我们看到 AI 在企业内部催生的效应也开始在各个企业之间显现。

NVIDIA 助力推动汽车市场的开放式创新

科技公司 NVIDIA（英伟达）的业务模式创新方法展示了如何将 AI 应用延伸至客户和业务合作伙伴。一些汽车企业并不具备自主研发的经验、硬件和数据。为了帮助这些企业应对自动驾驶汽车领域的巨量计算挑战，NVIDIA 正在创建一系列由 AI 驱动的共享功能：

- 覆盖多个客户的公共数据平台
- 用于训练和测试的模拟
- 视觉任务的通用处理。

汽车制造商可以根据自己的需求和现有能力来选择租用自动驾驶汽车硬件，基于更庞大的数据集来训练其模型，或者也可以选择使用 NVIDIA 预先训练的模型。无论是哪种情况，汽车制造商都无需投入大量资本来发展硬件和 AI 开发能力，而是可以将技术能力转化为运营支出，并且随着硬件和软件的逐步改进而不断受益。

事实证明，AI 在各个机构中的应用最终发展成为了推动经济发展的强大力量。

许多蓬勃发展的 B2C 平台（连同硬件/软件公司、学术机构和政府）投入了大笔资金来推动 AI 研究，此类 AI 项目通常针对公共领域。此外，这些平台也一直在稳步培养其高管团队和远程知识工作者团队（自由工作者）中可迁移性的技能，从而充分利用他们在 AI 采用前沿获取的直接经验。

新冠疫情期间，办公模式发生了剧变，从而加速释放了专业人才流动的潜能。疫情催生的“大离职潮”反而释放出了推动改变全球经济力量的强大动能。

面临哪些障碍？

知识扩散（知识和人才的传播）并非总是与“吸收能力”（组织适应和整合这些洞察及技能的能力）协同一致。¹⁴ 制度障碍、僵化管理以及对变化的抵制都可能会阻碍企业采用 AI。

只有在传统的企业领导者摒弃对过去工作方式的怀旧观念时，AI 的变革性价值才能成为现实，才能充分发挥 AI 的财务、经济和社会影响力。他们需要从战略层面做出深谋远虑的思考，并落实到实际行动上，才能充分把握未来的创新机遇。

要迈出第一步，最重要的一点就是拨开迷雾，摒弃固有观念，认识到并积极拥抱 AI 的最新真实现状。

延伸阅读

自 2020 年底以来，IBM 商业价值研究院发布一系列关于构建顶级 AI 能力的研究报告。这一系列报告从全局和企业的观点探讨了 AI 发展趋势，并融合了许多必要的相关主题，深入剖析了采用 AI 可以实现哪些财务和经济价值。

每一篇研究报告都提出了多项针对特定主题的具体建议，并针对在采用 AI 业务实践方面具备不同成熟水平的企业进行了量身定制。

我们建议您查阅这些切实可行的行动指南，其中凝聚了数十个 AI 项目的成功经验以及数百位 AI 从业者及相关专家的专业知识和智慧结晶。具体列表如下：

- **战略与愿景:** [重新思考 AI 方法](#)
- **数据和技术:** [应对 AI 数据困境](#)
- **工程和运营:** [扩展 AI 的公认概念](#)

关于 作者



Nicholas Borge

FutureTech 研究员,
麻省理工学院计算机科学与人工智能实验室
njborge@mit.edu
linkedin.com/in/nicholasborge

Nicholas Borge 是麻省理工学院 FutureTech 项目团队的成员, 致力于研究 AI 的经济性和未来型工作。Nick 曾在 Sony Music 担任智能自动化总监, 创立过一家 AI 初创公司, 并在面向财富 500 强企业的战略和技术咨询领域拥有超过 11 年的经验。Nick 拥有麻省理工学院的工程和管理硕士学位, 并且是麻省理工学院系统设计和项目管理项目的研究员。

Subhro Das, PhD

MIT-IBM Watson
人工智能实验室研究员
subhro.das@ibm.com
linkedin.com/in/subhrodas/

Subhro Das 是 IBM 研究院 MIT-IBM Watson 人工智能实验室的研究员。作为实验室的首席研究员, 他负责与麻省理工学院合作开发新型 AI 算法。他的研究兴趣广泛涉及机器学习优化方法、强化学习、可信赖的机器学习以及以人为中心的 AI 算法。他拥有卡内基梅隆大学电气和计算机工程专业的硕士和博士学位。

Martin Fleming, PhD

Varicent 首席收入科学家
martin@fleming41.com
linkedin.com/in/flemingmartin

Martin Fleming 在 Varicent 担任首席收入科学家。Varicent 是总部位于多伦多的一家销售绩效管理软件提供商。Martin 还是生产力研究所 (Productivity Institute) 的研究员, 这是一个由八所英国大学组成的联盟组织。他的研究方向是技术、生产力和经济学的交叉领域。他即将出版的《Breakthrough, A Growth Revolution》一书的作者。此前, Martin 曾担任 IBM 的首席经济学家和首席分析官。

关于 作者



Brian Goehring

人工智能全球研究总监
IBM 商业价值研究院
goehring@us.ibm.com
linkedin.com/in/brian-c-goehring-9b5a453/

Brian Goehring 是 IBM 商业价值研究院的副合伙人，负责领导 AI 商业研究议程，与学者、客户和其他专家合作共同发展数据驱动的思想领导力。他拥有 20 多年的战略咨询经验，涵盖大多数行业和业务职能的高级客户。他拥有普林斯顿大学的哲学学士学位，并持有认知研究和德语证书。

Neil Thompson, 博士

FutureTech 总监，
麻省理工学院计算机科学与人工智能实验室
neil_t@mit.edu
linkedin.com/in/neil-thompson-5724a614

Neil Thompson 是麻省理工学院计算机科学和人工智能实验室的 FutureTech 研究项目总监，也是麻省理工学院数字经济计划的首席研究员。此前，他曾担任麻省理工学院斯隆管理学院创新与战略助理教授以及哈佛大学创新科学实验室客座教授。他曾在劳伦斯利弗莫尔国家实验室、贝恩公司、联合国、世界银行和加拿大议会等组织任职。他拥有加州大学伯克利分校的商业和公共政策博士学位和计算机科学和统计学硕士学位，以及伦敦经济学院的经济学硕士学位。

合作者

Adam Bogue
IBM 研究院业务开发主管

Alex Gorman
IBM 软件部客户支持项目总监

Cathy Reese
IBM Consulting 高级合伙人、业务总监

Shannon Todd-Olson
IBM Consulting 高级合伙人

致谢

作者和合作者诚挚感谢 MIT-IBM Watson 人工智能实验室及其联合总监 Aude Oliva 和 David Cox 为本项目提供资金，以及 Seth Dobrin、Glenn Finch 和 Sriram Raghavan 提供的大力支持。



附录

详细案例研究的范围和规模

各个行业和职能领域的案例研究
以及所采用的机器学习技术。

企业访谈

名称	行业	职能	业务解决方案	机器学习技术
BESTSELLER	消费品	时尚设计	从目录图像中提取产品属性，从而提高设计和销售效率	视觉
Boston Scientific	工业品	医疗器械设计	使用迁移学习自动对支架执行视觉检查，从而降低人工成本	视觉
Crédit Mutuel	银行	客户服务	使用分层 NLP 为客户顾问提供更好的答复建议，从而加快通话流程	语言
Global Bank	银行	内部审计	运用“即时校对器”改善文档质量，从而提升审计能力	语言
IFFCO-Tokio	保险	理赔自动化	通过自动评估直接向索赔人付款，从而减少保险支出	视觉
KPMG	专业服务	税收抵免	通过文档搜索改进文档编写，从而增加税收抵免	语言
Marketing Platform	专业服务	精准广告	通过公开试验的边际激励来降低训练精准定位模型的成本	其他
McCormick	消费品	研发/产品设计	为实验提供初始口味特征建议，从而提高研发效率	其他
Navtech	信息技术	销售	构建计算机视觉平台来支持访问数字产品目录	视觉
NVIDIA	信息技术	自动驾驶	汇集数据并提供 AV 技术即服务，从而开创全新的商业模式	视觉
Suncor	能源	站点运营	预测不利的加工条件，建立针对柴油生产问题的早期预警	其他
Zzapp	信息技术	公共卫生	使用卫星图像识别积水以消灭疟蚊	视觉

信息来源: Neil Thompson 博士: <http://www.neil-t.com/teaching-cases/>

附录

BESTSELLER

Boston Scientific

Crédit Mutuel

Global Bank

IFFCO-Tokio

KPMG

Marketing Platform

McCormick

Navtech

NVIDIA

Suncor

Zzapp

BESTSELLER (绫致时装)

使用 AI 发挥企业数据的价值

概述

想要进行需求预测，就必须为算法提供产品特征信息。可用的信息越多，可捕获的各种历史需求模式变化就越多，预测准确性就越高。

但是，产品属性的颗粒度越细，其评估过程就越困难、越耗时。深度学习提供了解决方案，可以快速准确地对产品分类，同时最大程度减少人工干预，从而增加可用于预测算法的特性。时装公司 BESTSELLER 的经历就是 AI 发挥作用的生动体现。

机遇：减少浪费，缩短周转时间

在时装行业，每年两个销售季可售出约 80% 的商品，剩下的商品大都以较大折扣卖出，或者最终捐赠或丢弃。这种过度生产会对利润产生不利影响，还会导致严重的可持续性问题。

BESTSELLER 为印度市场设计、制作和销售服装。针对旗下的 4 个品牌，他们设计并制作 3500 个样品，但最终仅有 1100 个入选进入生产流程。这些成功的候选产品被分为 5000-6000 个最小存货单位 (SKU)，例如，按颜色和尺码等分类，最终生产出 150 万件成品。BESTSELLER 可以售出约 78% 的产品，这个成绩在时装行业已经算相当不错了，但仍有机会更有效地按照客户喜好调整生产，百尺竿头更进一步。该企业计划到当年底，将品牌组合增加一倍以上（由 4 个增加到 9 个），销售率的提高可对盈利能力产生显著影响。

BESTSELLER (续)

挑战:可用于分析的设计元素有限

BESTSELLER 希望更深入地了解有助于推动特定产品销售的因素。这样就可以在设计过程中利用这些信息，还可以确保销售的产品数量与生产的产品数量相匹配，从而提高销售效率。不仅如此，这些信息还有助于提高设计效率。然而，他们对产品属性（例如，颜色和尺码）、库存和地点的数据进行初步分析后发现，根本没有足够的信息可用于做出有意义的推断。该团队需要更丰富的数据集。

衣服可以按照廓形、剪裁、面料、款式和各种设计要素来描述。事实上，BESTSELLER 使用的分类法包括 7000 多种设计样式和 4000 多种颜色。这些特征中的大部分都可通过观察产品图像来识别，但这类信息很少在产品主数据中得到标记。BESTSELLER 需要一种方法，快速高效地提取这些信息。

解决方案:分析图像以丰富可用特征

解决方案是使用计算机视觉技术，直接从图像中提取更多特征。BESTSELLER 拍摄了 10000 张照片（一个销售季目录），并为所有 4 个品牌开发了模型。在短短三周内，该公司的团队成功开发并训练了卷积神经网络 (CNN)，能够根据各种特征对图像分类。这些深度学习衍生的功能可与传统分析方法（例如，回归分析或主成分分析）结合，以便更好地了解推动销售的因素。

成果:改进设计采样与销售效率

即使在全球销售因疫情而持续低迷之际，BESTSELLER 在过去一年半中的销售和设计效率还是实现了显著提升。销售效率提升至 82%（从 78% 增加了 4%），企业为每个品牌创建的设计样本数量减少了 15%，而最终选择的设计数量并没有任何减少。这样一来，设计人员就能够关注于数量更少、采用可能性更大的设计，从而提高采样效率。

附录

BESTSELLER

Boston Scientific

Crédit Mutuel

Global Bank

IFFCO-Tokio

KPMG

Marketing Platform

McCormick

Navtech

NVIDIA

Suncor

Zzapp

Boston Scientific (波士顿科学)

避免迁移学习的陷阱

概述

迁移学习 (transfer learning) 是指改变最初为一个任务而训练的模型的用途, 并将其用于其他任务。从这个意义上说, 在解决一个问题时获得的“知识”可用于解决另一个相关问题——例如, 用于识别小轿车的模型可应用于识别卡车。

迁移学习有助于节省工作, 降低训练成本, 但也可能造成 AI 能力大幅下降 (最高可达 45%), 因此适合这种方法的应用范围比较有限。然而, Boston Scientific 的经历表明, 通过“逐步缩小”问题范围, 迁移学习仍可以让 AI 有惊艳的表现, 能够实现超过 99% 的模型能力, 并节省超过 500 万美元的人工成本。

机遇: 心脏支架的质量检查对于确保患者安全意义重大, 但成本也非常高昂

Boston Scientific 致力于为各种外科应用领域生产心脏支架。团队需要对心脏支架进行检查, 确保不存在连接断裂或表面缺损等缺陷。准确的检查对于临床治疗的成败关系重大, 因此美国食品和药物管理局根据对患者的风险严格监管逃逸率 (也就是漏检的缺陷部件的比例)。

传统做法是由人类专家完成大部分检查, 但这并非最佳选择。Boston Scientific 的数据科学经理 Eric Wespi 评论说: “人员目测检查通常速度缓慢, 成本高昂, 还可能带来不必要的质量风险。”这个道理浅显易懂: 如果要求人们长时间集中注意力, 以发现小概率事件, 他们的表现通常不太理想。此外, 做出的判断因人而异, 专家的时间成本也很高。Boston Scientific 有大约 3000 名专家负责检查, 每年要花费数百万美元。

挑战: 图像分类需要大量数据支持训练

Boston Scientific 实施了一个基于规则的自动化系统, 可以使用尺寸测量和其他方法来捕捉常见问题。该团队将系统调整为相对保守, 即设定可忽略不计的误报率。然而, 5%-10% 的误报率仍然过高。许多符合质量要求的零件被标记为缺陷品, 要留待人工检查员处理。

Boston Scientific (续)

卷积神经网络 (CNN) 特别适合图像分类, 但这样的模型需要大量的数据以支持训练。在许多情况下 (尤其对于比较新和比较罕见的缺陷), 团队没有足够的数据用于从头开始训练这些模型。收集或生成这些数据不切实际, 而且成本过高。

解决方案: 应用迁移学习, 逐步缩小问题范围

团队想知道, 如果使用预先经过训练的模型, 是否有助于改善结果。他们应用了以下方法:

1. **缩小问题范围:** 针对每个缺陷, 将检查工作划分为粒度更小、范围更窄的任务 (例如, “图像的这部分是否包含连接?”, 或者 “这个连接是否断了?”)。
2. **定制现有模型:** 使用数个开源 CNN, 例如 VGG16、EfficientNet (B0 到 B7)、Mask R-CNN、YOLOv3、ResNet-50 和 Inception-v3。对于每种情况, 该团队首先根据开源模型的预先训练的权重, 定制最后几个网络层, 然后使用自己的数据重新训练模型。
3. **测试数据需求:** 该团队发现, 他们只需远少于预期的样本 (例如, 每个缺陷只需 100-1000 个样本, 另外需要 5000-6000 个无缺陷心脏支架的样本) 就可以超越人工水平。

为了提高模型的强健性, 他们还通过扰动因素生成额外的样本, 以扩充训练数据。这种扰动只是简单的调整, 不会影响分类, 例如, 调整亮度或增加噪声。

所有工作都是在相对有限的 5 万美元的预算内由 3 人小团队完成的。模型训练速度快, 成本低。对于 9 个模型, 仅需 1-2 秒就可处理一张图片, 在单个 GPU 上训练每个模型只需 2-10 小时。

成果: 模型表现优异, 人工成本降低

所有模型的最终准确度均高于 90%, 较小的网络 (例如 VGG16) 在解决简单问题时的表现甚至更为出色。通过使用更复杂的模型和更多的数据, 准确性可进一步提升, 例如, 对于使用 100 个示例的 B0 网络, EfficientNet 的准确度可达 97%; 对于使用 1000 个示例的 B7 网络, 准确度可达 99% 以上。

这种表现水平大大超过我们对迁移学习结果的传统认知。通常, 迁移学习的准确性会显著下降, 需要更多数据来抵消负面影响。而在本例中, 在将现有模型应用于较为简单的问题时, 似乎已不需要这种操作。

部署这 9 个模型相当于直接节省了 500 万美元的人工成本, 因为由系统标记为缺陷而需要人工检查的部件数大幅减少, 从而可将多名专家解放出来, 重新安排至其他高价值项目。

Boston Scientific 的成功经验表明, 迁移学习在适当的条件下效果良好:

- 可以使用现有的通用模型。对于图像处理任务, 即使任务有较大的差异, 这种网络的早期层的可迁移性似乎也非常高。
- 使用系统来处理较简单的问题, 然后在网络上进行优化, 可以有效避免迁移学习引起的性能下降。

附录

BESTSELLER

Boston Scientific

Crédit Mutuel

Global Bank

IFFCO-Tokio

KPMG

Marketing Platform

McCormick

Navtech

NVIDIA

Suncor

Zzapp

Crédit Mutuel

(法国国民互助信贷银行)

使用 AI 为客户顾问提供适当的信息

概述

提升客户服务效率和加深客户关系往往是“鱼与熊掌不可兼得”。专业人员能够比较有效地解决与特定产品或服务相关的问题，但在面对超过单一领域的更广泛的客户关系时，他们可能缺乏必要的客户服务背景信息。Credit Mutuel 使用 AI 为客户代表提供精心整理的信息，为涉及多个产品的客户扩展专用的联系点。

机遇:提高人类顾问的服务质量

在 Crédit Mutuel 银行，每位客户都有一名专门的顾问提供服务。顾问作为第一联系人，帮助客户就各种产品（例如支票、储蓄、抵押贷款和投资）协调与该银行的关系。如果顾问能够更迅速、更方便地访问相关信息，他们就能更快地响应客户需求，从而省出更多时间，为其他客户提供服务。该银行每月会接到约 300 万个电话和 700 万封电子邮件，对解决时间的任何改进都会产生重大影响。

挑战:各种产品和团队的文档不一致

让一名顾问负责众多产品的挑战在于，他们需要迅速获得相关产品的必要信息，而这会带来沉重负担。为了解决客户查询，顾问（通常是多面手）使用内部搜索引擎或电话来获取有关特定产品的答案。但是，Crédit Mutual 银行网络中各个银行的信息组织方式各不相同，这加大了搜索的复杂程度。此外，语言和术语也可能不同。这意味着传统的现成语言模型不足以确定提供给这些顾问的信息的优先顺序。

Crédit Mutuel (续)

解决方案: 自定义文字嵌入和分层分类

为了创建为产品自定义的语言搜索, Crédit Mutuel 银行首先收集了客户顾问在 3-4 个月内面临的所有问题, 然后整理这些问题的答案 (又花费了 4 个月的时间), 针对当前生产环境中的所有 11 个业务领域重复这项工作。然后, 他们训练了一个用于自定义文字嵌入的深度学习模型, 并用它为每个领域训练一个单独的支持向量机 (SVM) 模型, 以选择最有可能解决每个问题的答案。该公司还构建了数万个对话步骤, 用于支持收集有关初始问题的任何缺失信息。初始域分类 (只关注简短、简单的开场问题) 使用 FastText15 模型进行开发, 该模型与下一次最佳尝试模型 (BERT) 的表现一样好, 但速度要快得多, 在每周只需 10-15 秒训练时间和 20-30 毫秒分类时间的情况下可取得 90% 的 F1 分数。以这种方式进行拆分有助于最大程度减少每个特定于领域的 SVM 模型中的类别数量。

成果: 提高回答质量, 加速解决来电

改进的语言模型提高了回答的质量和速度。虚拟助手现在能够为 85% 的客户案例提供满意的答案 (每年还额外为客户回答 200 万个问题), 而问题解决时间也从平均 3 分钟缩短到 1 分钟。每月为客户和顾问节省的时间总计达到数万小时。

本例充分说明: AI 并非用于提供答案, 而是作为人类主导的工作流程中不可或缺的一部分, 生成一系列更小但更有针对性的建议输出, 而人类可根据主观判断来使用它们。

附录

BESTSELLER

Boston Scientific

Crédit Mutuel

Global Bank

IFFCO-Tokio

KPMG

Marketing Platform

McCormick

Navtech

NVIDIA

Suncor

Zzapp

Global Bank (全球银行)

AI 可以有力补充现有流程，并尽可能减少中断，从而增强处理能力

概述

在银行业等高度监管的行业中，管理和维护流程文档是一项关键任务。这让银行能够更加稳定一致地执行重复性流程，并在外部审计期间证明合规性。为了帮助确保准确性、完整性以及增强理解，银行团队需要对文档进行内部审计，包括审查流程控制以及尝试重复运用流程控制措施。然而，手写记录流程通常比较复杂，采用自由文本等非结构化形式，而且手动审计也非常耗时。一家全球性银行展示了如何利用深度学习来扩展这项手动任务，同时补充而不是替代现有的人员和流程。

机遇:通过审计保证提高控制质量

银行有许多流程，从开设新储蓄账户到转账，每个流程都存在欺诈或洗钱等风险。管理这些风险至关重要，因为金融业受到严格监管，违规行为将受到严厉处罚。此外，随着在线服务提供商和其他新兴机构的兴起，传统银行面临日益激烈的竞争形势。客户更换银行从未如此简单，而唯有信任才是留住客户的关键因素。

银行可以通过严格的控制系统来降低这些风险。一些控制系统已实现自动化，但大多数控制系统仍然采用手动流程。为了帮助确保这些控制措施设计得当，并能够继续发挥效用，银行的内部审计 (IA) 部门需要通过抽检测试来验证控制措施的有效性。如果银行发现问题 (例如，开设不合理的账户)，则会实施纠正行动计划 (CAP)，例如增加刷新管制实体列表的频率。银行检查的控制措施越多 (检查频率越高)，就能够为企业提供更高层次的保障。

挑战:提高文档质量以促进高效审计

为了重复运用控制措施并评估其有效性，内部审计部门需要依赖于高质量的文档。这至少应当包含足够的信息来确定需要做什么、应当采用何种方式以及可以实现什么样的预期成果。如果缺少其中任意信息，则审计人员可能需要与控制负责人/归档人员讨论修订方案，而这会增加审计所需的工作量。此外，文档对于支持监管机构了解所应用的控制措施至关重要。因此，文档还应当重点呈现职责、时间安排和其他流程级信息。

Global Bank (续)

审计流程是一项重要的手动工作。Global Bank 每年要执行大约 1000 次审计工作，每次检查大约 10 项控制措施，每次平均耗时三个小时。Global Bank 将继续提高其处理能力（审计人员数量将增长 30%），并且已经拥有全球最大的内部审计部门之一，因此优化这些资源的生产力至关重要。

解决方案:使用 NLP 主动标记信息中的潜在缺失项

Global Bank 运用“即时校对器”来帮助文档编写者提高文档质量，从而提高审计流程的效率。其理念是开发一个自然语言处理 (NLP) 模型，该模型可以自动基于 5W 测试（内容、原因、对象、时间、地点）来标记控制文档中可能缺失哪些重要信息。Global Bank 可以在最初编写文档时使用该系统，也可以扫描现有文档以识别可能存在问题的文档。

Global Bank 使用预先训练的基于变换器的双向编码器表示 (BERT) 模型（一种 NLP 技术）构建了概念验证。该银行使用此模型进行命名实体识别，用于尝试识别 5W 中的每个要素。为了识别专用的 5W 要素，Global Bank 需要对模型进行调优。而且，由于即将开展新的审计工作，并计划向其他部门推出这项功能，因此需要对模型进行多次重新训练。然而，BERT 是一个庞大且复杂的模型，Global Bank 需要投入大量的计算资源来对其进行重新训练。此外，出于安全原因，Global Bank 只能使用本地部署硬件。

该解决方案侧重于两个方面。首先，Global Bank 构建了一个更易于重新培训的新模型。Global Bank 与 IBM 合作构建了此模型，基于之前的合作成果快速上线，并运用 IBM Watson® Studio 完成了本地部署实施。其次，Global Bank 着手增加了这种新模型的可用数据。此前，Global Bank 使用 Python 构建了一个直接连接至其内部审计平台的标记系统，可支持审计人员在审计工作中添加新的注释。该公司使用原始 BERT 模型增强了此系统，可为审计人员提供即时反馈，并让这些数据可供新的 IBM 模型使用。¹⁸

成果:提高审计效率和审计能力

此系统具备三项关键优势。第一，针对控制描述中的缺失项提供即时反馈，可在文档撰写过程中提高文档的完整性和准确性，这也让新加入者能够更快上手。第二，Global Bank 增强了手写记录流程与公司标准的一致性。第三，更高的质量减少了审计人员与控制负责人之间的来回传递，从而提高所有人员的效率。

在短短四个月内，此系统的快速采用就已经实现了可观的成效。50 位活跃用户在 5000 多个控制项中总共输入了 12000 个条目（自定义注释），并且每周还会新增数百个条目。随着审查流程效率的提高，Global Bank 能够节省大约 30000 小时的工作量，并安排这些人员来开展额外的质量保证工作，而任何其他方式都不可能实现如此惊人的效果。

附录

BESTSELLER

Boston Scientific

Crédit Mutuel

Global Bank

IFFCO-Tokio

KPMG

Marketing Platform

McCormick

Navtech

NVIDIA

Suncor

Zzapp

IFFCO-Tokio (IFFCO 东京通用保险公司)

通过改进 AI 流程，生成更理想的客户激励措施

概述

与许多行业一样，保险公司在提供适当的激励措施方面也面临着挑战。一般来说，如果保险公司支付账单，投保人就没有动力去寻找最优惠的服务价格。但有时 AI 可以改变流程，帮助他们实现这种价值。

机遇：直接向客户付款

IFFCO-Tokio 每年在印度处理约 50 万起机动车损坏索赔。客户通常选择在私营修车厂维修，这些维修厂向 IFFCO-Tokio 提供报价以获得批准，但这个过程存在一些问题。维修厂为了利益，往往会夸大预估成本，而客户也没有动力去争取最优惠的价格。这导致保险公司和维修厂在费用和维修上产生分歧。许多客户依靠车辆谋生，而每次索赔延误最长可达 20-30 天，给他们带来严重的不利影响。有时他们无法承担等待的代价，被迫放弃维修，选择驾驶不安全的车辆。

IFFCO-Tokio 决定直接向客户支付维修费用，为他们提供更好的激励措施，让他们担负起维修责任。但存在两个挑战：一是确保足够快的报价周期，二是在没有维修厂报价的情况下正确预估维修费用。

挑战：人工工作和低质量的数据

最初，IFFCO-Tokio 通过人工方法完成这项工作。他们开发了一个智能手机应用，供客户上传车辆损坏照片，接收预估成本的报价，自行决定是否接受报价，并独立于维修时间表接收赔付。一切都由客户自己负责。

事实证明，新方法非常受客户欢迎，但缺点是耗时而且不精确。该流程的每次索赔周期最长可达 5 个小时，大部分时间是由评估人员检查零部件以及填写有关“维修/更换”决定和每个零件预估成本的表格所花费的。他们的工作也受到低质量照片的影响。以前，车辆受损照片是在专业维修厂的受控环境中拍摄的，而客户提交的照片往往存在角度不正确、光线不佳和强光等问题。

IFFCO-Tokio (续)

解决方案: 指导式图像捕获与可变的训练数据量

IFFCO-Tokio 曾希望通过由机器学习自动执行每个零部件的初步预估, 以加快这个过程, 但他们意识到, 图像质量和一致性是关键所在。为了提高图像捕获的质量, 该团队使用相机模板增强应用的功能, 指导客户进行照片构图, 并增加了其他说明。针对强光或反射造成难以辨认损坏的零部件类型等问题, 他们还增加了训练数据量, 例如, 金属零部件的训练图像数量增加了 3 倍, 玻璃零部件的训练图像数量增加了 5 倍。这种改进图像质量措施的组合为推理和训练提供了更出色的数据, 支持使用深度学习技术对汽车型号、受损的零部件和损坏类型进行分类。在此基础上, 系统可以确定零部件是否需要维修或更换, 并估算成本。

通过实现决策过程中各部分的自动化, 人为监督工作会相应地减少。这可能会导致滥用, 例如对赔偿过的损坏情况提出重复索赔。在本例中, IFFCO 预计欺诈行为会增加, 于是他们构建了欺诈检测引擎, 以识别使用过的图像。然而, 他们发现总体系统被滥用的可能性相当低, 因为最终总是由经验丰富的评估人员做出决定。

成果: 缩短时间, 节省成本留住/争取客户

该系统在有意和无意间取得了巨大成功。评估人员的工作量显著减少。每次索赔的端到端时间平均缩短到 30 分钟 (包括与客户协商的时间), 新系统在不到一年的时间内就收回了成本。更令人惊讶的是, IFFCO-Tokio 的结算价格也下降了 40%, 接受率从 30% 提高到 65%。此外, 该系统还提高了弹性, 客户在修车厂因疫情关闭期间, 仍然能够收到赔付。最后, 或许也是最重要的一点, 新系统为提高客户满意度以及更好地留住和争取客户做出了直接的贡献。AI 不仅推动了效率的提高, 也增加了企业收入。

附录

BESTSELLER

Boston Scientific

Crédit Mutuel

Global Bank

IFFCO-Tokio

KPMG

Marketing Platform

McCormick

Navtech

NVIDIA

Suncor

Zzapp

KPMG (毕马威)

机器学习方法的选择由任务复杂性而非数据可用性驱动

概述

自然语言处理 (NLP) 和文本挖掘方法通常依赖于深度学习来实现命名实体识别 (NER)。这些方法能够从句子和段落中提取含义，并且通常在复杂任务中表现出色。然而，KPMG 在文档搜索和分类方面的经历表明，如果特定的子任务足够简单，传统的机器学习可能是最理想的方法。

机遇: 更好地证明研发税收抵免

在美国，研发 (R&D) 方面的税收优惠是非常可观的。税收抵免最高可占到符合条件的增量研发支出的 11%-15.8%；许多州还提供自己的研发优惠政策，因此收益可能进一步增加。尤其是对于规模较小的研发组织而言，这些税收抵免对于使研发项目具备商业可行性或者确保投资者首先为启动项目提供资金具有决定性作用。KPMG 受聘于各种规模的企业，负责记录他们完成的研发工作，帮助确保企业最大程度享受研发税收抵免。

挑战: 人工作业过多

美国国税局 (IRS) 通过包含 4 个部分的检验，评估研发税收抵免申请的可行性。该检验评估提出申请的活动是否：

- 涉及到创建新的业务组件或改进现有业务组件
- 具备技术性质
- 发现有助于消除与业务组件的方法、功能或设计相关的不确定性的新信息
- 包含通过模拟、建模或测试等方式进行试验的过程

显然，这种评估具有主观性，因此提供强有力的证据是取得良好结果的关键。

证据通常从组织的文档中收集，可以采取多种形式。这包括演示文稿、电子邮件、会议记录、实验室报告、测试记录和工程图纸等。内容通常是非结构化信息，数量庞大，存储在众多不同位置。在某些情况下（例如，在敏捷或持续改进/持续交付环境中），证据内容可能非常有限。无论如何，法规并没有规定什么样的证据符合“充分”这个标准，因此必须审查尽可能多的可用信息，并尽力提供相关和高质量的信息。

KPMG (续)

KPMG 凭借自己在发现流程方面的丰富经验, 为客户提供审计准备工作方面的支持。从传统上而言, 该流程包含自上而下的方法, 从项目列表开始, 搜索与这些项目相关的文档存储库, 手动搜索文档 (例如, 使用关键字), 阅读文档并标记满足 4 个检验过程中每一个的特定部分。这是一项繁重的人工作业, 需要进行一些优先排序工作, 存在将某些有价值的证据排除在外的风险, 同时还要占用客户科学家和工程师的宝贵时间来支持这项工作。KPMG 想知道, 机器学习能否帮助他们做得更好。

解决方案: 基于规则的方法优于机器学习

KPMG 发起了一场内部黑客马拉松活动, 四个团队使用各种不同的方法, 就问题的一部分 (文档分块) 展开竞争。这些团队获得了 1000 份含标记部分的文档, 要求针对每份文档, 给出有关与 4 项检验中每一项的相关性的置信度分数。

通过对单词和句子标记, 将文档分块为多个部分。团队尝试了一系列方法, 包括统计学习 (例如, 正则表达式、支持向量机、决策树和随机森林)、用于命名实体识别的深度学习 (DL) 以及基于规则的方法。他们发现, 使用即开即用的文档发现软件的准确率约为 55% (与人工搜索关键字的有效性差不多), 深度学习的准确率超过 70%。而基于规则的方法的准确率最高, 超过 85%。这可能是由于文档格式的标准化程度相对较高, 所以相对简化了文档分块任务。

成果: 更好的证据增强了申请抵免的能力

该系统现已应用于多家 KPMG 客户, 取得了非常理想的效果。它每月可处理超过 5000 份文档, 值得注意的是, 搜索方式从选择性的自上而下变为详尽的自下而上。税法相对稳定, 因此该系统只需最小程度的维护和改进, 同时能够更好地利用人类专业知识。还有一些传闻也可以证明其显著影响。例如, 某 KPMG 客户通过使用机器学习来审查研发项目文档, 以确定税收抵免的资格, 额外获得了 40% 的研发税收抵免。

值得思考的是深度学习相较于其他方法的表现。比较进一步证明, 即使有足够大的数据集, 深度学习也往往只有在数据或问题极其复杂时才更胜一筹。在本例中, 简单的规则和关键字足以识别每项检验的相关信息, 同时还能提供更出色的可解释性。

附录

BESTSELLER

Boston Scientific

Crédit Mutuel

Global Bank

IFFCO-Tokio

KPMG

Marketing Platform

McCormick

Navtech

NVIDIA

Suncor

Zzapp

Marketing Platform

通过提高可视性, 控制 AI 成本

概述

在本地基础架构中, 企业往往会忽视训练和运行机器学习模型的成本。例如, 模型产生的价值与实现该价值所需的工程和计算资源之间存在脱节。

Marketing Platform 案例研究表明, 让计算需求和成本之间的联系透明化能够起到激励作用, 以有意义的方式降低成本。

机遇: 迁移上云, 更好地利用现有数据

Marketing Platform 通过预测最有可能接受每项促销活动的客户群体, 帮助零售商和非营利组织提高营销工作的回报。准确性方面改进一小步, 只要规模够大, 销售额或获捐额就有可能前进一大步, 因此非常有吸引力。

可用于这些模型的数据集非常庞大。Marketing Platform 运营着一个包含数千成员的数据合作社区, 其中 25-40% 的成员定期提供有关交易、捐赠或订阅等方面的数据。这些数据与编译的第三方数据 (例如人口统计、人口普查和家庭收入数据) 相结合。经过特征工程处理后, 数据包含 12000 个变量, 几乎覆盖了全美人口。

处理如此丰富的数据集已达到 Marketing Platform 现有本地基础架构的极限能力。该团队只能针对内部数据训练模型, 即使这样, 一次也只能使用一小部分数据 (例如, 5 万-10 万个样本) 进行训练。Marketing Platform 知道, 如果可以使用更多的数据, 就有更大的机会实现更多的价值。

Marketing Platform (续)

挑战:初始成本非常高

迁移上云 (IBM Cloud Pak® for Data) 提高了组织管理数据和利用所有数据资产 (无论是线下还是线上) 的能力。计算资源的可扩展性显著提高, 因此他们能够针对 60 万条记录 (之前最多 10 万条) 和 800 个特征 (之前最多 150-200 个) 训练模型。通过与机器学习工具 (例如 XGBoost16) 结合使用, 使响应速度提高了 20%-30%, 从而显著增加了客户的投资回报。

最初, 计算资源的扩展 (以及初始迁移成本和学习曲线) 导致总成本开始增加。在完全本地的环境中, 基础架构的成本与利用率无关, 因此数据科学家能够运行他们想要的任何内容, 只受到计算资源可用性的限制。在拥有几乎无限的可扩展性之后, 数据科学家必须更审慎缜密地设计试验。

解决方案:与收入激励机制挂钩有助于抵消额外的成本

幸好, 迁移上云也帮助平台领导者更好地了解并最终优化支出。该团队现在能够生成每个模型的计算成本, 并在构建数据探索和分析时结合这些收入激励机制。

该团队还通过改进集群分配、数据流和整体建模管道, 提高计算资源的使用效率。他们还测试机器学习模型本身, 优化了约 100 个必需的模型参数, 并根据有效性修复其中的几个参数, 从而最大程度减少每次需要调整的参数数量。

成果:降低了训练成本, 实现更广泛的推广

最终, 每次训练的成本从 1500 美元大幅降至数百美元, 而且模型性能提高了 30%。

该平台在这个领域的成功推动自身的数据科学能力实现转型, 在一年内, 他们在美国的从业人员数量从 40 人增加到 4000 人。

其的经历表明:

- 通过在每个模型的运行层面了解训练成本, 有助于使迁移到新机器学习方法的成本低于预期。
- 数量非常重要: 即使准确性只是改进了一小步, 但如果大规模应用, 也会让组织的收益前进一大步。

附录

BESTSELLER

Boston Scientific

Crédit Mutuel

Global Bank

IFFCO-Tokio

KPMG

Marketing Platform

McCormick

Navtech

NVIDIA

Suncor

Zzapp

McCormick (味好美)

借助 AI 指导的探索，增强研发创造力

概述

AI 通常用于根据过去的成功经验提出建议，但这可能导致解决方案大同小异。McCormick 的经历表明，AI 也可用来探索解决方案领域，形成全新的创造性解决方案组合 — 可能是以前从未尝试过的方法。通过这种方式，AI 可帮助增强和加速创造性试验。

机遇：使用 AI 加速开发新的口味特征

McCormick 生产各种各样的产品，包括调味品、酱汁和辛香料等（其中一些也会出售给 B2B 客户以添加到第三方产品中）。配方是精确比例的特定配料的组合，通过实现标准化以保持一致性，并提供可用于描述饮食体验的口味特征。为 B2B 产品创建口味的过程竞争激烈，有多家公司根据客户要求生产配方。为了提高新口味的成功率，McCormick 研究了两个重要因素：

- 丰富食品科学家的经验。初级调味师通常经过 7 年的学徒制学习，期间积累了丰富经验并收获了有价值的洞察。正是这种积累的经验带来创造能力；了解什么有效，什么无效，以及二者之间的自由度。
- 提高试验过程的效率。要为客户研发候选的口味，需要执行几个步骤。调味师设计一系列不同的口味特征，制作样品并进行试验（单独试验以及在试验厨房中烹饪后试验），然后不断改进，直到觉得产品值得提交给客户为止。¹⁷ 提高这个过程的效率有助于加快产品面世速度，以及更好地利用食品科学家的时间。

McCormick 想知道 AI 能否提供帮助。如果调味师能够从经验中凭直觉获得洞察，那么通过分析可能会更快获得更多洞察。捕获这些洞察并进行系统化处理，可帮助调味师更好地探索口味。这包括两种方式。第一种是在众所周知的领域内寻找最佳口味。第二个是在尚未探索但很有前景的新领域内寻找最佳口味。最终的结果是更快的研发速度和更好的研发质量。

挑战:配额分配和较大的研究空间

该公司利用有关十多年来创建的约 35 万种配方的数据,这涵盖多种产品属性,包括类别(如,烘焙食品和咸味零食)、形式(如,调味品、佐料和干湿酱料)、数量、类型以及成功指标(如产品品尝分数)。他们还收集了 4 万种原料的功能属性(例如,货架稳定性和流通速度)以及非功能性属性(例如,颗粒度、钠含量和 FEMA18 号*)。由于数据集的维度如此之广,因此团队需要一种方法来压缩问题,使其便于管理。

解决方案:图形表示与数据降维

该公司新开发了一个 SAGE 深度学习系统,用于生成新的口味特征。它采用两个用户定义的主要输入:(1) 一个种子配方(例如,韩国烧烤的口味特征),以及(2) 输出配方中需要的任何特定限制条件(例如,“必须有芒果”)。然后,系统生成与种子存在不同程度偏差的配方:4 个配方仅针对预期口味进行微调,另外 4 个配方的自由度更大但仍受限制,剩下的 4 个配方则表现出显著差异。这为调味师提供了一系列迭代改进的选项,具体取决于所期望的口味新颖程度。

为了完成这项工作,团队需要执行一些步骤。首先是数据降维,例如,将 4 万种不同的原料分为 3000 组。其次,他们从 35 万个语料库中抽取了 3000 个标有“成功”评级的配方样本,用于训练模型。最后,他们将模型表述为图形问题,定义原材料之间的距离指标,其中每个配方都用向量表示。

成果:表现相当于人类的 20 年经验

McCormick 发现,使用该系统的初级食品科学家的表现大致与具有 20 年经验的食物科学家相当,而且大大减少了所需的试验次数。他们还发现,该系统有助于更加充分地利用全球知识。有一次,该系统向一位没有加拿大市场经验的美国调味师推荐了加拿大口味的产品。

附录

BESTSELLER

Boston Scientific

Crédit Mutuel

Global Bank

IFFCO-Tokio

KPMG

Marketing Platform

McCormick

Navtech

NVIDIA

Suncor

Zzapp

Navtech

Navtech 深度学习创造全新平台

概述

仅当有一个集中式提供商进行所需的前期投资，并将成本分散到足够广泛的客户群中之时，某些创新成果才具有经济可行性。机器学习越来越符合这种情况，尤其是基于感知（例如，图像识别）自动执行某些任务的场景，因为这些模型的开发和维护可能需要大量数据和计算资源。对于许多个人和企业而言，这项技术根本遥不可及。

Navtech 发现了商机，通过创建模型，并将其作为“即服务”提供给客户，为全球众多钻石零售商带来先进的计算机视觉技术。这是双赢的商业案例，也是证明 AI 不仅可以提高效率和性能，还可在此过程中解锁新的能力和业务模式的绝佳示例。

机遇：数字化产品目录帮助提高销售额

据估计，仅印度就有 30 万家钻石珠宝零售商。其中许多是规模较小的公司，库存容量有限，通常借助定制珠宝来增加产品销量。与仅提供自己库存中的商品相比，增加定制珠宝选项会使客户转化率（找到自己喜欢的商品并购买的客户数量）翻倍。

直观的产品目录对于销售过程至关重要。每个零售商都会为自己的商品库存保留一个产品目录，并用其他珠宝的照片作为补充，以激发客户定制产品的想法。这些目录一般都是实体小册子或杂志，只能呈现有限数量的商品，无法频繁更新。

数字产品目录消除了许多这些限制，但也带来其他挑战。工作人员编辑各种来源（例如，库存、网络和制造商目录）的图片，并手动将其分类到文件夹中。这个过程很慢（每张图片需要 30-60 秒，最多可达一百万张图片），容易出错（例如，有很多重复图片，很难记住自己已经看到过的内容），并且分类不够详细（例如，只能分为戒指和项链）。如果能够自动创建目录，支持客户根据其他条件进行搜索，就理想不过了。

挑战:对于单个零售商而言,计算机视觉技术的成本过于高昂

利用深度学习对图片分类的计算机视觉系统有助于提高速度和准确性,但现实情况是,对于大多数零售商来说,它们根本遥不可及。深度学习属于资源密集型技术;训练和实施模型都需要大量数据和计算资源。而且该系统的使用频率并不高,不足以证明其成本的合理性,特别是在劳动力的比较成本(即人工分类)较低的情况下。例如在印度,零售员工的起步工资每月只有100美元左右。因此,对于任何一家零售商而言,计算机视觉的商业案例都不太可能具有吸引力。

解决方案:一次构建,提供“即服务”

Navtech 新兴技术总经理 M.I.M.Loya 博士提出一个想法:构建计算机视觉系统并将其作为“即服务”提供给客户。这是双赢局面。Navtech 拥有进行初始投资的资源,并以可承受的收费提供对系统的访问;而零售商则受益于持续、低成本的系统访问。

Navtech 为试点项目选择了三个属性,并为每个属性分别创建一个深度学习模型:

- 对于产品类别(例如,戒指和手镯)和款式,他们使用 VGG16 网络对图片进行分类。通过以自定义方式训练网络的头、第一层和第二层,对由 ImageNet 训练的开源主干进行微调。
- 对于钻石切割(例如圆形、方形),他们改用 Mask RCNN 进行对象检测和分类(使用 VGG16 的准确率仅为 55-56%)。该模型的训练数据由一名实习生标记,他手动绘制围绕每颗钻石形状的多边形掩模。

成果:以可控的成本更广泛地访问前沿机器学习技术

该系统支持零售商更快构建更大的数字产品目录。它每分钟最多可分类 100 张图片(人工仅能分类 1-2 张),产品类别和款式的准确率达到 90-93%,钻石切割的准确率达到 85-86%(人工标记的准确率为 80%)。让人略感惊讶的是,该团队还是在使用相对较小的数据集(每个模型只使用 3000 张标记图片)的情况下实现这样的成绩的。Navtech 还进行了一些后处理,某个模型(例如,款式)的结果可用于提高另一个模型(例如,钻石切割)的预测置信度。这么小的数据集之所以能够产生如此大的作用,也可能是由于问题复杂性相对于 ImageNet 较低(比方说,珠宝图片的变化少于猫的图片变化)。

这个经历说明了深度学习的成本与效益之间的权衡,某些用例只能由可以为更广泛的受众提供服务的大型集中式企业实现。由于需要大规模交付,此类系统必须作为更广泛的产品和服务架构的一部分提供,它们以传统软件开发为基础,但更具成本效益。

附录

BESTSELLER

Boston Scientific

Crédit Mutuel

Global Bank

IFFCO-Tokio

KPMG

Marketing Platform

McCormick

Navtech

NVIDIA

Suncor

Zzapp

NVIDIA (英伟达)

为自动驾驶汽车构建计算和数据平台

概述

为了给自动驾驶汽车 (AV) 构建安全可靠的 AI 模型, 需要大量的计算能力和训练数据, 以及与这种规模相配的技能、资源和专业知识。在这种情况下, 我们可能会看到更大平台的出现: 汇集多个参与方的数据, 汇总足够数量的需求以证明所需大量投资的合理性, 并最终实现新的业务模式 — 向汽车制造商和车队运营商提供“AV 软件即服务”。

机遇: 解决复杂问题, 推动 AV 的大规模采用

驾驶员支持系统越来越普及, 现在可以执行泊车、紧急制动、变换车道等功能。车辆实现完全自动驾驶后, 会出现多种多样的应用, 包括货运、公共交通以及按需运输 (如机器人出租车)。2021 年, 全球自动驾驶汽车系统市场规模为 820 亿美元, 预计到 2030 年将增长至 7700 亿美元, 2022 年到 2030 年的复合年增长率 (CAGR) 为 39.1%。²⁰

挑战: 巨大的数据和计算需求

由于需要执行一系列规划和控制任务 (例如, 寻找行人、检测道路标记和交通信号灯), 因此全自动自动驾驶汽车的数据需求巨大。这些功能必须足够强大, 可以应对不同的环境状况 (例如, 天气、语言环境), 并能处理罕见的突发事件 (例如, 意外失去动力)。目前的系统在这些功能的数量和有效性方面不断改进; 但是, 证明 AV 能够有效避免伤亡的可靠性标准非常高。RAND 公司预计, 为了达到和人类驾驶相当的错误率, 自动驾驶汽车可能需要 110 亿英里的测试, 相当于 100 辆汽车连续试驾 500 多年。²¹ NVIDIA 估计, 要让某些 AV 任务达到理想的性能, 需要约一百万个场景的训练示例。

考虑到每个场景都需要众多传感器收集的数据, 因此自动驾驶汽车带来的计算挑战非常巨大。为了创建完整 AV 技术栈所需的感知模型, NVIDIA 估计一个高效的开发团队可能需要约 5000 个专用 GPU。²² 单个模型要在 32 个 GPU 上运行 3-6 天, 每个任务需要进行 25-50 次深度学习试验。单个汽车企业通常没有自行开发这些系统所需的技能、经验、硬件和数据资源。

解决方案:在多个客户之间共享数据和计算平台

NVIDIA 通过以下方式应对这些挑战:

- 扩展覆盖多个客户的公共数据平台: 汇集多家企业的数据有助于增加可用于训练的数据量, 并实现更出色的模型性能, 特别是对于边缘案例。数据的质量由用于规定传感器规格和放置等标准的参考架构保证。
- 用于训练和测试的模拟: 可模拟数亿个驾驶场景, 以补充现实世界的的数据, 帮助引导模型进行静默上路测试和迭代, 例如, 在车辆中运行 AI, 比较它与驾驶员实际行为的差异。
- 视觉任务的通用处理: NVIDIA 通过在基于 ResNet 的单一模型架构上联合训练多个任务, 最大程度减少所需的计算。训练完整模型后, 可针对每个特定任务优化模型的头部 (后期层), 而无需重新训练模型的主干 (早期层)。请注意, 这种计算并不比单个任务所需的计算多出多少, 说明可能存在许多通用处理, 这在计算机视觉领域非常容易理解。

成果:支持在技术栈范围开展竞争的新型业务模式

以这种方式集中管理数据为 AV 技术带来了新的可能性。根据他们的需求和现有能力, 汽车制造商参与方可以租用 AV 硬件, 基于更大的数据集训练自己的模型, 也可以使用 NVIDIA 预先训练的 AV 模型。无论是哪种情况, 汽车制造商都无需在硬件和开发能力方面进行大量资本投资, 而是可将 AV 技术作为运营支出, 并且随着硬件和软件的逐步改进而不断受益。

这也代表了新市场格局的开端。一方面是垂直整合的汽车制造商 (例如特斯拉), 他们可以共同设计软件和硬件, 以提供更加无缝的体验。另一方面, 越来越多的模块化汽车制造商在硬件质量方面参与竞争, 并从 NVIDIA 等集中式参与方那里购买软件, 这大大降低了进入 AV 市场的成本, 并有可能刺激更激烈的竞争。无论是哪种模式, 要取得成功, 都离不开 AV 软件的质量, 后者对于整体体验起到至关重要的作用。

附录

BESTSELLER

Boston Scientific

Crédit Mutuel

Global Bank

IFFCO-Tokio

KPMG

Marketing Platform

McCormick

Navtech

NVIDIA

Suncor

Zzapp

Suncor (森科能源)

性能和可解释性并不总是需要权衡取舍

概述

深度学习通常在预测非线性（小变化可能产生巨大影响）和高度耦合（要素之间的依赖关系较多）的流程方面表现出色。通常，如果采用深度学习可以实现足够的性能改进，企业愿意在自己模型的可解释性方面做出一些牺牲（例如黑匣子）。但是，Suncor 的经历表明，如果利益足够高，可解释性也非常重要。

机遇：更有效地预测问题，管理产品质量

Suncor Energy 专业从事从油砂中提取合成原油的生产业务。特别是柴油生产，涉及到通过氢化处理去除硫氮，在高温和高压下将直馏柴油与氢气（以及钴等固体金属催化剂）混合。这个工艺非常复杂，包含多个变量（压力、温度和流速），它们相互作用，影响到产品的最终质量。因此，企业必须严格监控这些要素，最大程度减少“意外”，避免产品质量偏离可接受的标准而无法出售。Suncor 的柴油日均产量达到 43000 桶，因此尽量避免意外的商业动机非常强烈。

挑战:具有高影响力的决策需要高水平的可解释性

生产质量的最终责任在于现场负责人,他们监督生产场地,并做出会影响生产质量的关键运营决策。他们主要根据个人经验和判断,确定要做出哪些调整及其产生的影响。这些具有高影响力的决策必须正当合理,因此需要使用明确的理由进行解释,这意味着用于决策支持的任何分析方法都必须透明,并且得到充分的理解。

解决方案:主成分分析的性能相当,而且可解释性更高

Suncor 计划通过开发一种“意外标记”模型,旨在提高预测能力,帮助发现新出现的不良状况,为采取纠正措施赢得足够时间。数据科学团队确定了与产品质量相关的 11 个需要实时评估的要素,并构建模型,其中包含来自传感器数据的 30 种不同的测量值。一开始,他们探索了多种先进但难以解释的机器学习方法,包括神经网络、长短期记忆网络、随机森林、梯度提升和决策树,其中 XGBoost (结合了决策树与梯度提升) 的表现最佳。

然而,当团队将 XGBoost 的表现与更简单的方法(例如,传统的统计方法)进行比较时,他们发现后者的表现大大超出预期。例如,主成分分析 (PCA) 的性能仅仅比 XGBoost 低 10%,但更易于解释。

成果:提供透明、理由正当的早期预警

在与现场负责人利益相关方共同对这两种方法进行彻底的测试之后,Suncor 认为,在这种情况下,更高的可解释性的权重要高于性能权衡。通过将 PCA 与预测结合使用,可以揭示每个要素背后的相关权重,也就是说明哪些要素在推动预测方面最重要。最终的系统能够提前一小时预测意外事件,准确率达到 80%,预测时间间隔为 5 分钟。

附录

BESTSELLER

Boston Scientific

Crédit Mutuel

Global Bank

IFFCO-Tokio

KPMG

Marketing Platform

McCormick

Navtech

NVIDIA

Suncor

Zzapp

Zzapp Malaria

从卫星图像学习, 抗击疟疾

概述

在卫星图像应用领域, 计算机视觉常用于识别可见对象, 卷积神经网络 (CNN) 是最常见的默认选择。即使对象本身并不清晰可见, 使用 CNN 的预测模型有时也可以根据其他特征 (例如, 相关对象周围的区域) 推断出对象的存在。然而, ZzApp Malaria 的经历表明, 并非总是需要使用该网络, 线性回归等传统方法可能足以应对这种情况。

机遇: 通过处理积水, 预防疟疾

2020, 疟疾造成约 62.7 万人死亡, 其中 96% 发生在非洲。传病媒介控制是预防疟疾传播 (携带疟疾病毒的按蚊叮咬) 的主要途径。目前的主要预防方式是使用有效的预防物 (如蚊帐), 或在室内喷洒杀虫剂, 但这些方法的有效性有限, 而且在室外不起作用。另一种方法是直接处理社区内的积水 (蚊子交配繁殖之地), 但这些计划不够系统或全面, 无法实现广泛的有效性。

挑战: 卫星图像无法显示较小的水体

积水处理的难点必须先发现积水, 才能进行管理。较大的水体在卫星图像中很容易看到, 现已开发计算机视觉算法进行自动识别。然而, 即使使用先进的卫星成像技术, 也很难探测到被覆盖或只在换季时出现的较小水体。如果能够更精确地找到这种积水, 那么就有可能改善直接喷洒效果以及蚊子种群控制。

Zzapp (续)

解决方案:通过地形推断是否存在积水

Zzapp Malaria 的成立就是为了解决这个问题,他们首先关注圣多美的疟疾多发地区。他们创建了一个应用,支持实地检查人员记录发现的水体位置,并持续跟踪水处理情况,从而建立一组正样本的训练集(即存在积水的位置)。他们还收集了卫星图像(例如,照片、红外和雷达图像),用于训练基于 CNN 的对象检测算法。该算法在检测大型水体时表现出色,但针对小型水体的有效性不高,特别是当水体被覆盖时。

作为替代方案,该团队从图像中提取 50 个基于图像的地形特征,并将其用于传统的线性回归方法,旨在确定地图每个部分出现积水的可能性。该方法的准确率达到 75%,与 CNN 的表现相当,但能够更加清晰透明地确定哪些因素可以改善预测。研究团队还发现,地形因素与地点密切相关,而线性回归比神经网络方法更容易应用于其他地点。

成效:可用于其他地点的透明且可迁移的方法

回归模型的表现相对出色,原因可能是可以利用积水特征,将其作为数据拓扑特征的函数,而不必使用 CNN 推断。无论如何,回归模型较高的透明度和可迁移性帮助 Zzapp 实现了雄心壮志:轻松将该方法扩展到圣多美以外具有不同地形的其他地区(加纳、桑给巴尔及其他地区)。

关于研究洞察

研究洞察致力于为业务主管就公共和私营领域的关键问题提供基于事实的战略洞察。洞察根据对自身主要研究调查的分析结果得出。要了解更多信息，请联系 IBM 商业价值研究院: iibv@us.ibm.com

选对合作伙伴, 驾驭多变的世界

IBM Consulting 积极与客户协作, 运用业务洞察和先进的研究方法与技术, 帮助他们在瞬息万变的商业环境中保持独特的竞争优势。

IBM 商业价值研究院

IBM 商业价值研究院(IBM Business Value Institute, IBV)成立于2002年。凭借 IBM 在商业、技术和社会交叉领域的独特地位, IBV 每年都会针对成千上万高管、消费者和专家展开调研、访谈和互动, 从中分析提炼出可信赖的、振奋人心和切实可行的洞察, 帮助领导者做出更明智的业务决策。

需要 IBV 最新研究成果, 请在 ibm.com/ibv 上注册以接收 IBV 的电子邮件通讯。您可以在 Twitter 上关注 [@IBMIBV](https://twitter.com/IBMIBV), 或通过 <https://ibm.co/ibv-linkedin> 在 LinkedIn 上联系我们。

访问 IBM 商业价值研究院中国官网, 免费下载研究报告: <https://www.ibm.com/ibv/cn>

备注和参考资料

- 1 信息来源: “Fast Start in cognitive innovation: Top performers share how they are moving quickly.” IBM Institute for Business Value. January 2017. <https://www.ibm.com/blogs/internet-of-things/fast-start-cognitive/> Unpublished data. C&A8. In general, where is your organization in its adoption of cognitive computing? Select the most advanced level for your organization; “Shifting toward Enterprise-grade AI: Confronting skills and data challenges to realize value.” IBM Institute for Business Value. September 2018. <https://www.ibm.com/thought-leadership/institute-business-value/report/enterpriseai> Unpublished data. AI1. In general, where is your organization in its adoption of artificial intelligence? Select the most advanced level for your organization; “The business value of AI: Peak performance during the pandemic.” IBM Institute for Business Value. November 2020. <https://www.ibm.com/thought-leadership/institute-business-value/report/ai-value-pandemic#> Unpublished data. S6. In general, where is your organization overall and your particular function in terms of adoption of artificial intelligence? 2022 Omdia AI Market Maturity survey <https://omdia.tech.informa.com/OM023919/AI-Market-Maturity-Survey--2022-Database> Q1. What is the state of AI deployment in your company? The rating scale in Omdia survey has been assumed to equivalent to IBM IBV rating scale in the following way: investigating technology and use cases = considering; Identified at least one use case and developing pilot = Evaluating; Currently piloting AI in at least one function or business = Piloting; Live AI deployment in at least one function or business unit = Implementing; Scaling AI deployment across multiple business functions or units = Operating/optimizing.
- 2 “The business value of AI: Peak performance during the pandemic.” IBM Institute for Business Value. 2020. <https://ibm.co/ai-value-pandemic>
- 3 “McCarthy, J; M.L. Minsky; N. Rochester; C.E. Shannon. “A proposal for the Dartmouth summer research project on artificial intelligence.” Accessed on July 13, 2022. <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>
- 4 Goodfellow, Ian; Yoshua Bengio, Aaron Corville. “Deep Learning.” The MIT Press. 2016. <https://www.deeplearningbook.org>
- 5 LeCun, Yann; Yoshua Bengio; Jeffrey Hinton. “Deep Learning.” Nature. May 28, 2015. <https://www.nature.com/articles/nature14539.pdf>
- 6 Burns, Ed. “Timeline of AI winters casts a shadow over today’s applications.” TechTarget. Accessed on July 13, 2022. <https://www.techtarget.com/searchenterpriseai/infographic/Timeline-of-AI-winters-casts-a-shadow-over-todays-applications>
- 7 Thompson, Neil C.; Kristjan Greenewald; Keeheon Lee; Gabriel F. Manso. “Deep Learning’s Diminishing Returns.” IEEE Spectrum. September 24, 2021. <https://spectrum.ieee.org/deep-learning-computational-cost>
- 8 Ibid.
- 9 Ibid.
- 10 World Health Organization malaria fact sheet. April 6, 2022. <https://www.who.int/news-room/fact-sheets/detail/malaria>

- 11 "Fast Start in cognitive innovation: Top performers share how they are moving quickly." IBM Institute for Business Value. January 2017. <https://www.ibm.com/blogs/internet-of-things/fast-start-cognitive/> Unpublished data. Q&A10 What are the important value drivers for cognitive computing? Select the top 5. "Shifting toward Enterprise-grade AI: Confronting skills and data challenges to realize value." IBM Institute for Business Value. September 2018. <https://www.ibm.com/thought-leadership/institutebusiness-value/report/enterpriseai> Unpublished data. AI2. What are the important value drivers for artificial intelligence/cognitive computing? Select top 5. "The business value of AI: Peak performance during the pandemic." IBM Institute for Business Value. November 2020. <https://www.ibm.com/thought-leadership/institute-business-value/report/ai-value-pandemic#> Unpublished data Q8. What are the most important value drivers for artificial intelligence? Select top 5.
- 12 "The business value of AI: Peak performance during the pandemic." IBM Institute for Business Value. 2020. <https://www.ibm.com/thought-leadership/institute-business-value/report/ai-value-pandemic>
- 13 Payraudeau, Jean-Stéphane; Anthony Marshall; Jacob Dencik. "Unlock the business value of hybrid cloud: How the Virtual Enterprise drives revenue growth and innovation." IBM Institute for Business Value. 2021. <https://ibm.co/hybrid-cloud-business-value>. Payraudeau, Jean-Stéphane; Anthony Marshall; Jacob Dencik. "Extending digital acceleration: Unleashing the business value of technology investments." IBM Institute for Business Value. 2021. <https://ibm.co/hybrid-cloud-business-value>
- 14 Fleming, Martin. "Breakthrough: A Growth Revolution." Business Expert Press. 2022
- 15 An open source NLP library developed by Facebook AI. <https://fasttext.cc>
- 16 XGBoost is a decision-tree-based ensemble ML algorithm that uses a gradient boosting framework. <https://towardsdatascience.com/https-medium-com-vishalorde-xgboost-algorithm-long-she-may-rein-edd9f99be63d>
- 17 McCormick has observed that each 5-10 years of experience halves the number of iterations.
- 18 The Flavor Extract Manufacturer's Association of the United States. FEMA numbers refer to ingredients generally recognized as safe and allowed in the United States. <https://www.femaflavor.org/>
- 19 VGG16 (also called OxfordNet) is a convolutional neural network architecture named after the Visual Geometry Group from Oxford. <https://blog.keras.io/how-convolutional-neural-networks-see-the-world.html>
- 20 Report Ocean press release. "Autonomous Vehicle System Market [(CAGR) of 39.1%] by Product Type, End-User, Application, Region – Global Forecast to 2030." July 14, 2022. https://www.marketwatch.com/press-release/autonomous-vehicle-systemmarket-cagr-of-391-by-product-type-end-user-application-region-global-forecast-to-2030-2022-07-14?mod=search_headline
- 21 Kalra, Nidhi and Susan M Paddock. "Driving to Safety: How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability?" Rand Corporation. 2016. https://www.rand.org/content/dam/rand/pubs/research_reports/RR1400/RR1478/RAND_RR1478.pdf
- 22 GPUs are arranged into purpose-built deep learning systems (for example, the NVIDIA DGX, which comprises 8 GPUs per server).

© Copyright IBM Corporation 2022

国际商业机器 (中国) 有限公司
北京市朝阳区金和东路 20 号院 3 号楼
正大中心南塔 12 层
邮编: 100020

美国出品 | 2022 年 8 月

IBM、IBM 徽标、ibm.com、IBM Cloud Pak for Data、IBM Research 和 Watson 是 International Business Machines Corporation 在世界各地司法辖区的注册商标。其他产品和服务名称可能是 IBM 或其他公司的商标。以下 Web 站点上的“Copyright and trademark information”部分中包含了 IBM 商标的最新列表：ibm.com/legal/copytrade.shtml。

本档为自最初公布日期起的最新版本，IBM 可能随时对其进行更改。IBM 并不一定在开展业务的所有国家或地区提供所有产品或服务。

本档内的信息“按现状”提供，不附有任何种类的（无论是明示的还是默示的）保证，包括不附有关于适销性、适用于某种特定用途的任何保证以及非侵权的任何保证或条件。IBM 产品根据其提供时所依据的协议条款和条件获得保证。

本报告的目的仅为提供通用指南。它并不旨在代替详尽的研究或专业判断依据。由于使用本出版物对任何企业或个人所造成的损失，IBM 概不负责。

本报告中使用的数据可能源自第三方，IBM 并未对其进行独立核实、验证或审查。此类数据的使用结果均为“按现状”提供，IBM 不作出任何明示或默示的声明或保证。



扫码关注 IBM 商业价值研究院



官网



微博



微信公众号



微信小程序

