

高盛 AI 生态报告：人工智能可解决药物研发、医保控费和医院运行效率难题

钛媒体注：2016 年年底，美国高盛集团发布了一份长达 99 页的重磅人工智能报告：《人工智能、机器学习和数据将推动未来生产力的发展》（AI, Machine Learning and Data Fuel the Future of Productivity）。报告围绕人工智能，阐述了人工智能的生态及未来，从医疗、农业、金融、零售业、能源业等领域描绘了人工智能对其带来的影响。

以下是动脉网对其翻译和整理之后的全文呈现。

这份报告认为，所谓 AI，即是制造以人类智能的方式学习并解决问题的智能机器和计算机程序的科学与工程。该领域包括自然语言处理与翻译、视觉感知与模式识别，以及决策制定等。

近年来，机器学习（Machine Learning, ML）和深度学习（Deep Learning, DL）的应用领域极速扩张，而数据、更快的硬件、更好的算法则是推动人工智能的进展的三大基石。下文中，动脉网为您节选了报告中关于人工智能对医疗领域的影响，带您一窥未来医疗的发展方向。报告指出，到 2025 年，医疗年均成本预计可节约 540 亿美元。

机器学习在医疗领域中有广泛的应用前景。医疗行业需要丰富且定义明确数据集，也需要随时随地对患者进行监督，而医疗结果也存在着极大的可变性。

机器学习可为其中不少的子行业提供获得高额回报的潜力，如药物发现、测试分析、治疗优化和患者监护等。随着人工智能和机器学习的不断整合，人们将有望在新药研发的过程中显著地实现“去风险”，不但将节约每年约 260 亿美元的研发成本，同时还将提高全球医疗信息领域的效率，节约的成本价值超过每年 280 亿美元。

机遇何在？

药物发现与开发。在药物开发过程中结合机器学习，有着提升开发效率的潜力。机器学习不但可以加速时间范围，还可以提高到达后期试验阶段药物的成功概率（probability of success, POS）。Medicxi Ventures 的合伙人 David Grainger 认为，错误发现率（False Discovery Rate, FDR）是一个统计学现象，而避免 FDR 则有可能将后期试验阶段的风险减半。

此外，在药物发现的早期阶段中，现有虚拟筛选的方法名为“高通量筛选”，而它非常容易受到 FDR 的影响。如果可以将第 3 阶段试验的风险减半，就可以为大型制药公司节约数十亿美元的成本，影响其超过 900 亿美元的研发经费并带来有意义的回报，使其能够腾出资源集中于寻找更有潜力的机会。

备注：虚拟筛选(virtualscreening, VS)也称计算机筛选，即在生物活性筛选之前，利用计算机上的分子对接软件模拟目标靶点与候选药物之间的相互作用，计算两者之间的亲和力大小，以降低实际筛选化合物数目，同时提高先导化合物发现效率。

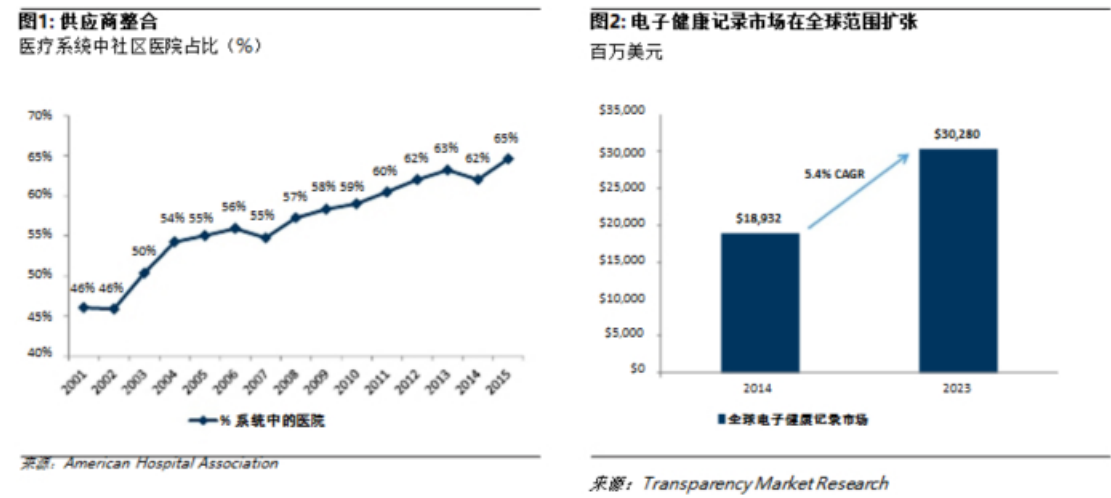
虽然与后期试验相关的巨额费用往往侧重于临床试验的设计元素，但我们认为，将 AI/ML 应用于优化后期阶段在选择标准、规模和研究长度等方面的决策，也可以实现有意义的效率提高。

医生/医院的效率。由于监管和分裂等原因，美国医疗体系在历史上对新技术的采用一直十分缓慢。除了需要应对系统的挑战，从药物发现到医生和诊所将新药应用于医疗实践之间的过程往往十分漫长且没有连续性。

美国市场研究咨询机构透明市场研究(Transparency Market Research)公司的数据显示，美国政府最近发布的一系列纳入《美国复苏与再投资法案》的法

令，已经推动了诸如电子健康记录等领域的快速增长，全球市场预计将在 2023 年达到约 300 亿美元。

数据的聚合，不断改进的数据捕获技术，以及独立医院的不断减少等，已经为数据的大规模利用创造了一个前所未有的机遇。这一切也将提高机器学习算法和人工智能的各项功能，以在医疗领域的各个方面改善速度、降低成本和提高精度。



总部设在伦敦的谷歌 DeepMind 正与英国国民健康服务(National Health Service, NHS)合作开发一款旨在监测肾脏疾病患者的 APP，以及一个前身名为“患者抢救”、旨在支持诊断决策的平台。

任何 AI/ML 系统的关键都是海量的数据，因此 DeepMind 和 NHS 达成了一个数据共享协议，NHS 将为 DeepMind 提供动态的新数据流和历史数据，以用于训练 DeepMind 的算法。

只有有了海量的数据，才有可能对临床数据进行实时分析。当然，如果 DeepMind 可以随时有效获取患者数据，它所能提供的见解将远远超出肾脏疾病的范围。

痛点何在？

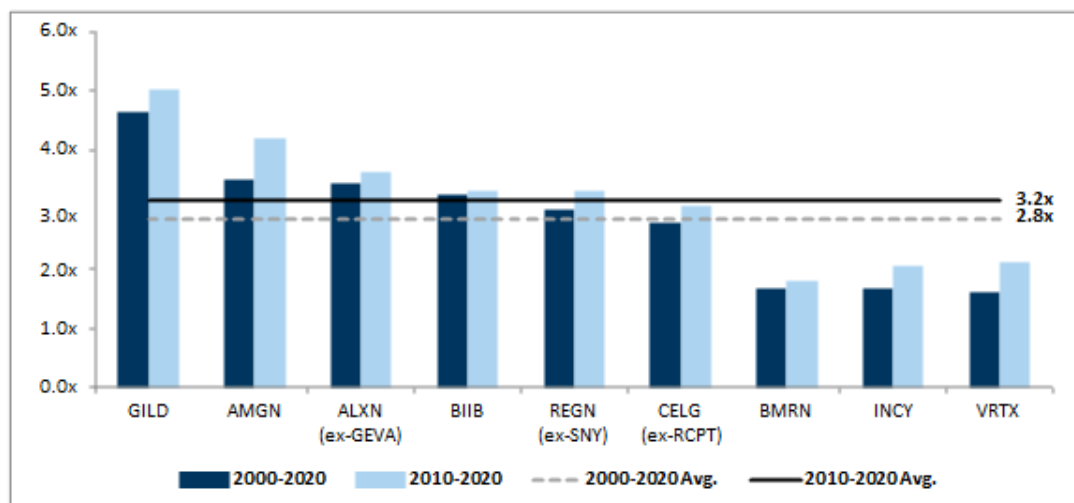
药物发现与开发。医疗领域的重要痛点之一，是药物发现与开发的时间和成本。根据塔夫特药物发展研究中心（ Tufts Center for the study of Drug Development ）的数据，一款新药的面市从药物发现到获得 FDA 批准平均大约需要 97 个月。

虽然对专业技术的持续聚焦可以帮助改善时间跨度，但新药研发的成本却仍在持续增加。德勤的数据显示，自 2010 年以来，12 家主要制药公司的获批药物开发成本已经增加了 33%，至约每年 16 亿美元。

研发回报。生物制药研发的生产力至今仍然是一个充满争议性的话题。开发一款成功药物的成本持续增加，但由于报销制度中的不利因素、患者量的降低和企业间的竞争等，新药研发的收入回报环境也不容乐观。

虽然我们预计 2010 - 2020 年的研发回报相对与 2000-2010 年会有所提高，但实际上二者之间的变化微不足道。此外，影响研发回报最重要的不利因素之一在于那些失败的研发产品，特别是那些已经达到后期试验阶段的药物；这些药物的成本每年估计就占到了 400 亿美元以上。

图 3: 20年与10年的总收益率
总收益率 (累计收入除以一般公认会计原则 (GAAP) 的研发支出)

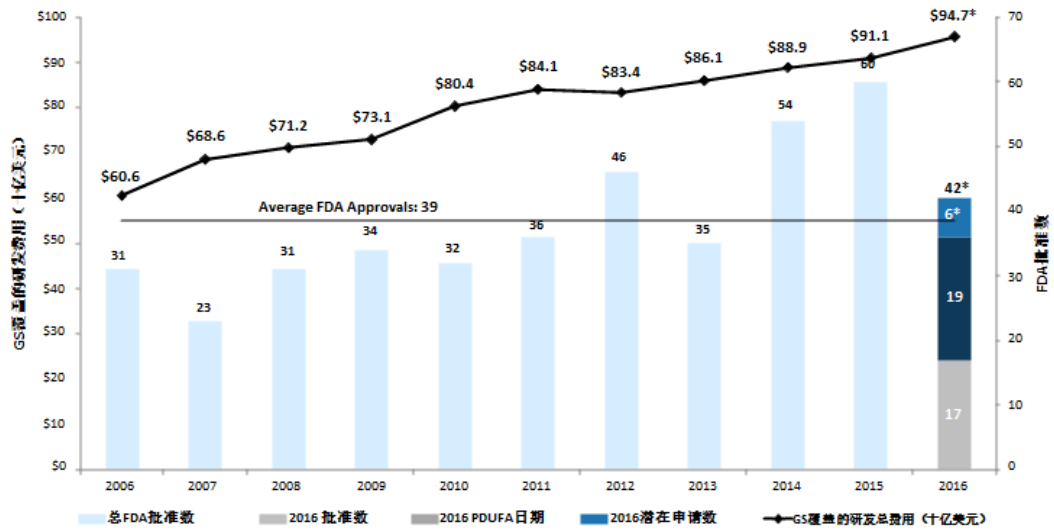


来源: 各公司数据、高盛全球投资研究部

医生/医院的效率。医疗领域的一项特别挑战，依然是医生的医疗实践明显滞后于新药和新治疗方法的获批。因此，许多医疗领域的机器学习和人工智能专家正不断鼓励主要的医疗服务供应商，让在其工作流程中融入现代的机器学习工具，以使其充分利用收集到的和已发表的海量医疗数据存储。

机器学习和人工智能有望降低药物发现和医疗实践之间的时间差；与此同时，它们还能对治疗进行优化。例如，从北美放射学会 2009 年对肝胆放射的研究可见，23%的第二意见会改变诊断结论，而这也是专注于医学影像的机器学习公司有望能解决的领域。此外，那些致力于利用机器学习在基因组层面进行疾病判断的公司，例如 Deep Genomics 等，正帮助供应商精确定位，以提供更有效和更有针对性的治疗。

图 4: 大型生物制药公司10年研发支出 vs. FDA批准数
十亿美元



来源: FDA, 各公司数据, 高盛全球投资研究部

注: 总研发支出包括美国大型生物科技公司 (ALXN, AMGN, BILB, CELG, GILD, REGN and VRTX), 美国大型制药公司 (ABBV post-ABBV split, BMY, JNJ, LLY, MRK and PFE) 和欧洲大型制药公司 (AZN, BAYG, GSK, NOV, NOV, ROG and SASI).

目前开展新药研发业务的通行方法为何？

目前，药物发现和开发业务是一个极为漫长的研究、测试和审批过程，持续的时间可达 10 年以上。据塔夫特药物发展研究中心报道，一款药物从第一阶段推进到通过 FDA 审批，平均需要花费 96.8 个月的时间。

新型治疗方法的研发之所以是一个独特的挑战，不仅是因为它所需要的漫长时间，而且还由于整个开发过程中各个阶段的 POS 都十分低。

药物发现始于最初的目标确定。一旦确定目标后，人们通常利用高通量筛选 (HTS) 来“命中发现”。HTS 成本十分昂贵，它由机器人自动完成，通过在同一时间进行数以百万计的试验，找出最有潜力达到目标的化合物，提高药物发现的“命中”几率。

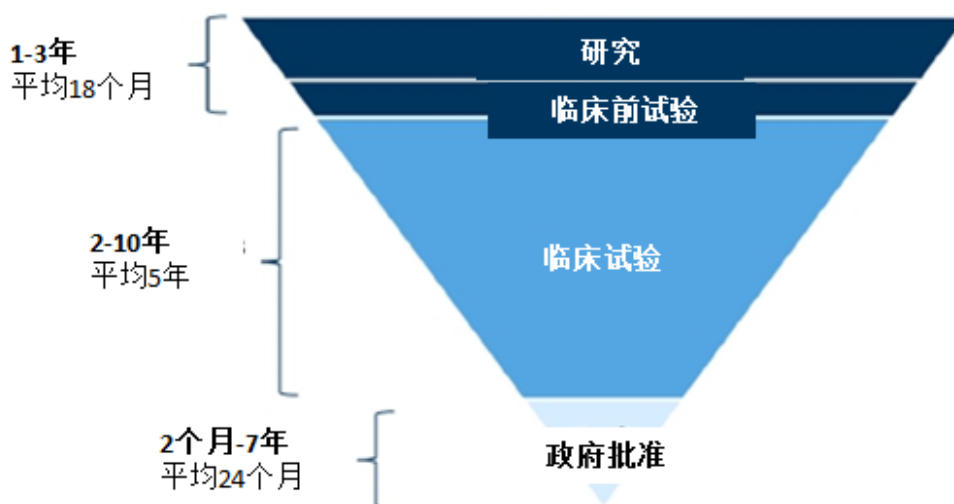
“命中”的结果经过优化成为导引化合物，然后再进一步深入优化，为进入临床前药物开发过程做好准备。当一款药物在进入第一阶段之前，这整个过程下来通常就需要 1-3 年，而它的 POS 却仅仅为 20%。

第一阶段：重点为安全性；健康的志愿者（POS 20%）。

第二阶段：重点为有效性；有某种疾病或健康状况的志愿者（POS 40%）。

第三阶段：进一步收集不同人群有关安全性和有效性、剂量和药物联用等的信息。志愿者数量为几百到数千名（POS 60%）

图 5: 药物发现与开发时间线



来源：冯嘉烈公主癌症基金会、高盛全球投资研究

AI / ML 如何起作用？

在医疗领域，完美利用机器学习和 AI 优势的案例范围很广。在那些案例中，决策和/或预测不是由人的理解或直觉，而是由数据、且是由那些远远超出人类能力范畴的各种影响因素考量而驱动的。深度学习更是展示了其独特的潜力，因为它可以利用在不同任务中学到的知识，来提高在其他任务中的绩效。

减少发现失败，增大 POS。人们将大量资本投入巨大的机会成本中，来探索、研究新的治疗方法，而在这一过程中，能达到第一阶段试验的成功概率

(POS) 大约只有 20%。因此，迄今为止，学者们已经纷纷倡导利用 AI/ML 开发有效和准确的虚拟筛选方法，以取代昂贵且耗时的高通量筛选过程。

最近，谷歌和斯坦福的研究人员正致力于利用深度学习开发虚拟筛选技术，以取代或增强传统的高通量筛选 (HTS) 过程，并提高筛选的速度和成功率。通过应用深度学习，研究人员能够实现跨越多个靶点的众多实验的信息共享。正如 Bharath Ramsundar 等在其一篇机器学习相关的论文中所称：

“我们的实验表明，深层神经网络优于所有其他方法……尤其在于，深层神经网络大大超越了所有现有的商业解决方案。

在许多靶点上，它都实现了接近完美的预测质量，使其尤其适合被用作虚拟筛选装置。总之，深度学习提供了建立虚拟筛选并将其作为药物设计管道中标准步骤的机会。” (Massively Multitask Networks for Drug Discovery, 2015/2/6)

2012 年,默克公司主持了一项由数据科学公司 Kaggle 发起的旨在确定虚拟筛选统计技术的挑战。现在，Kaggle 已经开始测试深度学习和 AI 的应用，并与 AI 药物发现初创公司 Atomwise 开展合作。

Atomwise 最近利用 AI 技术，在不到一天的时间内对现有的 7000 多种药物进行了分析测试，为寻找埃博拉病毒治疗方案做出了贡献。根据该公司的统计，如果利用传统方法，这项分析需要花费数月甚至数年才能完成。

提高医生/医院效率。在改善诊断(Enlitic, DeepMind Health)、分析放射学结果 (Zebra MedicalVision, Bay Labs)、基因组医学(Deep Genomics)等领域，甚至利用 AI 治疗抑郁症、焦虑和 PTSD (Ginger.io) 等方面，我们已经看到了应用机器学习的一些早期成功范例。

由于医疗数据数字化和数据聚合的不断发展，医疗数据将变得更易于访问。这使得 AI/ML 不仅可以削减与过程任务相关的成本，而且还可以利用算法使得过去不相交的数据集互通，以改善医疗护理本身。最终，由于 AI/ML 能够做出超出人类能力的因素考量，使得其可以帮助供应商以更高的效率进行诊断和治疗。

量化机遇

药物发现失败的成本。根据我们的分析，通过实施机器学习和人工智能，在以下情况人们有望将药物开发与发现的相关风险减半：

图6: AI/ML可削减260亿美元的研发成本
百万美元

ML/AI为药物发现/研发成本带来的机遇	
失败药物的开发成本	\$37,175
获批药物数量	43
每件获批药物的失败成本	\$865
FDA批准数	60
失败药物的总成本	\$51,872
ML/AI削减成本的可能性 (\$mn)	\$25,936
每年获批药物成本 (\$mn)	\$2,567
失败药物成本	\$865
在总成本中占比 (%)	34%

来源：德勤、FDA、高盛全球投资研究部

- 获批药物的平均年度开发成本为 16 亿美元，其中包括与失败药物相关的成本（德勤）。
- 失败药物的年度成本为 300 亿美元，这笔资金完全可以平均分配给获批的药物群（德勤）。

2015 年，FDA 报告了 60 种获批药物。这意味着算上失败药物的研发成本，该年度每种获批药物的成本约为 6.98 亿美元，其中就有将近 420 亿美元用在了失败药物上。我们认为，机器学习和人工智能可以将新药研发过程中的风险减半：到 2025 年，全球制药行业每年即可节省约 260 亿美元。

加速获得转型至电子健康记录的收益。当前，仅在美国一地，医疗信息技术人员的年度薪酬就已达到约 70 亿美元。根据 BLS 的数据，由于人口老龄化及政府向数字化转型需求的推动，预计在 2014-2024 年，医疗信息技术人员的就业前景将实现大幅度提高：相较其他所有职业 7% 的增长率，此类职业的增长将达到令人咋舌的 15%，远远高于平均水平。

然而，考虑到这项职业中的许多工作任务很容易被自动化和软件等替代，我们认为机器学习和 AI 有可能将取代几乎所有这类工作。

BLS 认为，医疗信息技术人员的任务是确保用于报销和/或研究的患者医疗数据的质量、准确性、可访问性和安全性，同时利用技术分析患者数据，以提高医护质量和控制成本。

医疗行业中 AI/ML 日益广泛的应用，可能会对这类职业产生严重影响。根据人均医疗支出和全球支出份额估计，AI/ML 有望在 2025 年在全球范围内削减超过 280 亿美元的年度成本。

图7: AI/ML能取代所有医疗信息技术员 (HIT) 的位置
百万美元。

医疗信息技术员 (HIT)	
美国中位年度岗位工资	\$51,636
美国工作岗位数量	218,776
美国年度成本(\$mn)	\$11,297
美国医疗支出 (\$mn)	
美国医疗支出 (\$mn)	\$2,998,469
全球医疗支出 (\$mn)	\$7,536,116
美国份额	40%
全球HIT成本 (\$mn)	
全球HIT成本 (\$mn)	\$28,392

来源: 人口统计局、世界银行、BLS、高盛全球投资研究部。

谁会被扰乱？

综上所述，机器学习和人工智能可以节约药物发现与开发的成本，提升 POS，为供应商和医疗设施的效率增益，因此，它们有潜力大幅度地改变大型制药公司以及整个医疗系统的前景。我们有理由相信，从长远来看，机器学习和人工智能技术必将激增，缩短研发时间，降低失败药物的损失，并使得药物开发中的竞争加剧。

此外，效率增益和自动化可能会对一些医疗专业人员和公司造成一定的混乱，特别是在解释医疗结果和诊断的人士与实际交付护理或执行手术的人士之间，例如放射科医生、提供第二意见的专家以及行政或支持人员等。我们认为这种混乱将会长期存在，因为现在许多技术仍处于早期开发阶段，并且采用这些技术的成本相对于其他改进机制可能稍显过高。

采用的挑战

虽然 AI/ML 在医疗领域的许多子领域均存在明显的机遇，但技术采用的障碍仍然存在。

成本。实施 AI/ML 需要配备必要工具和能力，但它们的成本可能十分昂贵。特别是在医疗行业中，医疗成本仍然是大众关注的焦点。为了确保 ML 算法能很好地利用数据，人们需要有意义资本和专门技术，而光是确保拥有足够的计算能力，就将花费不菲的金钱。

可解释性。算法需要梳理多个数据集，而这往往会生成一些所谓的黑盒。以前一直受到严格监管的医疗行业有可能因此会推迟 AI/ML 应用的发展。

人才。采用 AI/ML 技术的障碍也可能来自相关领域人才的聚合。2013 年，Google 支付了超过 4 亿美元收购了 DeepMind Technologies；而根据新闻报道，彼时该团队的成员大概只有十几个人。聚合这样一群高级人才的难度以及由此产生的高昂成本可能会令人望而却步。

数据。虽然美国政府已经颁布法令帮助实现电子健康记录的数字化，但将普遍使用纸张的系统转变为完全电子化的过程仍然存在挑战。此外，虽然许多机构已经迈过了“有意义使用”的门槛，但是重要患者数据的碎片化和缺乏可获得性仍然阻碍着改革的进一步发展。