

中国人工智能产业 发展报告



新智元
AI ERA

前言

60年前，达特茅斯会议召开，人工智能作为一门新兴学科在美国肇始。经过六十年的演进，人工智能即将成为改变人类经济、社会的新技术引擎。2016年9月3日，习近平主席在二十国集团工商峰会发表主旨演讲，指出以互联网为核心的新一轮科技和产业革命蓄势待发，人工智能、虚拟现实等新技术日新月异，虚拟经济与实体经济的结合，将给人们的生产方式和生活方式带来革命性变化。随着人工智能从学术课题研究全面步入产业经济爆发阶段，目前全球正迈向以互联网+、人工智能、3D打印、自动驾驶等技术代表的第四次工业革命。中国是全球创新舞台上的佼佼者，也是迈向智能时代最快的国家之一。

新智元作为专注人工智能行业的资讯和智库平台，利用自身积累优势和外部专家资源，在人工智能60年，暨首届世界人工智能大会召开之际撰写了《2016中国人工智能产业发展报告》，报告分为四个部分：产业篇、技术篇、应用热点篇与投融资篇。

产业篇主要对全球与中国的人工智能产业进行概括性介绍，概览全球产业战略与布局。技术篇聚集处于人工智能领域前沿也是产业核心关键地位的深度学习技术，由具有多年实践经验的一线技术专家撰写。热点篇关注2016年以来发展态势迅猛的智能驾驶，从技术、企业、竞争等角度展开。投融资篇除了对全球与中国人工智能领域投融资情况进行综述，特别推出“新智元100”分析报告。新智元100是新智元子品牌，关注人工智能行业的投融资情况。过去3个月的时间中新智元对中国100多家人工智能初创企业进行调研，搜集到各公司团队、技术人员占比、融资额、估值、专利数量等核心数据，经过分析写成“新智元100”分析报告整体上看，这100家企业几乎覆盖了目前中国AI创业公司中所有的发展方向和技术选型，无论在数量还是类型上都具有较强的代表性。另外，新智元评选的“2016最具竞争力人工智能创业企业”名单也将在报告中公布。

通过本报告，新智元希望提供关于中国人工智能产业的全局概览和近期发展趋势洞见。

本报告的撰写得到了许多业内人士的帮助，在此一并表示感谢。报告已尽量提供准确数据与资料，并进行多次校对，但也难免会有疏漏，如有错误，欢迎指正。

目 录

一、产业篇	1
(一) 人工智能 60 年	1
1. 人工智能：感知 + 理解 + 决策	1
2. 人工智能三“起”三“落”后迎来爆发	2
(二) 全球人工智能产业发展态势	3
1. 全球人工智能产业将进入快速增长期	3
2. 全球人工智能企业竞争日趋激烈	4
3. 人工智能已上升为国家战略	8
(三) 我国人工智能产业发展态势	10
1. 我国人工智能产业将在 2018 年突破 200 亿元	10
2. 百家人工智能企业助推产业转型升级	10
3. 国家和地方政策助力人工智能产业健康快速发展	13
二、技术篇——深度学习	15
(一) 技术演进	15
1. 深度学习：像人脑一样思考	15
2. 深度学习的三个里程碑	17
(二) 深度学习发展现状	19
1. 多家巨头力推产业布局	19
2. 三大领域技术革新	21
3. 三大开源框架促进技术落地	21
(三) 深度学习未来展望	23
三、应用热点篇——自动驾驶	25
(一) 概述	25
1. 四个等级两种路径	25
2. 无人车商用时间线：3-4 年之后	26
(二) 人工智能与自动驾驶（自动驾驶中的人工智能技术）	26
1. 感知（数据）	27

2. 决策 (计算).....	32
3. 地图.....	34
4. 车联网.....	35
(三) 无人驾驶产业.....	37
1. 国外自动驾驶发展趋势.....	37
2. 中国自动驾驶的发展趋势.....	41
3. 趋势：智能出行公司的平台优势明显，成有力武器.....	42
四、投融资篇——新智元 100 报告.....	44
(一) 全球 AI 创业公司投融资市场概览.....	44
(二) 中国人工智能创业与投融资概览.....	46
(三) 新智元 100 分析报告.....	47
(四) 新智元 100 最具竞争力榜单 Top 10.....	51
附录：新智元 100 评选榜单.....	55

一、产业篇

人工智能经过 60 年的发展，已逐渐从技术走向应用。近几年，在深度学习的推动下，人工智能取得了飞速发展。世界各国纷纷将人工智能作为国家战略，积极推动产业发展，企业将人工智能作为未来的发展方向积极布局，围绕人工智能的创新创业也在不断涌现。未来，人工智能将深刻改变人类的生产、生活方式。

（一）人工智能 60 年

1. 人工智能：感知 + 理解 + 决策

自 1956 年达特茅斯会议提出“人工智能”这个词以来，业界对“人工智能”的认知也在不断发生变化。参考 Stuart Russell and Peter Norvig 的定义，对人工智能的认知可以按思考还是行动、像人还是理性两个维度分为四种，即像人一样行动、像人一样思考、合理地思考以及合理地行动。



图表 1 对人工智能的不同认知及其特点

资料来源：《Artificial Intelligence: A Modern Approach》3rd Edition

前三种认知方式下的人工智能由于技术受限和其它一些原因，尚未实现大规模产业化应用。像人一样行动以阿兰·图灵在 1950 年提出的图灵测试为代表，强调人工智能应该像人一样行动。近年来，又有人提出全面图灵测试，增加了视觉信号和物理操纵需求，从而使图灵测试覆盖了自然语言处理、知识表示、自动推理、机器学习、计算机视觉和机器人六大学科。因此 60 年后，图灵测试对于验证一个系统是否具备智能，仍然有效。像人一样思考基于认知建模，更加强调像人一样思考。人的大脑一直是一个未解之谜，目前美国和欧盟均在开展人脑研究，一旦破解了大脑思考方式这个世界难题，类人的智能研究将取得重大突破。现阶段的人脑研究尚不足以支撑人工智能建立像人一样思考的系统。合理地思考则是逻辑主义流派提倡的通过制定规则使智能系统合理地思考。这种基于规则的认知方式开发出的人工智能系统精准度较高。但对非形式的知识制定规则并不容易，因此仅适用于规则清晰的专业领域，在通用领域中难以得到大规模应用。

以合理地行动为代表的人工智能带动了新一轮的人工智能浪潮。合理地行动现阶段以基于深度学习的人工智能为代表，强调通过感知+理解+决策来实现，是建立在大量先验知识的基础上做出的相对合理的判断和决策。尽管基于深度学习的人工智能需要基于大量先验知识做出判断，无法实现完美合理性，但在海量数据的支撑下，“感知+理解+决策”的人工智能有望不断接近完美合理，使得人工智能技术在产业上得到大规模应用。

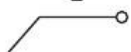
2. 人工智能三“起”三“落”后迎来爆发

人工智能研究始于 20 世纪 40 年代，从人工智能概念的诞生至今已有 60 年。根据人工智能技术及产业发展的整体形势，我们将其分为三个阶段。

第一阶段（20 世纪 50 年代中期到 80 年代初期）：深耕细作，30 年技术发展为人工智能产业化奠定基础。在 1956 年之前，人工智能就已经开始孕育。神经元模型、图灵测试的提出以及 SNARC 神经网络计算机的发明，为人工智能的诞生奠定了基础。1956 年的达特茅斯会议代表人工智能正式诞生和兴起。此后人工智能快速发展，深度学习模型以及 AlphaGo 增强学习的雏形——感知器均在这个阶段得以发明。随后由于早期的系统适用于更宽的问题选择和更难的问题时效果均不理想，因此美国、英国相继缩减经费支持，人工智能进入低谷。

第二阶段（20 世纪 80 年代初期至 21 世纪初期）：急功近利，人工智能成功商用但跨越式发展失败。80 年代初期，人工智能逐渐成为产业，第一个成功的商用专家系统 R1 为 DEC 公司每年节约 4000 万美元左右的费用。截止到 20 世纪 80 年代末，几乎一半的“财富 500 强”都在开发或使用“专家系统”。受此鼓励，日本、美国等国家投入巨资开发第 5 代计算机——人工智能计算机。在 90 年代初，IBM、苹果推出的台式机进入普通百姓家庭中，奠定了计算机工业的发展方向。第 5 代计算机由于技术路线明显背离计算机工业的发展方向，项目宣告失败，人工智能再一次进入低谷。尽管如此，浅层学习如支持向量机、Boosting 和最大熵方法等在 90 年代得到了广泛应用。

第三阶段（21 世纪初期至今）：量变产生质变，人工智能有望实现规模化应用。摩尔定律和云计算带来的计算能力的提升，以及互联网和大数据广泛应用带来的海量数据量的积累，使得深度学习算法在各行业得到快速应用，并推动语音识别、图像识别等技术快速发展并迅速产业化。2006



年，Geoffrey Hinton 和他的学生在《Science》上提出基于深度信念网络（Deep Belief Networks, DBN）可使用非监督学习的训练算法，使得深度学习在学术界持续升温。2012 年，DNN 技术在图像识别领域的应用使得 Hinton 的学生在 ImageNet 评测中取得了非常好的成绩。深度学习算法的应用使得语音识别、图像识别技术取得了突破性进展，围绕语音、图像、机器人、自动驾驶等人工智能技术的创新创业大量涌现，人工智能迅速进入发展热潮。

未来，人工智能的热度将可能会有所回落，但人工智能技术的发展将深入到金融、交通、医疗、工业等各个领域，逐渐改变人类的生产生活方式。



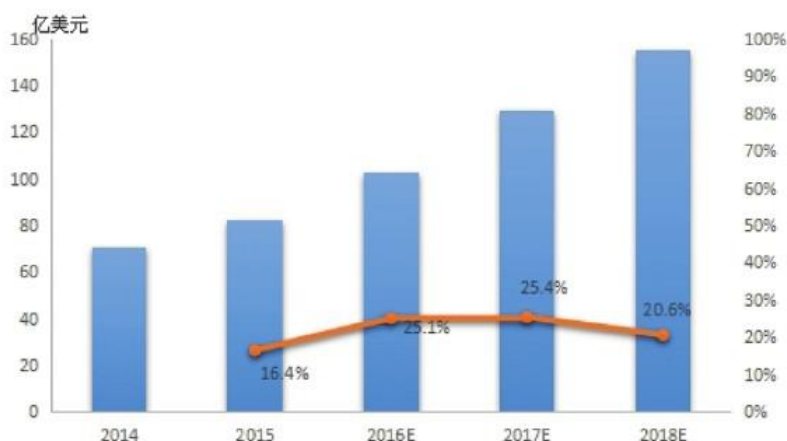
图表 2 人工智能发展历程

（二）全球人工智能产业发展态势

经过 60 年的发展，人工智能在深度学习、海量数据和高性能计算的支撑下，现已进入产业化应用初期。2016 年，基于深度学习的智能语音、图像识别、智能驾驶等技术开始向各个应用领域渗透，全球人工智能产业规模快速增长。为抢占人工智能高地，谷歌、微软、IBM、Facebook 等企业在人工智能领域的战略布局进一步突出，围绕人工智能的创新创业进一步繁荣。美国、日本等国家也先后出台人工智能相关政策及国家计划，为产业发展创造良好的生态环境。

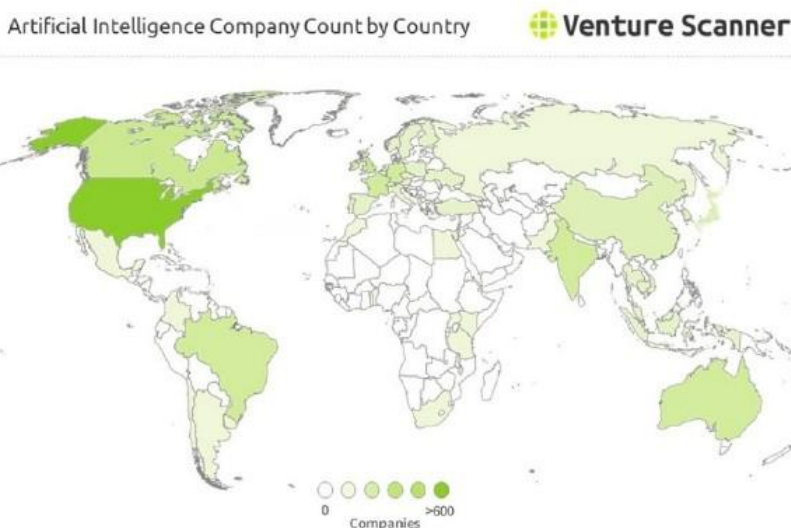
1. 全球人工智能产业将进入快速增长期

在深度学习技术和开源平台的推动下，人工智能技术门槛逐渐降低，受到全球下游应用需求的迫切倒逼，人工智能赢得了加速发展的黄金期，围绕人工智能的应用和创新不断涌现。2015 年，全球人工智能产业规模达到 82.2 亿美元，预计 2016 年将突破 100 亿美元。据 BBC 预计，2020 年全球人工智能市场规模将达 183 亿美元。在未来 10 年甚至更久的时间里，人工智能将是众多智能产业技术和应用发展的突破点。



图表3 2014-2020年全球人工智能产业规模

除产业规模快速增长外，围绕人工智能的创业企业数量也大幅提升。根据 Venture Scanner 对全球 71 个国家人工智能公司的统计，截至到 2016 年第三季度，全球人工智能创业公司数量已有 1287 家，其中 585 家获得投资，投资金额总计达到 77 亿美元，其中美国投资金额超过 31 亿美元。



图表4 2016年全球人工智能企业区域分布

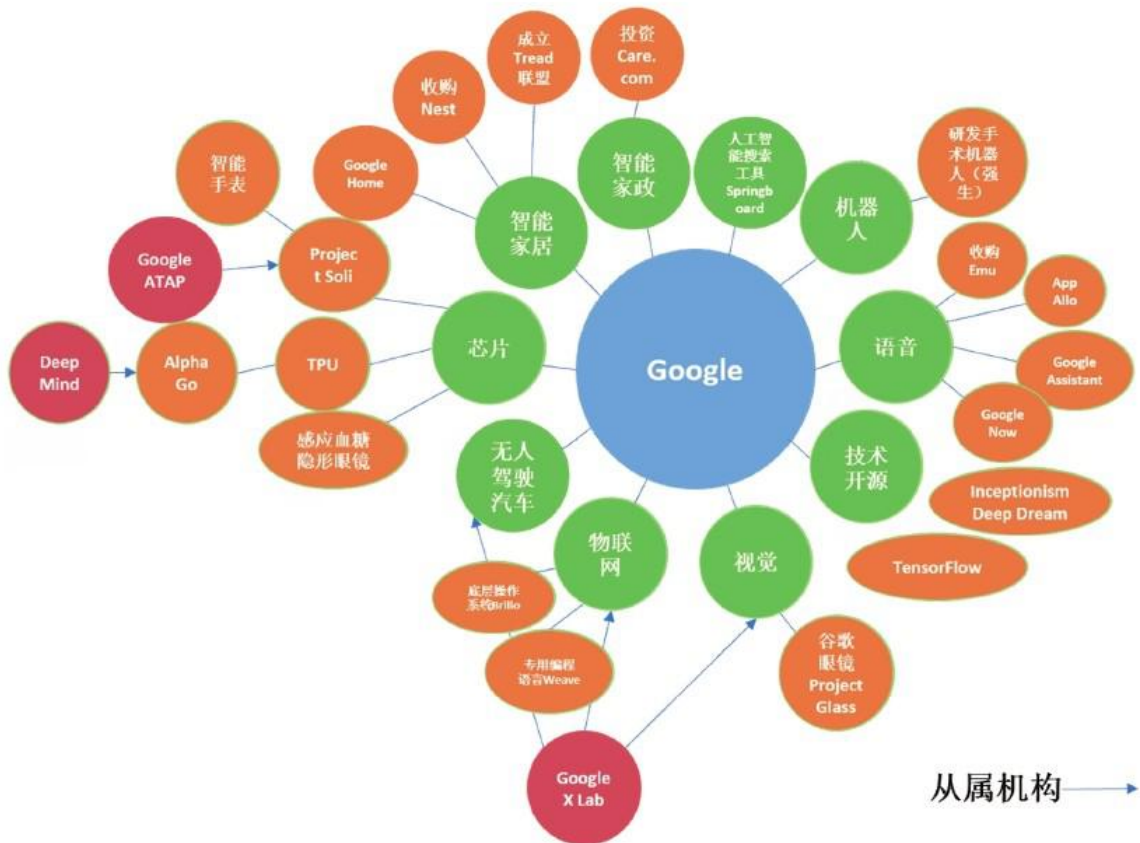
资料来源：Venture Scanner

2. 全球人工智能企业竞争日趋激烈

谷歌、微软、IBM、Facebook 等企业凭借自身优势，积极布局整个人工智能领域。各大企业通过加大研发投入力度、招募高端人才、建设实验室等方式加快关键技术研发；同时，通过收购等方式吸收人工智能优秀中小企业来提升整体竞争力；此外，各大企业还积极开放、开源技术平台，构建围绕自有体系的生态环境。

2016 年 4 月，谷歌 CEO Sundar Pichai 第一次明确提出将 AI 优先作为公司大战略。谷歌以深度

学习技术为依托，涉足人机交互、语言理解、机器人等人工智能核心技术应用领域，全方位布局人工智能产业。技术方面，谷歌通过加强自身技术水平，提升谷歌传统搜索、翻译和社交业务；推动集视、听、说、感知和控制于一体的无人驾驶汽车；并先后开源了第二代机器学习平台 TensorFlow 以及自然语言理解软件 SyntaxNet 的源代码，引领互联网巨头在人工智能领域开源的趋势改为浪潮。谷歌通过对 DeepMind 等人工智能行业创业企业的并购以及与强生、福特等传统产业巨头的合作，实现人工智能领域的全面布局及纵深式发展。



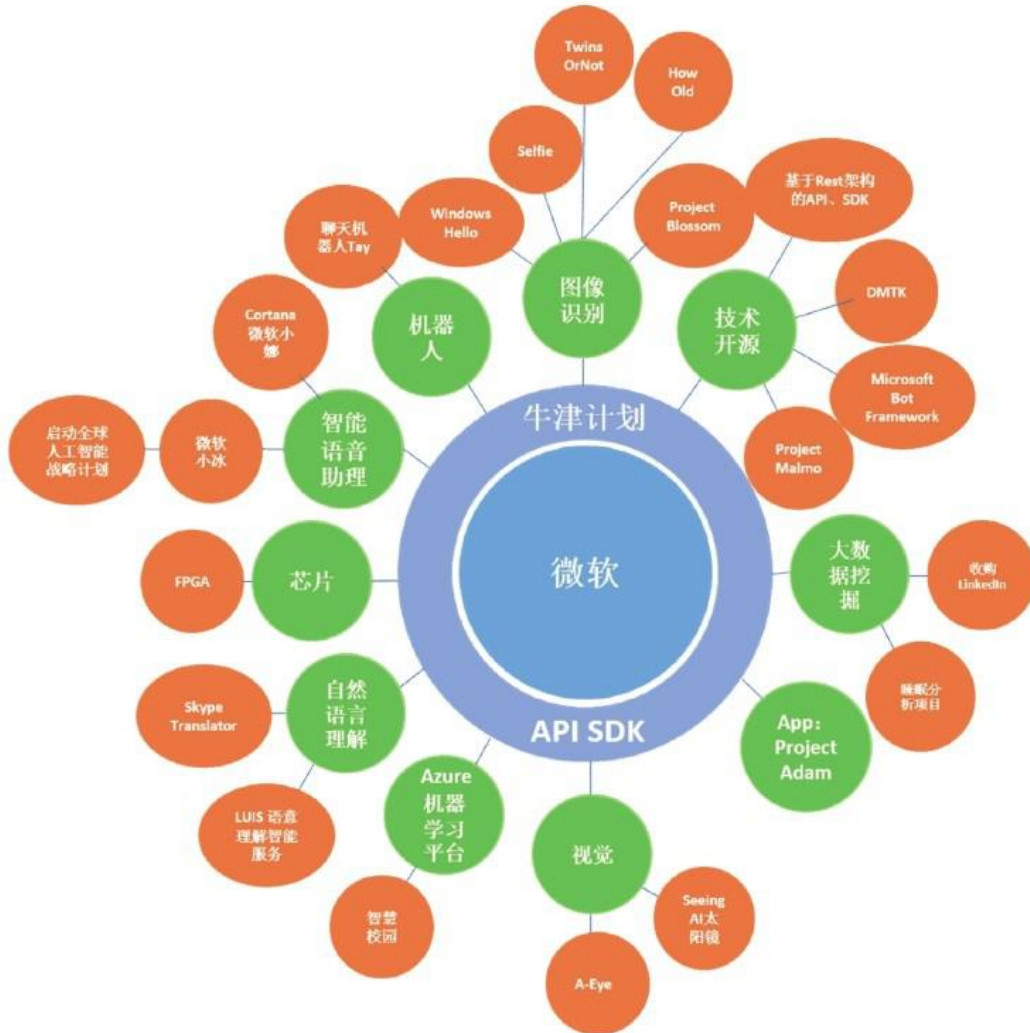
图表 5 谷歌的人工智能布局

微软在人工智能领域动作不断。微软研究院是最早开始从事人工智能研究的，发布了 Cortana 和 Skype Translator 等一系列产品。微软在 2015 年 5 月初发布了人工智能领域的牛津计划，由一系列基于云端的机器学习相关的 API、SDK 和相关服务等组成，旨在让开发人员们不需要繁复的机器学习背景也能开发跨平台的更智能和更交互的应用。当时，牛津计划率先开源了人脸识别、语音处理和计算机视觉三个部分。

2016 年 1 月 25 日，微软在 GitHub 上发布了其深度学习工具包——Computational Network Toolkit (简称 CNTK)。CNTK 是一个统一的深度学习工具包，它通过一个有向图将神经网络描述为一系列计算步骤。在有向图中，叶节点表示输入值或网络参数，边表示输入之上的矩阵运算。CNTK 使得实现和组合前馈型神经网络 DNN、卷积神经网络 (CNN) 和循环神经网络 (RNN/LSTM) 变得非常容易。

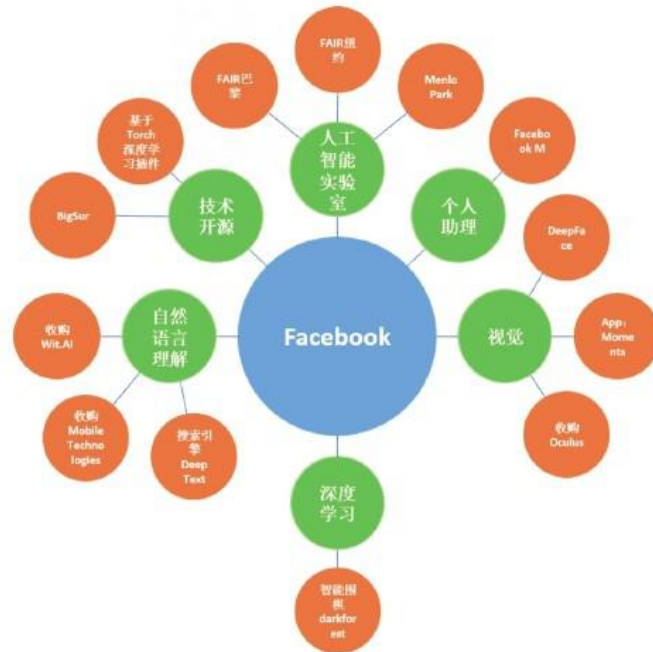
在 2016 年的 Window Build 开发者大会上，微软开放了 Microsoft Bot Framework，开发者可以

直接接入 Microsoft Bot Framework 来开发类似微软小冰这样的聊天机器人。9 月，公司更宣布成立 5000 人的人工智能部门，与 Windows 和 Office、云计算等部门并列。



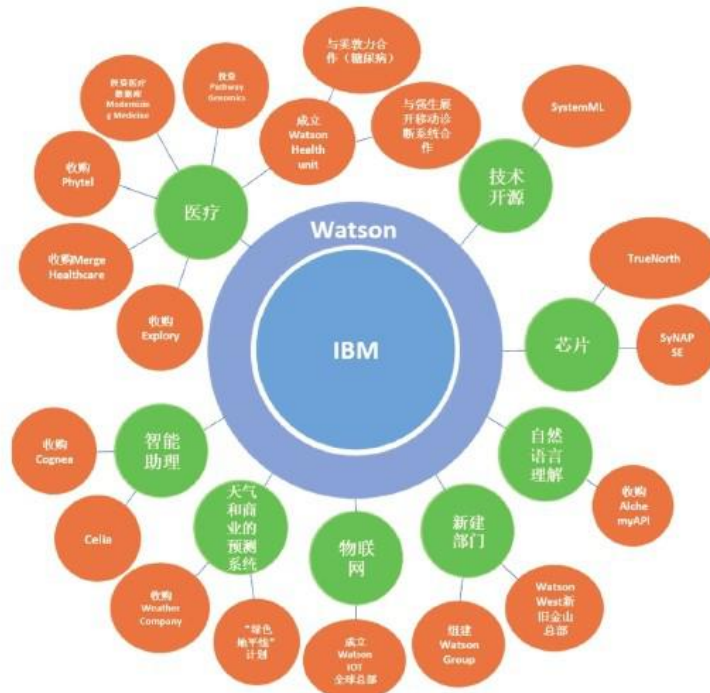
图表 6 微软人工智能布局

Facebook 积极组建人工智能实验室，并通过开源技术平台等方式来获取更大的成功。Facebook 现有两大实验室，其中一个是重点发展基础研究的 Facebook AI (FAIR) 项目，由 Yann LeCun 负责，主要专注于基础科学和长期研究。FAIR 现有三个实验室，分别位于美国纽约、加州门洛帕克以及法国巴黎，拥有超过 130 位人工智能专家。AML (应用机器学习部门) 则专注于人工智能产品应用，由西班牙裔机器学习专家 Joaquin Candela 负责。AML 正试图为排名、广告、搜索、语言翻译、语音识别、自动产生视频字幕以及自然语言理解等领域开发更好的算法以提升 Facebook 的基础。为了进一步弥补在语音技术方面的短板，2015 年初，Facebook 收购了语音指令创业公司 Wit.AI，之后建立了语言技术部门。为了进一步提升技术水平，Facebook 先后开源 fbcunn——用于在 Torch 上更快速地训练神经网络的模块、人工智能硬件平台 Big Sur 等十余个项目。Facebook 将其 AI 系统进行开源，有助于整个产业加速发展。



图表 7 Facebook 人工智能布局

IBM 在认知计算平台 Waston 项目上持续投入，并成立专门部门推动 Watson 商业化，目前 Waston 海量内容的分析能力已在医疗和金融领域率先使用。2014 年 9 月，IBM 发布基于自然语言处理的数据分析服务沃森分析 (Watson Analytics)，并向公众提供免费试用。Watson Analytics 云服务支持用户上传数据并通过英语语句查询分析结果，还能根据不同的分析场景向用户建议数据清洗步骤，以及最佳的数据可视化方法。2015 年 5 月，IBM 宣布 14 家来自美国和加拿大的癌症治疗机构将开始部署沃森 (Watston) 计算机系统，该套系统将能根据病人肿瘤的基因指纹选择出适合的治疗方案。



图表 8 IBM 人工智能布局

人工智能发展条件的成熟同时催生了大量人工智能创业企业。Venture Scanner 将 1139 家人工智能公司划分为 13 个细分行业，包括深度学习 / 机器学习（通用）、深度学习 / 机器学习（应用）、自然语言处理（通用）、自然语言处理（语音识别）、计算机视觉 / 图像识别（通用）、计算机视觉 / 图像识别（应用）、手势控制、虚拟私人助手、智能机器人、推荐引擎和协助过滤算法、情境感知计算、语音翻译、视频内容自动识别 13 个细分行业。其中机器学习（应用）分类以约 300 家企业的数量遥遥领先，自然语言处理公司数量位列第二。

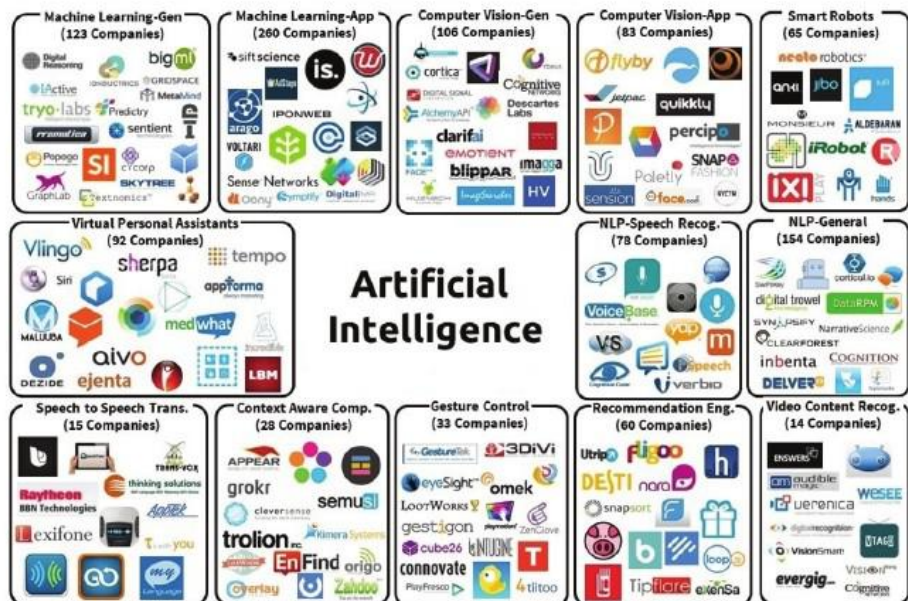


图 9 2016 年第一季度全球人工智能创业企业图谱

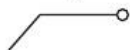
资料来源：Venture Scanner

随着人工智能的持续发展，围绕人工智能的竞争将日趋激烈。纵观智能语音、智能图像、自然语言处理、智能驾驶等人工智能技术的广泛应用，人工智能已成为推动产业升级、创新发展的关键动力。未来将会有越来越多的企业融入到人工智能产业之中，推动新一轮的技术浪潮。人工智能在众多领域的应用潜力，也会给行业发展带来无限的想象力。

3. 人工智能已上升为国家战略

新一轮的人工智能浪潮受到各国政府的高度关注，美国、日本、韩国等国家近几年纷纷出台多项战略、计划积极推动人工智能发展，人工智能已逐渐上升为国家战略。

美国于 2013 年启动创新神经技术脑研究（BRAIN）计划，由美国国立卫生研究院（NIH）、国家科学基金会（NSF）、DARPA、白宫科技政策办联合承担，并计划 10 年投入 45 亿美元。DARPA 于 2015 年召开“未来技术论坛”，预测未来 30 年的技术发展。美国白宫于 2016 年 5 月宣布成立“人工智能和机器学习委员会”，用于协调全美各界在人工智能领域的行动，并将在奥巴马任期结束前多用人工智能提高政府办公效率。



图表 10 美国人工智能相关战略、计划

时 间	战略、计划	备 注
2013 年 4 月	“推进创新神经技术脑研究计划”(BRAIN)	政府拨款 1.1 亿美元
2014 年	NIH 小组制定未来十年详细计划	十年总投资 45 亿美元
2015 年 10 月	DARPA “未来技术论坛”	未来 30 年技术发展预测
2015 年 11 月	CSIS 发布《国防 2045：为国防政策制定者评估未来的安全环境及影响》报告	指出人工智能是影响未来安全环境的重要因素
2016 年 2 月	DARPA 表示正在发展人工智能技术，以奠定其理论基础	支撑美国第三次“抵消战略”
2016 年 5 月	美国白宫成立人工智能和机器学习委员会	探讨制定人工智能相关政策和法律

日本政府近年来高度重视人工智能技术。2015 年 1 月，日本发布“新机器人战略”，希望通过发展机器人技术，推动工业生产力的提高。同年，日本政府先期投入 10 亿日元在东京成立“人工智能研究中心”，集中开发人工智能相关技术。2015 年底，日本政府发布第五个科学与技术基础五年计划。计划中，日本政府提出了名为“超级智能社会（super smart society）”的未来社会构想，要发展信息技术、人工智能以及机器人技术，整个计划的预算为 26 万亿日元。日本还于 2016 年制定高级综合智能平台计划（AIP），是为实现日本第五个科技基础五年计划、建设“超级智能社会”而提出的人工智能、大数据、物联网、网络安全综合发展计划。

图表 11 日本人工智能相关战略、计划

时 间	战略、计划	备 注
2015 年 1 月	新机器人战略	通过发展机器人技术，推动工业生产力的提高。
2015 年	人工智能研究中心	前期投入 10 亿日元。
2015 年 12 月	第五个科学与技术基础五年计划	提出名为“超级智能社会（super smart society）”的未来社会构想，发展信息技术、人工智能以及机器人技术，预算 26 万亿日元。
2016 年	高级综合智能平台计划（AIP）	人工智能、大数据、物联网、网络安全综合发展计划。

2013 年 5 月，韩国启动 Exobrain 计划，计划由韩国未来创造科学部（MSIP）主持，历时 10 年，总预算为九千万美元，计划的目标是开发专业领域人机交流的自然语言对话系统。2014 年，韩国发布的第二个智能机器人总规划（2014-2018）希望能将机器人产业与其他制造业和服务业相结合，保持在机器人技术及相关重点产业的的优势。2015 年，韩国未来创造科学部（MSIP）发布了 Star Lab 软件研发项目，人工智能是五大关键领域之一。

图表 12 韩国人工智能相关战略、计划

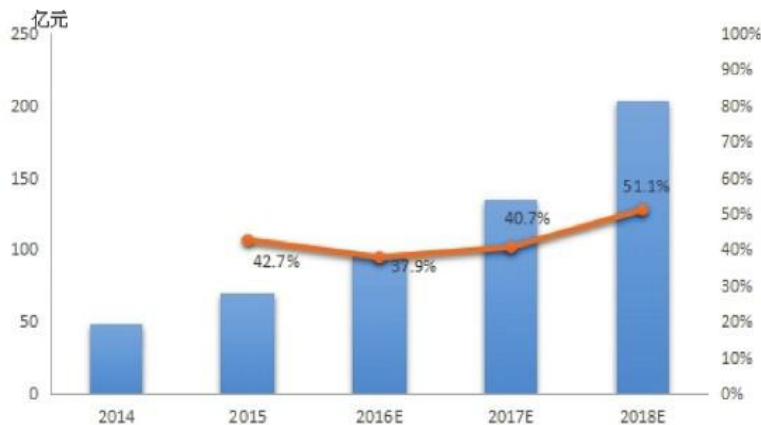
时 间	战略、计划	备 注
2013 年 5 月	Exobrain 计划	历时 10 年，总预算为九千万美元，计划的目标是开发专业领域人机交流的自然语言对话系统。
2014 年	第二个智能机器人总规划（2014-2018）	将机器人产业与其他制造业和服务业相结合，保持在机器人技术及相关重点产业的的优势。
2015 年	AI Star Lab	人工智能是项目五大关键领域之一。

（三）我国人工智能产业发展态势

在全球人工智能浪潮下，我国人工智能产业正在积极健康发展。2016年，我国人工智能产业规模将进一步提升。产业规模增长的背后是百度、阿里巴巴、腾讯、科大讯飞等企业在人工智能领域的不懈努力，以及围绕人工智能积极创新的中小企业。此外，国家和地方政府通过战略指引、政策支持等方式积极推动我国人工智能产业做大做强。

1. 我国人工智能产业将在 2018 年突破 200 亿元

人工智能技术在我国移动互联网、智能家居、无人驾驶等领域的应用继续不断深入，人工智能产业规模持续高速增长。根据新智元统计，2015年，中国人工智能产业规模进一步扩大，达到69.33亿元，同比增长42.65%。预计2016年，中国人工智能产业规模将达到95.61亿元。此后，在无人驾驶及机器人等应用的推动下，人工智能产业规模快速增长，预计2018年将突破200亿元，并带动相关产业规模增长超过1000亿元。



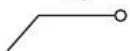
图表 13 2014-2020 年中国人工智能产业规模

2. 百家人工智能企业助推产业转型升级

我国人工智能产业快速增长的背后，是不断壮大的人工智能企业和不断涌现的围绕人工智能的创新创业。百度、阿里巴巴、腾讯、科大讯飞等将人工智能作为整体战略提前布局。围绕人工智能的创业企业瞄准细分市场深耕细作，希望在未来人工智能竞争中占据一席之地。

2.1 百度

总体来看，百度希望依托人工智能技术进行全面转型，建立完整的人工智能生态体系。在国内的互联网巨头公司中，百度最早开始人工智能战略布局。在底层基础资源支撑方面，百度拥有网络搜索引擎核心业务积累的丰厚的数据资源，全国多个计算中心和国内最大的 GPU 集群，为其提供了高性能的计算能力。在此基础上，百度重点发力人工智能技术的自主研发，先后成立了深度学习研究院、大数据研究院、硅谷人工智能实验室以及硅谷智能驾驶团队，并聘请吴恩达等人工智能顶级专家，开展机器学习、深度学习、机器人、图像识别、语音识别、无人驾驶等各个人工智能领域的技术研究。



在技术研发的基础上，百度也积极将实验室中的技术投入产品进行实践。基础功能方面，百度基于智能语义、图像识别技术推出语音搜索、百度识图等产品，并在百度主流产品中均加入了这部分基础功能。商业实践方面，百度的人工智能技术与百度外卖、百度糯米等应用深度融合，通过深度学习算法，利用海量的 O2O 线上数据进行推算，从而帮助用户规划时间、路线，提升工作效率。传统产业方面，百度目前已经将图像识别、数据风控技术用于信贷产品的审批当中，在提高审批效率的同时有效控制风险。相对来说，金融、医疗、教育等行业的应用还处于起步阶段。新兴技术产业方面，无人车成为百度的重点方向，投入较大。由于无人车的发展和普及需要法律法规、交通等全面配套，因此短期内难以获利。



图表 14 百度人工智能布局

2.2 阿里巴巴

2016 年之前，阿里巴巴重点开放计算资源及人工智能共性技术，并将人工智能统一到云服务中进行宣传推广，以实现错位营销。阿里巴巴从 2012 年开始组织团队从事人工智能研究，经过多年的厚积薄发，2015 年推出可视化人工智能平台 DT PAI，集成了阿里的核心算法库。在此技术基础上，阿里推出了虚拟助理阿里小蜜以及智能程序阿里小 Ai。2016 年 8 月，阿里在小 Ai 的基础上推出 ET 机器人，ET 机器人拥有智能语音识别、图像或视频识别、情感分析等技术。阿里云首席科学家周靖人介绍，ET 目前处于 1.0 阶段，已初步具备听、说、看的感知能力，未来能够在交通、工业生产、健康等领域输出决策。此外，阿里将其人工智能技术与电商平台、大数据、云计算等原有业务相融合，提升各板块的技术水平，并且成立智能生活事业部，整合电商、数据、平台资源，与其他厂商合作，推出智能家居产品。



图表 15 阿里巴巴人工智能布局

2.3 腾讯

尽管人工智能起步相对较晚，但腾讯通过收并购、基础技术研究合作以及三大硬件平台积极布局，实现了人工智能的跨越式发展。收并购方面，自 2013 年起，腾讯先后投资搜狗、ScaledInference、Skymind、CloudMedx、iCarbonX 以及 Diffbot，总投资额超过 5 亿美元。基础技术研究合作方面，腾讯成立 WHAT LAB、优图实验室、微信北京研发中心以及智能计算与搜索实验室，并于 2016 年成立 AI lab，聚焦自然语言处理、语音识别、机器学习、计算机视觉等四大发展方向。围绕上述技术研究，腾讯推出文智自然语言处理技术、机智机器学习平台、优图人脸识别技术，以及机器人微宝和虚拟机器人 DreamWriter 等各种人工智能技术产品。腾讯还积极打造 QQ 物联、微信智能硬件以及 TOS+ 战略，在实现“人与人”连接的基础上，进一步实现“人与物”及“物与物”的连接。

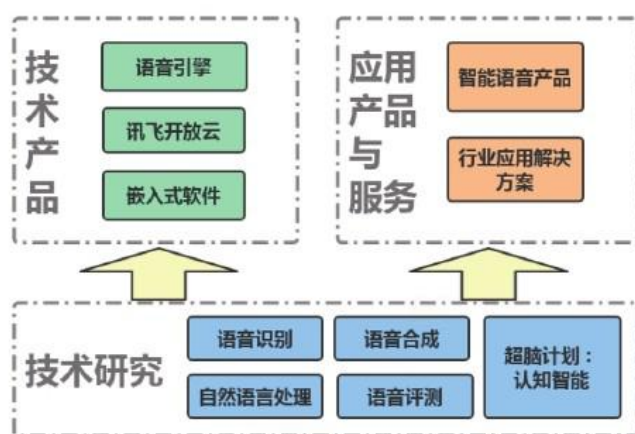


图表 16 腾讯人工智能布局



2.4 科大讯飞

不同于互联网巨头依托数据资源和技术基础发展人工智能技术的路径，科大讯飞以语音技术为产业化方向，核心技术集中于语音识别、语音合成、自然语言处理等几个方向。在此基础上，科大讯飞构建了全面的语音技术平台，推出语音技术引擎、讯飞语音云以及嵌入式软件，提供语音技术服务。此外，科大讯飞开发了多种语音应用软件，在电信、公共安全、教育、智能建筑等领域提供相关行业应用解决方案。作为国内最大的语音技术供应商，科大讯飞与电信、金融、家电、汽车等多个主流行业的龙头企业建立合作关系，形成对语音市场的绝对占有。



图表 17 科大讯飞人工智能布局

科大讯飞超脑计划

科大讯飞目前正在执行的讯飞超脑计划，就是想实现包括感知智能和认知智能在内的全面突破。在感知智能领域，语音识别、手写识别方面每年保证 30% ~ 50% 的错误率的下降。讯飞在认知智能上的研究目标，关键是让机器能理解会思考，这必须要突破语言理解、知识表示、联想推理，自主学习等多个方面。

美国华盛顿图灵中心在研究如何让机器人通过美国高中生物测试。日本国立情报研究所在研发能够考上东京大学的高考机器人。我国 863 计划也正在开展一个题为“类人答题系统”的科研项目，其目的就是要研发高考机器人。该项目包括 9 个课题组，全国共有 31 家单位参加联合攻关，其中科大讯飞作为牵头单位。

除龙头企业外，我国近几年围绕人工智能的创新创业大量涌现。新智元对 14 个应用领域的人工智能创业企业从营收规模、估值等角度进行梳理分析，具体内容将在投融资篇呈现。

3. 国家和地方政策助力人工智能产业健康快速发展

我国政府近几年持续关注人工智能发展，并针对人工智能制定多项国家战略。2015 年 5 月，国务院发布的《中国制造 2025》中明确提出“加快发展智能制造装备和产品”。2015 年 7 月，《国

务院关于积极推进“互联网+”行动的指导意见》中将人工智能作为重点布局的11个领域之一。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）》指出将重点突破新兴领域人工智能技术等。2016年5月，发改委、科技部等4部门联合印发《“互联网+”人工智能三年行动实施方案》。方案提出，到2018年，打造人工智能基础资源与创新平台，人工智能产业体系、创新服务体系、标准化体系基本建立，基础核心技术有所突破，总体技术和产业发展与国际同步，应用及系统级技术局部领先。在重点领域培育若干全球领先的人工智能骨干企业，初步建成基础坚实、创新活跃、开放协作、绿色安全的人工智能产业生态，形成千亿级的人工智能市场应用规模。

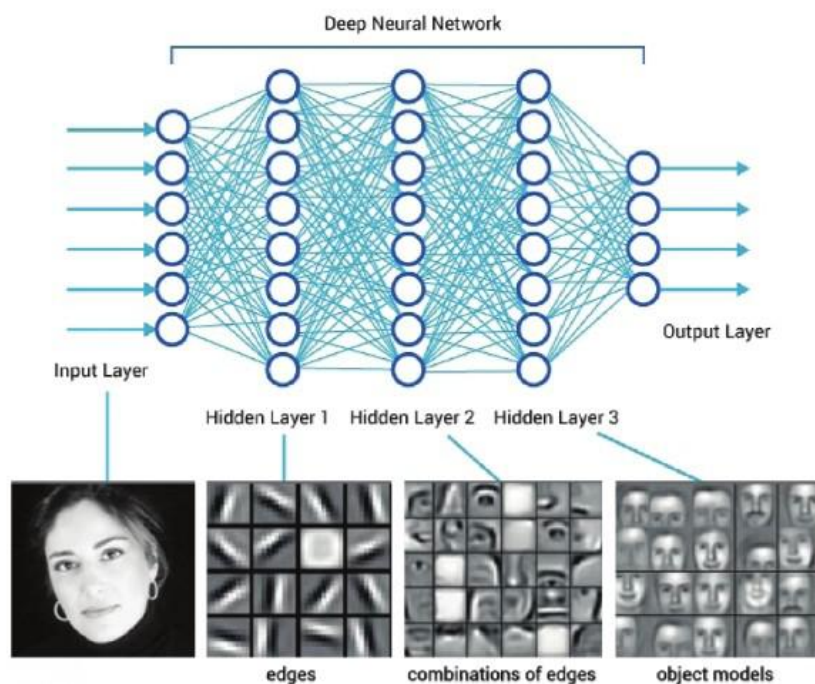
图表 18 我国人工智能相关政策

时 间	发布单位	战略、规划	备 注
2015年5月	国务院	《中国制造2025》	明确提出“加快发展智能制造装备和产品”
2015年7月	国务院	《国务院关于积极推进“互联网+”行动的指导意见》	明确提出人工智能作为重点布局的11个领域之一。
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）》	重点突破新兴领域人工智能技术等。
2016年5月	发改委	《“互联网+”人工智能三年行动实施方案》	到2018年，打造人工智能基础资源与创新平台，人工智能产业体系、创新服务体系、标准化体系基本建立，基础核心技术有所突破，总体技术和产业发展与国际同步，应用及系统级技术局部领先。

二、技术篇——深度学习

(一) 技术演进

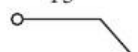
1. 深度学习：像人脑一样思考



图表 19 深度学习结构示意图¹

深度学习 (Deep Learning)，是目前人工智能领域最流行的技术。具体来讲，深度学习模型由一系列相互关联的神经元组成，经训练后得到关联权重，数据通过整个网络便可自动得到更具有语义

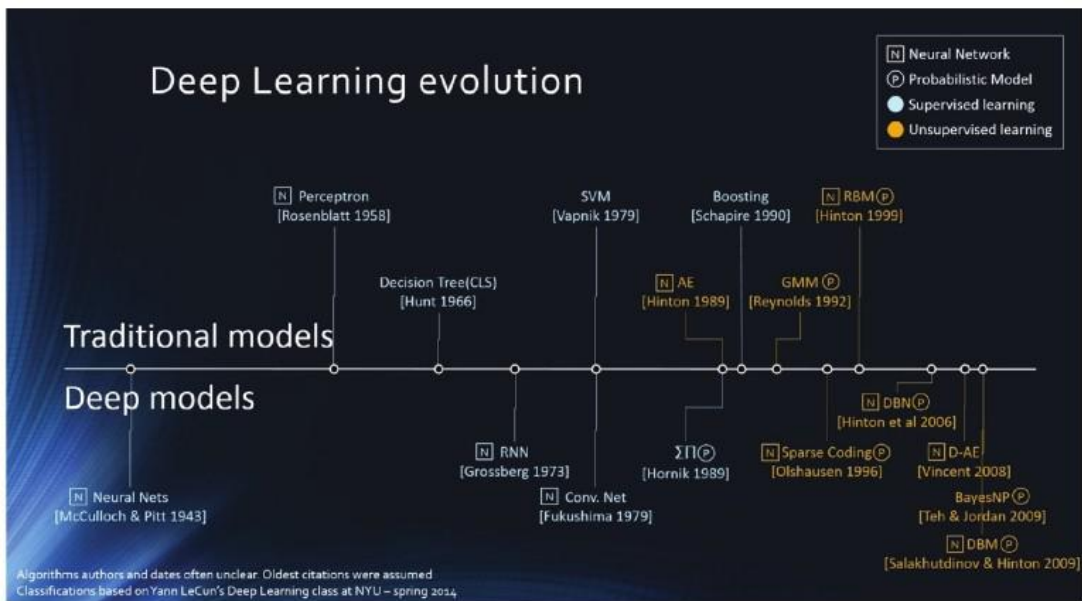
¹ <http://www.amax.com/blog/?p=804>



的特征表示，进一步可以直接用于分类图像和语音，甚至是控制无人机或是无人车。

深度学习试图使用包含复杂结构或由多重非线性变换构成的多个处理层对数据进行高层抽象，与目前的人脑模型接近，符合人类层次化地组织概念、由简至繁分层抽象的认知过程。深度学习可以模拟人脑从外界环境中学习、理解甚至解决模糊歧义的过程。与浅层学习相比，深度学习最大的不同在于利用简单的神经网络实现更层次化的特征表示，取代人工挑选的复杂特征（即特征工程）表示，并能够在具体任务上达到更好的效果。

深度学习的概念并不新颖，但直到近年才得到认可。业界普遍认为，是超大规模训练数据、复杂的深层模型和分布式并行训练造就了今天深度学习的繁荣。具体来讲，包含多个隐藏层的神经网络，利用现在的高性能计算机和人工标注的海量数据，通过迭代得到超过浅层模型的效果。深度学习带来了模式识别和机器学习方面的革命。



图表 20 神经网络发展时间线

目前的深度学习模型可以分为神经网络和深度模型两条发展路径如下表所列。

图表 21 神经网络方面的发展

1943 年 McCulloch 和 Pitt 设计的人工神经元
1958 年 Rosenblatt 提出感知机
1979 年 Fukushima 设计最早的卷积神经网络，已经具备了多层卷积和池化，但是训练方法上还存在缺陷
1985 年 Hinton 将反向传播引入神经网络训练，发现可以得出带有语义的网络结构
1985 年业界提出认知学问题：人类是否也是类似的依赖神经元连接来进行认知
1989 年 Hinton 提出自编码 (Auto Encoder)
1999 年 Hinton 等人提出受限玻尔兹曼机

图表 22 深层模型发展

1965 年乌克兰数学家 Ivakhnenko 发表深度前向多层感知器;
1971 年 Ivakhnenko 设计八层网络;
1986 年 Dechter 将深度学习引入到机器学习领域;
2000 年 Aizenberg 等将其与神经网络结合;

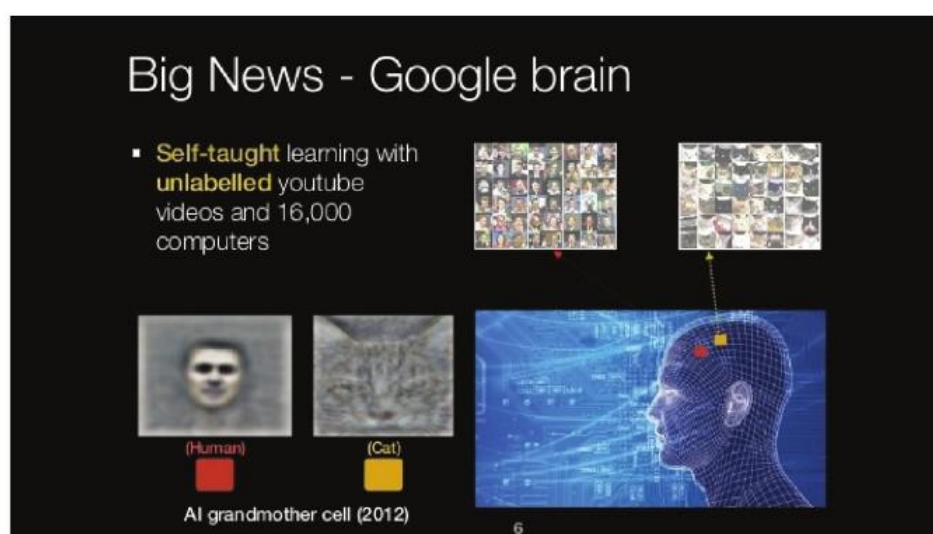
在应用方面，第一个成功的例子是 1989 年 LeCun 利用卷积神经网络结合反向传播训练方法进行手写体识别，后期推广到全美支票识别。然而由于人工智能行业整体的不景气，训练速度和规模受限等问题，包括 1997 年 Hochreiter 提出的 LSTM 等技术的应用不如 1995 年 Vapnik 提出的支持向量机方法广泛。

21 世纪初，伴随着信息革命的发展，计算机性能大幅提升，大规模并行计算设备如 GPU 等的出现，使得深度学习模型的训练过程可以提速千倍，在同等海量数据规模的情况下，其表现超过了支持向量机等浅层模型，并随着数据的增加可获得持续改进，业界对其应用前景的评估日益乐观，开始源源不断地投入深度学习的改进中。

2. 深度学习的三个里程碑

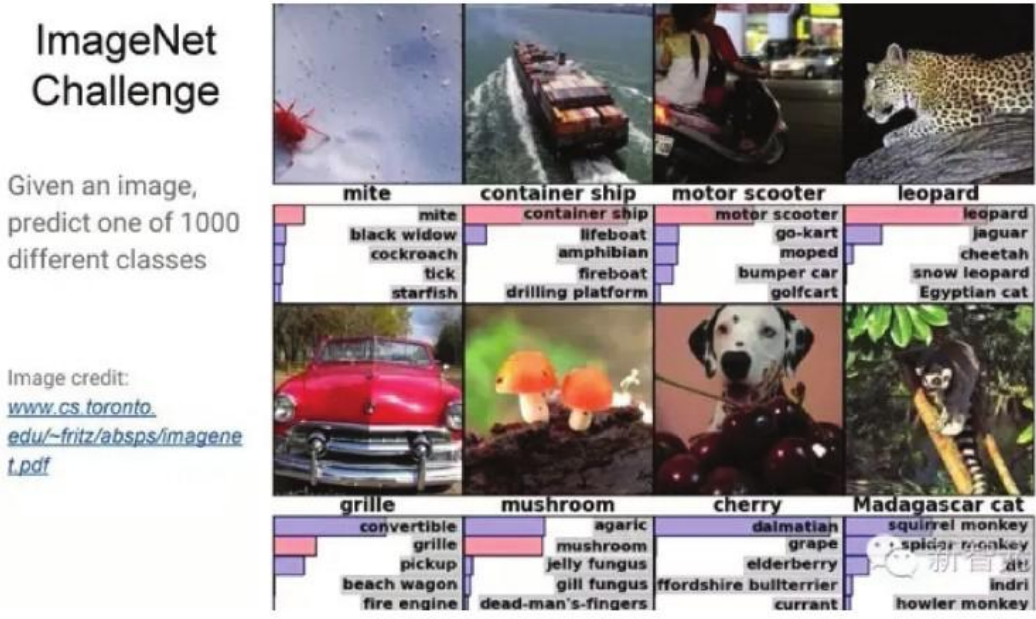
任何技术的发展过程，都伴随着一些重要的突破节点，在深度学习技术方面，可划分为模型初步、大规模尝试和遍地开花三个阶段。

2006 年前后，深度模型初见端倪，这个阶段主要的挑战是如何有效训练更大更深层次的神经网络，曾一度因为梯度方法容易造成层次间信息损失而效果受到影响。业界尝试利用逐层预训练的方法（还有一种尝试方法是上文提到的 LSTM）来解决该问题，预训练首先通过无监督学习得到一些比较稳定的特征，而后用监督学习稍加调整便可得到较好效果，最早是 1992 年 Schmidhuber 提出，而后 Hinton 等人在 2006 年改进为前向反馈。



图表 23 Google 大脑计划

2011 年底，大公司逐步开始进行大规模深度学习的设计和部署。“Google 大脑”项目启动，由时任斯坦福大学教授的吴恩达和 Google 首席架构师 Jeff Dean 主导，专注于发展最先进的神经网络。初期重点是使用大数据集以及海量计算，尽可能拓展计算机的感知和语言理解能力。该项目最终采用了 16000 个 GPU 搭建并行计算平台，以 YouTube 视频中的猫脸作为数据对网络进行训练和识别，引起业界轰动，此后在语音识别和图像识别等领域均有所斩获。



图表 24 ImageNet 挑战赛 资料来源：新智元

近年来深度学习获得了非常广泛的关注，其进展的一个直观的体现就是 ImageNet 竞赛。在这个竞赛中参赛算法在数千个种类图像和视频的大规模数据上测试检测和分类的正确率快速上升。

Neural Nets: Rapid Progress in Image Recognition

Team	Year	Place	Error (top-5)	ImageNet challenge classification task
XRCE (pre-neural-net explosion)	2011	1st	25.8%	
Supervision (AlexNet)	2012	1st	16.4%	
Clarifai	2013	1st	11.7%	
GoogLeNet (Inception)	2014	1st	6.66%	
Andrej Karpathy (human)	2014	N/A	5.1%	
BN-Inception (Arxiv)	2015	N/A	4.9%	
Inception-v3 (Arxiv)	2015	N/A	3.46%	



图表 25 2011 年至 2015 年 ImageNet 图片分类结果 资料来源：新智元

近年来，随着 GPU 的提速，深层网络的很多训练过程已无需预训练就可直接进行，如 2011 年 Ciresan 等利用深度卷积神经网络赢得多项文本识别、交通标志识别和医学图像识别等竞赛。2012 年之前，竞赛中物体的识别率一直提升得非常缓慢（低于 70%）。在 2012 年引入深度学习之后，Hinton 等人赢得 ILSVRC-2012 ImageNet，识别率一跃升至 80%，现在已超过 95%。这标志了人工特征工程正逐步被深度模型所取代。从排名来看，深度学习目前具有压倒性优势，已经取代了传统视觉方法在这一竞赛中的地位。

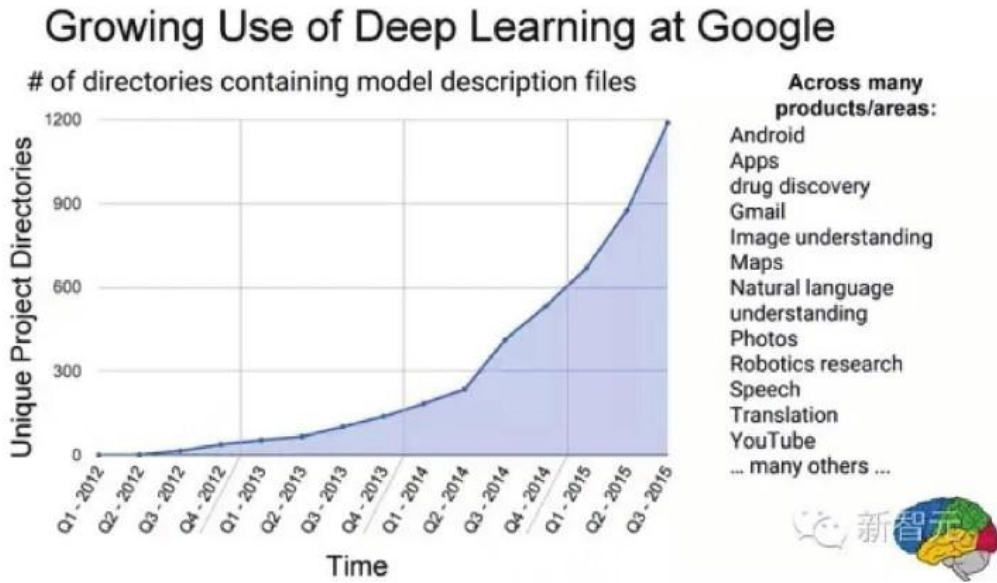
进一步地，强化学习也重新成为焦点，2016 年 Google 子公司 DeepMind 研发的基于深度强化学习网络的 AlphaGo，与人类顶尖棋手李世石进行了一场“世纪对决”，最终赢得比赛。AlphaGo 突破了传统程序，搭建了两套模仿人类思维的深度学习：价值网络承担棋局态势评估，策略网络选择如何落子。

（二）深度学习发展现状

1. 多家巨头力推产业布局

技术的发展离不开产业巨头的推进，深度学习也不例外。Facebook 是最早使用 GPU 加速 DNN 训练的公司之一。深度网络和 GPU 在 Facebook 人工智能研究院（FAIR）推出的专门用于深度学习训练的计算平台中扮演着重要的作用。Facebook 寄希望以此推动机器智能的发展并帮助人们更好的交流，目前在信息推荐、过滤攻击言论、推荐热门话题、搜索结果排名等等已经使用了大量人工智能和深度学习的技术。在 Facebook，还有许多应用能够从机器学习模型中获利，然而对于很多没有深厚机器学习背景的工程师来说，想要利用这些机器学习基础设施，还是有很大的困难。2014 年末 Facebook 发布 FB Learner Flow，试图重新定义自己的机器学习平台，把人工智能与机器学习中最先进的算法以最便利的方式提供给 Facebook 的工程师。这一平台能方便地在不同的产品中重复使用多种算法，并可以延伸到成千上万种模拟的定制试验操中，轻松地对实验进行管理。这一平台提供了创新性的功能，比如从流水线定义和对 Python 编码进行自动化平行移用中自动生成用户界面（UI）试验。目前，超过 25% 的 Facebook 工程开发团队都在使用 FB Learner Flow。其未来的改进计划包括效率、速度和自动化等。

上文提到，Google 也在深度学习领域投入了大量力量。以 2011 年“Google 大脑”为契机，Google 逐年在多个内部重点项目上如广告系统、YouTube、Gmail 和机器翻译等部署深度学习，为 Google 进一步改善产品质量和用户体验提供了强有力的支持。此外，鉴于开源 Android 的成功经验，Google 在深度学习模型和工具方面也积极拥抱开源，目前正在大力推动 TensorFlow 开源深度学习平台。Google 目前有能力快速搭建和训练基于海量数据的模型，解决真实世界中的实际问题，并在不同的平台（比如移动端、GPU、云端）部署生产模型，相关技术都以云服务 API 的形式提供。



图表 26 Google 产品中深度学习的应用在迅速增长

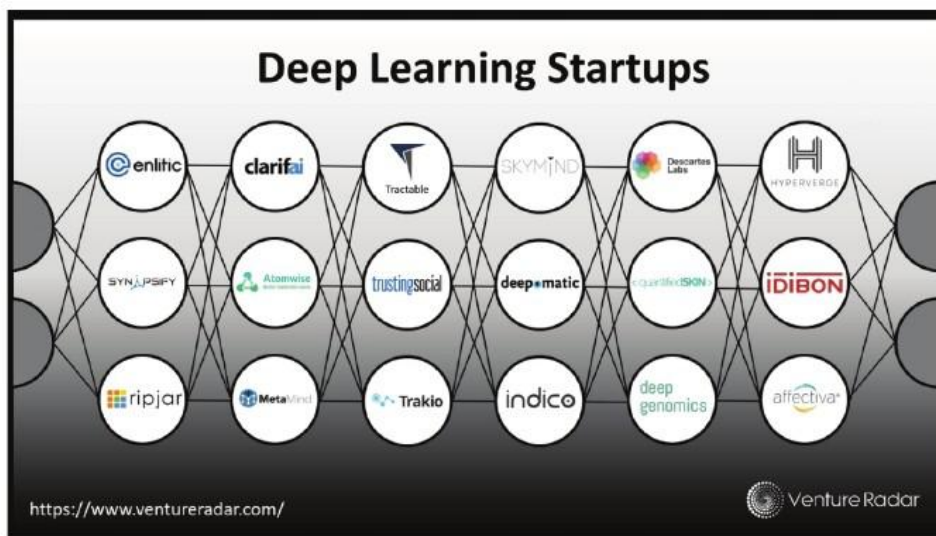
就国内来讲，百度是较早引入并大规模实践深度学习的典范。2013 年，百度成立深度学习实验室，是百度历史上首个正式成立的前瞻性研究机构，致力于“让计算机像人脑一样智能”的科学研究，并推出“百度大脑”计划，支撑百度相关产品线模型的改进，近年来在搜索技术、自然语言处理、机器翻译、机器学习、数据挖掘、推荐及个性化、语音、多媒体等领域取得了大量成果，并广泛应用于百度 PC 和移动端产品当中。

国内另一家使用深度学习技术的典型企业是科大讯飞。科大讯飞从 2010 年开展 DNN 语音识别研究，2011 年上线首个中文语音识别 DNN 系统，2013 年语种识别首创 BN-ivec 技术，2016 年将注意力模型神经网络应用于认知智能。



图表 27 科大讯飞的深度学习之路，来源新智元

2. 三大领域技术革新



图表 28 基于深度学习技术的创业公司

在产业界巨头和学术界精英的不断推动下，基于深度学习的人工智能终于迎来了“春天”，数不清的应用和创业公司出现，试图利用深度模型和垂直领域的积累来掘得第一桶金，主要聚集的领域有图像识别、语音识别和自然语言处理。

图像识别领域最早尝试深度学习，从早期 LeCun 的卷积神经网络，到 Hinton 的更深层次网络，再到百度以图搜图和人脸识别等技术，整个领域逐渐从人工特征工程和浅层学习模型，转向基于深度学习和大数据的方法。如 Face++ 以云服务方式提供基于深度学习的人脸识别技术，并且在金融、安防、零售领域分别开始了商业化探索。

语音识别领域，深度学习技术能够更好描述特征状态空间，尤其可以引入更高维的表示，能更好描述特征间相关性，取代了长久以来占据垄断性地位的混合高斯模型。如百度采用深度学习技术，进行声音建模的语音识别系统相比于传统的 GMM 语音识别系统而言，相对误识别率能降低 25%，早在 2012 年就上线了基于深度学习技术的语音搜索系统。

自然语言处理领域是近两年来深度学习逐步渗透的一个领域，比起图像和语音，自然语言单词带有较强的语义内涵，需要更细致更扩展的表示方法才能极大程度上保留信息。2003 年，加拿大蒙特利尔大学教授 Yoshua Bengio 等提出用 embedding 的方法将词映射到一个矢量表示空间，然后用非线性神经网络来表示 N-Gram 模型，此后语义变得可计算。通过大规模未标注文本和无监督学习，可以自动学习出字、词、句子的语义表示，一举摆脱知识库、词法、句法等传统自然语言障碍。Google 自然语言翻译的最新进展。

3. 三大开源框架促进技术落地

深度学习目前表现出来的趋势，不光是技术，还有商业模式的转变。过去几个月，所有巨头都

将自己的深度学习 IP 开源。核心目的是为了吸引用户、扩大市场，吸引人才、加速创新。开源会使技术发展更快，但主宰市场的仍将是巨头。

深度学习算法和平台本身的设计与实现是较难的问题，也不应该是应用开发者所需要过分关注的事情。近年来依托开源运动，全球互联网巨头从 2015 年起掀起新一轮最前沿技术平台开源共享的风潮。例如谷歌和 Facebook 分别将深度学习平台 TensorFlow 和 Torchnet 全面开源，在全球范围内大幅推进人工智能与深度学习技术的应用普及，扩大产业整体的市场规模，极大加速了深度学习在应用领域的迅速推广。这些软件所具有的共同设计特征是：容易表达、可扩展、多平台适用、可重复使用和快速见效。

TensorFlow 是 Google 的第二代机器学习系统，用来理解学习大规模数据和模型。这一系统非常灵活，可适用于图像、语音、文字理解等不同应用。Google 使用了上千块 GPU 并在性能上较同等 CPU 极大的提升。其优点主要有三个，一是高度扩展的设计，更快的实验速度加速研究进程；二是容易分享模型，开发代码应用到可重用的效果；三是通过同一个系统把研究工作直接用于生产环境。

Paddle 是百度的深度学习系统，已经在内部运行和推广多年，多次获得百度最高奖荣誉，并且已经做出了一些实际的产品，较为成熟。在性能和各项指标上都有优点，如代码简洁、设计干净，无过多抽象、速度较快，显存占用小、可多机多卡并行，支持异构、文档翔实等等，推出后 Paddle 也获得了较高的肯定，是一个不错的深度学习工具，在国内有较大的应用潜力。

另一个值得一提的是 Caffe，是一个清晰而且高效的深度学习框架，以易用性、扩展性和速度快迅速得到业界的认可，并且在文档撰写和代码迭代过程中非常注重与用户的互动，时至今日已拥有一个庞大的开发社区。

GitHub 根据上述框架衍生出数十个开源项目，构成良好的深度学习开发社区氛围，进一步推动了该技术的推广和落地。下表列举了星级超过 1500 的大型项目。

图表 29 GitHub 深度学习开源排名

项 目	星级	应用项目
Deep Dream	9042	一款图像识别工具。
Keras	7502	一款由 Python 实现的深度学习库，包括卷积神经网络、递归神经网络等。运行在 Theano 和 TensorFlow 之上。
Roc AlphaGo	7170	由学生主导的一个独立项目，重新实现了 DeepMind 在 2016 发表于 Nature 论文 "Mastering the game of Go with deep neural networks and tree search(用深度学习和树搜索学习围棋)" (Nature 529, 484-489, 28 Jan 2016)。
Neural Doodle	6275	运用深度学习将涂鸦变为优雅的艺术品，从照片生成无缝纹理，转变图片风格，进行基于实例的提升，等等。(语义风格传递的实现)
CNTK	5957	计算网络工具 (Computational Network Toolkit, CNTK)
TensorFlow Examples	5872	面向初学者的 TensorFlow 教程和代码示例
OpenFace	4855	基于深度学习网络的面部识别。



(续表)

项 目	星级	应用项目
Nupic	4364	智能计算的 Numenta 平台 (Numenta Platform for Intelligent Computing, Nupic): 一个脑启发式的计算智能和机器智能平台, 基于皮层学习算法的生物精确神经网络模型。
Leaf	4281	面向黑客的开源机器智能框架。
Char RNN	3820	基于 Torch 开发的多层递归神经网络的字符级别语言模型。
Neural Talk	3694	一个 Python+numpy 项目, 用多模式递归神经网络描述图像。
deeplearning4j	3673	基于 Hadoop 和 Spark 的 Java, Scala & Clojure 深度学习工具。
TFLearn	3368	深度学习库, 包括高层次的 TensorFlow 接口。
OpenAI Gym	3020	一种用于开发和比较强化学习算法的工具包。
Magenta	2914	用机器智能生成音乐和艺术
Colornet	2798	用神经网络模型给灰度图上色。
Synaptic	2666	基于 node.js 和浏览器的免架构神经网络库。
Neural Talk 2	2550	Torch 开发的图像简介生成代码, 运行在 GPU 上。
Image Analogies	2540	使用神经匹配和融合生成相似图形。
Deep Learning Flappy Bird	1721	使用深度强化学习破解 Flappy Bird 游戏。

(三) 深度学习未来展望

技术的发展, 尤其是到大规模落实阶段, 都难免会发现局限性, 也正是这些局限性, 不断促进业界思考和改进技术, 进而得到未来的发展方向。

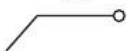
就深度学习而言, 首先, 缺乏理论支持。对于深度学习架构, 业界存在一系列的疑问: 卷积神经网络为什么是一个好的架构 (事实上其存在梯度散射等缺点), 深度学习的结构需要多少隐层, 在一个大规模的卷积网络中到底需要多少有效的参数 (很多权重相互之间似乎都存在冗余), 随机梯度下降方法优化权重得到一个局部最优值如何解决。虽然深度学习在很多实际的应用中取得了突出的效果, 但这些问题一直困扰着深度学习的研究人员。深度学习方法常常被视为黑盒, 大多数的结论确认都由经验而非理论来确定。不管是为了构建更好的深度学习系统, 还是为了提供更好的解释, 深度学习都还需要更完善的理论支撑。随着神经网络框架变得越来越复杂而精密, 在感知、语言翻译等方面的大部分最先进的神经网络框架正在发展并且不再仅仅关于简单前馈式 (feed forward) 框架或者卷积式框架 (convolutional)。特别地, 它们正在混合并匹配不同的神经网络技术如 LSTM、卷积、自定义目标函数、多皮层柱 (multiple cortical columns) 等等。此外, 神经网络超参数和网络结构的设计目前多依赖人工经验, 如何能够自动且高效地得到优化是值得关注的方向。

其次, 缺乏推理能力。深度学习技术缺乏表达因果关系的手段, 缺乏进行逻辑推理的方法。解决这个问题的一种典型方法是将深度学习与结构化预测相结合。目前几个带有结构化预测模块的增强的深度学习系统已经被提出来用于 OCR、身体姿态检测和语义分割等任务中。总的来说, 需要更多新的思路以应用于需要复杂推理的任务中。尽管深度学习和简单推理已经应用于语音和手写字

识别较长时间，仍需要在大的向量上使用新的范式来代替基于规则的字符表达式操作。最终，那些结合了复杂推理和表示学习的系统很可能为人工智能带来巨大的进步。深度学习和强化学习的交叉应用会继续，不仅仅聚焦在图像、声音和文本数据。如端对端学习控制无人车和机器人，使用深度学习和强化学习来完成原始传感器数据到实际动作执行器的直接映射。深度学习模型正从过去的只是分类一步步发展到试图理解如何在方程中加入规划控制相关的变量。

第三，缺乏短时记忆能力。人类的大脑有着惊人的记忆功能，不仅能够识别个体案例，更能分析输入信息之间的整体逻辑序列。这些信息序列富含大量的内容，信息彼此间有着复杂的时间关联性。例如在自然语言理解的许多任务（例如问答系统）中，需要一种方法来临时存储分隔的片段，正确解释视频中的事件，并能够回答有关它的问题，这需要网络具备记住的视频中发生的事件抽象表示的能力。然而包括递归神经网络在内的深度学习系统，都不能很好地存储多个时间序列上的记忆。这使得研究人员提出在神经网络中增加独立的记忆模块，如 LSTM，记忆网络（Memory Networks），神经图灵机（Neural Turing Machines），和 Stack 增强 RNN（stack-Augmented RNN）。虽然这些方法思路很直观，也取得了一定的成果，但在未来仍需要更多的尝试和新的思路。

最后，缺乏执行无监督学习的能力。无监督学习在人类和动物的学习中占据主导地位，我们通过观察能够发现世界的内在结构，无需被告知每一个客观事物的名称。在机器学习领域，神经网络的复兴恰恰是在无监督学习取得不断进步的 2005 年左右，虽然无监督学习可以帮助特定的深度网络进行“预训练”，但最终绝大部分能够应用于实践的深度学习方法都采用了纯粹的有监督学习。这并不能代表非监督学习在深度学习中没有作用，反而具有非常大的潜力，因为业界拥有的非标记数据比标记数据多很多，只是尚未找到很合适的非监督学习算法。非监督学习在未来存在巨大的研究空间，今后计算机视觉的进步将有赖于在无监督学习上取得突破，尤其是对于视频的理解。



虎嗅注：本文为新智元《中国人工智能产业发展报告》第一部分，第二部分已于同日在虎嗅网研究报告栏目发布。

2016 年 10 月 21 日