

推荐 (维持)

“人工智能+”时代呼啸而来

2016年09月07日

IT宿命系列之：人工智能专题深度之二

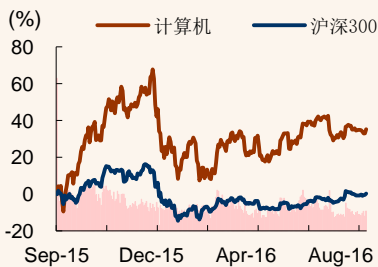
上证指数 3091

行业规模

		占比%
股票家数 (只)	144	5.0
总市值 (亿元)	21845	4.4
流通市值 (亿元)	13433	3.5

行业指数

%	1m	6m	12m
绝对表现	2.4	18.4	42.1
相对表现	-1.9	10.3	42.8



资料来源：贝格数据、招商证券

相关报告

1、《魔鬼的较量：量子计算 VS 量子通信—量子计算专题（二）》  
2016-08-21

2、《IT宿命系列之：Fintech 智能投顾专题（二）—大数据+AI，成就智能投顾的两大核心基石》  
2016-08-14

3、《工信部发布区块链技术和产业发展论坛未来两年规划点评—登堂入室，区块链“最官方”规划出台》  
2016-08-12

刘泽晶

liuzejing@cmschina.com.cn  
S1090516040001

周楷宁

010-57601791  
zhoukaining@cmschina.com.cn  
S1090516060001

研究助理

宋兴未  
songxingwei@cmschina.com.cn

我们看好人工智能开源平台百舸争流争夺 AI 生态圈会极大加速产业发展。我们看好人工智能芯片不断进化下的 AI 算法效率大幅提升。“人工智能+”时代呼啸而来，我们先期看好专用人工智能的普及，中长期看好通用人工智能的普及。

- **第三次浪潮：人工智能奇点临近。**人工智能在两次高潮和低谷之后迎来第三次浪潮，各国纷纷把人工智能上升到国家战略层面，美国白宫组织四场研讨会讨论人工智能，日本提出“超智能社会”，中国发布《“互联网+”人工智能三年行动实施方案》。仿佛一夜间，各国对人工智能热情大增，我们看到人工智能领域巨头们的卡位战争已经上升到开源平台、芯片以及应用。相关一级创投金额 5 年增长了 12 倍，我们预计二级市场的机会随时来临。
- **开源平台：超越 Android，AI 生态圈之争极大加速产业发展。**我们看到各大科技巨头和科研机构纷纷争夺开源人工智能技术，如 Google 的 TensorFlow、微软的 DMTK、Facebook 的 Torchnet、OpenAI、Baidu 的 Paddle 等，百舸争流千帆竞，以此来获得大量的用户需求和开发人员，建立开放共享、互利共赢的人工智能生态圈，我们认为人工智能开源平台有望成为引爆产业的导火线，从而比 Android 更深远地影响世界。
- **人工智能芯片：从通用到专用，类脑计算不断进化，效率大幅提升。**我们认为人工智能芯片效率的提升和突破将带领产业实现跳跃式发展。GPU 王者 Nvidia，与其合作的组织 2 年增长 30 多倍。FPGA 兼顾性能与灵活性，与 GPU 强强对决。我们预计人工智能专用芯片在竞争中有望实现超越摩尔定律的发展速度，硬件的春天即将来临。
- **应用：“人工智能+”时代呼啸而来，改变 IT 的命运。**我们先期看好专用人工智能的普及，中长期看好通用人工智能的普及。人工智能在未来几乎会给各行各业带来升级重构，包括安防、制造业、金融、交通、教育、法律、医疗等行业，AI 的触角伸向每个行业 and 每个人，从而改变 IT 未来的走向。
- **投资建议：建议从两个角度选择 A 股相关标的：1、数据能力；2、应用能力。**  
**重点推荐：**“人工智能+安防”：东方网力；“人工智能+金融”：同花顺；“人工智能+医疗”：思创医惠；“人工智能+底层操作系统”：中科创达；“人工智能+大数据”：神州泰岳。**关注：**科大讯飞、佳都科技、汉王科技。

□ **风险提示：**政策支持力度低于预期；核心技术发展遭遇瓶颈。

**重点公司主要财务指标**

	股价	15EPS	16EPS	17EPS	16PE	17PE	PB	评级
东方网力	24.63	0.77	0.48	0.65	51	38	10	强烈推荐-A
同花顺	67.85	1.78	2.14	2.82	32	24	19	强烈推荐-A
思创医惠	30.01	0.34	0.52	0.70	58	43	9	强烈推荐-A
中科创达	56.17	1.17	0.34	0.56	166	101	18	强烈推荐-A
神州泰岳	11.41	0.18	0.25	0.33	45	35	4	强烈推荐-A
科大讯飞	30.52	0.33	0.40	0.55	76	55	6	强烈推荐-A
佳都科技	9.13	0.34	0.19	0.27	48	34	5	审慎推荐-A
汉王科技	25.67	0.02	0.09	0.12	275	220	7	审慎推荐-A

资料来源：公司数据、招商证券

## 正文目录

一、人工智能第三次浪潮：从开源平台、芯片到应用的竞争白热化，创投金额 5 年间增长了 12 倍 .....	6
1、人工智能发展历程：在两次高潮和低谷之后迎来第三次浪潮 .....	7
2、人工智能全球热潮上升到国家战略层面 .....	11
3、人工智能巨头竞争白热化：从开源平台、芯片到应用 .....	12
4、人工智能领域的投资并购如火如荼，创投金额 5 年间增长了 12 倍 .....	13
二、人工智能开源平台：百舸争流千帆竞，争夺生态圈话语权，比 Android 更深远地影响世界 .....	18
1、Google TensorFlow：最受欢迎的人工智能开源平台 .....	18
2、微软 DMTK：机器学习算法+大数据的强大交叠 .....	22
3、Facebook Torchnet：轻量化机器学习开源工具集 .....	24
4、OpenAI：非盈利性开源人工智能组织 .....	25
5、Baidu Paddle：厚积薄发的深度学习开源平台 .....	28
6、其他开源工具：Caffe、Theano、MXNet、CNTK、DL4J .....	30
7、人工智能开源平台争夺生态圈话语权，争夺下一个时代的“Android” .....	31
三、人工智能芯片：从通用到专用，类脑计算的进化 .....	33
1、类脑计算的进化 .....	33
2、GPU 王者 Nvidia，在深度学习上使用 Nvidia 的组织 2 年增长了 30 多倍，为绝大部分的人工智能平台提供计算能力 .....	34
3、FPGA：性能与灵活性的折衷，人工智能领域 GPU 的强劲对手 .....	40
4、人工智能专用芯片：大幅提升人工智能算法运行效率，推动人工智能产业加速发展 .....	43
5、芯片是人工智能时代的战略制高点 .....	49
四、人工智能应用：“人工智能+”时代呼啸而来 .....	51
1、“人工智能+”时代呼啸而来 .....	51
2、安防：聚焦生物识别技术 .....	58
3、制造业：人工智能有望推动半自动化生产到全自动化生产的飞跃 .....	60
4、金融：关注投资决策辅助、信用风控、与智能支付 .....	60
5、无人驾驶：技术突破、立法加快，人工智能决定可靠性 .....	62
6、法律：案件预判与文案编纂，解放律师双手 .....	66

7、智能家居：天花板尚远，人工智能有望成为核心	67
8、医疗：为健康诊断和药品研发插上高飞的翅膀	68
五、投资建议	70

## 图表目录

图 1：人工智能从开源平台、芯片到应用的产业链	6
图 2：人工智能发展历程第三次浪潮	7
图 3：人工智能发展历程：从萌芽到爆发	8
图 4：深度学习的效果随着数据量的增大显提升	10
图 5：各国纷纷在战略层面布局人工智能	11
图 6：各个行业对人工智能投资的顶级公司	13
图 7：人工智能领域的并购：2012-2016	14
图 8：Google 成为收购人工智能创业公司最活跃的巨头	14
图 9：2011-2016 人工智能风险投资领域活跃度排名	15
图 10：人工智能公司融资历史	16
图 11：人工智能领域的创投金额 5 年间增长了 12 倍	16
图 12：美国各个行业对人工智能的投资占比	17
图 13：TensorFlow 核心在于阐释高层次机器学习运算过程	18
图 14：TensorFlow 中包含张量的运算数据流	19
图 15：TensorFlow 高度匹配深度学习运算需求	19
图 16：TensorFlow 支持分布式计算	20
图 17：TensorFlow：最受欢迎的人工智能开源平台	21
图 18：TensorFlow 版本更新历史	21
图 19：DMTK 核心架构	22
图 20：DMTK 的分布式机器学习框架——Multiverso	23
图 21：DMTK 三大核心研究方向	24
图 22：Torchnet 在 Github 平台上的用户贡献排名	25
图 23：OpenAI Gym 三大核心研究方向	26
图 24：OpenAI Gym 简单计算任务测试环境	27
图 25：OpenAI Gym MuJoCo 机器人测试环境	27
图 26：Spark 异构系统计算平台	28

图 27: Paddle 平台主要特点.....	29
图 28: 基于 Paddle 平台的相关基础工具和应用案例 .....	29
图 29: 人工智能芯片分类.....	33
图 30: 人工智能大脑的进化 .....	33
图 31: GPU VS CPU .....	34
图 32: 采用 GPU 加速与只采用 CPU 训练 CNN 的性能比较 .....	34
图 33: GPU 发展历程 .....	35
图 34: GPU 实现加速计算的原理 .....	36
图 35: nvidia 游戏领域收入 .....	38
图 36: nvidia 可视化领域收入.....	38
图 37: nvidia 数据中心领域收入 .....	39
图 38: nvidia 无人驾驶领域收入 .....	39
图 39: 在深度学习上使用英伟达的组织 2 年增长了 30 多倍 .....	39
图 40: NVIDIA 为绝大部分的人工智能平台提供计算能力 .....	40
图 41: FPGA 的结构 .....	40
图 42: FPGA 、GPP、ASIC 比较 .....	41
图 43: FPGA 深度学习研究里程碑 .....	42
图 44: 谷歌 TPU 芯片 .....	43
图 45: Nvidia Tesla P100 芯片 .....	44
图 46: IBM 集成 16 块 TrueNorth 芯片的电路板.....	44
图 47: IBM 随机相变神经元组成的人工神经网络 .....	45
图 48: 英特尔 Xeon Phi 家族成员.....	46
图 49: 中星微电子“星光智能一号”芯片和主板 .....	47
图 50: DianNao 结构 .....	48
图 51: PuDianNao 结构.....	49
图 52: 通用人工智能和专用人工智能的关系 .....	51
图 53: 人工智能应用分类.....	52
图 54: 人工智能项目按应用方向分类 .....	53
图 55: 人工智能项目按成熟度分类 .....	53
图 56: 美国人工智能应用案例占比 .....	54
图 57: 各类型人工智能公司数量.....	55

图 58: 美国人工智能公司的地理分布 .....	55
图 59: 人工智能细分应用领域投资热度分布 .....	56
图 61: 2007-2020 年全球生物识别技术市场规模 (亿美元) .....	59
图 62: 通过 V 字手进行恐怖分子身份识别 .....	59
图 63: 人工智能将催生全自动化智能生产工厂 .....	60
图 64: AlphaSense 智能搜索帮助提高投资决策效率 .....	61
图 65: Lending Club 的智能风控模式 .....	62
图 66: 无人驾驶技术发展历程的四个阶段 .....	63
图 67: 无人驾驶的关键技术 .....	64
图 68: 截至 2015 年, 美国已经有 16 个州启动无人驾驶立法 .....	65
图 69: NVIDIA 具有学习功能的自动驾驶系统 .....	65
图 70: Lex Machina 协助律师进行案件分析与预判 .....	66
图 71: 2012-2020 年我国智能家居市场规模 .....	67
图 72: 智能家居产品分类 .....	68
图 73: 智能健康管理公司 Welltok 近年融资额不断创新高 .....	68
图 74: 人工智能应用于药品研发 .....	69
图 75: 计算机行业历史 <a href="#">PEBand</a> .....	71
图 76: 计算机行业历史 <a href="#">PBBand</a> .....	71
表 1: 人工智能开源平台对比 .....	31
表 2: GPU 通用计算标准 .....	37
表 3: GPU 主流生产厂商 .....	38
表 4: 人工智能芯片对比 .....	49
表 5: 人工智能可能的重构的领域与方式 .....	57
表 6: 重点公司主要财务指标 .....	70



# 一、人工智能第三次浪潮：从开源平台、芯片到应用的竞争白热化，创投金额 5 年间增长了 12 倍

人工智能（Artificial Intelligence, AI）是对人的意识、思维的信息过程的模拟，它是一门研究、开发用于模拟、延伸和扩展人的智能的理论及应用系统的技术科学，该领域的研究包括机器人、语言识别、图像识别、自然语言处理和专家系统等，除了计算机科学以外，人工智能还涉及信息论、控制论、自动化、仿生学、生物学、心理学、数理逻辑、语言学、医学和哲学等多门学科。

人工智能本质上是為了研制出具有类人智能的机器，能够模拟、延伸和扩展人类智能的理论、方法、技术以及应用形式。

如果把人工智能与人对比，要实现以下功能：听说读写、行动思考学习等。

语音语义：语音识别、说话人识别、机器翻译；语音合成、人机对话等

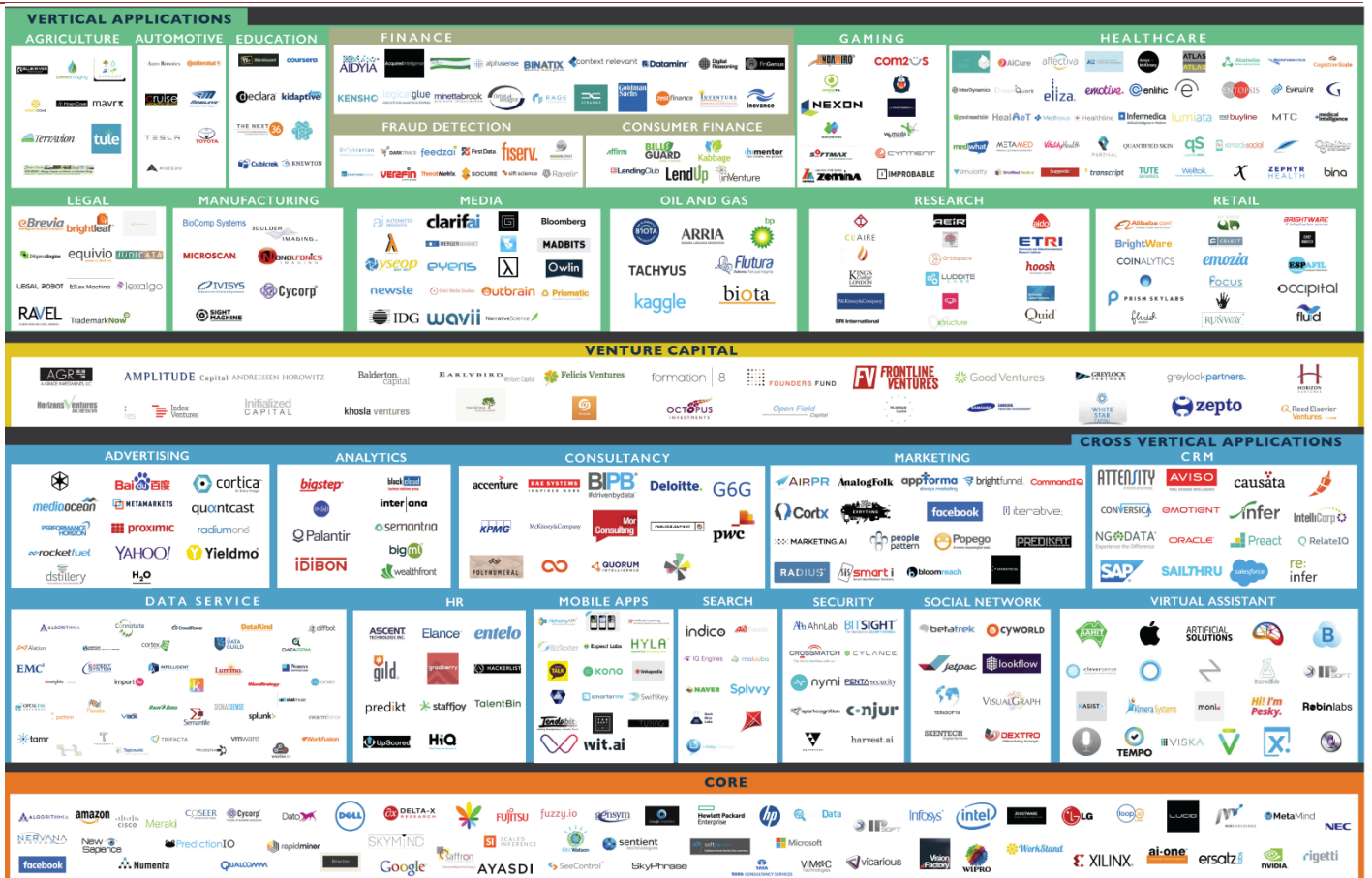
计算机视觉：图像识别、人脸识别、虹膜识别、文字识别、车牌识别等

行动：智能机器人、无人驾驶汽车、无人机等

思考：人机对弈、定理证明、自动推理和搜索方法、医疗诊断等

学习：机器学习、知识获取、知识处理、知识表示等

图 1：人工智能从开源平台、芯片到应用的产业链



资料来源：Pitchford，招商证券

敬请阅读末页的重要说明

## 1、人工智能发展历程：在两次高潮和低谷之后迎来第三次浪潮

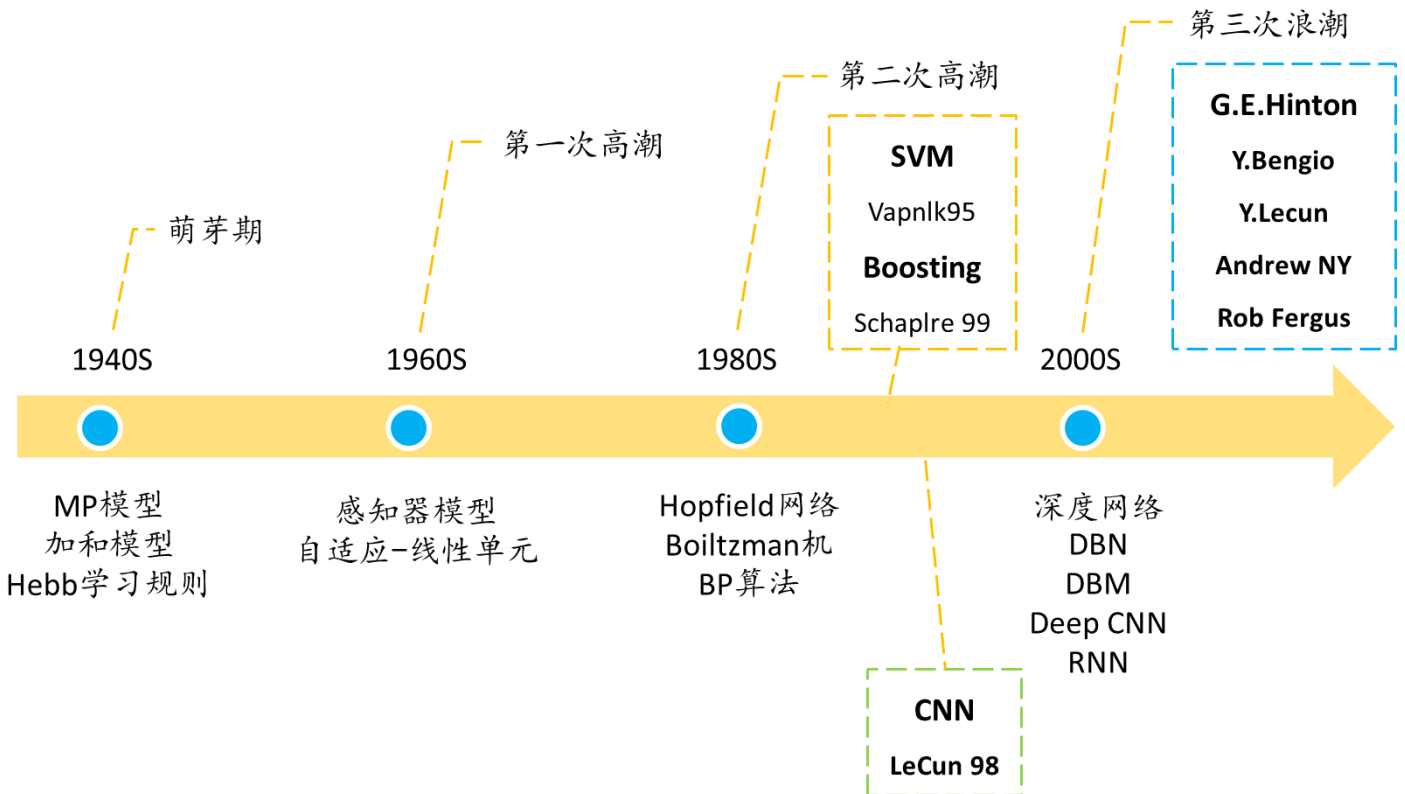
人工智能作为一门学科诞生至今已有 60 年的历史，期间经历了 2 次高潮和低谷。而从 2010 年到现在又迎来人工智能发展的第三次浪潮。

人工智能 60 年的发展，道路虽然起伏曲折，但进展也可谓硕果累累。无论是基础理论创新，关键技术突破，还是规模产业应用，都是精彩纷呈，使我们每一天无不享受着这门学科带来的便利。人工智能因其十分广阔的应用前景和对一个国家的重大战略意义，近年来日益得到政府部门、学术界的高度关注。

### (1) 1956 年达特茅斯会议：人工智能学诞生

1956 年夏，达特茅斯学院助教约翰·麦卡锡（John McCarthy）、哈佛大学马文·闵斯基（Marvin Minsky）、贝尔电话实验室克劳德·香农（Claude Shannon）、IBM 公司信息研究中心纳撒尼尔·罗切斯特（Nathaniel Rochester）、卡内基梅隆大学艾伦·纽厄尔（Allen Newell）和赫伯特·西蒙（Herbert Simon）等先驱在美国达特茅斯学院行了以此为其两个月的学术讨论会，从不同学科的角度探讨用机器模拟人类智能等问题，并首次提出了人工智能的概念，达特茅斯会议上 AI 的名称和任务得以确定，同时出现了最初的成就和最早的一批研究者，因此标志着人工智能学科的诞生。

图 2：人工智能发展历程第三次浪潮



资料来源：招商证券整理

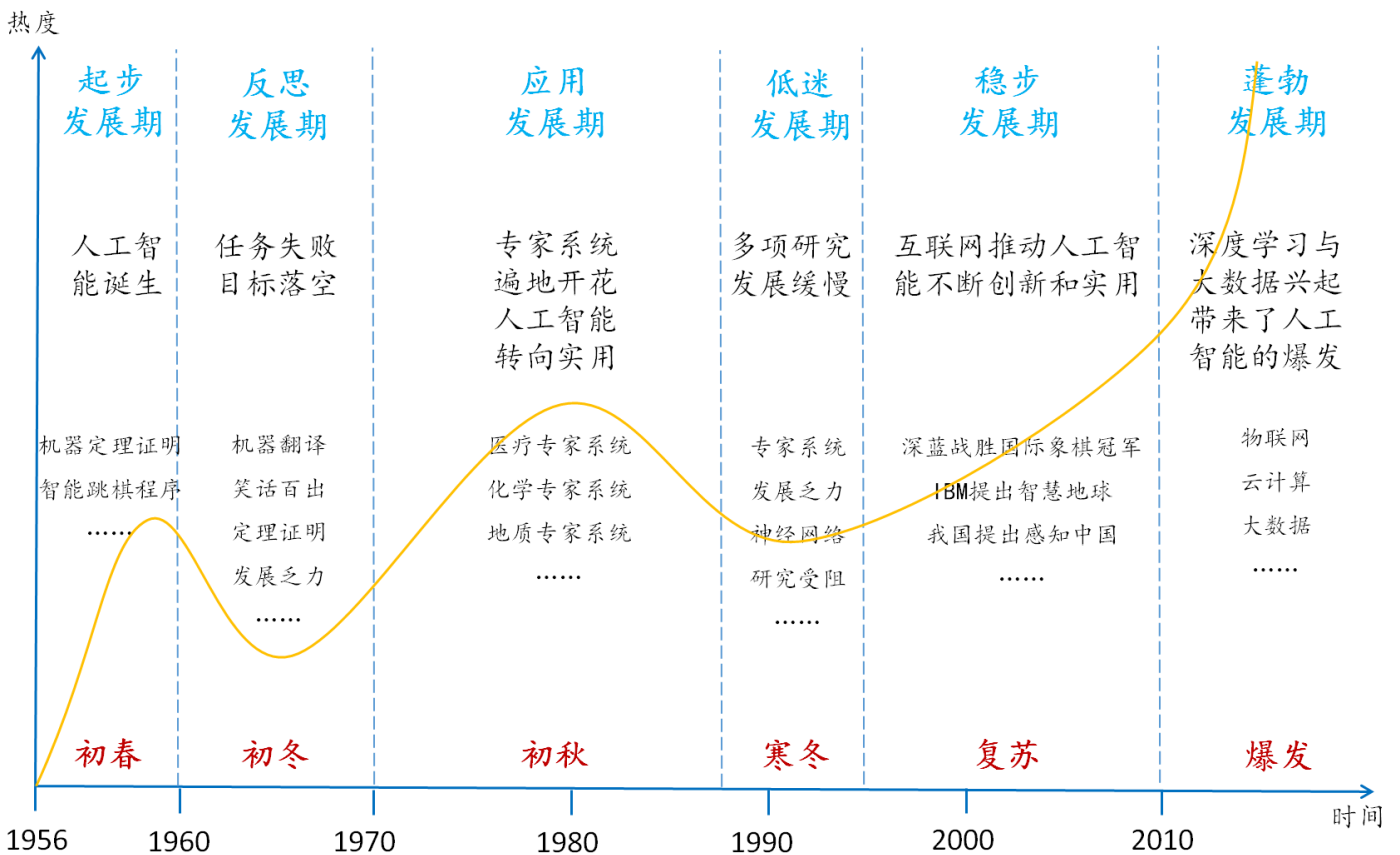
### (2) 1956 年至 1974 年：人工智能的第一次大发展

1956 年达特茅斯会议之后的十几年是人工智能的黄金年代。从 50 年代后期到 60 年代

涌现了大批成功的 AI 程序和新的研究方向，其中最有影响力的包括搜索式推理、自然语言、微世界等。在这段时间内，计算机被用来解决代数应用题、证明几何定理、学习和使用英语。初期研究取得了显著的成果，这些成果在得到广泛赞赏的同时也让研究者们对开发出完全智能的机器信心倍增。

1963 年 6 月，MIT 从新建立的 ARPA（即后来的 DARPA，国防高等研究计划局）获得了 220 万美元经费，用于资助 MAC 工程，其中包括 Minsky 和 McCarthy 五年前建立的 AI 研究组。此后 ARPA 每年提供三百万美元，直到七十年代为止，在麻省理工、卡内基梅隆大学、斯坦福大学、爱丁堡大学建立的人工智能项目都获得了来自 ARPA 等政府机构的大笔资金，在接下来的许多年间，这四个研究机构一直是 AI 学术界的（和经费）中心。不过，这些投入却并没有让当时的乐观预言得以实现。

图 3：人工智能发展历程：从萌芽到爆发



资料来源：中科院，招商证券整理

### (3) 1974 年至 1980 年：人工智能的第一次低谷

70 年代初，人工智能遭遇了瓶颈。由于计算机性能的瓶颈、计算复杂性的指数级增长、数据量缺失等问题，AI 研究者们遭遇了无法克服的基础性障碍。例如，在今天已经比较常见的机器视觉功能在当时找不到足够大的数据库来支撑程序学习，机器无法吸收足够的数据量，因此很难实现视觉方面的智能化。

由于缺乏进展，对人工智能提供资助的机构（如英国政府，DARPA 和 NRC）对无方向的 AI 研究逐渐停止了资助。到了 1974 年已经很难再找到对人工智能项目的资助，研究



经费被转移到那些目标明确的特定项目上。人工智能遭遇了 6 年左右的低谷。

1969 年 Minsky 和 Papert 出版了著作《感知器》，书中暗示感知器具有严重局限，而感知器是神经网络的一种形式，它最终将能够学习，做出决策和翻译语言。这本书的影响是破坏性的，联结主义的研究因此停滞了十年。

到 1970 年代后期，尽管遭遇了公众的误解，人工智能在逻辑编程、常识推理等一些领域还是有所进展。

#### (4) 1980 年至 1987 年：人工智能的第二次大发展

在 80 年代，一类名为“专家系统”的 AI 程序开始为全世界的公司所采纳，知识库系统和知识工程成为了 80 年代 AI 研究的主要方向。专家系统是一种程序，能够依据一组从专门知识中推演出的逻辑规则在某一特定领域回答或解决问题。专家系统仅限于一个很小的知识领域，从而避免了常识问题；其简单的设计又使它能够较为容易地编程实现或修改，实践证明了这类程序的实用性。

1980 年卡内基·梅隆大学为 DEC (Digital Equipment Corporation, 数字设备公司) 设计了一个名为 XCON 的专家系统，这套系统在 1986 年之前能为公司每年节省四千万美元。全世界的公司都开始研发和应用专家系统，到 1985 年它们已在 AI 上投入十亿美元以上，大部分用于公司内设的 AI 部门。为之提供支持的产业应运而生，其中包括 Symbolics, Lisp Machines 等硬件公司和 IntelliCorp, Aion 等软件公司。

1981 年，日本经济产业省拨款 8.5 亿美元支持第五代计算机项目，目标是制造出能够与人对话、翻译语言、解释图像，并且能像人一样推理的机器。随后，英国、美国也纷纷响应，开始向 AI 和信息技术领域的研究提供大量资金。

1986 年，人工智能领域著名的 BP Algorithm (Error Back Propagation Algorithm, 误差反向传播算法) 被 Rumelhart、McClelland 等大师提出，这使 1970 年以来一直遭人遗弃的联结主义重获新生。

#### (5) 1987 年至 1993 年：人工智能的第二次低谷

1987 年 AI 硬件市场需求突然下跌。Apple 和 IBM 生产的台式机性能不断提升，到 1987 年时其性能已经超过了 Symbolics 和其他厂家生产的昂贵的 Lisp 机。老产品失去了存在的理由：一夜之间这个价值五亿美元的产业土崩瓦解。XCON 等最初大获成功的专家系统维护费用居高不下，暴露出各种问题，专家系统的实用性仅仅局限于某些特定情景。

到 80 年代晚期，战略计算促进会大幅削减对 AI 的资助。1991 年，人们发现十年前日本人宏伟的“第五代工程”并没有实现。这些事实情况让人们从对“专家系统”的狂热追捧中逐步走向失望。人工智能研究再次遭遇寒冬。

尽管遇到各种批评，这一领域仍在不断前进。来自机器人学这一相关研究领域的 Rodney Brooks 和 Hans Moravec 提出了一种全新的人工智能方案，号召“自底向上”地创造智能，他们认为感知运动技能对于常识推理等高层次技能是至关重要的。

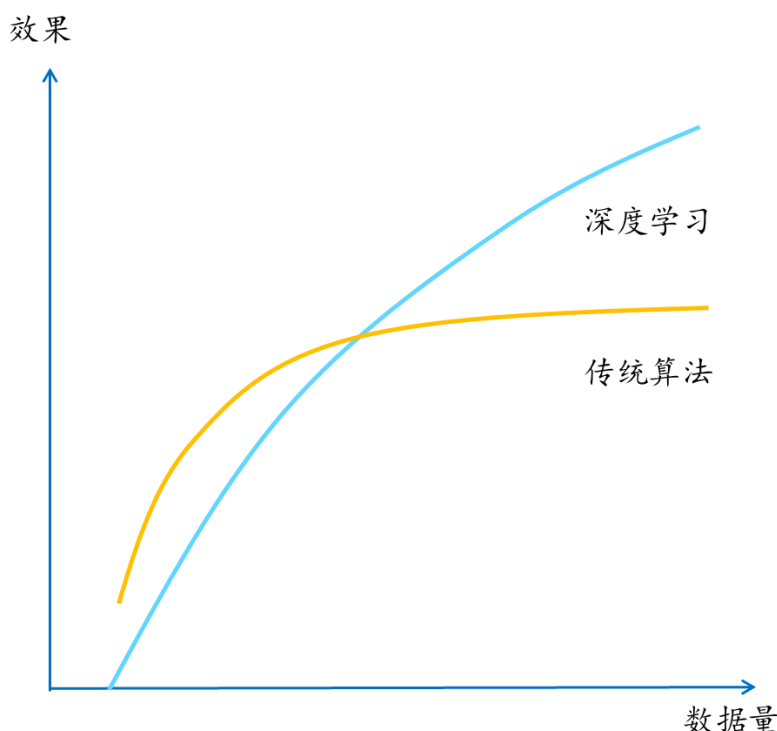
### (6) 1993 年至 2010 年：人工智能复苏期

1993 年到 2010 年这一阶段，人工智能处于稳步发展时期，互联网推动人工智能不断创新和实用。

人工智能已被成功地用在技术产业中，取得了一些里程碑式的成果：1997 年 5 月，IBM 研发的计算机“深蓝”(Deep Blue)战胜了国际象棋冠军卡斯帕罗夫；2005 年，Stanford 开发的一台机器人在一条沙漠小径上成功地自动行驶了 131 英里，赢得了 DARPA 挑战大赛头奖；2009 年，洛桑联邦理工学院发起的蓝脑计划声称已经成功地模拟了部分鼠脑。

2008 年 11 月 IBM 提出“智慧地球”概念，数字化、网络化和智能化，被公认为是未来社会发展的大趋势，而与“智慧地球”密切相关的物联网、云计算等，更成为科技发达国家制定本国发展战略的重点。自 2009 年以来，美国、欧盟、日本和韩国等纷纷推出本国的物联网、云计算相关发展战略。

图 4：深度学习的效果随着数据量的增大显提升



资料来源：公开资料、招商证券

### (7) 2010 年到现在：人工智能进入爆发式增长期

大数据、云计算支撑人工智能产业爆发，人工智能将成为下一轮技术变革的核心。

人工智能新一轮的爆发包括大数据、云计算和算法三个核心要素。

第一，数据的急剧增长。得益于互联网、社交媒体、移动设备和廉价的传感器，这个世界产生的数据量指数型增长。大数据是人工智能发展的助推剂，这是因为有些人工智能技术使用统计模型来进行数据的概率推算，比如图像、文本或者语音，有了大量的数据才能为模型的训练提供原材料。

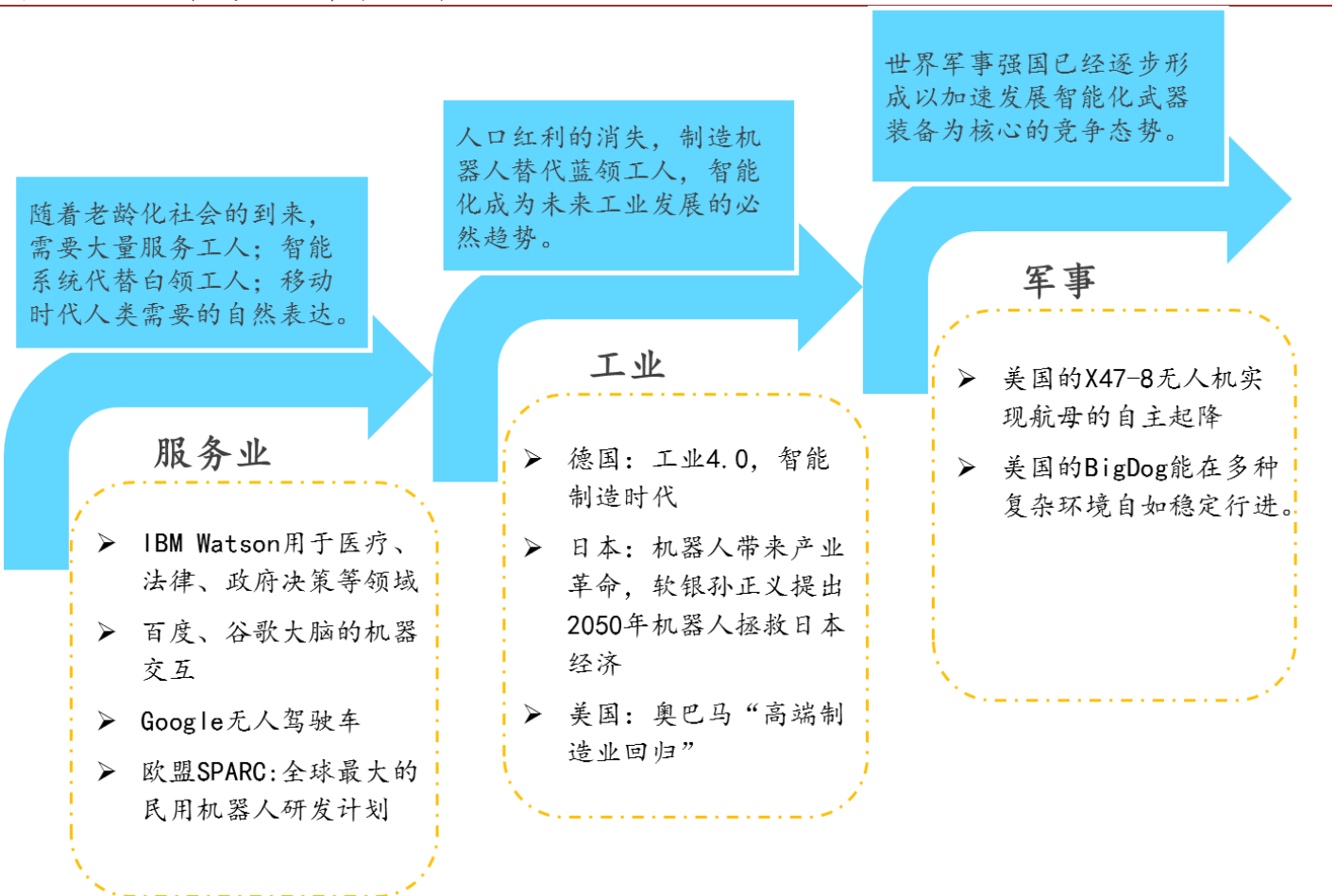
第二，计算能力的进步。云计算和大规模 GPU 并行计算的发展为深度学习的应用提供了计算基础。当人们把原本用于游戏中处理高质量画面的 GPU 拿来运行深度学习算法后，计算机可以几百倍地加快模型的训练速度。目前，AlphaGo 的计算能力是 IBM 深蓝的近 3 万倍。未来随着速度更快的芯片乃至量子计算的发展，运算速度会进一步提高，成本也会大幅降低。

第三，算法的发展，特别是深度学习的应用。算法是解决一个设计程序或完成任务的路径方法。与大部分传统算法相比，深度学习的效果随着数据量的增大有显著的提升，因而更大量的数据可以提高算法的表现。深度学习增加了神经网络的层数，因此网络才有了更强大的特征挖掘能力。机器学习算法目前被开源使用，这种情形将促成更大进步，因为在开源环境下开发人员可以补足和增强彼此的工作。

## 2、人工智能全球热潮上升到国家战略层面

人工智能已经成为国家服务业、工业和军事的核心竞争力，因此世界各国制定了国家级发展战略：

图 5：各国纷纷在战略层面布局人工智能



资料来源：招商证券整理

2016年5月美国白宫计划组织四场研讨会讨论人工智能。白宫还成立了人工智能委员会，用于协调全美各界在人工智能领域的行动。美国交通部宣布历时10年投资40亿美元的提案，旨在实现无人驾驶汽车上路。

日本从2016年开始执行的“第五期科学技术基本计划”中，日本政府列入总额约26万亿日元的研发经费，重点研发物联网及人工智能系统，提出要实现领先于世界的“超智能社会”（即 Society5.0）。

2016年5月25日我国四部委发布《“互联网+”人工智能三年行动实施方案》，到2018年形成千亿级的人工智能市场应用规模。2016年8月8日发布的《“十三五”国家科技创新规划》多次讲到人工智能。

### 3、人工智能巨头竞争白热化：从开源平台、芯片到应用

#### （1）开源人工智能基础平台

同时各大科技巨头纷纷开源人工智能技术，以此来获得大量的用户需求和开发人员，建立开放共享、互利共赢的人工智能生态圈，典型的例子有：

谷歌开源了人工智能基础平台 TensorFlow；

Facebook 开源了人工智能基础平台 Torchnet；

微软开源了人工智能基础平台 DMTK；

IBM 开源了人工智能基础平台 SystemML；

雅虎开源 CaffeOnSpark；

百度开源 AI 人工智能 Warp-CTC 源代码；

特斯拉创始人 Elon Musk 等共同创立非营利开源人工智能公司 OpenAI 等。

#### （2）布局人工智能芯片

人工智能芯片定义了 AI 产业链和生态圈的基础计算架构，核心芯片是人工智能时代的战略制高点。Google、IBM、英伟达、Intel、中科院都在积极布局人工智能芯片。

2016年8月17日，英特尔表示将开发人工智能技术的专用芯片；

IBM 正在设计基于大脑结构的芯片 TrueNorth；

今年5月，谷歌自主研发新型芯片支持人工智能；

英伟达推出人工智能超级芯片特斯拉 P100 GPU；

中科院研发出全球首个能够“深度学习”的“神经网络”处理器芯片“寒武纪”；

Facebook、微软以及 Twitter 都在通过设计新的芯片加强人工智能研发。

### (3) 研发人工智能核心技术

Google、IBM、Microsoft、Facebook、Amazon、百度等企业巨头充分认识到人工智能技术引领新一代信息产业发展的战略意义，纷纷投入重金收购企业、招募人才和研发核心技术，力图掌握人工智能时代的主动权：

Google 正在研发自动驾驶汽车、Deepmind 人机博弈；

IBM 投资 10 亿美元组建 Watson 人工智能部门，在医疗金融等领域推广应用；

微软推出聊天机器人小冰和人工智能助理小娜；

Facebook 专注于人工智能助理、图像视频识别等技术；

百度开发无人驾驶、语音识别、百度度秘等人工智能多领域的应用产品。

图 6：各个行业对人工智能投资的顶级公司

SOFTWARE AND IT	FINANCIAL SERVICES	MANUFACTURING	AUTOMOTIVE	INDUSTRIAL AUTOMATION
Google Microsoft Amazon Facebook Mitre IBM Deloitte	Bloomberg American Express TD Ameritrade Deutsche Bank Goldman Sachs BNP Paribas	Abb National Instruments Toshiba GE	Tesla Ford GM Toyota	Bosch Siemens Rockwell Automation Honeywell
HEALTHCARE	TELECOMMUNICATIONS	RETAIL	SEMICONDUCTORS	INTERNET
GE Healthcare Nuance Mayo Clinic Siemens Healthcare McKesson Massachusetts General Hospital	Huawei Nokia BT Group Orange S.A. Nippon Verizon	Walmart Gamestop Target Rakuten Best Buy Barnes & Noble	Intel Texas Instruments Microchip Technology Altera Imagination Technologies ARM Mellanox Qualcomm	Google Facebook LinkedIn Amazon

资料来源：SPIDERBOOK，招商证券

## 4、人工智能领域的投资并购如火如荼，创投金额 5 年间增长了 12 倍

### (1) Google 成并购人工智能创业公司最活跃的收购巨头

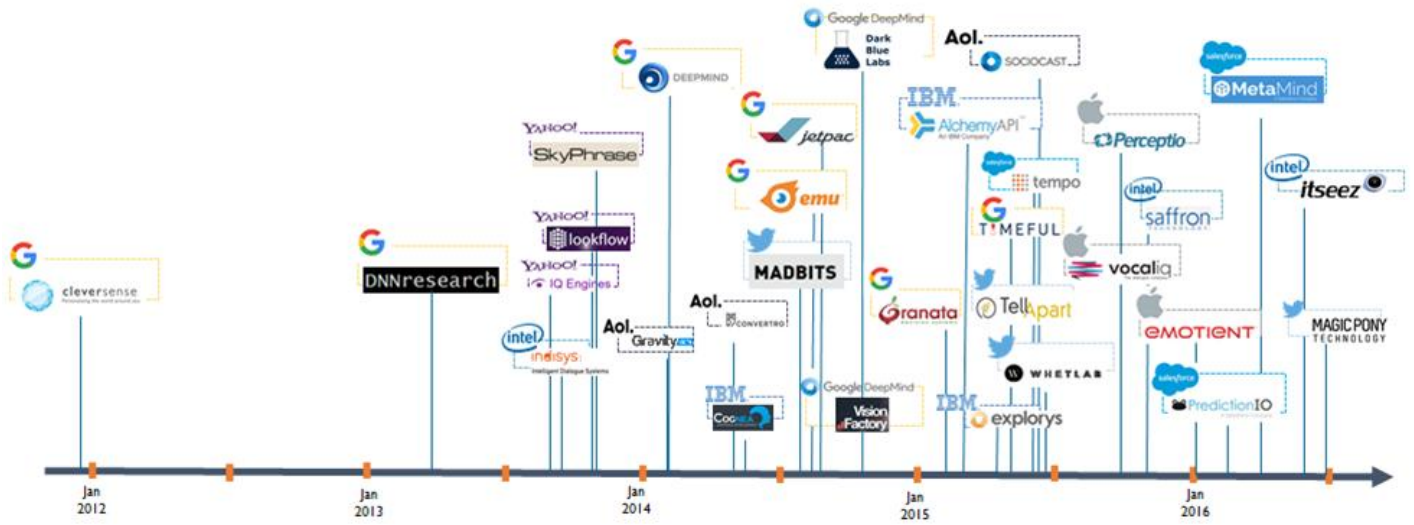
当前,全球范围内人工智能产业化应用蓬勃发展。谷歌、IBM、Intel、苹果、Yahoo 等国际巨头纷纷抢滩布局人工智能产业链,力图掌握人工智能时代的主动权。从 2013 年起,巨头对人工智能公司的并购呈现密集化趋势。

Google 成为人工智能公司收购热潮中的领军者,自 2011 年起,Google 先后收购了近 10 家 AI 领域的创业公司, Twitter 紧追其后,将 Magic Pony、 Madbits、 Whetlab 等公司收入囊中,而苹果公司和 Salesforce 自去年开始以收购方式布局人工智能后各并购 3 家 AI 创业公司。此外, Intel、IBM、Yahoo 等也都在积极收购 AI 创业公司。



图 7: 人工智能领域的并购: 2012-2016

Race For AI: Most Active Acquirers In Artificial Intelligence



资料来源: CB Insights, 招商证券

图 8: Google 成为收购人工智能创业公司最活跃的巨头

Top Acquirer Of AI Startups 2011-2016 year-to-date	
Rank	Acquirer
1	Google
2	Twitter
3	Apple Inc.
3	Intel
3	Salesforce
3	AOL
3	IBM
3	Yahoo

资料来源: CB Insights, 招商证券

(2) Intel 成为人工智能风投领域最活跃的机构投资者

图 9: 2011-2016 人工智能风险投资领域活跃度排名

Artificial Intelligence: Most Active Corporate Investors  
2011-2016YTD (as of 6/15/2016)

Investor	Rank	Select Investments
Intel Capital	1	DataRobot, prelert, lumiata, MAANA, inçoming saffron, PERFANT, EMOTIGT, Reflektion, Parallel Machines, MindMeld, smartzip, api.ai, indisyS, COLDLIGHT
Google Ventures	2	BUILDING ROBOTICS, clarifai, KENSHO, FRAMED, ZEPHYR HEALTH, Unbabel, tamr, MindMeld, Orbital Insight, Predilytics
GE Ventures	3	SIGHT MACHINE, ARTERYS, AYASDI, BITSTEW, MedAware, PingThings, stem, Predixion, MAANA
Samsung Ventures	4	vicarious, sentiance, Maluuba, IDIBON, jibo, MindMeld, ai
Bloomberg Beta	4	deep genomics, context relevant, AVISO, howdy, DOMINO, Orbital Insight, DigitalGenius, DIFFBOT
In-Q-Tel	6	MindMeld, CYLANCE, celect, INTERSET, Digital Reasoning, DOMINO
Tencent	7	DIFFBOT, CLOUDMEDX, SCALED INFERENCE, iCarbonX, skymind
Nokia Growth Partners	8	rocketfuel, WorkFusion, rapidminer, indix
Microsoft Ventures	8	BUILDING ROBOTICS, NEURA, insidesales, CrowdFlower
Qualcomm Ventures	8	clarifai, Predilytics, Welltok, tempo
Salesforce Ventures	8	DigitalGenius, insidesales, sense
AXA Strategic Ventures	8	NEURA, BI-BEATS, medlanes, pricemethod
New York Life Insurance Company	8	context relevant, DataRobot, Skycure, capricity

资料来源: CB Insights, 招商证券

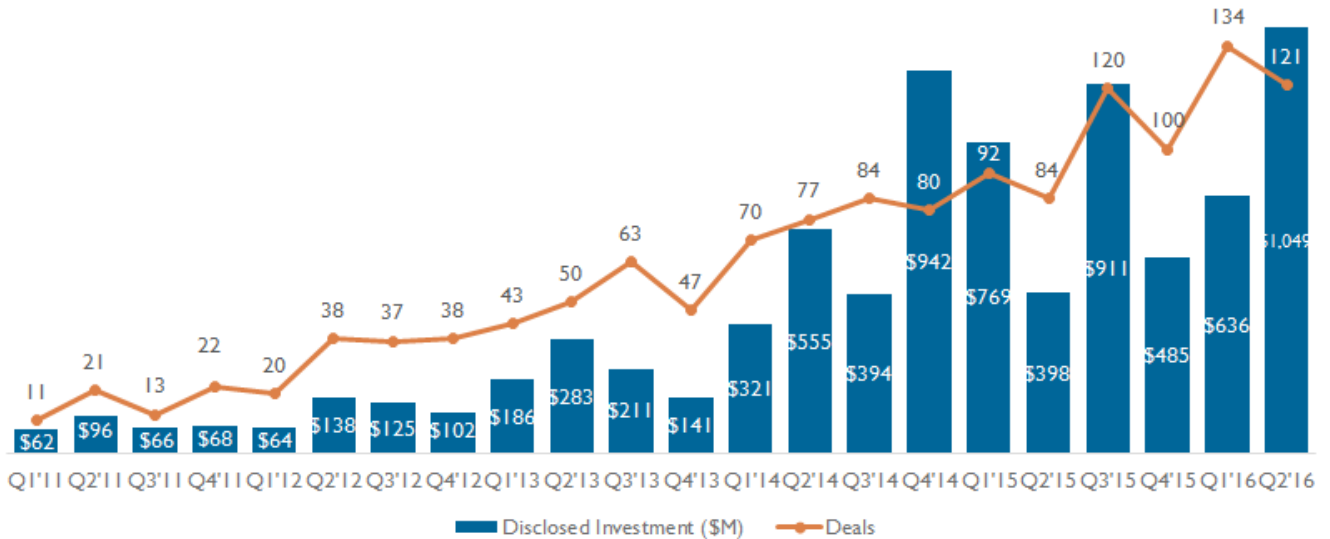
从 2011 年至今, 在风投领域, Intel 是最活跃的机构投资者。在 Intel Capital 的投资组合中有医疗健康领域的 Lumiata, 机器学习平台 DataRobot 以及图片视频公司 Perfant 等等。谷歌也在 AI 领域投资了 10 家公司, 包括 Building Robotics、Clarifai、Kensho 等等。谷歌在人工智能领域无疑是领先者, 通过进一步分析我们可以发现在并购市场上谷歌主要收购的对象是人工智能的底层技术, 而在风投市场主要投资的是人工智能在特定场景下的应用。同时, 在谷歌内部, 深度学习也被广泛应用。谷歌的许多产品如 Android、Gmail、地图、翻译、YouTube 等都有深度学习在背后支撑。

近日, Intel 宣布收购计算机视觉处理器公司 Movidius, 加速进军无人机、机器人以及 VR 领域。Movidius 的杀手级产品就是花费 9 年时间完全从零开始自主研发的全新架构低功耗视觉处理器。

(3) 人工智能领域的创投金额 5 年间增长了 12 倍

图 10: 人工智能公司融资历史

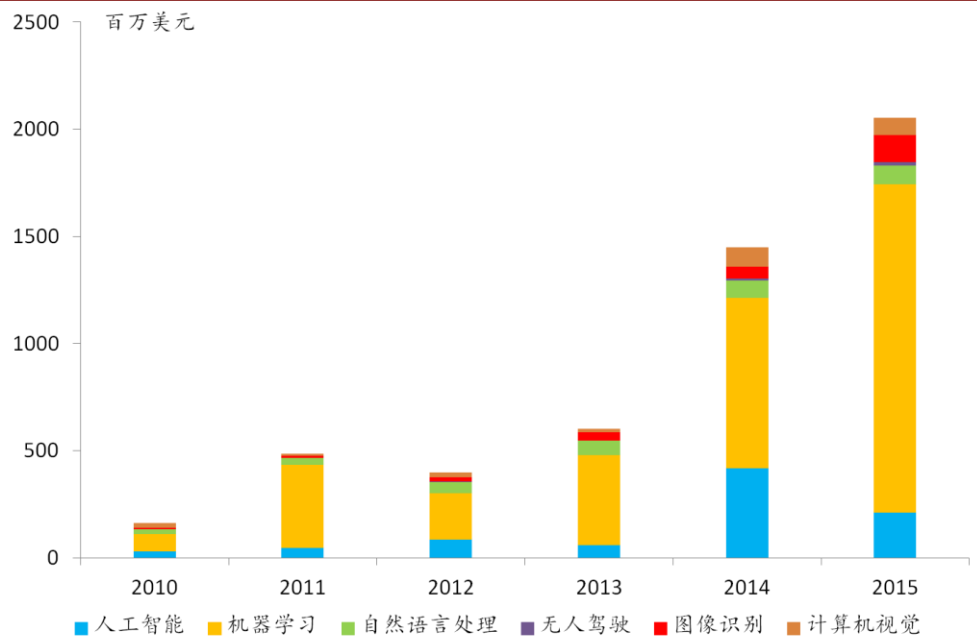
AI Funding History (Ex-Incubator/Accelerator)  
Q1'11-Q2'16



资料来源: CB Insights, 招商证券

从已经披露的人工智能领域投资来看, 2011 年到 2016 年的五年间, 无论是总体融资额还是交易次数都在逐年上升。随着人工智能在各个专业领域的技术突破, 投资者对于人工智能产业的兴趣和信心逐渐增加, 都在试图抓住这一热潮, 掌握更多的主动权。

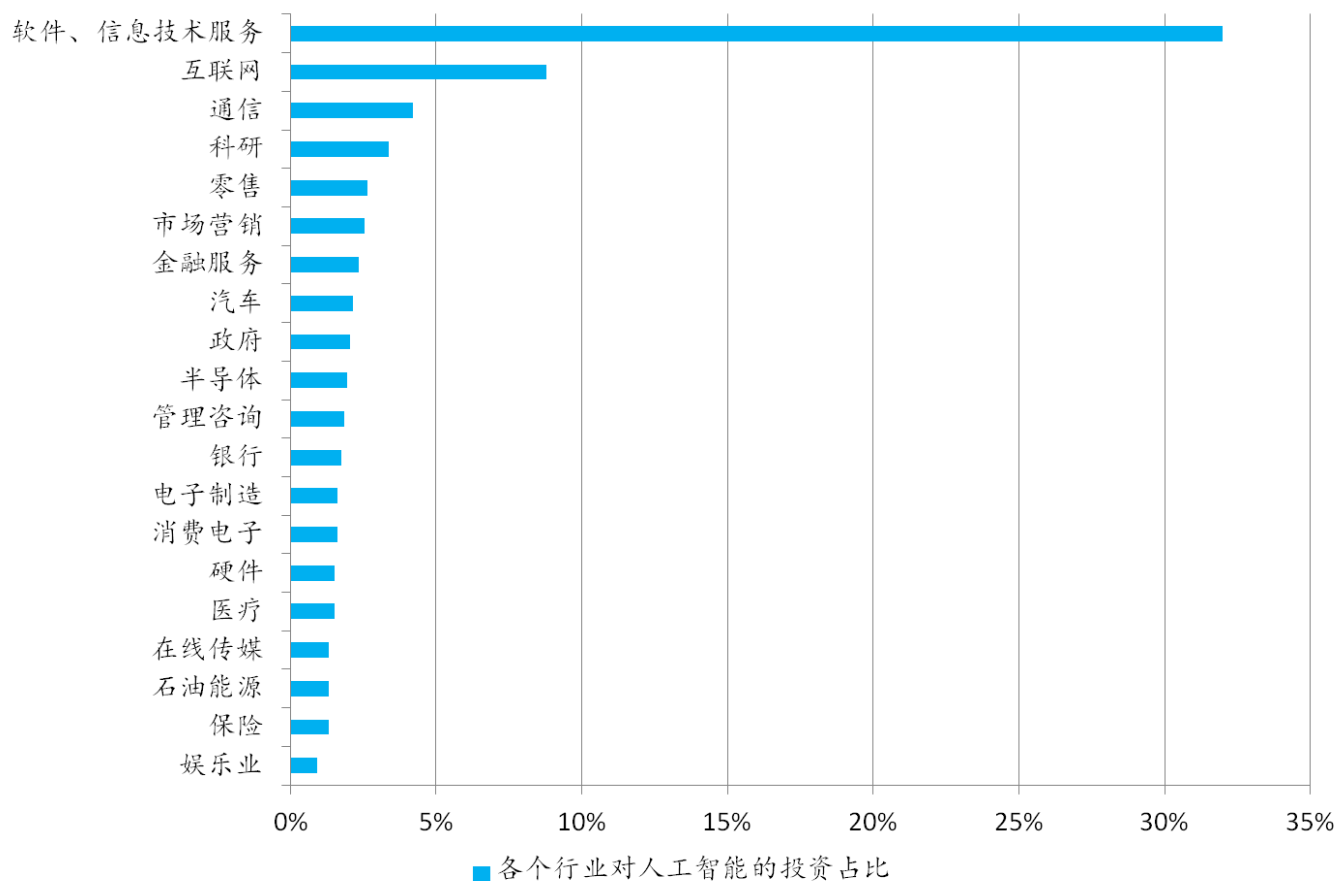
图 11: 人工智能领域的创投金额 5 年间增长了 12 倍



资料来源: Facebook, 招商证券

对使用人工智能算法的创业公司的交易与投资在 2016 年创了记录。这一趋势正在变革包括医疗、商业智能、广告在内的数个产业。特别是创业公司使用机器学习和自然语言处理解决金融方面的问题在过去几年吸引了投资者的兴趣。

图 12: 美国各个行业对人工智能的投资占比



资料来源: SPIDERBOOK, 招商证券

## 二、人工智能开源平台：百舸争流千帆竞，争夺生态圈话语权，比 Android 更深远地影响世界

同时各大科技巨头和科研机构纷纷开源人工智能技术，以此来获得大量的用户需求和开发人员，建立开放共享、互利共赢的人工智能生态圈，可以说人工智能开源平台，会比 Android 更深远地影响世界。

### 1、Google TensorFlow：最受欢迎的人工智能开源平台

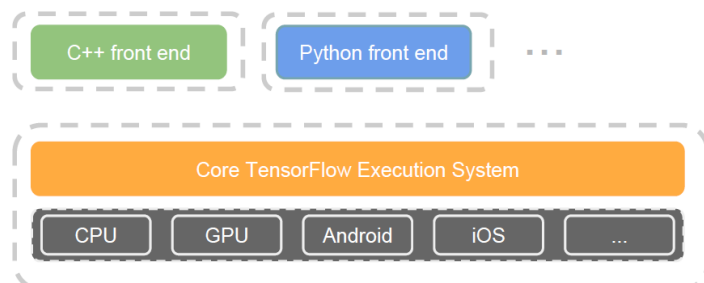
作为 Google 第二代深度学习系统，TensorFlow 被寄予厚望。TensorFlow 是 Google 在 2015 年 11 月宣布开源的第二代深度学习系统，主要用于编译执行机器学习算法的代码，是 2011 年所推出的第一代人工智能系统 DistBelief 的升级版。DistBelief 是一款非常成功的深度学习系统，Google 在此系统上建立起了庞大的神经网络模型并取得不少令人瞩目的技术成就，比如 DistBelief 曾从 Youtube 上没有被标记的海量视频中识别出包含猫的视频，将 Google 应用中的语音识别水平提高 25% 等，但 DistBelief 仅以神经网络为目的，在实际配置及应用中存在明显局限性。

TensorFlow 的设计架构克服了 DistBelief 通用性差、不够灵活等缺点，在运行速度、可扩展性和与产品可衔接性方面均有大幅改善，而且整个系统完全开源。作为 Google 所推出的第二代开源深度学习系统，TensorFlow 所具备的优异性能有望进一步降低深度学习的使用门槛，成为深度学习领域标杆式开发平台，加速推进人工智能行业新产品的研发进度。

图 13: TensorFlow 核心在于阐释高层次机器学习运算过程

### TensorFlow: Expressing High-Level ML Computations

- Core in C++
  - Very low overhead
- Different front ends for specifying/driving the computation
  - Python and C++ today, easy to add more



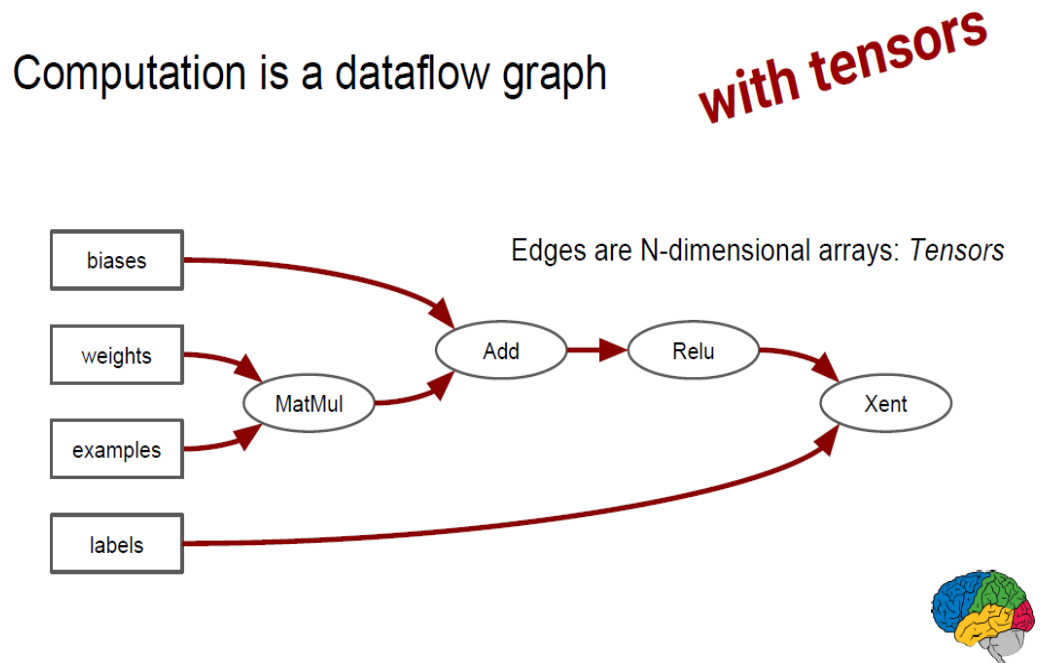
资料来源：Google，招商证券

**TensorFlow 专注解决深度学习问题。**TensorFlow 的命名生动体现了其运算模式的架构理念：Tensor 一词本意为张量，代表一个 N 维数组，比如该数组为 1 维数组时就表示向量，2 维时表示矩阵，更高维的数据流比如图像则可以用三维张量（行、列、色彩）



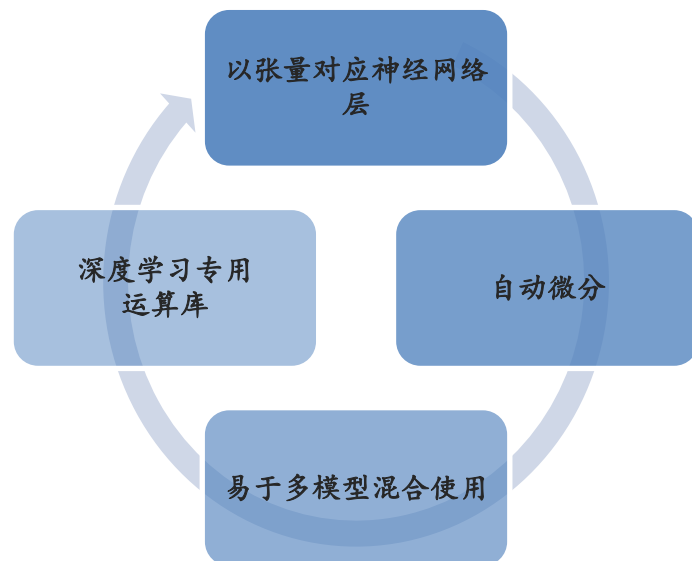
来代表；Flow（流动）则意味着基于数据流图的运算流程。TensorFlow 则意味着张量在数据流图上流动过程，也即运算处理过程。TensorFlow 的基本架构从多个角度考虑到了对深度学习的适应性，具体而言包括以下四个方面：1) 用一个或多个张量(Tensor)神经网络中的每一层(Layer)；2) 建立专用于深度学习的运算库，包括：2D、3D 卷积、Pooling、Softma 等高阶运算；标准损失函数，如 Corss Entropy、L1、L2 等；多种优化方法，如 Gradient Descent、AdaGrad、L-BFGS 等；3) 自动计算相关微分导数；4) 易于混合使用多种不同模型进行训练，包括 LSTMs、卷积模型、注意力模型，增强学习、类图灵机模型等。

图 14: TensorFlow 中包含张量的运算数据流



资料来源：Google，招商证券

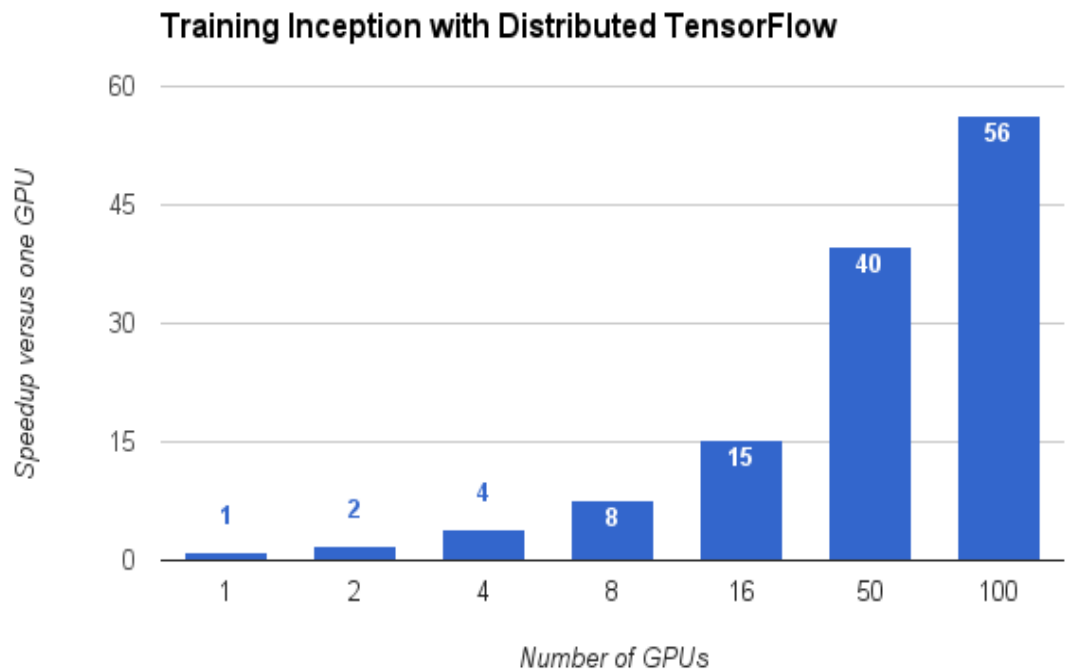
图 15: TensorFlow 高度匹配深度学习运算需求



资料来源：Google，招商证券

背靠 Google 强大技术支撑，TensorFlow 整体性能显著提升。TensorFlow 可以被架设在各类常见计算平台上，无论是智能手机、计算机、大型服务器，还是由 GPU 组成的分布式系统，都能实现通用；在编程语言方面，TensorFlow 支持包括 Python、Cuda、C++ 等在内多种程序语言，系统自带的深度学习模型库功能丰富且方便易用。在 TensorFlow 在建立和训练神经网络速度等主要指标方面，TensorFlow 要比 DistBelief 快 5 倍左右。此外，TensorFlow 配备的芯片是 Google 自主研发的人工智能专用处理芯片 TPU，同时能够调用 Google 庞大的计算资源和海量用户基数，在处理大规模分布式训练方面具有无可比拟的明显优势。

图 16: TensorFlow 支持分布式计算

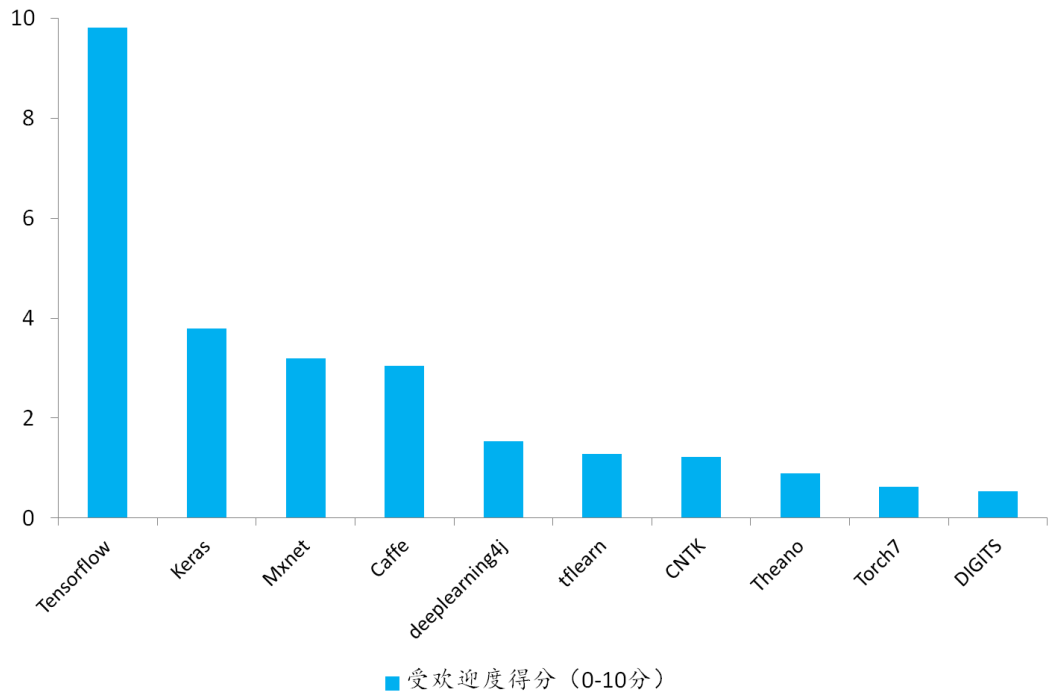


资料来源：Google，招商证券

2016 年 5 月，经过长达半年的测试与评估过程，DeepMind 宣布放弃此前一直使用的开源机器学习平台 Torch7，转而使用 TensorFlow 系统，这也意味着 AlphaGo 未来的测试训练将在 TensorFlow 上展开。

由于具有优异的性能以及 Google 方面庞大的维护团队，Tensorflow 一经面世就广受欢迎，HackerNews 曾开展一个深度学习工具受欢迎程度的投票，时间范围为 7 月 15 日到 8 月 15 日，结果显示 TensorFlow 受欢迎程度高居第一，且得分遥遥领先于第二名。

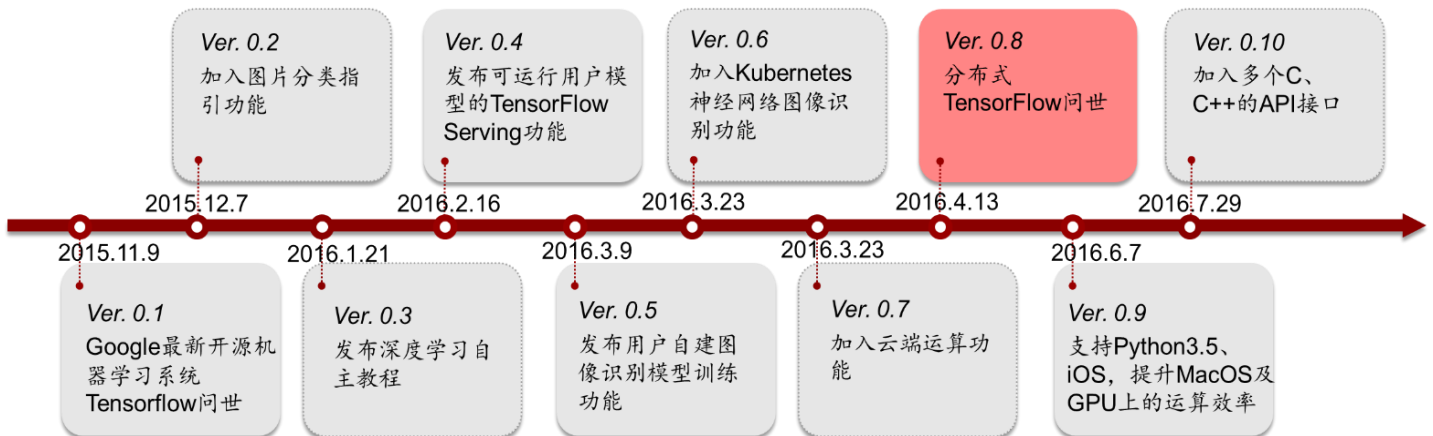
图 17: TensorFlow: 最受欢迎的人工智能开源平台



资料来源: HackerNews, 招商证券

**TensorFlow 更新及时, 开源程度不断加深。**TensorFlow 自从 2015 年 11 月发布 0.1 创始版本以来, 短短 9 个月内已经完成了 9 次更新, TensorFlow 系统正在加速完善成熟。最具有里程碑意义的一次更新时今年 4 月推出的 TensorFlow0.8 版, 实现了分布式联机功能。虽然 Google 内部所用 TensorFlow 可在数以千计的服务器上同时运行, 但对外发布的低于 0.8 版本的系统却只能在单台机器上运行, 这使得 TensorFlow 在大规模并行应用方面受到了极大限制, 这也是此前 TensorFlow 受业界诟病的主要原因之一。而随着 0.8 版本 TensorFlow 的发布, 这一缺陷已弥补, 标准着 TensorFlow 开源透明程度的巨大提高。

图 18: TensorFlow 版本更新历史



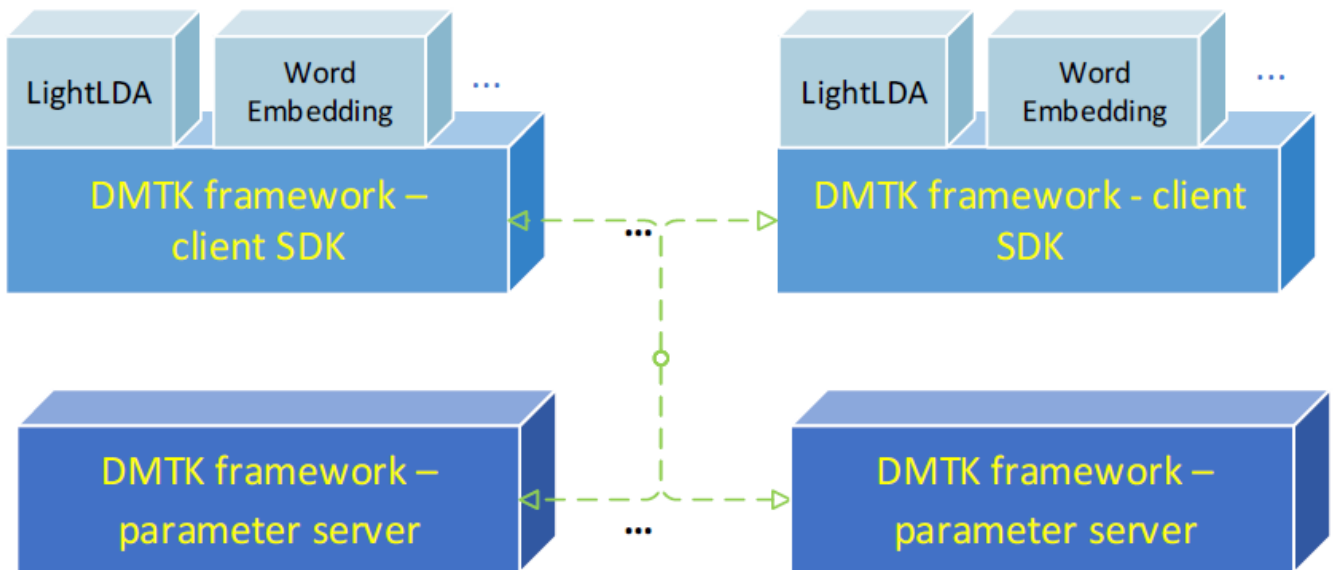
资料来源: 招商证券整理

## 2、微软 DMTK：机器学习算法+大数据的强大交叠

**DMTK，微软机器学习开源工具的扛鼎力作。**在 Google 宣布开源 TensorFlow 系统的同一天（2015.11.9），微软也宣布将分布式机器学习工具包 DMTK 开源。DMTK 由一个服务于分布式机器学习的框架和一组分布式机器学习算法构成，是一个将机器学习算法应用在大数据上的强大工具包，适用于在超大规模数据上灵活稳定地训练大规模机器学习模型。此外，DMTK 拥有丰富易用的 API 接口，能有效降低分布式机器学习的门槛，研发人员只需要专注于数据、模型和模型训练等机器学习的核心逻辑部分。DMTK 目前的工具包主要包含以下 3 个部分：

1) **分布式机器学习框架 (Multiverso)**：由参数服务器 (Parameter Server) 和客户端软件开发包 (Client SDK) 两部分构成。参数服务器支持存储混合数据结构模型，接受并聚合工作节点服务器的数据模型更新、控制模型同步逻辑，在原有基础上从性能和功能上都得到了进一步提升。客户端软件开发包 (Client SDK) 则支持维护节点模型缓存（与全局模型服务器同步）、节点模型训练和模型通讯的流水线控制、以及片状调度大模型训练等；

图 19：DMTK 核心架构



资料来源：Microsoft，招商证券

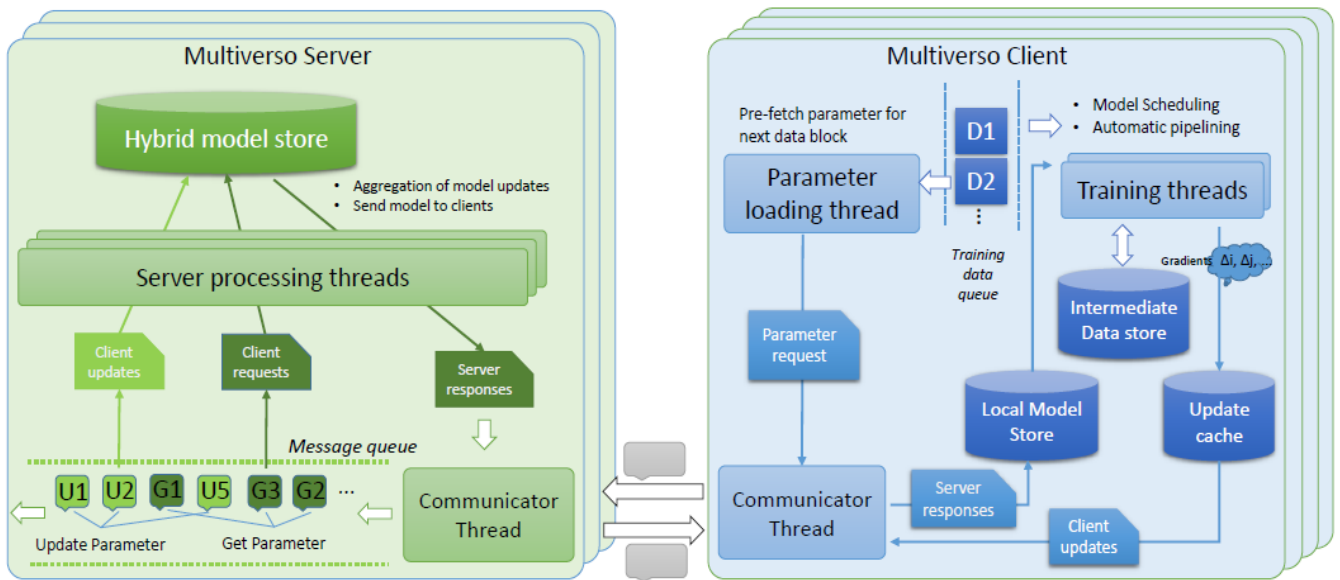
2) **LightLDA 主题模型**：LightLDA 主题模型是一种全新的计算复杂度与主题数目无关的高效算法，能够在一个普通计算机集群上处理超大规模的数据和模型，主要应用于互联网社交平台上的海量文本信息挖掘，得出相关的主题 (Topic)，为后续的机器学习和文本理解提供基础素材。据微软团队介绍，DMTK 的 LightLDA 算法是目前世界上唯一一款能在 2 千亿训练样本(Token)上训练 100 万个主题的机器学习算法，其相关参数已达到万亿级别，却仅使用了一共搭载 300 余个 CPU 内核的 20 台服务器就完成了运算过程，而其他同类算法若要实现上述运算过程一般需部署数千台计算机集群；

3) **分布式词向量训练模型 (Distribute Word Embedding, DWE)**：词向量模型通过挖掘文本数据，赋予一个词若干个指标维度（语义相关匹配），即将普通文本处理为词向

量。近来被普遍地应用于计算词汇的语义表示，它可以用作很多自然语言处理任务的词特征。目前 DMTK 为两种计算词向量的算法提供了高效的分布式实现：一种是标准的 word2vec 算法，另一种是可以对多义词计算多个词向量的新算法。

图 20: DMTK 的分布式机器学习框架——Multiverso

## Multiverso Framework



资料来源：Microsoft，招商证券

DMTK 有望助力微软在线产品实现性能跃升。DMTK 中包含的 LightLDA+DWE 组合所拥有的高效复杂处理能力，以及分布式开源架构，将 DWE 模型推升了一个前所未有的全新高度，有望给整个搜索、语义处理等相关产业带来颠覆性变革。目前在微软内部，DMTK 已经被应用到 Bing 搜索引擎、聊天机器人、广告服务等多款在线产品当中，借助 DMTK 的强大且高效的算力，有望给上述产品带来实质性的优化改善。以微软聊天机器人小冰为例，在短短 1 年时间内，第三代小冰与数千万人类用户平均每次对话轮数已经达到了 18 轮，而此前最领先同类聊天机器人平均对话轮数仅有 1.5 至 2 轮，目前小冰 3 还具备了自我进化功能，表明借助 DMTK 微软在语义分析和理解方面已经取得显著突破。



图 21: DMTK 三大核心研究方向

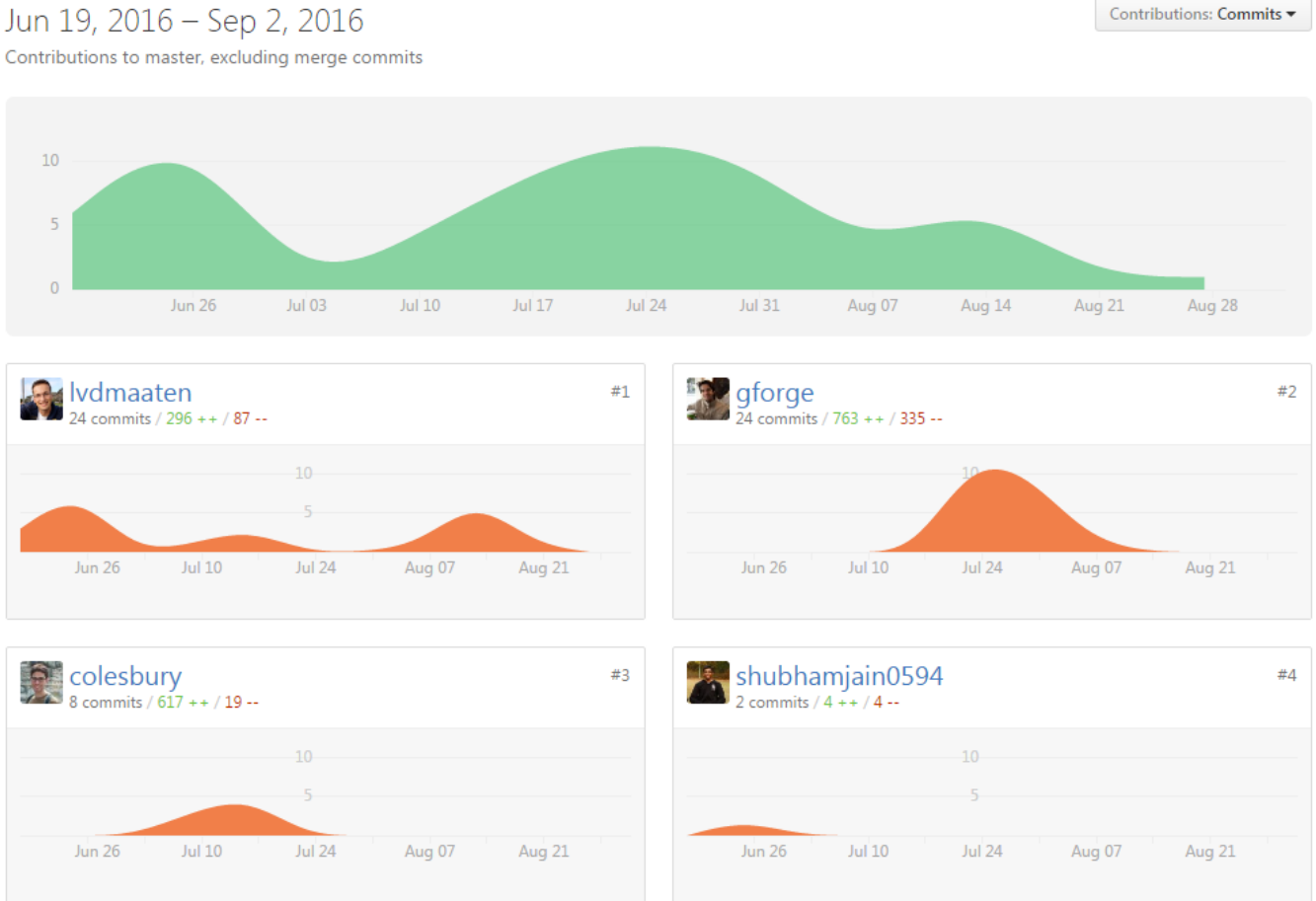


资料来源：招商证券整理

### 3、Facebook Torchnet：轻量化机器学习开源工具集

“小而美”的 Torchnet。2016 年 6 月，Facebook 人工智能研究实验室（FAIR）发表一篇论文，首次介绍该团队开发了一款基于 Torch7 的轻量级机器学习开源工具 Torchnet，目前该平台的源代码已经在 Github 开源。Facebook 没有像 Google、Microsoft、Tesla 等公司那样从头开始构建机器学习平台，而是直接在机器学习框架 Torch 7 的基础上实现。Torch 7 作为主流封装脚本语言的轻量级学习框架，针对常见代数运算效率很高，但是它不能执行抽象和公式化（Boilerplate）的代码实现。针对这一缺陷，Torchnet 采用了可抽象和公式化逻辑的开源框架，提供种类丰富的样本代码（Boilerplate Code）、关键抽象概念（Key Abstractions）和参考执行代码（Reference Implementations）等基本概念和代码集，且允许模块化编程和代码重复使用，从而减少 Bug 几率；此外，由于 Torchnet 采用的是 Torch 上通用的 Lua 编程语言，能直接对接异步数据输入和多 GPU/标准 x86 芯片运算，能有效提高运算效率。Facebook 目前还没有公布 Torchnet 的具体应用，但预计是涉及图像识别、自然语言处理、机器聊天等与 Facebook 现有核心业务相关的任务。

图 22: Torchnet 在 Github 平台上的用户贡献排名



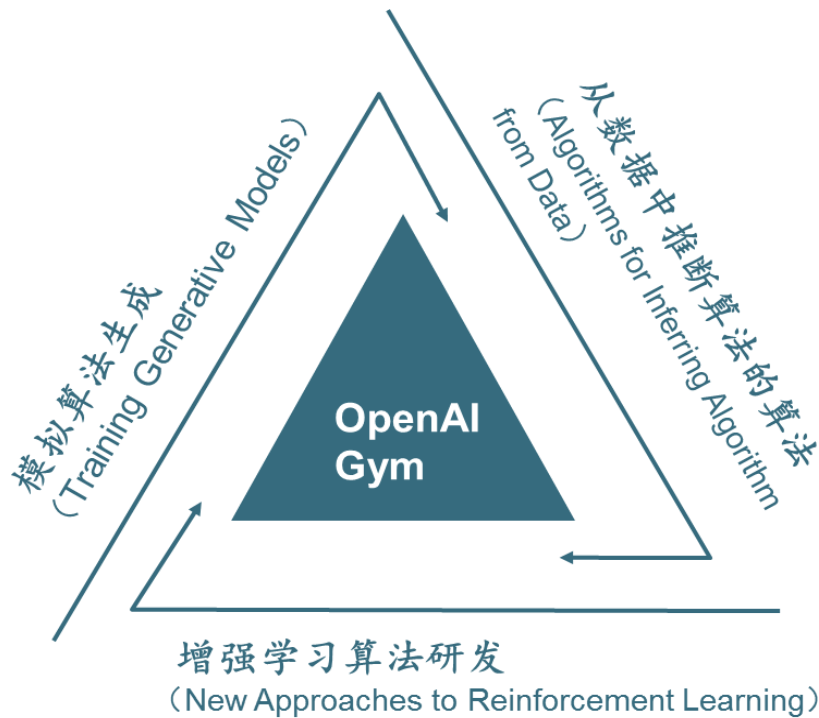
资料来源: github, 招商证券

#### 4、OpenAI：非盈利性开源人工智能组织

**OpenAI 的诞生初衷：既为战更为止战。** OpenAI 成立于 2015 年 12 月，投资者包括特斯拉 CEO Elon Musk、创业孵化器 Y Combinator 总裁 Sam Altman、PayPal 的联合创始人 Peter Thiel 以及 Amazon Web Service 等知名人士或公司，累计投资金额达 10 亿美元。作为行业内少见的非营利性组织，OpenAI 成立的初衷即是为了解决 Elon Musk 所担心的“人工智能威胁论 (the Threat of Bad AI)”，其长期目标和最终使命是以最大限度造福全人类的方式发展利用人工智能。根据 Musk 的“人工智能威胁论”，人类对人工智能的认知程度仍停留在初级阶段，对其发展方向的掌控能力也极为有限，一旦人工智能自我进化出邪恶意识将形成巨大威胁，因此极有必要全面深入理解乃至掌控人工智能发展；另一方面，目前主流研究团体如 Google、Facebook、Microsoft 等都是盈利性公司，难免会出于商业竞争目的而滥用、乱用、误用人工智能技术，加剧人工智能威胁的潜在风险。因此，设立一家非盈利性开源人工智能组织，分享全世界范围内的最新研究成果，消除企业界不完全公开研究细节的弊端，是发展人工智能的最佳实践方式之一。由于没有商业竞争和盈利的需求，OpenAI 或将会对自己的技术优势完全开源透明而不会有所保留。

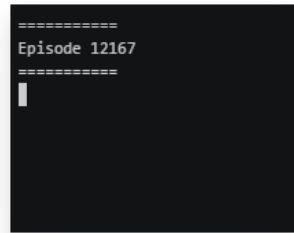
专注非监督式学习和增强学习，完备的环境集成是最大亮点。据 OpenAI 研发团队介绍，其研究内容主要为三个方向：1) 模拟算法生成 (Training Generative Models)，主要任务是快速感知与学习不同类型数据并生成新的信息，既包含监督学习，也有无监督学习；2) 从数据中推断算法的算法 (Algorithms for Inferring Algorithms from Data)，即神经网络图灵机，主要目的是让机器能够学习算法和程序等更为复杂信息，从而具备更强大推理能力；3) 增强学习算法研发 (New Approaches to Reinforcement Learning)，重点在于实现机器的自我学习能力，这也是 AlphaGO 所采用的技术手段。目前 OpenAI 已发布其第一款增强学习算法测试开源工具包 OpenAI Gym，其中最大的亮点在于包括了种类丰富研发测试环境集成，包括模拟 RL-Glue、RL-Py、Arcade Learning Environment 等，主要适用于研发和比较增强学习算法，OpenAI Gym 也整合了加州大学伯克利分校关于深度强化学习算法基准测试方面的最新研究成果。未来 OpenAI 还将继续发布非监督式学习和增强学习方面开源工具。

图 23: OpenAI Gym 三大核心研究方向

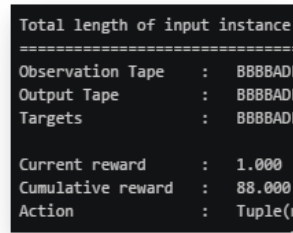


资料来源：招商证券整理

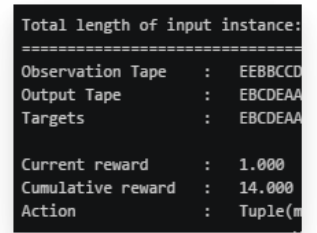
图 24: OpenAI Gym 简单计算任务测试环境



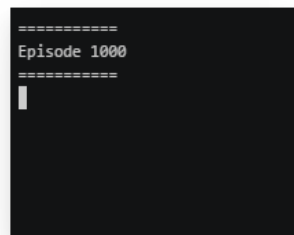
Copy-v0  
Copy symbols from the input tape.



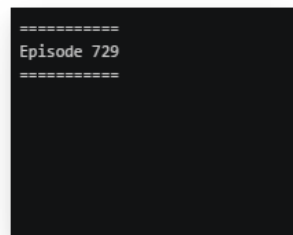
RepeatCopy-v0  
Copy symbols from the input tape multiple times.



DuplicatedInput-v0  
Copy and deduplicate data from the input tape.



ReversedAddition-v0  
Learn to add multi-digit numbers.



ReversedAddition3-v0  
Learn to add three multi-digit numbers.



Reverse-v0  
Reverse the symbols on the input tape.

资料来源: OpenAI Gym, 招商证券

图 25: OpenAI Gym MuJoCo 机器人测试环境



InvertedPendulum-v1  
Balance a pole on a cart.



InvertedDoublePendulum-v1  
Balance a pole on a pole on a cart.



Reacher-v1  
Make a 2D robot reach to a randomly located target.



HalfCheetah-v1  
Make a 2D cheetah robot



Swimmer-v1  
Make a 2D robot swim.



Hopper-v1  
Make a 2D robot hop.

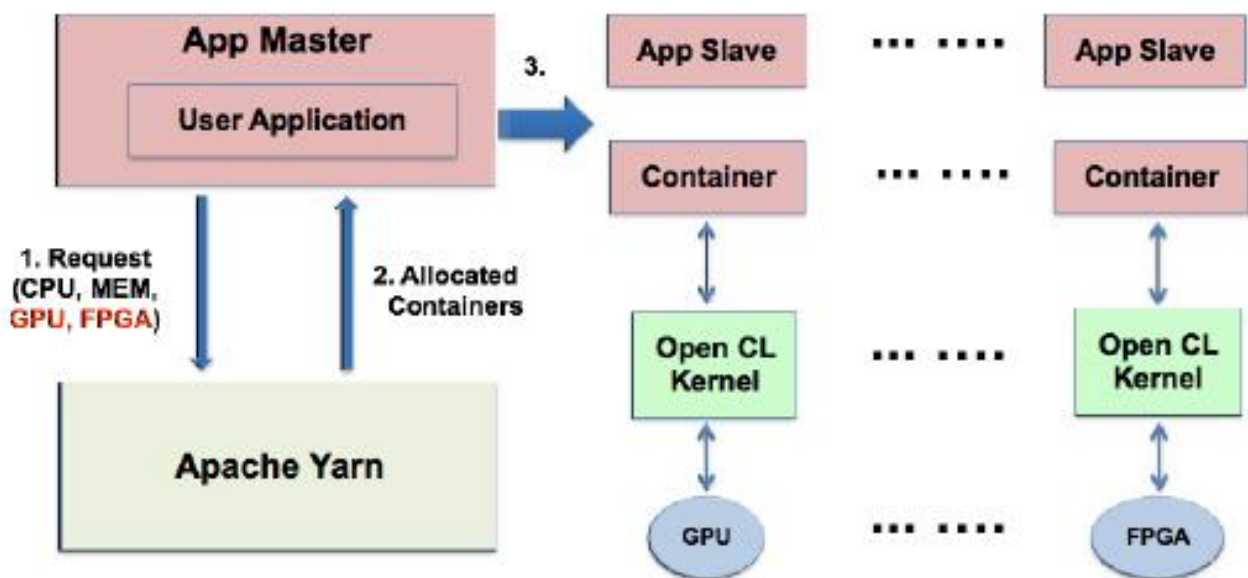
资料来源: OpenAI Gym, 招商证券

## 5、Baidu Paddle：厚积薄发的深度学习开源平台

**厚积薄发，广获肯定。**早在 2013 年，百度深度学习实验室(( Institute of Deep Learning, IDL)) 发现深度神经网络在训练数据方面相比传统单 GPU 训练平台有绝对优势，为此实验室搭建了分布式深度学习平台 Paddle(Parallel Asynchronous Distributed Deep Learning)。但创世版 PADDLE 是一个独立的深度学习平台，不能很好地支持其他平台数据的接入需求，限制了其性能发挥。为解决这一性能瓶颈，IDL 团队开发了 Spark on PADDLE 平台，让 PADDLE 变成百度 Spark 生态系统的一个功能模块，提高其在跨平台应用上的通用性，目前 Paddle 已经发展成为 Baidu 成熟的分布式深度学习平台，广泛应用于图像处理、语音识别、自然语言理解、无人驾驶等领域。自 2016 年 8 月 31 日，最新版本 Paddle 宣布开源以来，迅速受到众多业内研究人员和相关机构的测试，在性能和各项指标上，Paddle 获得了测试人员的广泛肯定。

**简洁、稳定、高效，Paddle 拥有出色架构。**Paddle 是一个云端托管的分布式深度学习平台，对于序列输入、稀疏输入和大规模数据的模型训练有着良好的支持，支持 GPU 运算，支持数据并行和模型并行。根据 Baidu 研发团队的介绍，Paddle 的架构是 Spark 异构系统计算平台，其运行流程主要分以下 3 步：1) App Master 启动用户应用程序 (Scala Driver)，向 Apache Yarn 请求其所需的资源，其中 GPU、FPGA 对应不同的资源类别；2) 用户应用程序取得所有资源，由 App Master 在相应的 App Slave 上启动 Container 运行用户程序的一个 Scala Worker；3) 按照程序 Scala Worker 的需求，调用相应的 C++ 的 OpenCL 程序，把所分配的数据传输到 GPU 或 FPGA 上，同时启动特定的 OpenCL Kernel 对输入数据进行处理并完成计算。Paddle 所拥有的出色架构，以及 Baidu 在多年的内部使用中对其不断优化，已经使之成为一款较为成熟的深度学习开发平台，尤其是具备代码简洁、设计干净、运行高效、支持拓展性强等亮点。

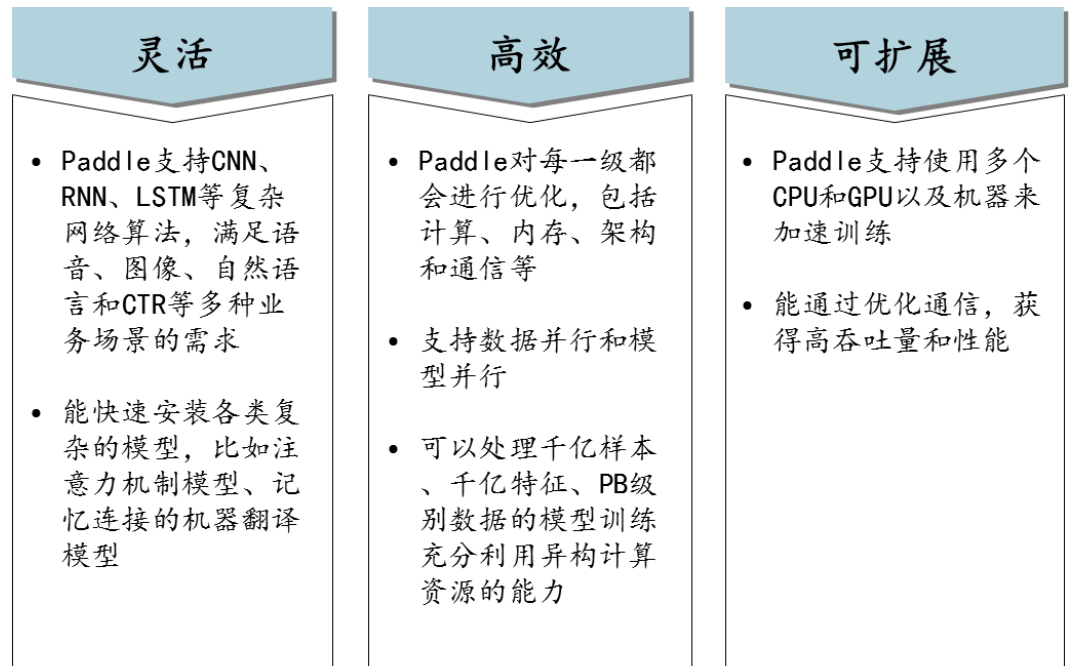
图 26: Spark 异构系统计算平台



资料来源：Baidu，招商证券



图 27: Paddle 平台主要特点



资料来源: github, 招商证券

图 28: 基于 Paddle 平台的相关基础工具和应用案例

基础工具:

 <p><b>百度机器学习 BML</b> 针对海量数据提供的云端托管的分布式机器学习平台，助力客户轻松使用最前沿的机器学习技术获得大数据预测分析能力。</p>	 <p><b>百度Mapreduce BMR</b> BMR是Apache Hadoop/Spark托管服务，方便您使用MapReduce、Spark、HBase、Hive、Pig、Kafka等进行大数据处理。</p>
 <p><b>百度OLAP引擎 Palo</b> PB级联机分析处理引擎，为客户提供稳定、高效、低成本的在线报表和多维分析服务。</p>	 <p><b>百度日志服务 BLS</b> 提供托管式日志收集、投递服务，助您从海量日志数据中获取洞察力，在大数据时代引领业务升级。</p>

应用案例:

 <p><b>百度魔图</b> 使用Paddle平台提取人脸面部特征，支持人脸相似度比对等功能。</p>	 <p><b>百度图片</b> 使用Paddle平台进行人脸识别的模型训练，支持人脸搜索等功能。</p>
 <p><b>百度语音</b> 使用Paddle平台进行大规模语音数据的处理，大大提高了语音识别的准确率。</p>	 <p><b>手机百度</b> 使用Paddle平台提取图片特征，支持图片诗慧指数等功能。</p>

资料来源: Baidu, 招商证券

与 Baidu 现有产品深度融合，未来应用场景广阔。作为 Baidu 在深度学习领域里程碑式的研究成果，Paddle 目前被用于 Baidu 旗下的各类的产品和服务中，已经拥有大量用户群体，具体应用场景包括广告点击率（CTR）预测、大规模图像分类、光学字符识别（OCR）、搜索排名、计算机病毒检测、用户精准推荐等。Paddle 在百度现有业务布局中几乎无处不在，且并大幅度地提升了这些产品的准确度，取得了良好效果。借助 Paddle 平台，Baidu 在未来有望开发出更灵活高效的大规模深度学习系统，并通过和各产品团队合作将其应用不同的产品中，打造出更为智能优质的产品和服务。

## 6、其他开源工具：Caffe、Theano、MXNet、CNTK、DL4J

### （1）Caffe

Caffe 诞生于加州伯克利分校，由贾扬清在该校从事博士后工作期间开发，于 2013 年 12 月宣布开源，目前该校视觉与学习中心（Berkeley Vision and Learning Center, BVLC）负责系统的维护工作。Caffe 全称是 Convolutional Architecture for Fast Feature Embedding，开发语言为 C++，主要针对卷积架构的网络提供解决方案，而对于循环神经网络、递归神经网络没有提供支持。Caffe 的特点是高性能、配置简单、适用平台多元等。Caffe 适用于图像分类、目标识别、图像分割等图像处理领域，同样也可用于处理非图像数据的分类、回归问题。

### （2）Theano

2008 年，Theano 诞生于蒙特利尔理工学院，派生了大量深度学习 Python 软件包，最著名的包括 Blocks 和 Keras。Theano 基于 Python 和 C++ 开发，支持全平台、CPU/GPU 切换运行，特点是使用灵活、功能强大，可以自己设计神经单元和神经连接，支持 MLP、CNN、RNN-LSTM 等多种深度神经网络。Theano 的缺点在于 C++ 代码的编译过程较为缓慢，Python 中 Theano 包的导入也偏慢。由于 Theano 的程序已经被打包，所以很难对 Theano 本身做出改进。

### （3）MXNet

MXNet 由 DMLC（Distributed（Deep）Machine Learning Community）组织开发，是一款快速高质量的开源深度学习工具。DMLC 该组织为一群来自多家高校和 IT 公司（包括卡耐基梅隆大学、MIT、华盛顿大学、微软、英特尔等）的极客发起成立，其成员曾是 CXXNet、Minerva、Purine 等项目主要开发者。MXNet 主要采用 C++ 编写，但同时也具备 R、python、julia 等语言的接口，该工具强调高内存使用效率，甚至能在智能手机上运行诸如图像识别等任务。

### （4）CNTK

Computational Network Toolkit (CNTK) 是微软研究院（Microsoft Research）出品的开源深度学习工具包，2016 年 1 月宣布开源，目前已经发展成一个通用、独立的深度学习系统。CNTK 是基于 C++ 并且跨平台的深度学习工具，目前还没有 Python 或者其他

高级语言的接口，支持 CPU 和 GPU 模式，部署简单、性能强大。在 CNTK 中，神经网络会被指定为向量运算的符号图结构，叶节点代表输入或者网络参数，其他节点代表算步骤，运算的组合则会形成神经网络的层。CNTK 不支持 ARM 架构，限制了其在移动设备上的部署能力。

### (5) DL4J

DL4J/Deeplearning4j (Deep Learning for Java) 是一款基于 Java 的深度学习开源工具，由创业公司 Skymind 于 2014 年 6 月发布，属于首个商用级别（非研究工具目的）深度学习开源工具。DL4J 与 Java、Scala 和 Clojure 均兼容，其深度学习框架包括分布式、多线程框架，以及普通的单线程深度学习框架，通过集群进行定型过程来实现快速处理大量数据的功能。DL4J 以即插即用为目标，可与 Hadoop 和 Spark 深度集成，并通过多项预设的使用方便开发者在快速集成深度学习功能，其用户包括埃森哲、雪佛龙、IBM 等知名公司。

## 7、人工智能开源平台争夺生态圈话语权，争夺下一个时代的“Android”

我们将人工智能开源平台对比整理如下。同时各大科技巨头和科研机构纷纷开源人工智能技术，以此来获得大量的用户需求和开发人员，建立开放共享、互利共赢的人工智能生态圈，可以说人工智能开源平台，会比 Android 更深远地影响世界。

表 1：人工智能开源平台对比

开源工具	研发机构	适用平台	开发语言	预训模型	分布式联机功能	主要特点
Caffe	BVLC	Ubuntu、OS X、AWS、Windows 等	C++、Python	有	不支持	主要针对卷积架构的网络提供解决方案，适用于图像分类、目标识别、图像分割等图像处理领域，高性能、配置简单、适用平台多元
CNTK	Microsoft	Windows、Linux 等	C++	无	支持	基于 C++ 并且跨平台的深度学习工具，部署简单、性能强大，但目前暂不支持在移动设备上部署
DMTK	Microsoft	Windows、Linux 等	C++	有	支持	机器学习算法与大数据处理深度融合，有丰富易用的 API 接口，算力强大且高效，实用性强
DL4J	Skymind	Linux、Ubuntu、Windows、OS X、Android	Java、Scala、C、C++	有	支持	属于首个商用级别（非研究目的）深度学习开源工具，以即插即用为目标，可与 Hadoop 和 Spark 深度集成，方便开发者在快速集成深度学习功能

<b>MXNet</b>	DMLC	Ubuntu、OS X、Windows、AWS、Android、iOS、JavaScript	C++、Python、Julia、Matlab、Go、R、Scala 等	有	支持	允许用户自由把图计算和过程计算混合, 支持自动多卡调度, 强调高内存使用效率、高使用性能
<b>OpenAI Gym</b>	OpenAI	Linux、OS X 等	Python 等	有	支持	专注非监督式学习和增强学习, 具有完备的训练测试环境集成, 承诺非营利性
<b>Paddle</b>	百度深度学习实验室	Windows、Linux、OS X 等	C++、Cuda 等	有	支持	云端托管的分布式深度学习平台, 以 Spark 异构计算平台为基础架构, 对于序列输入、稀疏输入和大规模数据的模型训练有着良好的支持, 代码简洁、运行高效、支持拓展性强
<b>TensorFlow</b>	Google Brain	Linux、OS X、Windows	C++、Python、Cuda 等	无	支持	有深度学习专用运算库, 易于混合使用多种不同模型进行训练, 运行速度快、可扩展性、与产品可衔接性强
<b>Theano</b>	蒙特利尔理工学院	Linux、Ubuntu、Windows 等	Python	有	不支持	可自行设计神经单元和神经连接, 支持 MLP、CNN、RNN-LSTM 等多种深度神经网络, 缺点是 C++代码的编译过程较为缓慢
<b>Torchnet</b>	Facebook FAIR	Linux、Android、OS X、iOS	C、Lua	有	不支持	采用可抽象和公式化逻辑的开源框架, 提供种类丰富的基本概念和代码集, 可执行抽象和公式化代码, 允许模块化编程和代码重复使用

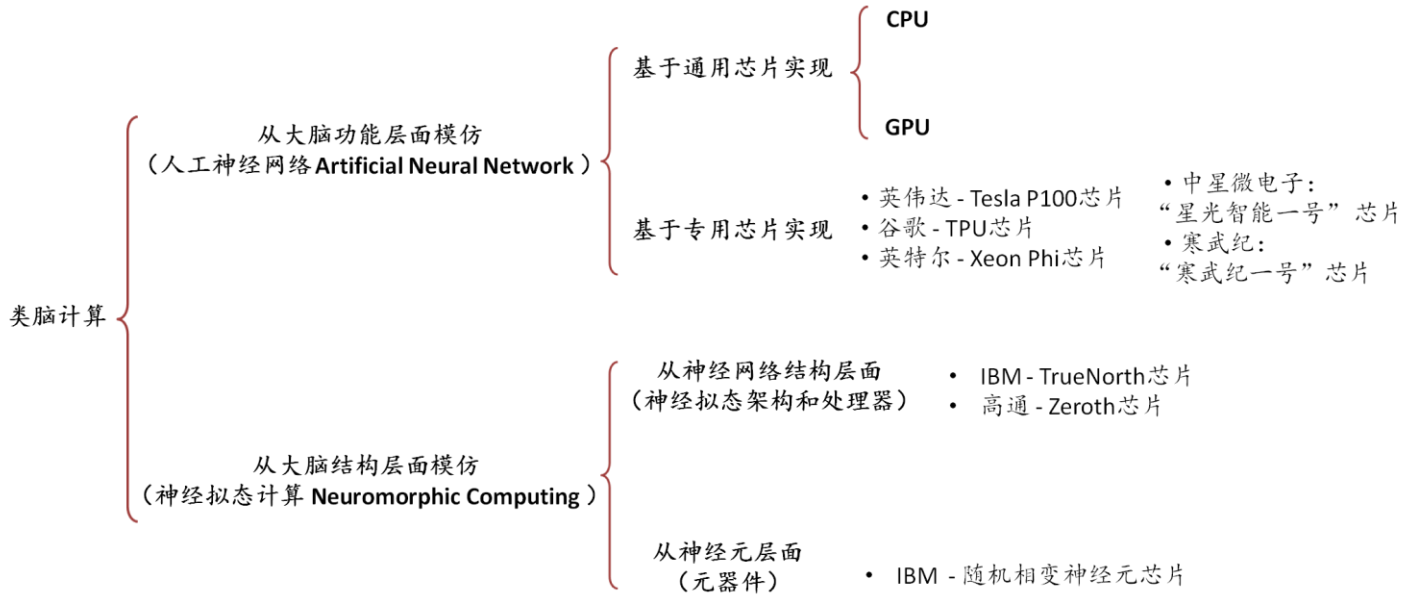
资料来源: 招商证券整理

### 三、人工智能芯片：从通用到专用，类脑计算的进化

#### 1、类脑计算的进化

类脑计算可以分为两个方向：一个是从人工神经网络从功能层面模仿大脑的能力；另一个神经拟态计算(neuromorphic computing)则是从结构层面去逼近大脑，其结构也有两个层次，一是神经网络层面，与之相应的是神经拟态架构和处理器，二是神经元层面，与之相应的是元器件。

图 29：人工智能芯片分类



资料来源：招商证券整理

图 30：人工智能大脑的进化



资料来源：招商证券整理

敬请阅读末页的重要说明



## 2、GPU 王者 Nvidia，在深度学习上使用 Nvidia 的组织 2 年增长了 30 多倍，为绝大部分的人工智能平台提供计算能力

### (1) GPU 的高并行处理

GPU (Graphics Processing Unit)，图形处理器，专门为 2d 和 3d 的图形、视频、可视计算和显示做出优化。

图 31: GPU VS CPU



资料来源: NVIDIA, 招商证券

GPU 具有高并行结构 (highly parallel structure)，在处理图形数据和复杂算法方面拥有比 CPU 更高的效率。对比 GPU 和 CPU 在结构上的差异，CPU 大部分面积为控制器和寄存器，GPU 拥有更多的 ALU (Arithmetic Logic Unit, 逻辑运算单元) 用于数据处理，而非数据高速缓存和流控制，这样的结构适合对密集型数据进行并行处理。CPU 执行计算任务时，一个时刻只处理一个数据，不存在真正意义上的并行，而 GPU 具有多个处理器核，在一个时刻可以并行处理多个数据。

图 32: 采用 GPU 加速与只采用 CPU 训练 CNN 的性能比较

批处理大小	CPU 训练时间	GPU 训练时间	GPU 加速
64 images	64 s	7.5 s	8.5 ×
128 images	124 s	14.5 s	8.5 ×
256 images	257 s	28.5 s	9.0 ×

- ImageNet 2012 年冠军模型
- 7 层模型
- 5 个卷积层 + 2 个全连接层
- Caffe
- 2 颗 10 核 Ivy Bridge CPUs
- 1 颗 Tesla K40 GPU
- CPU 版本使用 Intel MKL BLAS 库
- GPU 版本使用 CUDA BLAS 库

资料来源: Nvidia, 招商证券整理

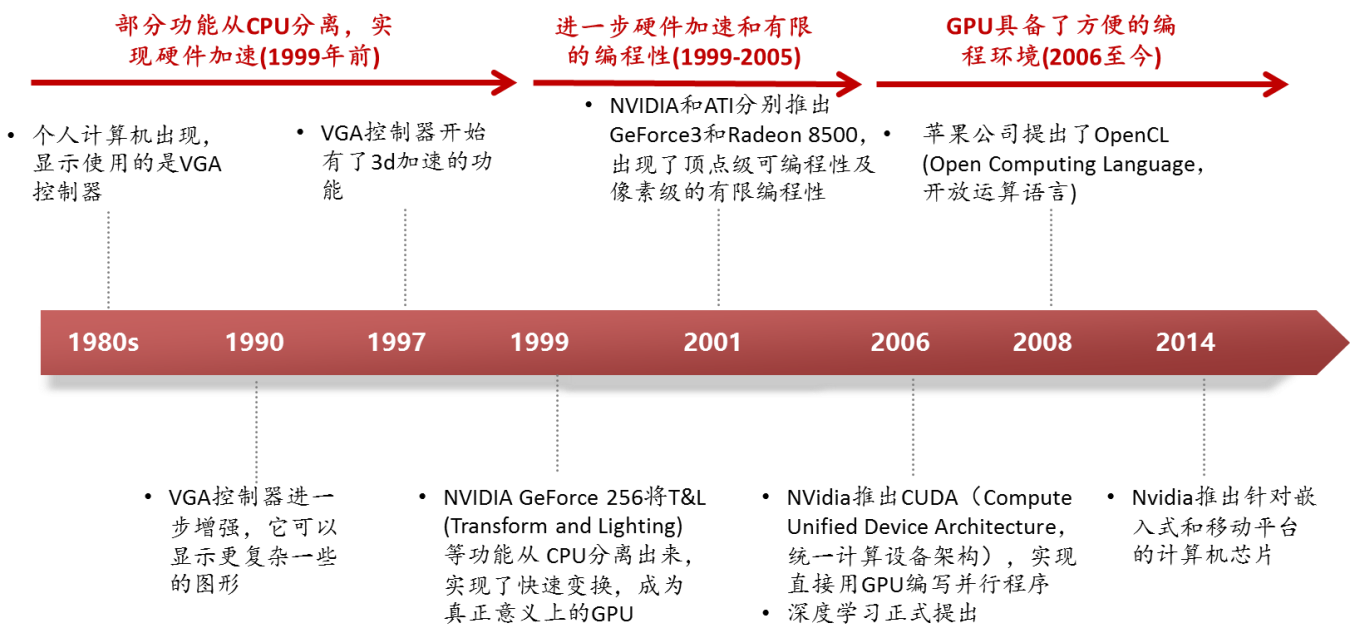
深度学习在神经网络训练中，需要很高的内在并行度、大量的浮点计算能力以及矩阵运算，而 GPU 可以提供这些能力，并且在相同的精度下，相对传统 CPU 的方式，拥有

更快的处理速度、更少的服务器投入和更低的功耗。

以 ImageNet 竞赛为例，基于 GPU 加速的深度学习算法，百度、微软和 Google 的计算机视觉系统在 ImageNet 图像分类和识别测试中分别达到了 5.98% (2015 年 1 月数据) 4.94% (2015 年 2 月数据)、4.8% (2015 年 2 月数据)、的错误率，接近或超过了人类识别水平。

## (2) GPU 发展历程：功能从 CPU 分离实现硬件加速和编程

图 33: GPU 发展历程



资料来源：招商证券整理

GPU 的发展经历了三个阶段：

第一代 GPU(1999 年以前)，部分功能从 CPU 分离，实现硬件加速，以 GE(Geometry Engine)为代表，只能起到 3D 图像处理的加速作用，不具有软件编程特性。

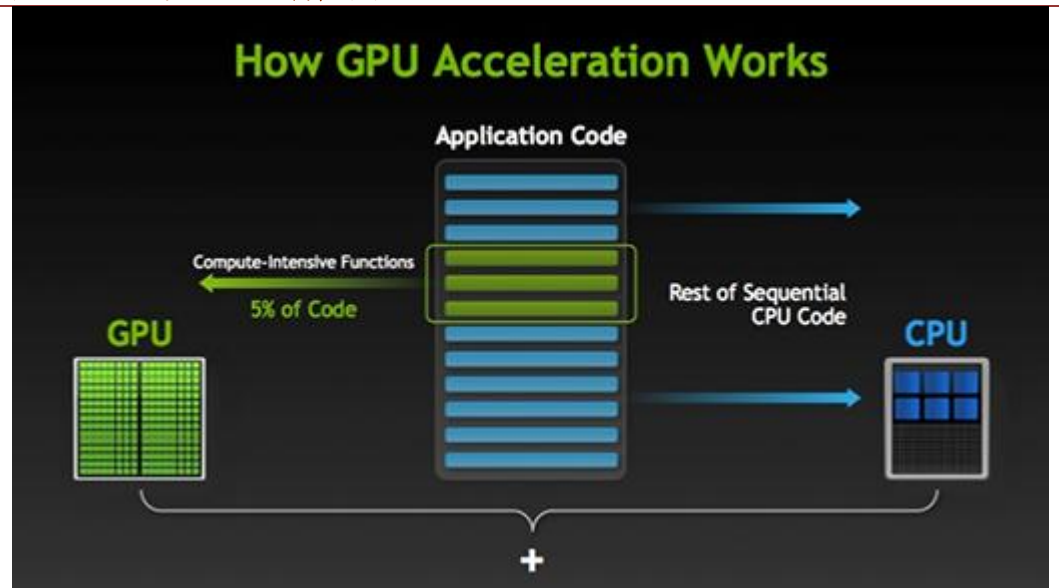
第二代 GPU(1999 年-2005 年)，实现进一步的硬件加速和有限的编程性，1999 年 NVIDIA GeForce 256 将 T&L(Transform and Lighting)等功能从 CPU 分离出来，实现了快速变换，也是真正意义上的 GPU 出现的标志；2001 年 NVIDIA 和 ATI 分别推出的 GeForce3 和 Radeon 8500，图形硬件的流水线被定义为流处理器，出现了顶点级可编程性，同时像素级也具有有限的编程性，但 GPU 的编程性比较有限。

第三代 GPU(2006 年以后)，GPU 实现方便的编程环境可以直接编写程序；2006 年 NVIDIA 与 ATI 分别为推出了 CUDA(Computer Unified Device Architecture, 统一计算架构)编程环境和 CTM(Close To the Metal)编程环境；2008 年，苹果公司提出一个通用的并行计算编程平台 OpenCL (Open Computing Language, 开放运算语言)，和 CUDA 绑定在 NV 的显卡上不同，OpenCL 和具体的计算设备没有关系。

### (3) GPU 的计算标准

随着 GPU 可编程性不断增强，特别是 CUDA 等编程环境的出现，使 GPU 通用计算编程的复杂性大幅度降低。由于可编程性、功能、性能不断提升和完善，GPU 已演化为一个新型可编程高性能并行计算资源，全面开启 GPU 面向通用计算的新时代。GPGPU (General Purpose CPU)，通用目的 GPU，它的计算原理是我们任何密集的计算所对应代码量可能往往连 5% 都不到，这部分代码，我们可以把它提取出来，放到 GPU 上去处理。

图 34: GPU 实现加速计算的原理



资料来源: cnnic, 招商证券

作为计算平台，人们往往将 CPU 和 GPU 联合起来，组成一个异构的平台。因为和 CPU 相比，GPU 所能处理的事情还比较单纯，比如 I/O、访问磁盘或网络，或者说数据采集和整理，或者一些串行的没法并行的工作，这些操作仍然需要 CPU 来辅助进行。

目前主要的通用计算标准有 CUDA (Compute Unified Device Architecture, 统一计算设备架构)、OpenCL (Open Computing Language, 开放计算语言)、Direct Compute。

CUDA 在 2006 年由 NVIDIA 提出，是一种将 GPU 作为数据并行计算设备的软硬件体系；CUDA 编程模型将 CPU 作为主机 (Host)，GPU 作为协处理器 (co-processor) 或者设备 (Device)，采用类 C 语言开发；目前支持 Windows、Linux、MacOS 三种主流操作系统，但仅支持 NVIDIA 开发的 GPU 芯片。

OpenCL 在 2008 年由苹果首先提出，苹果拥有其商标权；随后苹果与 AMD, IBM, Intel 和 NVIDIA 技术团队合作初步完善 OpenCL 标准，2010 年 OpenCL 1.1 发布。OpenCL 是第一个面向异构系统通用目的并行编程的开放式、免费标准，此异构平台可由 CPU、GPU 或其他类型的处理器组成；OpenCL 提供了基于任务分区和数据分区的并行计算机制，也是基于 C 的一个程式语言；不同于 CUDA 仅支持 NVIDIA 开发的 GPU 芯片，OpenCL 支持一系列硬件，包括 GPU、GPP、现场可编程门阵列 (FPGA) 和数字信号处理器 (DSP) 多种 GPU 芯片。

表 2: GPU 通用计算标准

标准名称	提出标准的公司	标准特征
CUDA (Compute Unified Device Architecture, 统一计算设备架构)	NVIDIA	1) 一种将 GPU 作为数据并行计算设备的软硬件体系 2) CUDA 编程模型将 CPU 作为主机 (Host), GPU 作为协处理器 (co-processor) 或者设备 (Device) 3) 采用类 C 语言开发 4) 支持 Windows、Linux、MacOS 三种主流操作系统; 仅支持 NVIDIA 的 GPU 芯片
OpenCL (Open Computing Language, 开放计算语言)	Apple 公司最初开发, 与 AMD, IBM, Intel 和 NVIDIA 技术团队合作完善	1) 第一个面向异构系统通用目的并行编程的开放式、免费标准, 此异构平台可由 CPU、GPU 或其他类型的处理器组成 2) OpenCL 提供了基于任务分区和数据分区的并行计算机制 3) OpenCL 也是基于 C 的一个程式语言 4) 支持多种 GPU 芯片
Direct Compute	Microsoft	1) 用于 GPU 通用计算的应用程序接口 2) 集成在 Microsoft DirectX 内, 允许 Windows 平台上运行的程序利用 GPU 进行通用计算 3) 和 OpenCL 一样, 都是开放标准

资料来源: 招商证券整理

Direct Compute 是一种由 Microsoft 开发和推广的用于 GPU 通用计算的应用程序接口, 集成在 Microsoft DirectX 内, 允许 Windows Vista 或 Windows 7 平台上运行的程序利用 GPU 进行通用计算。和 OpenCL 一样, Direct Compute 是开放标准。

**(4) GPU 王者 Nvidia: 在深度学习上使用 Nvidia 的组织 2 年增长了 30 多倍, 为绝大部分的人工智能平台提供计算能力**

目前市场上, 比如在 PC 领域的 GPU, 基本上只能见到 Nvidia, Intel, AMD 这几家公司的产品。

Nvidia 在这几家显卡生产厂商中实力最强, 有很多的产品线, GeForce, 主要是为桌面游戏准备的; Quadro, 它是专门为显示级的工作站准备的; TSELA, 它专门为科学计算准备的; TEGRA, 它是为移动端、自动驾驶平台等准备的; GRID, 它是专门为服务器集成虚拟化使用; NVs, 它是对特别大屏幕由多个显示器拼在一起的这种用途而专门定制的。

表 3: GPU 主流生产厂商

厂商名称	GPU 芯片名称	主要用途
NVIDIA	GeForce	游戏
	Quadro	显示级的工作站
	TESLA	科学计算
	TEGRA	移动端、自动驾驶平台
	GRID	服务器集成虚拟化
	NVs	多屏显示
Intel	集成显卡	图形处理
AMD (ATI)	Radeon	台式机、笔记本
	FirePro	工作站、服务器
Matrox		专业 2d 绘图领域

资料来源：招商证券整理

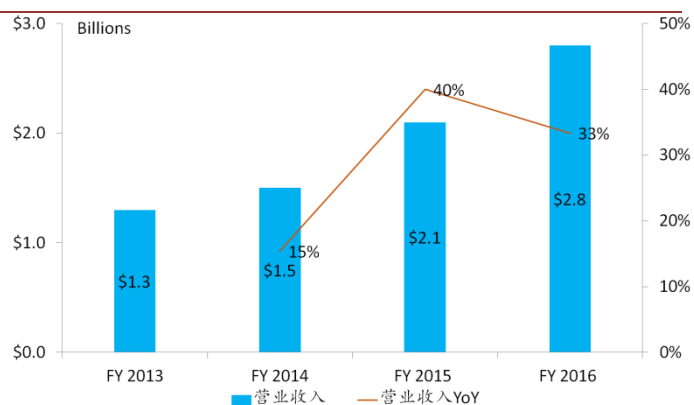
Intel 主要是集成显卡，集成在 Intel 的 CPU 中，如 i5, i7。现在最新的英特尔集成显卡，基本上达到了 NV 的入门级显卡的水平，满足绝大多数人的日常需求，比如上网、文字处理、看视频以及玩小游戏等。Intel 集成显卡最大优点，就是非常省电，同时直接利用这种内存作为显存，也能大大的降低了这种显卡的生产成本。

AMD 的前身 ATI，是专业的显卡生产厂商，后来被 AMD 收购。AMD 的主要产品有面向普通的桌面、移动平台的 Radeon 系列，为专业工作站平台准备的 FirePro 系列等。

Matrox，目前已经淡出主流市场，在一些专业的 2d 绘图领域还有用到。

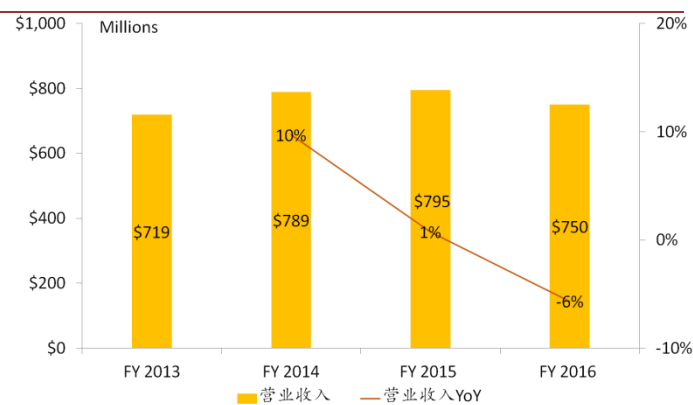
GPU 出色的浮点计算性能提高深度神经网络计算的有效性。

图 35: nvidia 游戏领域收入



资料来源：nvidia，招商证券

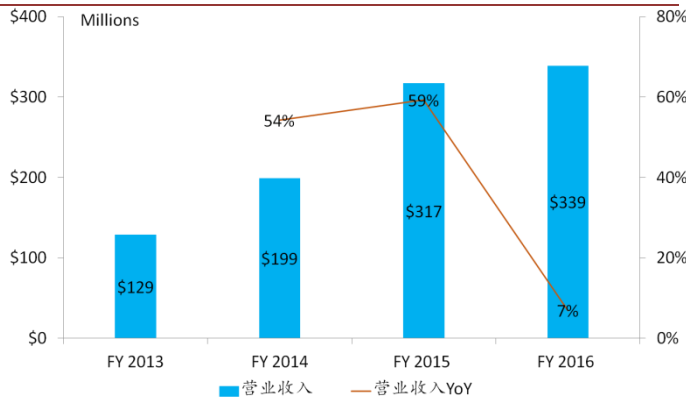
图 36: nvidia 可视化领域收入



资料来源：nvidia，招商证券

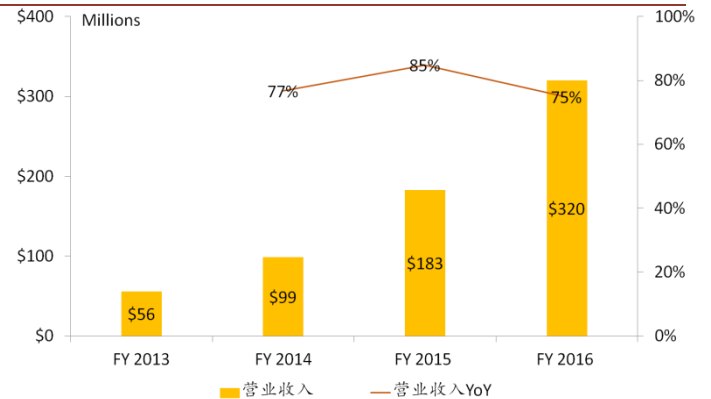


图 37: nvidia 数据中心领域收入



资料来源: nvidia, 招商证券

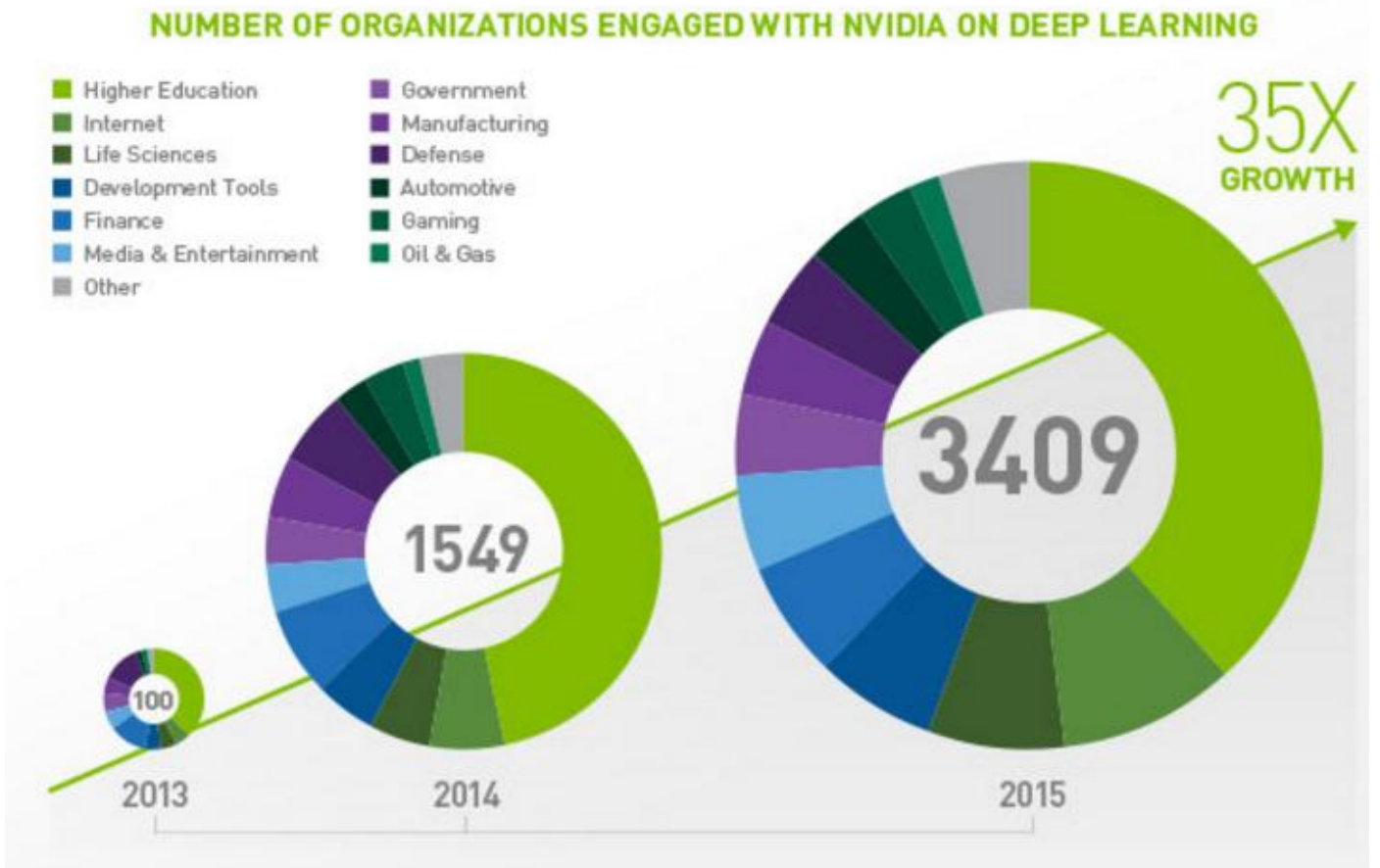
图 38: nvidia 无人驾驶领域收入



资料来源: nvidia, 招商证券

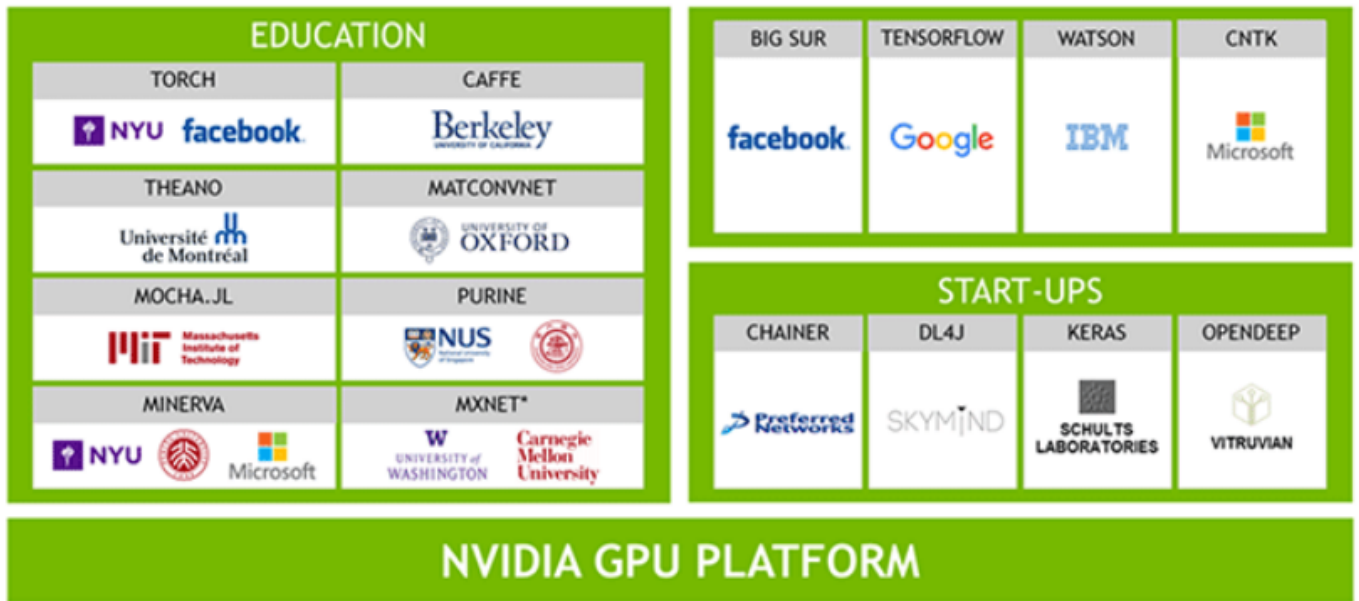
通过改进机器学习算法和升级计算硬件, 各行各业都在采用深度学习技术来处理爆炸性增长的数据量, 帮助他们找到新方法利用随手可得的数据财富开发新产品、服务和流程, 从而创造巨大的竞争优势, 在深度学习上使用英伟达的组织 2 年增长了 30 多倍。

图 39: 在深度学习上使用英伟达的组织 2 年增长了 30 多倍



资料来源: nvidia, 招商证券

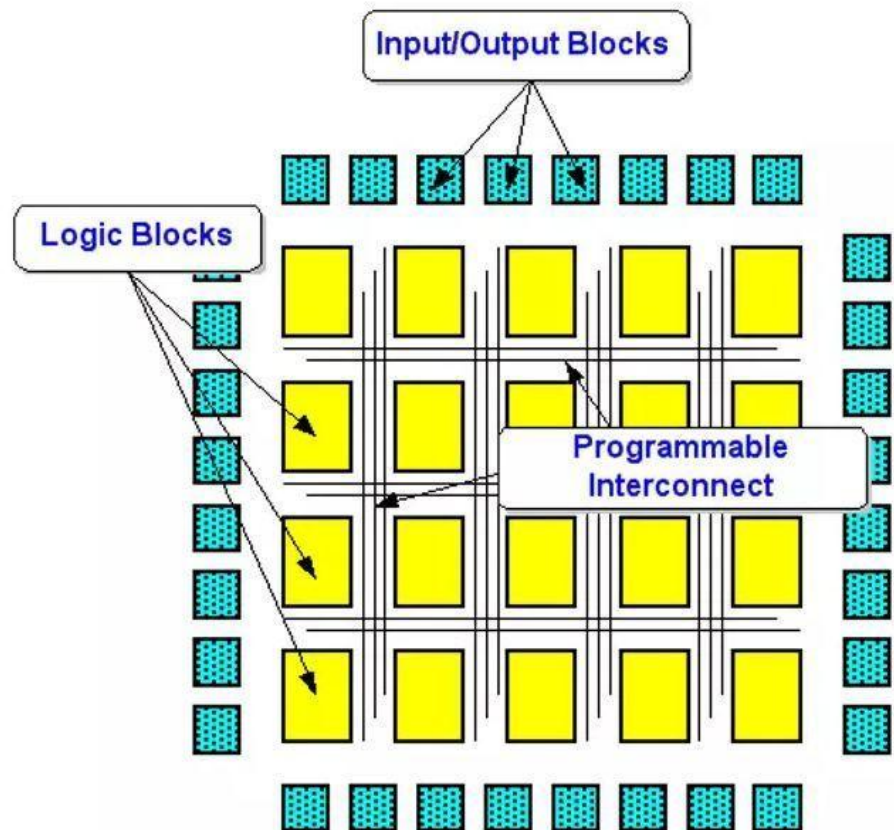
图 40: NVIDIA 为绝大部分的人工智能平台提供计算能力



资料来源: nvidia, 招商证券

### 3、FPGA：性能与灵活性的折衷，人工智能领域 GPU 的强劲对手

图 41: FPGA 的结构



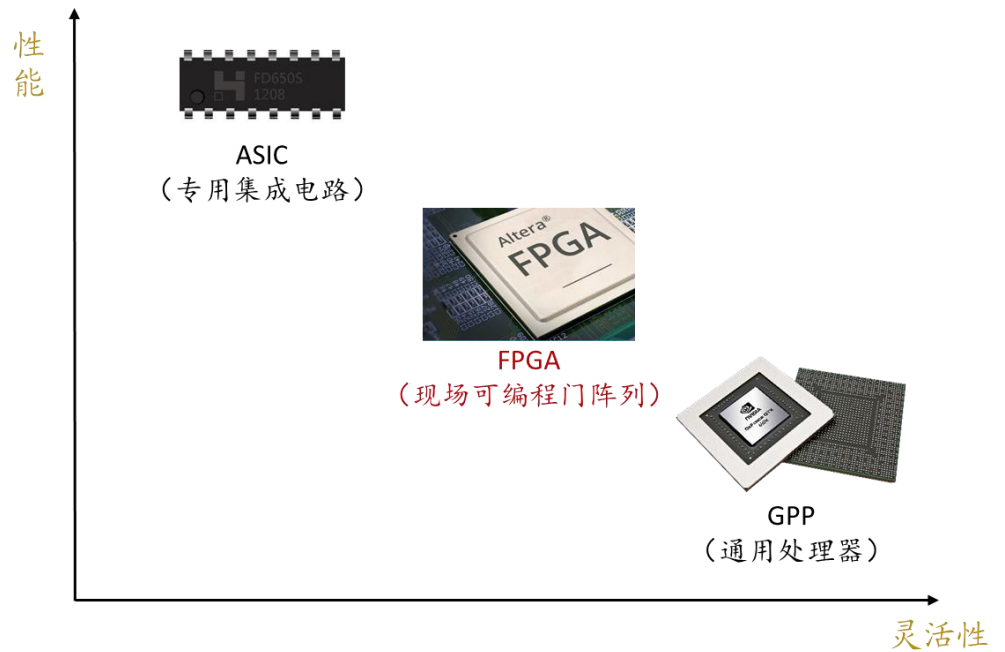
资料来源: 百度图片, 招商证券

FPGA (Field-Programmable Gate Array), 即现场可编程门阵列, 它是在 PAL、GAL、CPLD 等可编程器件的基础上进一步发展的产物。它是作为专用集成电路 (ASIC) 领域中的一种半定制电路而出现的, 既解决了定制电路的不足, 又克服了原有可编程器件门电路数有限的缺点。FPGA 以并行运算为主, 以硬件描述语言来实现。

作为 GPU 在算法加速上强有力的竞争者, FPGA 是否立即支持不同硬件, 显得尤为重要。FPGA 与 GPU 不同之处在于硬件配置灵活, 且 FPGA 在运行深度学习中关键的子程序 (例如对滑动窗口的计算) 时, 单位能耗下通常能比 GPU 提供更好的表现。最近, FPGA 工具开始采用包括 OpenCL 在内的软件级编程模型, 使其越来越受经主流软件开发训练的用户青睐。

在评估硬件平台的加速时, 必须考虑到灵活性和性能之间的权衡。一方面, 通用处理器 (GPP) 可提供高度的灵活性和易用性, 但性能相对缺乏效率。这些平台往往更易于获取, 可以低廉的价格生产, 并且适用于多种用途和重复使用。另一方面, 专用集成电路 (ASIC) 可提供高性能, 但代价是不够灵活且生产难度更大。这些电路专用于某特定的应用程序, 并且生产起来价格昂贵且耗时。

图 42: FPGA、GPP、ASIC 比较



资料来源: 招商证券整理

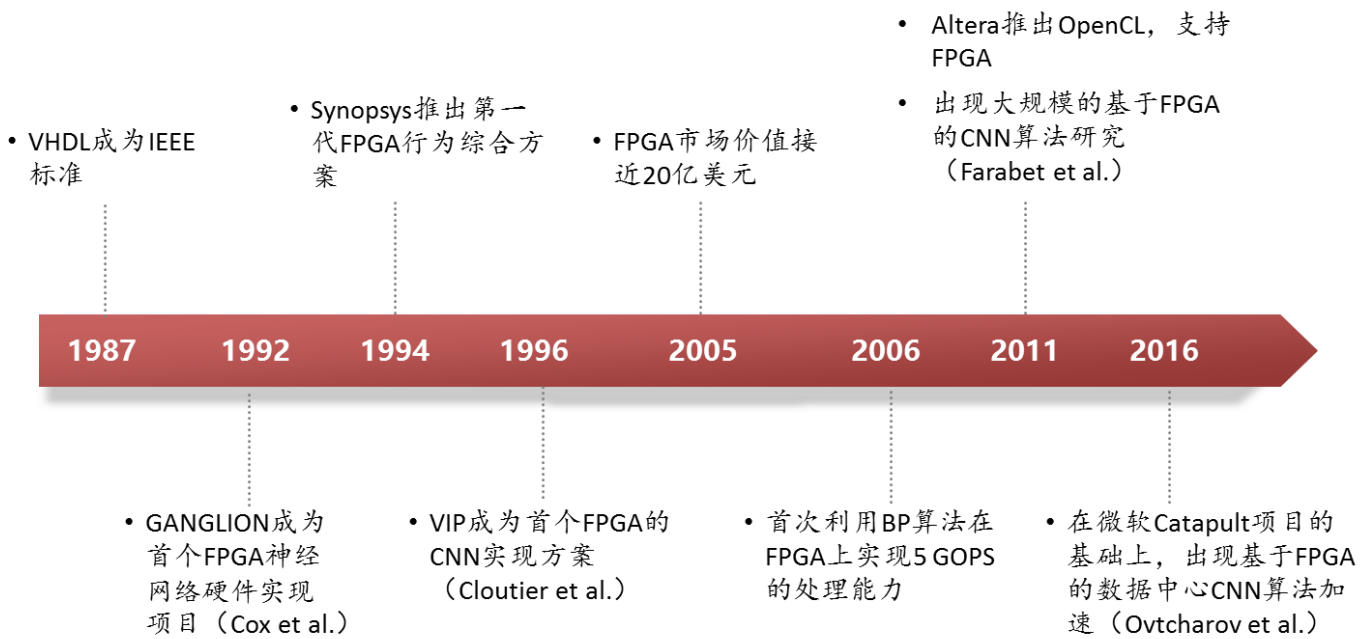
FPGA 是这两个极端之间的折中。FPGA 属于一类更通用的可编程逻辑设备 (PLD), 并且简单来说, 是一种可重新配置的集成电路。因此, FPGA 既能提供集成电路的性能优势, 又具备 GPP 可重新配置的灵活性。FPGA 能够简单地通过使用触发器 (FF) 来实现顺序逻辑, 并通过使用查找表 (LUT) 来实现组合逻辑。现代的 FPGA 还含有硬化组件以实现一些常用功能, 例如全处理器内核、通信内核、运算内核和块内存 (BRAM)。另外, 目前的 FPGA 趋势趋向于系统芯片 (SoC) 设计方法, 即 ARM 协处理器和 FPGA 通常位于同一芯片中。

对于深度学习而言, FPGA 提供了优于传统 GPP 加速能力的显著潜力。GPP 在软件层面的执行依赖于传统的冯·诺依曼架构, 指令和数据存储于外部存储器中, 在需要时再取出; 该架构的瓶颈是处理器和存储器之间的通信, 这严重削弱了 GPP 的性能, 尤其

影响深度学习经常需要获取的存储信息技术。相比较而言，FPGA 的可编程逻辑原件可用于实现普通逻辑功能中的数据和控制路径，而不依赖于冯·诺伊曼结构。它们也能够利用分布式片上存储器，以及深度利用流水线并行，这与前馈性深度学习自然契合。现代 FPGA 还支持部分动态重新配置，当 FPGA 的一部分被重新配置时另一部分仍可使用。这将对大规模深度学习模式产生影响，FPGA 的各层可进行重新配置，而不扰乱其他层正在进行的计算。这将可用于无法由单个 FPGA 容纳的模型，同时还可通过将中间结果保存在本地存储以降低高昂的全球存储读取费用。

最重要的是，相比于 GPU，FPGA 为硬件加速设计的探索提供了另一个视角。GPU 和其它固定架构的设计是遵循软件执行模型，并围绕自主计算单元并行以执行任务搭建结构。由此，为深度学习技术开发 GPU 的目标就是使算法适应这一模型，让计算并行完成、确保数据相互依赖。与此相反，FPGA 架构是为应用程序专门定制的。在开发 FPGA 的深度学习技术时，较少强调使算法适应某固定计算结构，从而留出更多的自由去探索算法层面的优化。需要很多复杂的下层硬件控制操作的技术很难在上层软件语言中实现，但对 FPGA 执行却特别具有吸引力。

图 43: FPGA 深度学习研究里程碑



资料来源：招商证券整理

对考察一系列设计工具的研究者而言，其对工具的筛选标准通常与其是否具备用户友好的软件开发工具、是否具有灵活可升级的模型设计方法以及是否能迅速计算、以缩减大模型的训练时间有关。随着 FPGA 因为高抽象化设计工具的出现而越来越容易编写，其可重构性又使得定制架构成为可能，同时高度的并行计算能力提高了指令执行速度，FPGA 将为深度学习的研究者带来好处。

对应用科学家而言，尽管有类似的工具级选择，但硬件挑选的重点在于最大化提高单位能耗的性能，从而为大规模运行降低成本。所以，FPGA 凭借单位能耗的强劲性能，加上为特定应用定制架构的能力，就能让深度学习的应用科学家受益。

FPGA 能满足这两类受众的需求，是一个合乎逻辑的选择。目前的 FPGA 市场由 Xilinx 和 Altera 主导，两者共同占有 85% 的市场份额。此外，FPGA 正迅速取代 ASIC 和应



用专用标准产品 (ASSP) 来实现固定功能逻辑。预计 FPGA 市场规模预计在 2016 年将达到 100 亿美元。

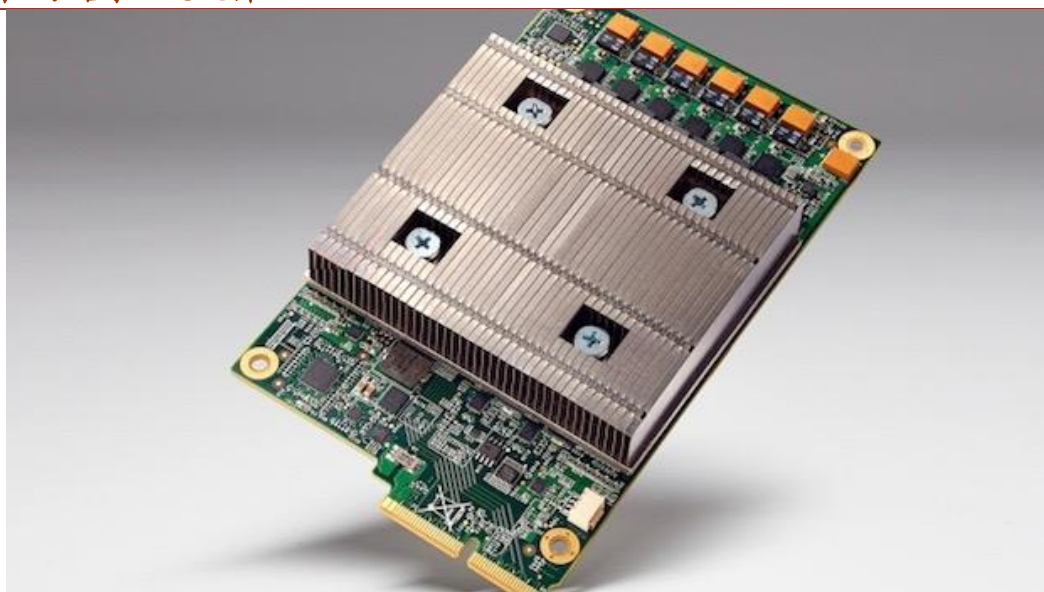
#### 4、人工智能专用芯片：大幅提升人工智能算法运行效率，推动人工智能产业加速发展

传统的算法在通用芯片(CPU 和 GPU)上效率不高，功耗比较大，因此从芯片的设计角度来说，通用型往往意味着更高的成本。于是出现了特定的算法加速器，来加速包括卷积神经网络、递归神经网络在内的各种神经网络算法，专用芯片的最大优势在于其成本和功耗降低。专用深度学习芯片将大幅提升人工智能算法运行效率，推动人工智能产业加速发展。

##### (1) 谷歌 TPU 芯片：匹配 Tensor Flow，专为人工智能而制

今年 5 月，谷歌展示了一款专门针对人工智能定制的新型芯片 TPU (Tensor Processing Unit); TPU 由谷歌与 OpenPower Foundation 合作开发，OpenPower Foundation 提供芯片设计。

图 44: 谷歌 TPU 芯片



资料来源：公开资料，招商证券

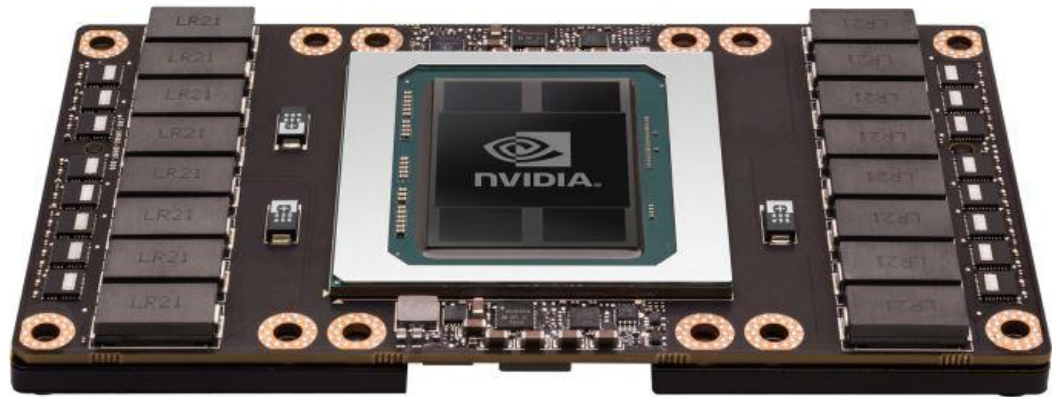
谷歌的 TPU(Tensor Processing Unit)是一种专用的加速器芯片，跟其深度学习软件 Tensor Flow 匹配。TPU 专门针对机器学习进行过裁减，运行单个操作时需要的晶体管更少，其研发目的是为了替代 GPU，实现更高效率的深度学习。

##### (2) Nvidia Tesla P100 芯片：专门用于加速深度学习

2016 年 4 月 6 日，英伟达 CEO 黄仁勋在其公司 GPTech 峰会上发布了一款支持深度学习的新芯片 Tesla P100。这是英伟达第一次设计一个专门用于该领域的芯片，专门用于加速人工智能，专门用于加速深度学习，并且英伟达决定全力投入人工智能。



图 45: Nvidia Tesla P100 芯片



资料来源：公开资料，招商证券

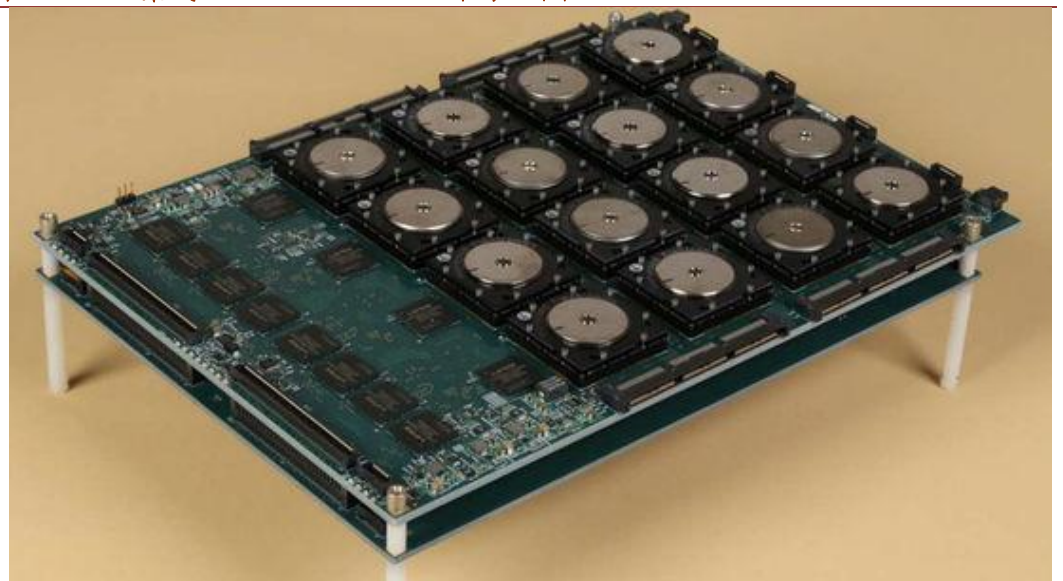
Tesla P100 芯片安装了 150 亿个晶体管，是目前市场上许多处理器、图形芯片的 3 倍；芯片面积为 600 平方毫米，双精度运算速度 5.3 万亿次，单精度运算速度 10.6 万亿次，半精度运算速度 21.2 万亿次。

Nvidia 同时发布了一款搭载了八个 P100 芯片、可用于深度学习的计算机 DGX-1。

### （3）IBM TrueNorth 芯片、随机相变神经元芯片：神经拟态计算，像大脑一样工作

2014 年 8 月 7 日，IBM 宣布研发出一款可以像大脑一样工作的计算机芯片 TrueNorth。

图 46: IBM 集成 16 块 TrueNorth 芯片的电路板



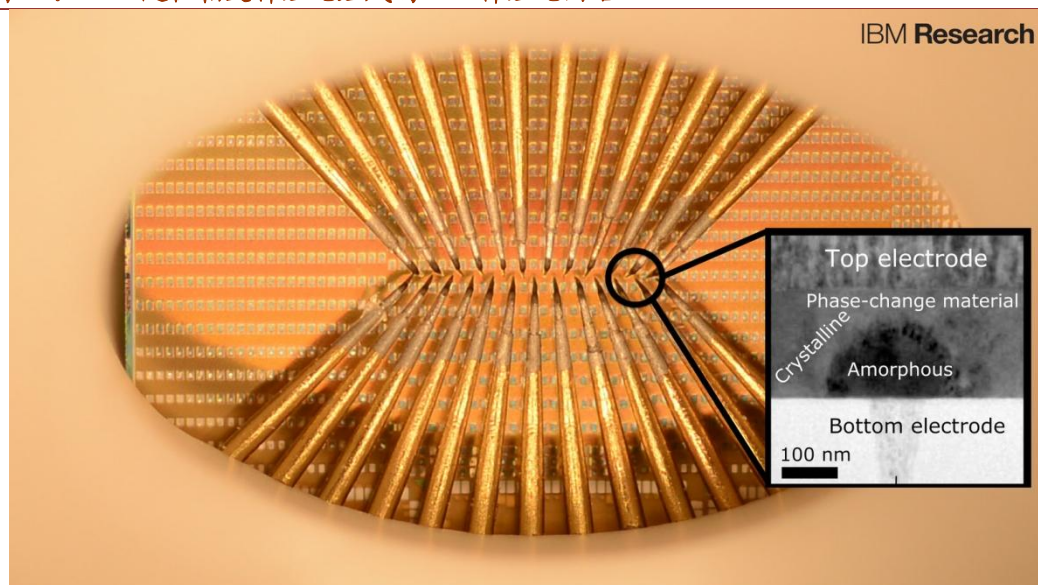
资料来源：公开资料，招商证券

TrueNorth 使用了 54 亿个晶体管,分成 4096 个“神经突触内核(neurosynaptic cores)”的结构;每一个“神经突触内核”结构都能使用 crossbar (交叉) 通讯模式来存储、处理并向其它结构传输数据,这些计算内核产生的效果相当于 100 万个神经元和 2.56 亿个突触。TrueNorth 芯片只要几厘米的方寸,功耗只有 65 毫瓦。

TrueNorth 是 IBM 参与 DARPA 的研究项目 SyNapse 的最新成果; SyNapse 全称是 Systems of Neuromorphic Adaptive Plastic Scalable Electronics (自适应可塑可伸缩电子神经系统,而 SyNapse 正好是突触的意思),其终极目标是开发出打破冯诺依曼体系的硬件

2016 年,IBM 苏黎世研究中心宣布,制造出世界首个人造纳米级随机相变神经元,该神经元能用于制造高密度、低功耗的认知学习芯片,可实现人工智能的高速无监督学习。该神经元尺寸最小能到纳米量级,因而信号传输速度极快,同时功耗较低,这就使得随机相变神经元具有生物神经元的特性。

图 47: IBM 随机相变神经元组成的人工神经网络



资料来源: IBM, 招商证券

IBM 相变神经元的整个架构包括输入端、神经薄膜、信号发生器和输出端,其中输入端类似生物神经元的树突,神经薄膜类似生物神经元的双分子层,信号发生器类似生物神经元的神经细胞主体,输出端类似生物神经元的轴突。而神经薄膜是整个神经元产生作用的关键物质,它类似生物神经细胞中的液态薄膜,当能量吸收到一定程度时就会产生信号并向外发射。这些信号经过输出端(轴突)传导,然后被其他神经元接收,以此循环形成信息处理过程

目前,IBM 已经构建了由 500 个该神经元组成的阵列,并让该阵列以模拟人类大脑的工作方式进行信号处理

对于类脑计算现在基本可以看到两个方向,人工神经网络从功能层面模仿大脑的能力,而神经拟态计算(neuromorphic computing)则是从结构层面去逼近大脑,其结构也有两个层次,一是神经网络层面,与之相应的是神经拟态架构和处理器,二是神经元层面,与之相应的是元器件。IBM 的 TrueNorth、随机相变神经元都是在神经拟态计算方向寻求突破,其中 TrueNorth 在神经拟态架构层面做出努力,刚宣布不久的人工神经元则是在神经元结构层面做出的努力。

#### (4) 英特尔 Knights Mill 芯片：专为机器深度学习设计，可独自充当处理器

2016年8月17日,英特尔数据中心集团执行副总裁戴安·布莱恩特在开发者大会(IDF)上宣布,将在2017年推出专为机器深度学习设计的芯片,Xeon Phi 家族新成员,代号Knights Mill。

Knights Mill 芯片可以独自充当处理器,不再需要单独的主机处理器和辅助处理器,可以直接接入RAM系统;英伟达的GPU和谷歌的TPU芯片,都是辅助处理器,必须和CPU一起工作。

图 48: 英特尔 Xeon Phi 家族成员



资料来源: Intel, 招商证券

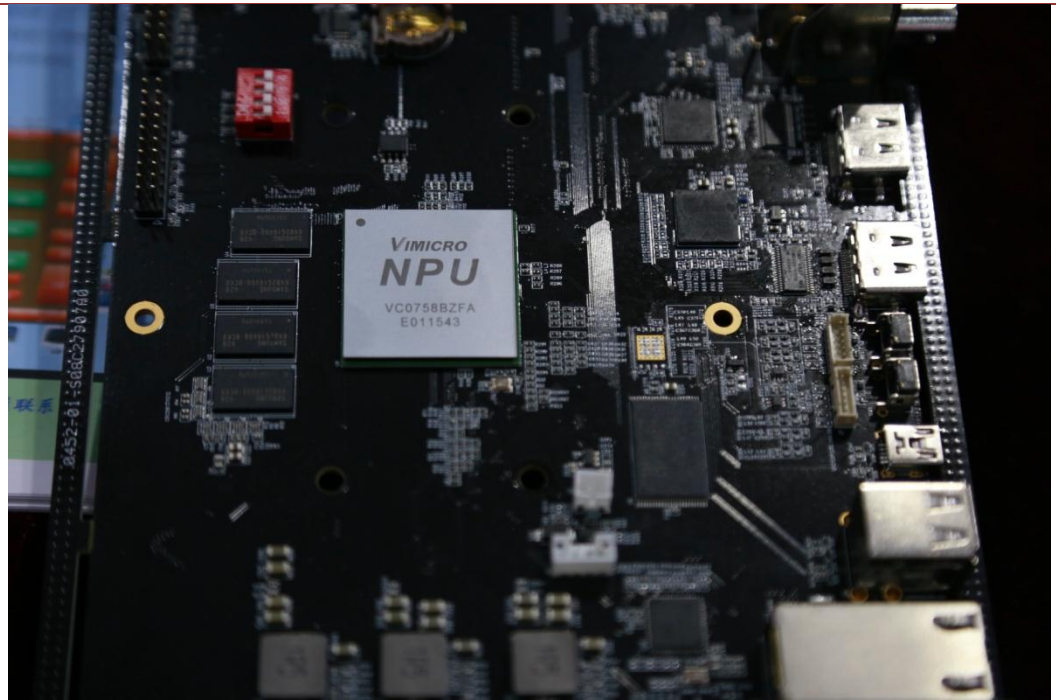
Xeon Phi 是 Intel 针对高性能计算市场推出的加速卡, 主要与 NVIDIA 的 Tesla、AMD 的 FirePro S 等产品竞争, 不过两者是基于 GPU 的, 而 Xeon Phi 是 X86 众核架构的。Xeon Phi 目前已经发展了三代, 第一代 Knights Corner, 22nm 工艺, 最多 61 个核心, 浮点性能 1TFLOPS; 第二代是 Knights Landing, 14nm 工艺, 最多 72 核心, 浮点性能 3+TFLOPS; 英特尔于 2014 年宣布第三代 Knights Hill, 制程工艺升级到 10nm。

#### (5) 中星微电子“星光智能一号”芯片：全球首颗具备深度学习人工智能的嵌入式视频采集压缩编码系统级芯片

中星微在今年 6 月 20 日, 率先推出中国首款嵌入式神经网络处理器芯片, 这是全球首颗具备深度学习人工智能的嵌入式视频采集压缩编码系统级芯片, 并取名“星光智能一号”。这款基于深度学习的芯片运用在人脸识别上, 最高能达到 98% 的准确率, 超过人眼的识别率。该芯片于今年 3 月 6 日实现量产, 目前出货量为十几万件。



图 49: 中星微电子“星光智能一号”芯片和主板



资料来源：公开资料，招商证券

NPU 采用了“数据驱动”并行计算的架构，单颗 NPU(28nm)能耗仅为 400mW，极大地提升了计算能力与功耗的比例，可以广泛应用于高清视频监控、智能驾驶辅助、无人机、机器人等嵌入式机器视觉领域。

中星微的“星光智能一号”与 IBM 人工神经元相比，共同之处是，都是属于受到生物人脑机理的启发而通过半导体电路与器件去实现的，都是采用 CMOS 半导体工艺来生产的。而不同之处在于，所采用的架构和实现方式不同，前者采用卷积型架构，用数据驱动的并行数字电路来实现，后者采用脉冲型架构，用数模混合电路来实现。

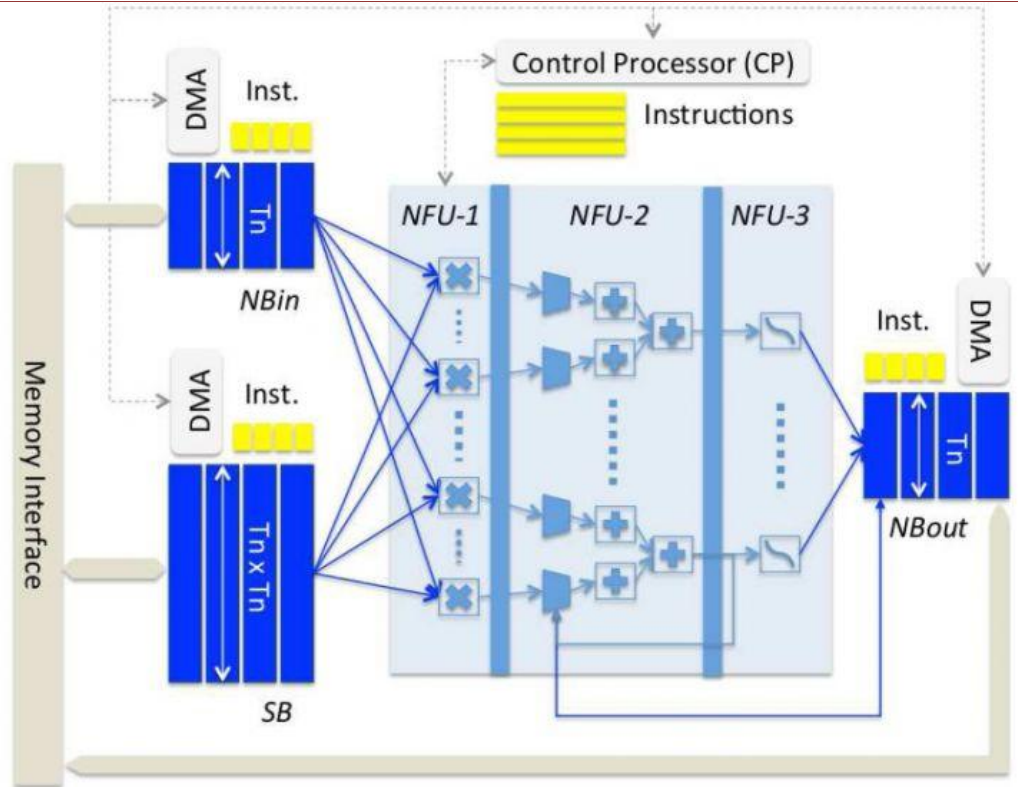
### (6) 寒武纪“DianNao”芯片：世界上第一款深度学习处理器芯片

2016 年 3 月，中国科学院计算技术研究所发布了全球首个能够“深度学习”的“神经网络”处理器芯片，名为“寒武纪”。

寒武纪不是用来代替 CPU 中央处理器的颠覆式革命，更像是一款针对智能认知等应用的专用芯片，优势集中在人脸识别、声音识别等人工智能方面。

目前，寒武纪系列已包含三种原型处理器结构：寒武纪 1 号（英文名 DianNao，面向神经网络的原型处理器结构）；寒武纪 2 号（英文名 DaDianNao，面向大规模神经网络）；寒武纪 3 号（英文名 PuDianNao，面向多种机器学习算法）。DianNao 是寒武纪系列的第一个原型处理器结构，包含一个处理器核，主频为 0.98GHz，峰值性能达每秒 4520 亿次神经网络基本运算，65nm 工艺下功耗为 0.485W，面积 3.02mm<sup>2</sup>。在若干代表性神经网络上的实验结果表明，DianNao 的平均性能超过主流 CPU 核的 100 倍，但是面积和功耗仅为 1/10，效能提升可达三个数量级；DianNao 的平均性能与主流 GPGPU 相当，但面积和功耗仅为主流 GPGPU 百分之一量级。

图 50: DianNao 结构



资料来源：寒武纪，招商证券

要降低处理器功耗，仅仅降低运算功耗是不够的，必须优化片上数据搬运。中科院计算所提出对神经网络进行分块处理，将不同类型的数据块存放在不同的片上 RAM 中，并建立理论模型来刻画 RAM 与 RAM、RAM 与运算部件、RAM 与内存之间搬运次数，进而优化神经网络运算所需的数据搬运次数。相对于 CPU/GPU 上基于 cache 层次的数据搬运，DianNao 可将数据搬运减少 10~30 倍。

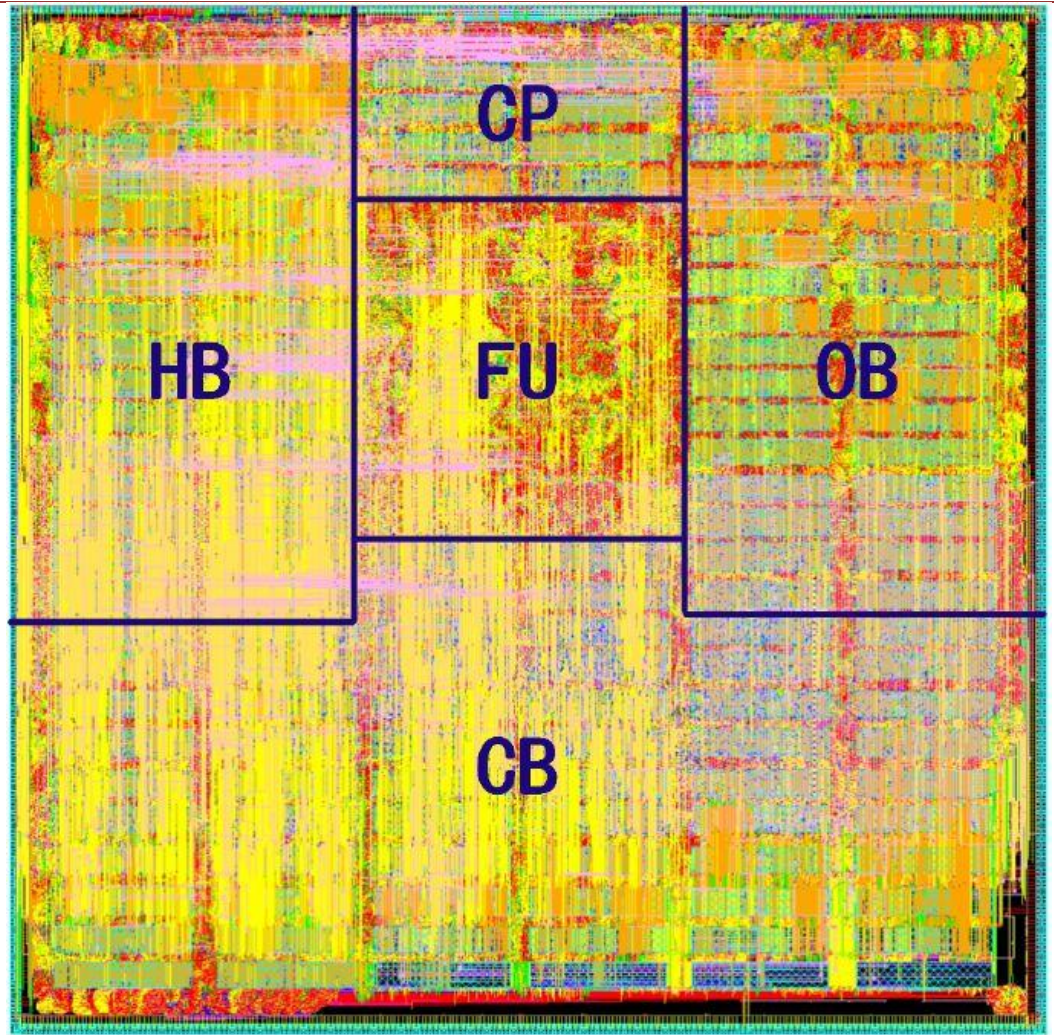
DaDianNao 在 DianNao 的基础上进一步扩大了处理器的规模，包含 16 个处理器核和更大的片上存储，并支持多处理器芯片间直接高速互连，避免了高昂的内存访问开销。在 28nm 工艺下，DaDianNao 的主频为 606MHz，面积 67.7 mm<sup>2</sup>，功耗约 16W。单芯片性能超过了主流 GPU 的 21 倍，而能耗仅为主流 GPU 的 1/330。64 芯片组成的高效能计算系统较主流 GPU 的性能提升甚至可达 450 倍，但总能耗仅为 1/150。

虽然神经网络已成为模式识别等领域的主流算法，但用户很多时候可能倾向于使用其他一些经典的机器学习算法。例如程序化交易中经常使用线性回归这类可解释性好、复杂度低的算法。在此背景下，寒武纪 3 号多用途机器学习处理器 PuDianNao 应运而生，当前已可支持 k-最近邻、k-均值、朴素贝叶斯、线性回归、支持向量机、决策树、神经网络等近十种代表性机器学习算法。

PuDianNao 的主频为 1GHz，峰值性能达每秒 10560 亿次基本操作，面积 3.51mm<sup>2</sup>，功耗为 0.596W（65nm 工艺下）。PuDianNao 运行上述机器学习算法时的平均性能与主流 GPGPU 相当，但面积和功耗仅为主流 GPGPU 百分之一量级。



图 51: PuDianNao 结构



资料来源：寒武纪，招商证券

### 5、芯片是人工智能时代的战略制高点

从传统芯片巨头到互联网巨头，以及一些新锐企业，都在大力发展人工智能芯片。因为芯片作为人工智能的上游产业将走在行业发展前沿，未来几乎所有的智能化应用场景，都将离不开人工智能芯片和软件。可以说，**核心芯片是人工智能时代的战略制高点**。

表 4: 人工智能芯片对比

芯片名称	生产厂商	芯片特点
TPU (Tensor Processing Unit)	谷歌	1) TPU 是一种专用的加速器芯片，跟其深度学习软件 Tensor Flow 匹配 2) TPU 专门针对机器学习进行过裁减，运行单个操作时需要的晶体管更少 3) TPU 意在取代 GPU

<p>Tesla P100</p>	<p>Nvidia</p>	<p>1) Tesla P100 芯片安装了 150 亿个晶体管，是目前市场上许多处理器、图形芯片的 3 倍 2) 芯片面积为 600 平方毫米，双精度运算速度 5.3 万亿次，单精度运算速度 10.6 万亿次，半精度运算速度 21.2 万亿次</p>
<p>TrueNorth</p>	<p>IBM</p>	<p>1) TrueNorth 使用了 54 亿个晶体管，分成 4096 个“神经突触内核 (neurosynaptic cores)”的结构 2) TrueNorth 芯片只要几厘米的方寸，功耗只有 65 毫瓦</p>
<p>随机相变神经元芯片</p>	<p>IBM</p>	<p>1) 芯片的神经元尺寸最小能到纳米量级，因而信号传输速度极快，同时功耗较低，这就使得随机相变神经元具有生物神经元的特性</p>
<p>Knights Mill</p>	<p>Intel</p>	<p>1) Xeon Phi 家族新成员，专为机器深度学习设计的芯片 2) 英伟达的 GPU 和谷歌的 TPU 芯片，都是辅助处理器，必须和 CPU 一起工作；Knights Mill 芯片可以独自充当处理器，不再需要单独的主机处理器和辅助处理器，可以直接接入 RAM 系统</p>
<p>星光智能一号</p>	<p>中星微电子</p>	<p>1) 中国首款嵌入式神经网络处理器芯片，全球首颗具备深度学习人工智能的嵌入式视频采集压缩编码系统级芯片 2) 基于该芯片的人脸识别，最高能达到 98% 的准确率，超过人眼的识别率</p>
<p>DianNao</p>	<p>寒武纪</p>	<p>1) 向神经网络的原型处理器结构 2) 包含一个处理器核，主频为 0.98GHz，峰值性能达每秒 4520 亿次神经网络基本运算，65nm 工艺下功耗为 0.485W，面积 3.02mm<sup>2</sup> 3) 实验结果表明，DianNao 的平均性能超过主流 CPU 核的 100 倍，但是面积和功耗仅为 1/10，效能提升可达三个数量级 4) DianNao 的平均性能与主流 GPGPU 相当，但面积和功耗仅为主流 GPGPU 百分之一量级</p>

资料来源：招商证券整理

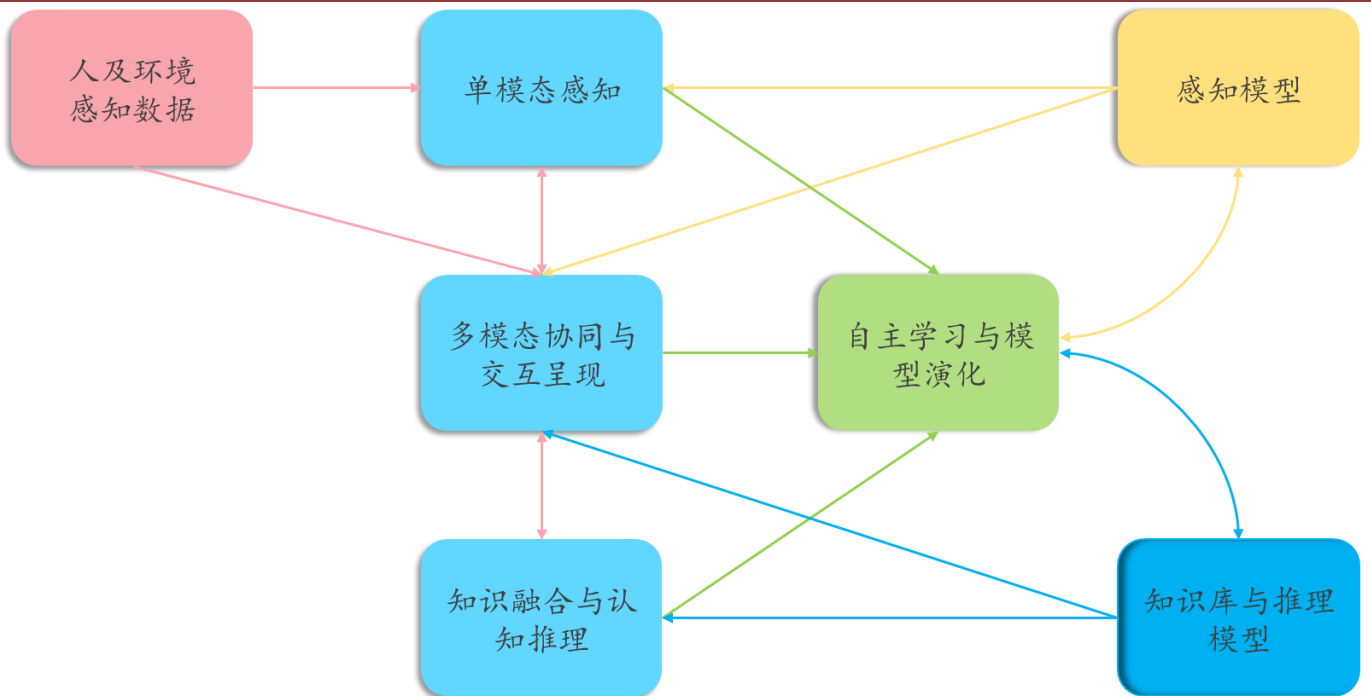
## 四、人工智能应用：“人工智能+”时代呼啸而来

### 1、“人工智能+”时代呼啸而来

人工智能就像一列火车，它临近时你听到了轰隆隆的声音，你在不断期待着它的到来。它终于到了，一闪而过，随后便远远地把你抛在身后。

专用人工智能在某种程度上是通用人工智能的基石。无论是多任务学习、迁移学习、在线学习、增强学习，都是专用人工智能向通用人工智能的有益尝试。专用人工智能与通用人工智能之间没有明确的界限，人工智能算法的专用性是相对的。目前深度学习能够解决的任务越来越多，其模型和学习算法越来越趋同，Deepmind 的强化学习能够学习多种多样的游戏，都可以认为是专用人工智能模式的通用化尝试。

图 52: 通用人工智能和专用人工智能的关系



资料来源：招商证券整理

当前，面向特定领域的人工智能，即**专用人工智能**，由于应用背景需求明确、领域知识积累深厚、建模计算简单可行，在局部智能水平的单项测试中可以超越人类智能，目前已经在多领域取得成功应用。2011 年，IBM 超级电脑 Watson 在美国智力竞赛节目 jeopardy 上战胜了人类冠军；2016 年，AlphaGo 以 4: 1 战胜围棋世界冠军李世石；ImageNet 图像识别精度屡创新高，目前领跑者微软亚洲研究院的计算机视觉系统已经超越人类进行对象识别分类的能力。



图 53: 人工智能应用分类



资料来源: Venture Scanner, 招商证券

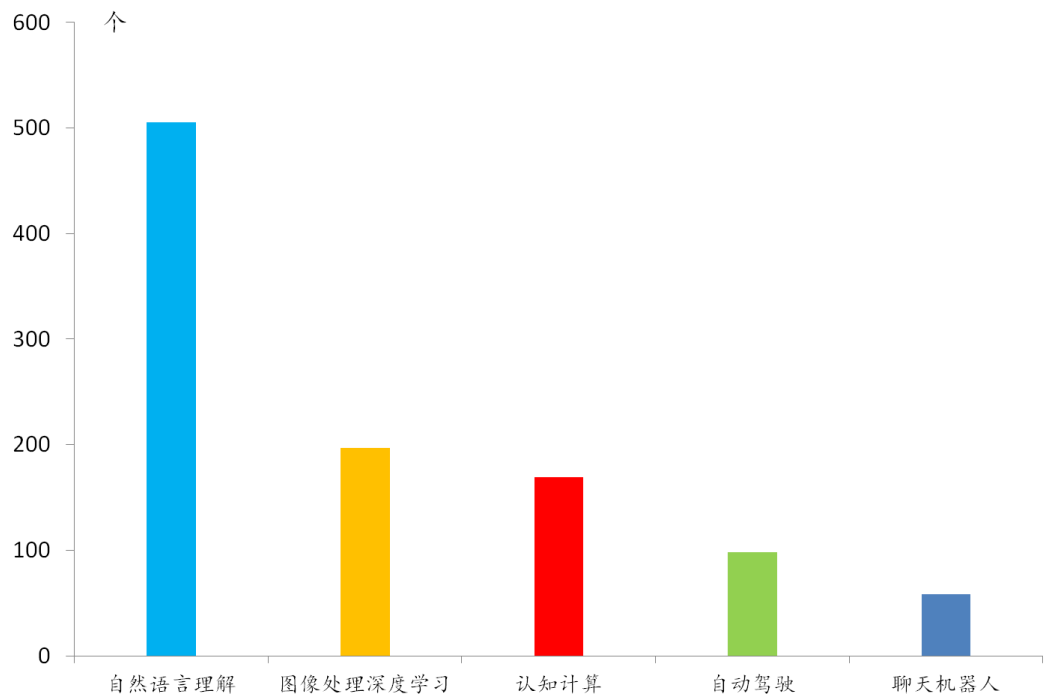
目前人工智能领域的公司根据行业共有 13 种类型, 包括致力于特殊领域的专用人工智能公司以及从事通用人工智能技术研发的公司, 分别为: 深度学习/机器学习 (通用)、深度学习/机器学习 (应用)、自然语言处理 (通用)、自然语言处理 (语音识别)、计算机视觉/图像识别 (通用)、计算机视觉/图像识别 (应用)、手势控制、虚拟私人助手、智能机器人、推荐引擎和协助过滤算法、情景感知计算、语音翻译、视频内容自动识别。

**深度学习/机器学习 (通用):** 这类公司主要建立可依靠现存数据进行学习的算法。典型例子包括预测数据模型与分析行为数据的软件平台。

**深度学习/机器学习 (应用):** 这类公司同样使用计算机算法, 但却是基于非常垂直的特殊案例中存在的数据库运行。典型例子有利用机器学习技术侦查金融诈骗或者识别最好的销售线索。

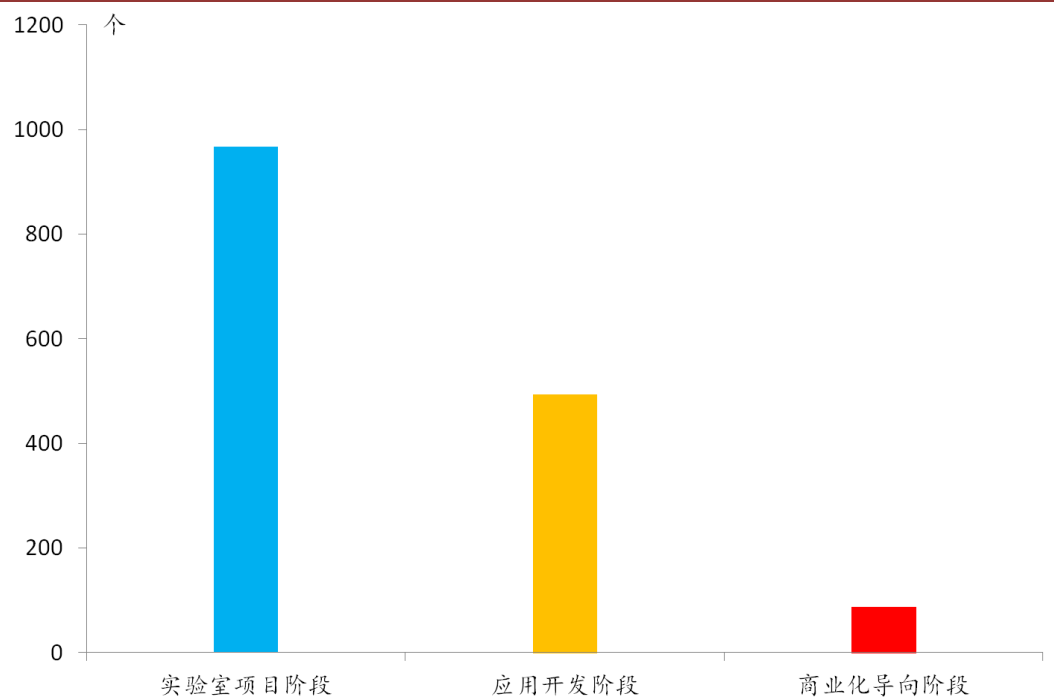
**自然语言处理 (通用):** 此类公司构建的算法能够处理输入的自然语言, 并将其转化为可理解的表达。例子包括文本自动生成以及文本挖掘生成数据。

图 54: 人工智能项目按应用方向分类



资料来源: SPIDERBOOK, 招商证券

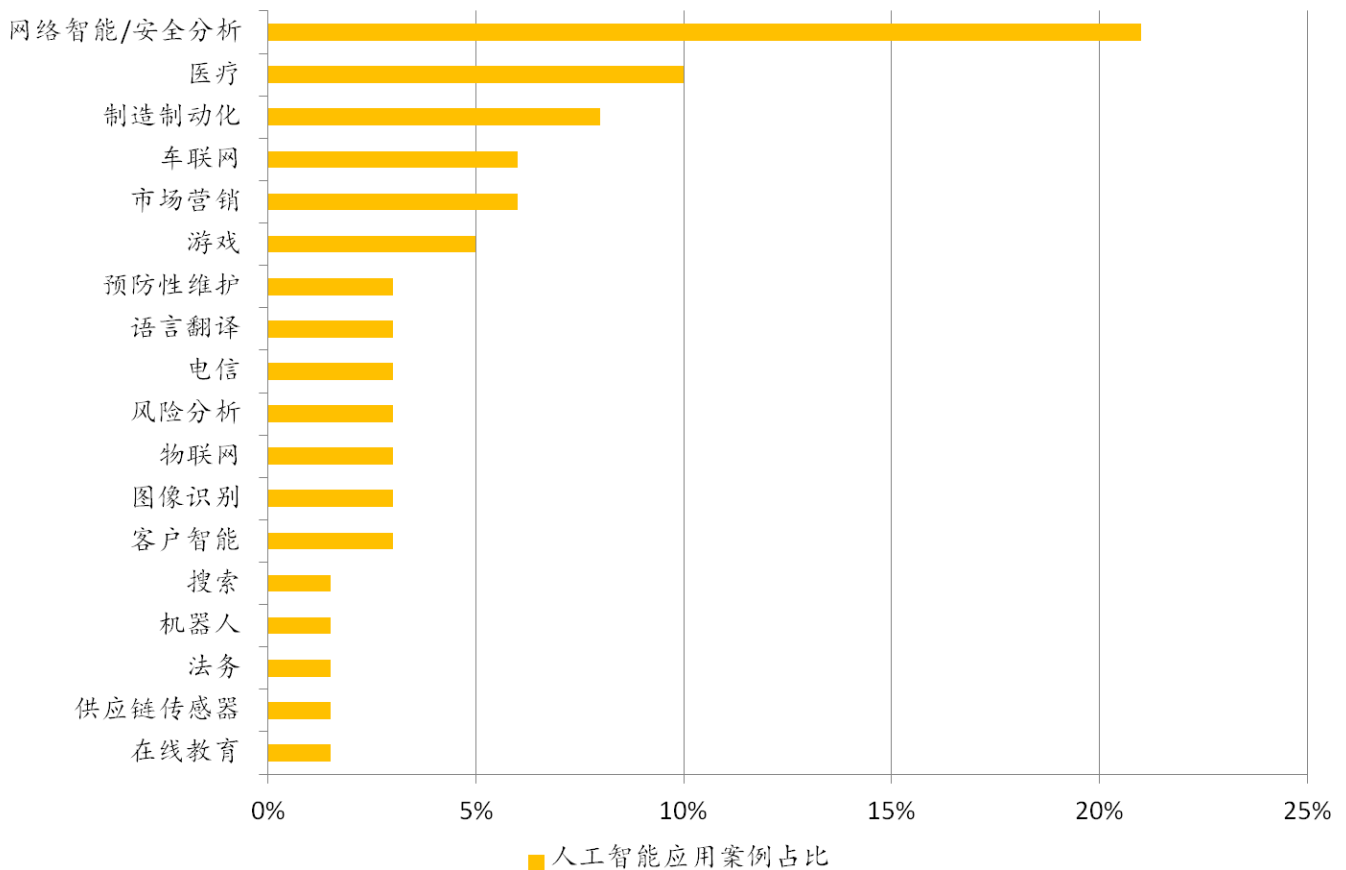
图 55: 人工智能项目按成熟度分类



资料来源: SPIDERBOOK, 招商证券



图 56: 美国人工智能应用案例占比



资料来源: SPIDERBOOK, 招商证券

**自然语言处理 (语音识别):** 公司产品能够处理人类语音的片段, 准确识别单词并推测含义。典型的例子是语音指令的检测并将其转化为可执行的数据。

**计算机视觉/图像识别 (通用):** 这类公司研发的技术主要是图像处理、分析, 可从中提取信息、识别图像中的物体。典型例子包括图像搜索平台和研发员使用的图像标签应用程序接口。

**计算机视觉/图像识别 (应用):** 这类公司是在非常垂直的案例中使用图像处理技术。典型案例包括面部识别软件和能让用户通过拍照搜索商品的软件。

**手势控制:** 公司产品可让用户通过手势与计算机互动或交流。典型例子包括让人们通过肢体动作控制游戏角色的软件以及仅用手势就能控制计算机和电视的软件。

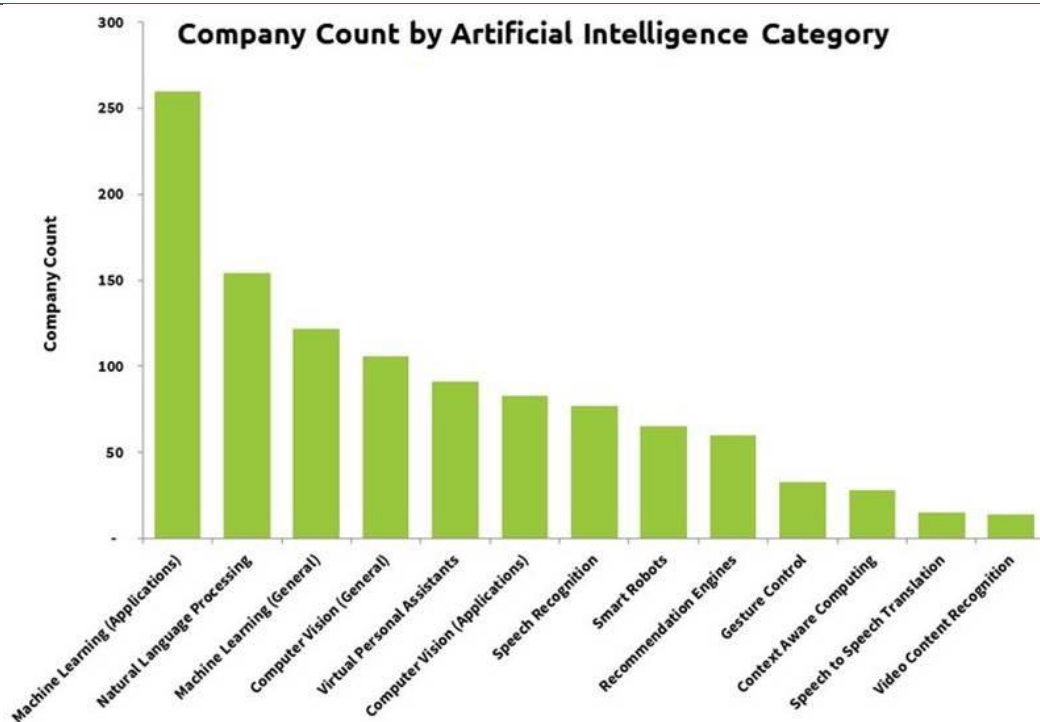
**虚拟私人助手:** 这是一类基于反馈和指令来为个体完成日常任务和服务的软件助理。典型例子有网络客服助理和个人助理 app, 管理个人日程安排等。

**智能机器人:** 可以进行经验学习并根据身边环境自主进行活动的机器人。典型例子有家庭机器人, 可以在互动中根据情感做出反应, 还有帮助人们找到商品的销售机器人。

**推荐引擎和协助过滤算法:** 软件能够预测用户对电影、餐厅等的偏好, 并推荐个性化的内容。典型例子有音乐推荐 app 和基于用户过去选择进行推送的美食推荐网站。

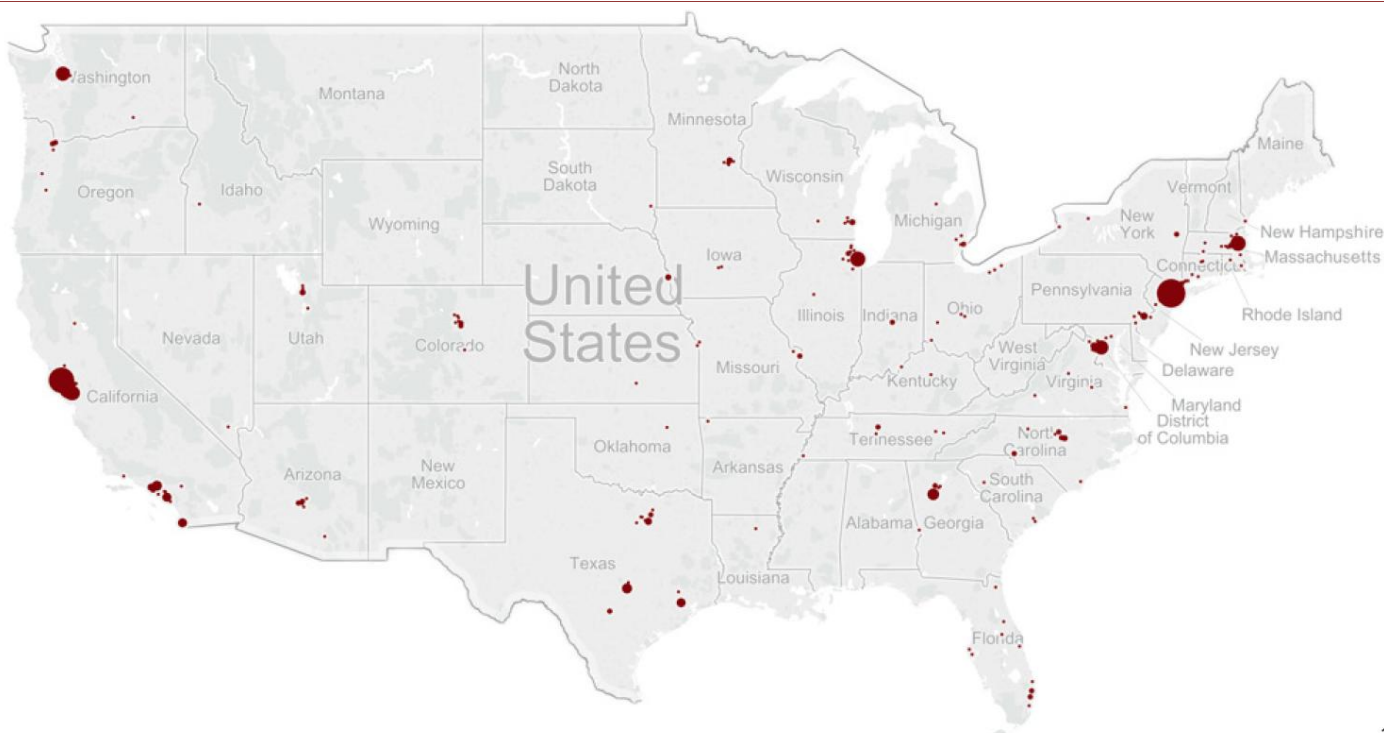
**情境感知计算:** 软件能够自动感知周围环境以及使用背景, 例如位置、方向、光度, 并以此调整行为。典型例子包括感知环境的黑暗度并调高亮度的应用。

图 57: 各类型人工智能公司数量



资料来源: Venture Scanner, 招商证券

图 58: 美国人工智能公司的地理分布



资料来源: Big Data Market, 招商证券

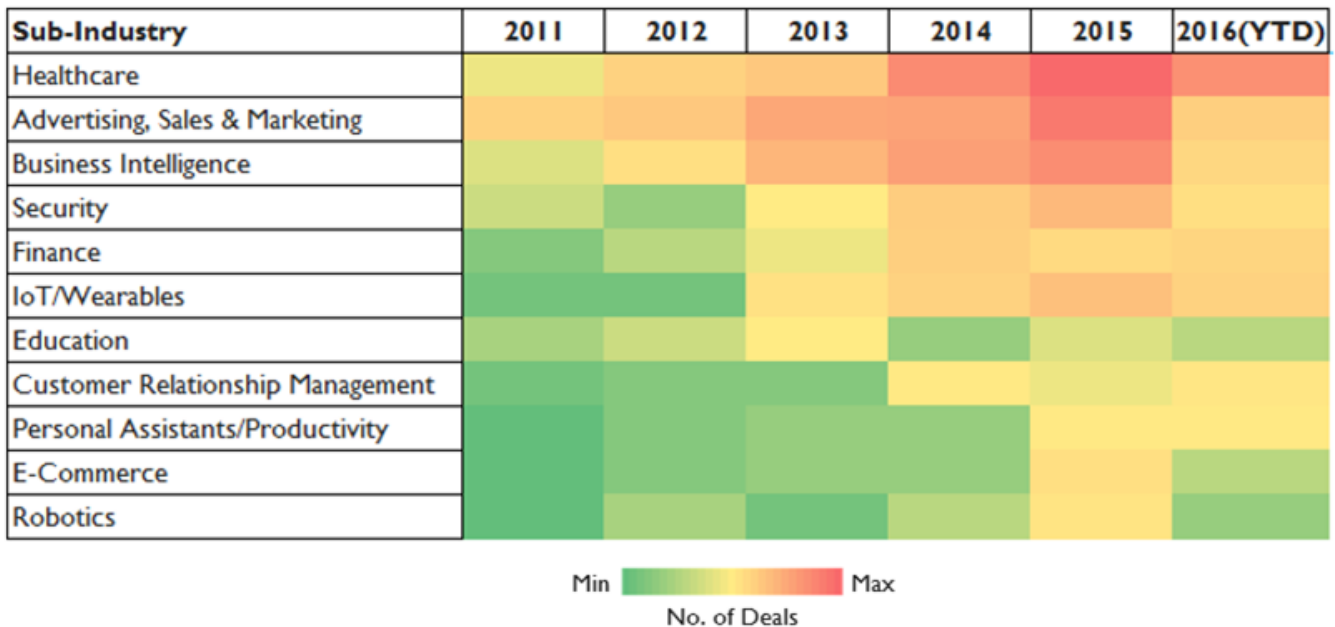
**语音翻译:** 识别人类语音并立刻自动从一种语言翻译至另一种语言的软件。典型例子是自动以及实时的将视频谈话或网络研讨会翻译为多种语言的软件。

**视频内容自动识别:** 这类软件可以将视频内容的一个样本与源内容文件相比较,通过它独特的特点识别内容。典型例子有对用户上传的视频与版权视频文件比较以侦测是否侵权的软件。

截至 2016 年第一季度,全球市场上,机器学习(应用)类人工智能公司数量最多,共有 263 家,而自然语言处理领域的人工智能公司数量次之,达到 154 家。

图 59: 人工智能细分应用领域投资热度分布

### Artificial Intelligence: Sub-Industry Heatmap 2011-2016 (as of 6/15/2016)



资料来源: CB Insights, 招商证券

在不远的将来,人工智能将在哪些领域有哪些施展拳脚的空间? 答案是人工智能会给几乎各行各业带来变革与重构,我们在此做一梳理,最先可能涉及的领域包括但不限于:安防、制造业、服务业、金融、教育、传媒、法律、医疗、家居、农业、汽车等。这些重构的领域之多、力度之大将是前所未有的,在这背后蕴藏的是人工智能可无限想象的市场空间。我们将选取安防、制造业、金融、无人驾驶、法律、智能家居、医疗这 7 个相对更具应用典型性的行业展开详细论述。

预计 2020 年,全球智能机器人市场规模约 830 亿美元,人工智能市场规模约 700 亿美元。预计 2025 年,工业机器人全球装机量将达到 1500 万-2500 万台,年均增速为 25%-30%,每年将产生 0.6 万-1.2 万亿美元的经济影响;医疗类、增强人体机能类和家用类等服务机器人每年将产生 1.1 万-3.3 万亿美元的经济影响。

表 5: 人工智能可能的重构的领域与方式

行业	智能重构方式
安防	<b>生物识别技术:</b> 通过计算机与光学、声学、生物传感器和生物统计学原理等手段结合, 利用人体固有的生理特性, 和行为特征来进行个人身份验证。可帮助进行安全监控和反恐等。
制造业	<b>智能生产:</b> 打通所有生产环节的数据, 实现机器与机器对话, 更好的预测产品需求并调整产能。代替部分人工, 有效提高生产效率并降低生产成本。
服务业	<b>机器服务:</b> 利用机器代替各类服务人员, 提高服务质量, 降低人员成本。
金融	<b>投资决策辅助:</b> 帮助金融从业人员迅速找到自己想要的信息。或利用大数据引擎技术、自然语义分析技术等准确分析与预测各市场的行情走向。提升投资决策效率。 <b>信用风险管控:</b> 投资项目的风险分析和决策、个人征信评级、信用卡管理。 <b>智能支付:</b> 利用图像识别或语音识别, 实现“刷脸支付”或“语音支付”。
教育	<b>个性化定制学习:</b> 利用人工智能实现针对每个学习者特质和需求的个性化定制学习。 <b>学习数据分析:</b> 基于已有的数据, 准确判断学生的知识掌握水平。
传媒	<b>智能用户分析:</b> 将人工智能用于精准营销, 判定用户最感兴趣的广告。 <b>信息快速编写:</b> 通过人工智能技术中的自然语言技术, 来进行稿件自动编写。
法律	<b>案件判决预测与分析:</b> 为法官、律师、法学教师、立法者或公民提供案件的判决预测与分析 <b>法律服务自动化:</b> 机器学习可以帮助研发合同模板系统、法律文书系统、风险防范系统及法律流程系统。
医疗	<b>健康诊断:</b> 形成诊断专家系统, 建立医学搜索引擎, 帮助实现医疗诊断及健康管理。 <b>新药开发:</b> 通过分析数据, 提高新药开发效率。
家居	<b>智能家居系统:</b> 提升家居安全性、便利性、舒适性, 并实现环保节能的居住环境。
农业	<b>农业数据分析:</b> 对土地营养情况、水利情况进行精细实时的跟踪与分析, 提供预测数据和决策建议。 <b>机器人技术:</b> 利用机器人技术代替传统的农民耕种, 实现效率更高的农业生产。
汽车	<b>自动驾驶:</b> 通过各类技术对周边环境进行探测, 根据所得数据进行自我学习并实现自动驾驶。

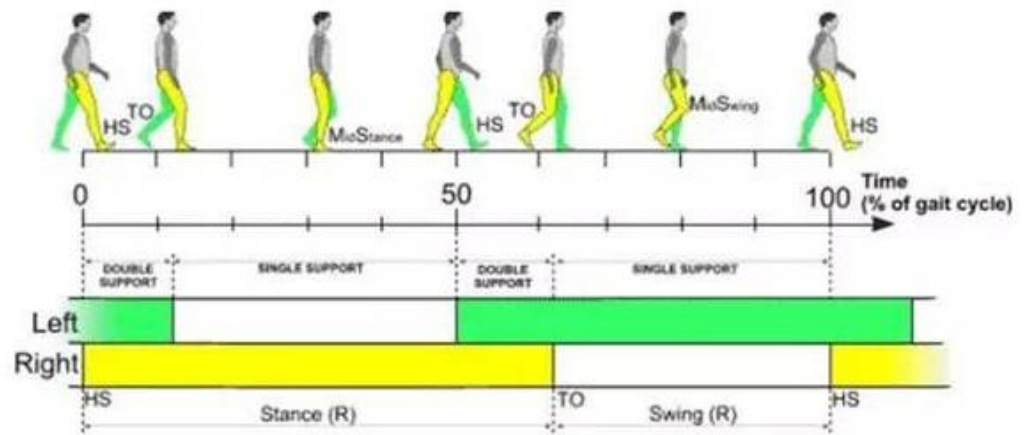
资料来源: 招商证券

## 2、安防：聚焦生物识别技术

**身份识别手段的多样性对于安防意义重大：**虽然现在的人脸识别率和语音识别率已经很高，但其前提是被验证身份的人愿意把脸或者声音暴露给你，而在很多犯罪案件甚至恐怖袭击事件中，犯罪分子会采取各种手段避免暴露自己的面部特征和声音。因此安防领域对于图像识别的要求更高，也要求更多的手段通过多维度来进行识别。**生物识别技术**是近年兴起的一种通过计算机与光学、声学、生物传感器和生物统计学原理等手段结合，利用人体固有的生理特性，和行为特征来进行个人身份验证的一种技术。指纹识别便是一种生物识别技术的应用，但观其未来的应用空间，远远不止指纹识别这一种方法。由于一种生物识别技术建立的前期需要对人的生理特征进行大量的采样学习，在这一阶段人工智能将发挥重要作用，因此人工智能技术的进步可以大大提高身份识别手段的多样性与准确率，对于安防的意义重大。

**碟中谍里展示的生物识别技术：**最近一部碟中谍电影里曾经展示过一个身份识别系统：人走过一条扫描走廊，通过扫描全身和人走路时的步态姿势来进行身份识别。这个看似玄而又玄的技术叫做步态识别技术，也属于生物识别技术的一个分支，是当今生物识别的一个热门研究领域。根据前瞻产业研究院的预测，2020年，全球生物识别技术的市场规模将达到250亿美元的水平。指纹识别占生物识别技术的份额最高，2013年占比60%，但占比将呈下降趋势，在未来越来越多的新颖生物识别技术将占比越来越高。

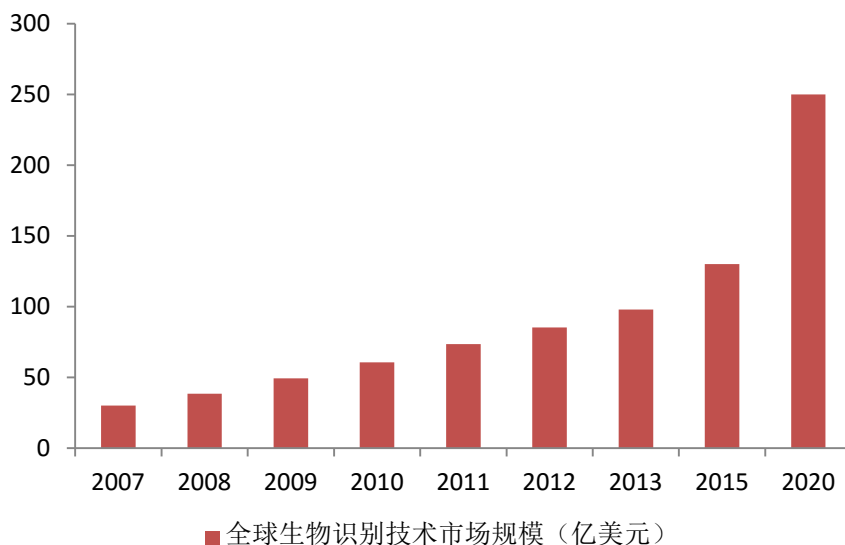
图 60：步态识别技术



资料来源：大数据文摘、招商证券



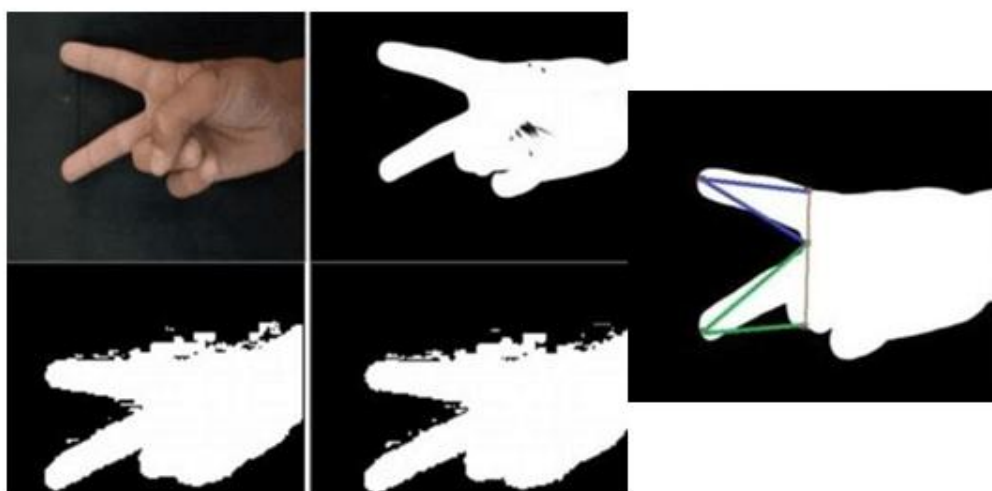
图 61: 2007-2020 年全球生物识别技术市场规模 (亿美元)



资料来源: 前瞻产业研究院、招商证券

**应用于反恐, 具有国家战略意义:** 恐怖袭击已经成为了全球性问题, 人工智能的应用可以帮助实现一些单靠人类不能实现的侦查手段, 对于国家来说具有战略意义。据 2016 年 3 月《每日邮报》的报道, 约旦的研究人员研发了一款基于机器学习技术通过手势识别恐怖分子的系统, 由于现在网络上恐怖分子会经常将行凶视频放到网络上, 视频中的恐怖分子往往通过蒙面、变声等手段掩饰自己的身份, 而这款系统通过分析恐怖分子经常比出的“V”字手势, 来确定其身份。这其实也是生物识别技术的一种应用。此外, 人工智能除了在生物识别技术上的应用外, 值得一提的还有美国赫赫有名的独角兽公司 Palantir, 有消息称其国家安防产品 Palantir Gotham 在美国政府 10 年追捕本·拉登的行动中发挥了情报分析的作用, 其产品主要致力于通过建立起政府手中掌握的海量财务信息、基因样本、图像影像片段等信息之间的联系、并借此寻找关键线索, 提前掌握恐怖份子可能发动袭击的消息。

图 62: 通过 V 字手进行恐怖分子身份识别

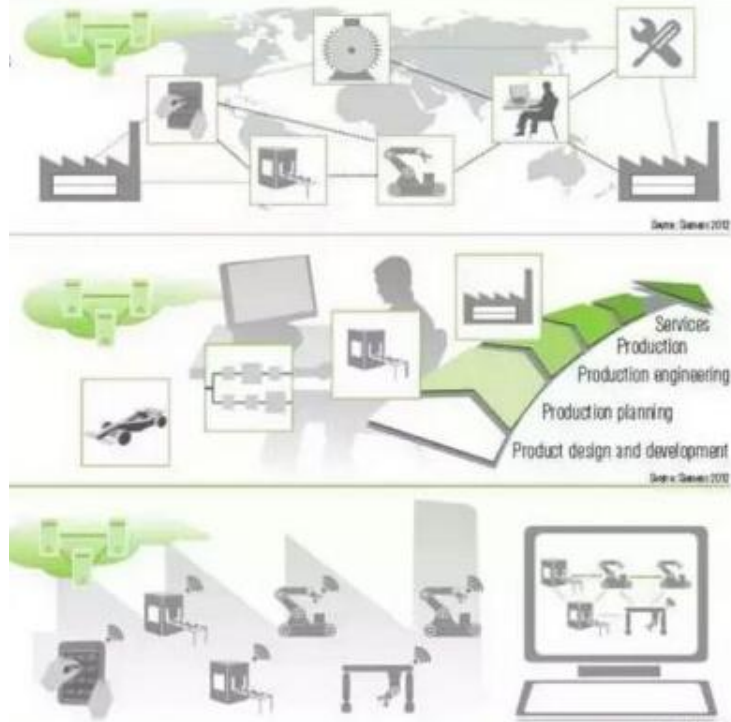


资料来源: Mu'tah University、招商证券

### 3、制造业：人工智能有望推动半自动化生产到全自动化生产的飞跃

**人工智能将催生全自动化智能生产：**经过几十年的发展，制造业已经实现了一定程度的自动化生产，但大多还只能归为半自动化生产的程度，无法完全脱离人工的参与。但人工智能技术的进步将为制造业的进一步自动化以及提升生产效率和准确性做出贡献，在未来有望彻底取代人工环节，实现半自动化智能工厂到全自动化智能工厂的质的飞跃。

图 63：人工智能将催生全自动化智能生产工厂



资料来源：知社学术圈、招商证券

**西门子安贝格工厂，最接近全自动化的智能工厂：**西门子安贝格电子制造工厂是被认为最接近工业 4.0 概念雏形的工厂。虽然早在 1989 年便已设立，但通过不断升级改造，截止 2015 年该工厂的产能较 1989 年提升了 8 倍，合格率提高了 40 倍。安贝格工厂超过 3 亿个元器件都有自己的“身份证”，每当一个元件进入烤箱时，机器会自动判断并使用相应的烘烤温度、烘烤时长，并判断哪一种元件应该在下一个进入烘箱，实现动态调节生产参数。值得一提的是，安贝格工厂自动化率已达近 75%，只有生产过程的开头部分需要人工参与，产品可与生产设备通信，然后通过 IT 系统集中控制和调节所有生产流程，使产品合格率达到 99.9988%。

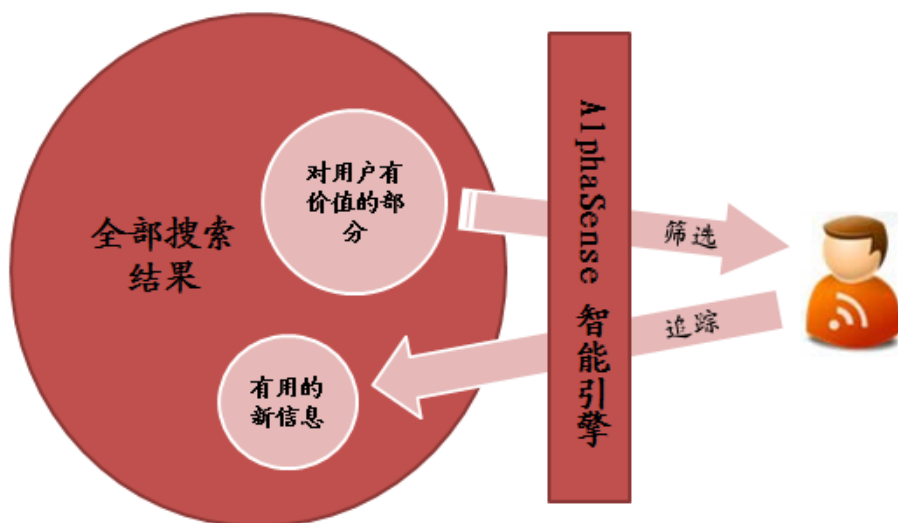
### 4、金融：关注投资决策辅助、信用风控、与智能支付

人工智能技术在金融领域的应用具有十分巨大的想象空间，并且很多应用还处在起步阶段，无论从技术角度还是普及程度看，都有很大的提升空间。可能的应用包括但不限于：

➤ **投资决策辅助，实现信息的智能筛选与处理：**核心价值在于帮助金融从业人员迅速

找到自己想要的信息，以及利用大数据引擎技术、自然语义分析技术等自动准确地分析与预测各市场的行情走向。典型的案例是美国的 AlphaSense 公司，它是一款针对金融从业人员的智能搜索引擎，可以帮他们排除不相关的谷歌搜索结果。AlphaSense 搜索数据库涵盖了自有客户的内部数据，同时还有数以千计的外部来源，包括超过 1,000 家卖方研究机构和超过 35,000 家上市公司，如券商研究，SEC 提交的文件和新闻稿特许财务数据。AlphaSense 可以利用自然语言处理和搜索技术，简化寻找和追踪最相关的信息。同时可以帮助用户快速查找和发现关键数据点，跟踪与智能提醒有用的新信息，可以有效提升决策效率。

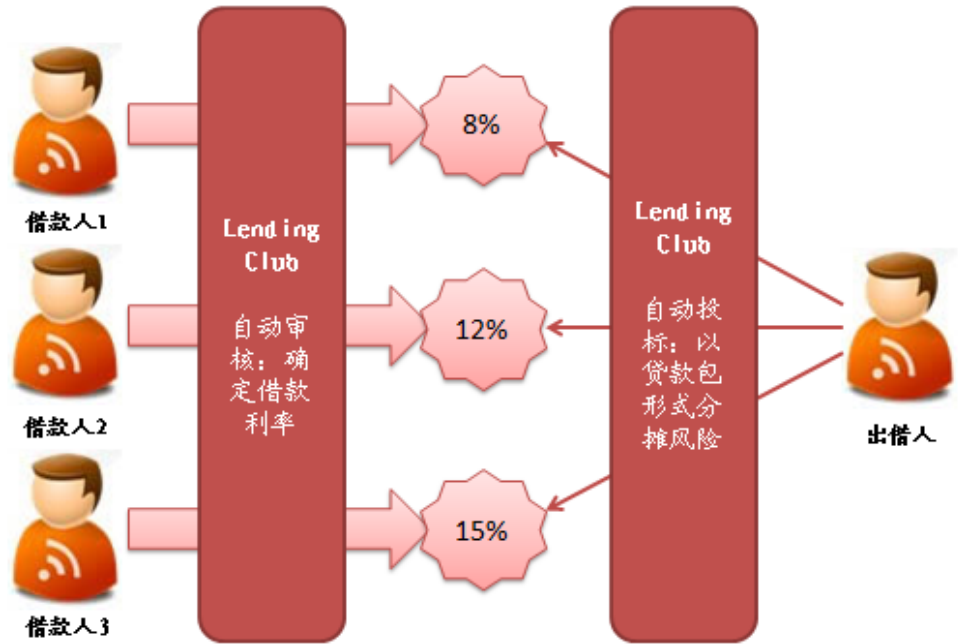
图 64: AlphaSense 智能搜索帮助提高投资决策效率



资料来源：招商证券

- **信用风险管控：**人工智能也能帮助建立金融风控平台，有助于投资项目的风险分析和决策、个人征信评级、信用卡管理。如 Lending Club 作为全世界著名的 P2P 网贷上市公司，在审核借款人信用等级时可以做到自动决策，如果借款人通过了审核，LC 会生成给出批贷或者不批贷的借款人，并且自动确定借款人的借款利率，然后借款人可以通过登录网站选择接受或者不接受贷款条件，在信用评级这一块，Lending Club 还运用了 Facebook 来挖掘用户的日常行为信息作为信用评级的考评手段之一。同时，在出借人端，Lending Club 还具有自动投标功能，用户只需提前设定好其希望投资的标的利率，然后平台自动将其资金投资到与用户设定的利率相匹配的“贷款包”中，以分散风险。

图 65: Lending Club 的智能风控模式



资料来源: 招商证券

- **智能支付:** 利用图像识别或语音识别, 实现“刷脸支付”或者“语音支付”。2015年3月, 马云在 CeBIT 开幕式上演示了一项叫做 Smile to Pay 的扫脸技术, 在购物的最后一个阶段扫描人脸即可支付, 识别率超过 99%, 这项技术是由蚂蚁金服与 Face++ Financial 合作研发的。2015年4月, 百度钱包在全球移动互联网大会上展示了语音支付功能, 在支付环节, 手机上会出现一串随机验证码, 用户用语音念出这串验证码后即可支付。

## 5、无人驾驶: 技术突破、立法加快, 人工智能决定可靠性

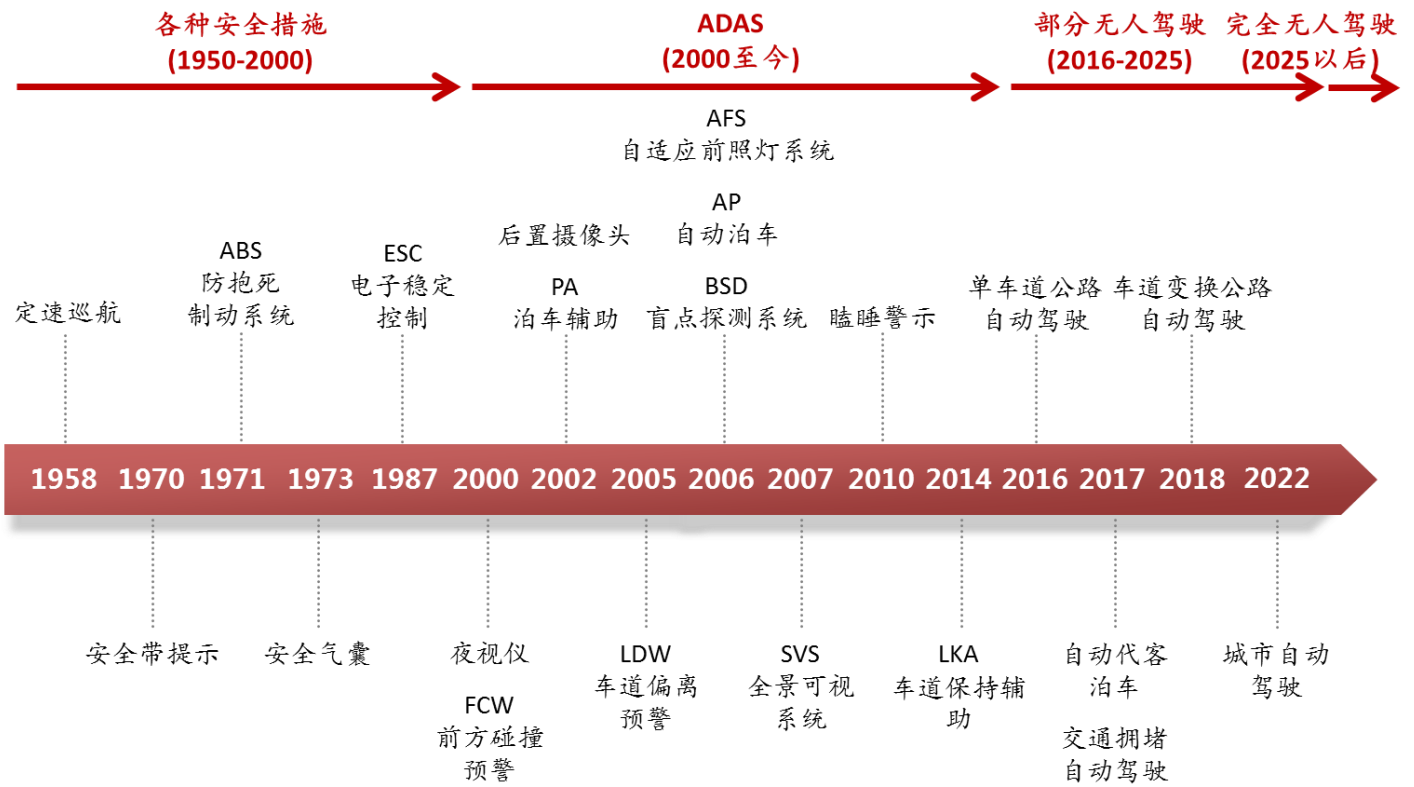
无人驾驶技术最初的发展从减少、防止机动车事故开始。

从 1950 年至 2000 年, OEM 车厂和 Tier 1 的零部件制造商对车辆发明、生产了许多结构性的改进。最核心的就是四大安全系统: 安全带, 防抱死制动系统 (ABS), 安全气囊和电子稳定控制 (ESC)。

从 2000 年至今, 汽车行业推出了一系列高级驾驶辅助系统 (Advanced Driver Assistant System, ADAS) 功能。从 2000 年的夜视仪、前方碰撞预警 (Forward collision warning, FCW), 到 2002 年的后置摄像头、泊车辅助 (Park assist, PA), 到 2005 年的车道偏离预警 (Lane departure warning, LDW), 到 2006 年的自适应前照灯系统 (Adaptive front lights, AFS)、自动泊车 (Automatic parking, AP)、盲点探测系统 (Blind spot detection, BSD), 到 2007 年的全景可视系统 (Surround view systems, SVS), 到 2008 年的前方碰撞辅助 (Forward collision assist, FCA), 到 2010 年的瞌睡警示 Drowsiness alert, 到 2014 年的车道保持辅助 (Lane departure assist, LKA)。通过一系列的 ADAS 功能进一步地减少、防止机动车事故。

2016 年以后，部分无人驾驶的功能如单车道自动驾驶、交通拥堵环境下的自动驾驶、车道变化自动驾驶、城市自动驾驶等等有望逐步实现。再进一步过多到完全无人驾驶。

图 66: 无人驾驶技术发展历程的四个阶段



资料来源: Google, IHS, BCG, 招商证券

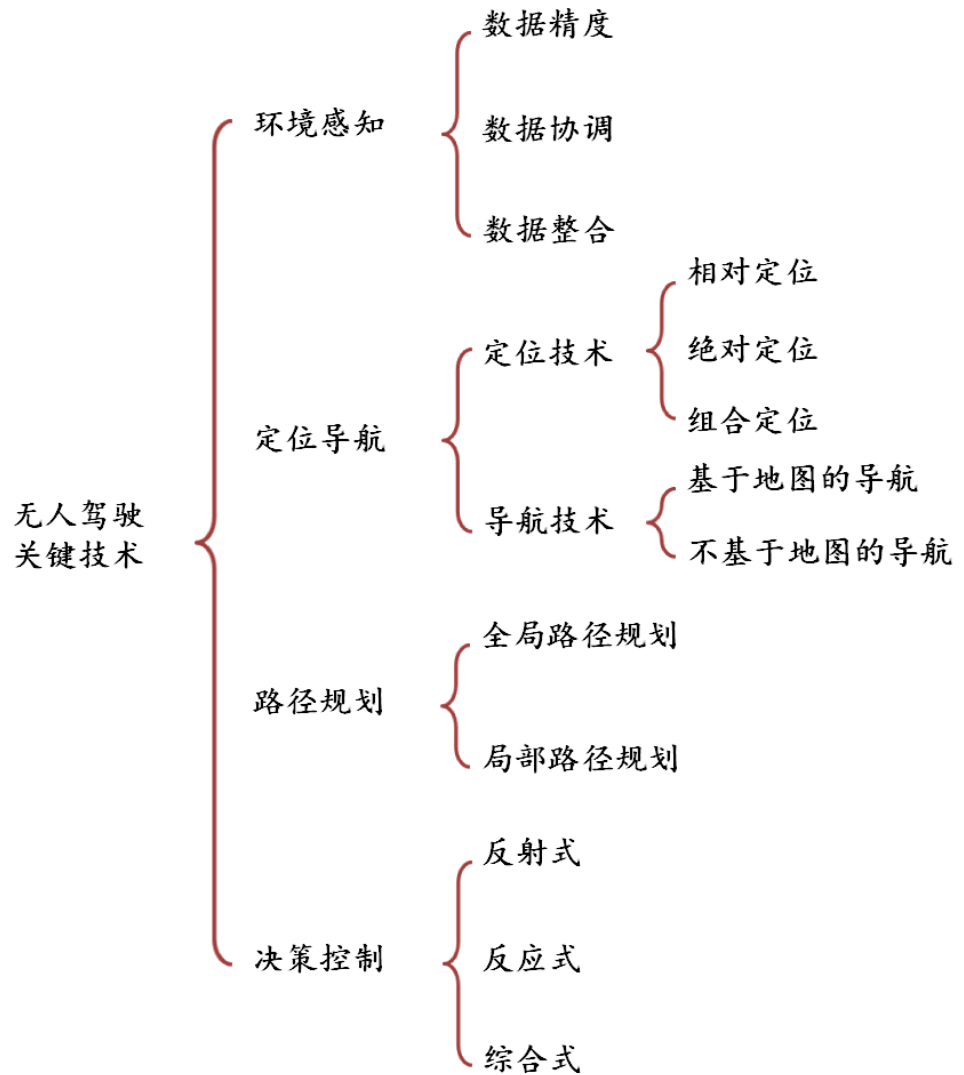
**无人驾驶关键技术正在突破:** 无人驾驶技术是一个涉及传感器、计算机、信息通讯、自动控制、导航定位、机器视觉、人工智能等诸多前沿学科的综合技术。根据无人驾驶的职能模块，可将无人驾驶的关键技术分为：环境感知技术、定位导航技术、路径规划技术和决策控制技术。

环境感知技术是通过多种传感器对车辆周围的环境信息进行感知。环境信息不仅包括了车辆自身状态信息，如车辆速度、转向度、位置信息、倾角、加速度等，还包括四周环境信息，如道路位置、道路方向、障碍物位置和速度、交通标志等。

定位导航技术主要包括定位技术和导航技术。定位技术可以分为相对定位（如陀螺仪、里程计算）、绝对定位（如 GPS）和组合定位。导航技术可以分为基于地图的导航和不基于地图的导航（如惯性导航）。其中高精度地图在无人驾驶的导航中起关键作用。



图 67: 无人驾驶的关键技术



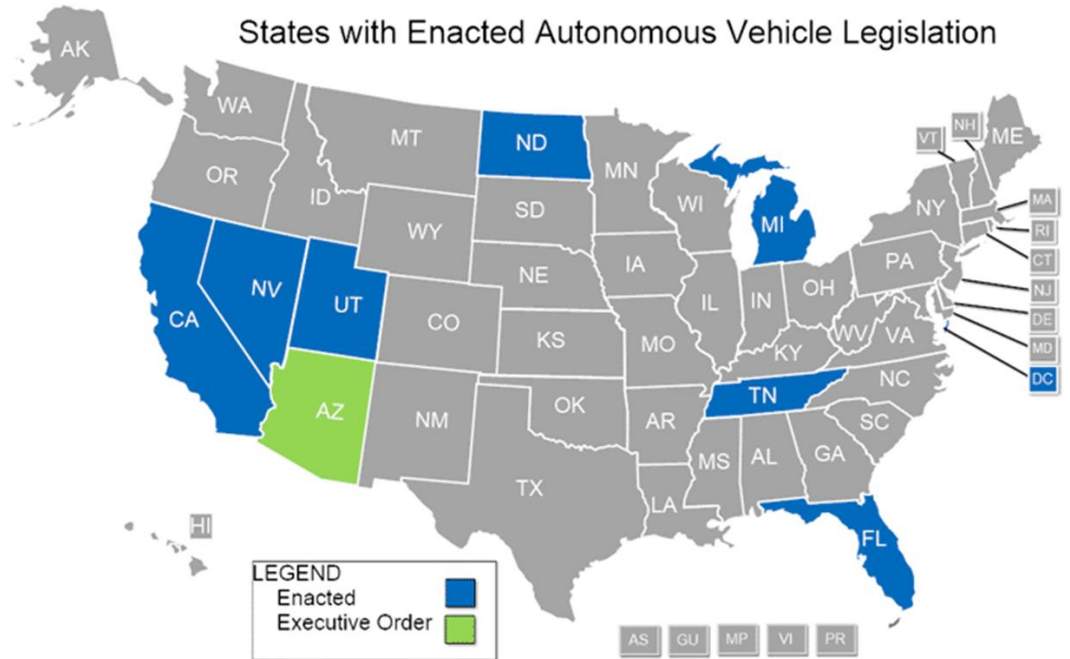
资料来源:《应用技术》, 招商证券

路径规划技术可以为无人驾驶提供最优的行车路径。无人驾驶车在行驶过程中, 行车路线的确定、如何躲避障碍物、路口转向等问题都需要通过路径规划技术完成。据适用范围不同, 路径规划技术通常可分为全局路径规划和局部路径规划。

决策控制技术相当于智能车的大脑, 它通过综合分析环境感知系统提供的信息, 对当前的车辆行为产生决策。决策技术还需要考虑车辆的机械特性、动力特性, 出合理的控制策略。常用的决策技术有机器学习、神经网络、贝叶斯网络、模糊逻辑等。根据决策技术的不同, 控制系统可分为反射式、反应式和综合式。

随着无人驾驶技术的不断发展演进, 对无人驾驶的立法变得十分重要。**截至 2015 年, 美国已经有 16 个州启动无人驾驶立法**; 截至 2014 年是 12 个州, 截至 2013 年是 9 个州, 截至 2012 年是 6 个州。

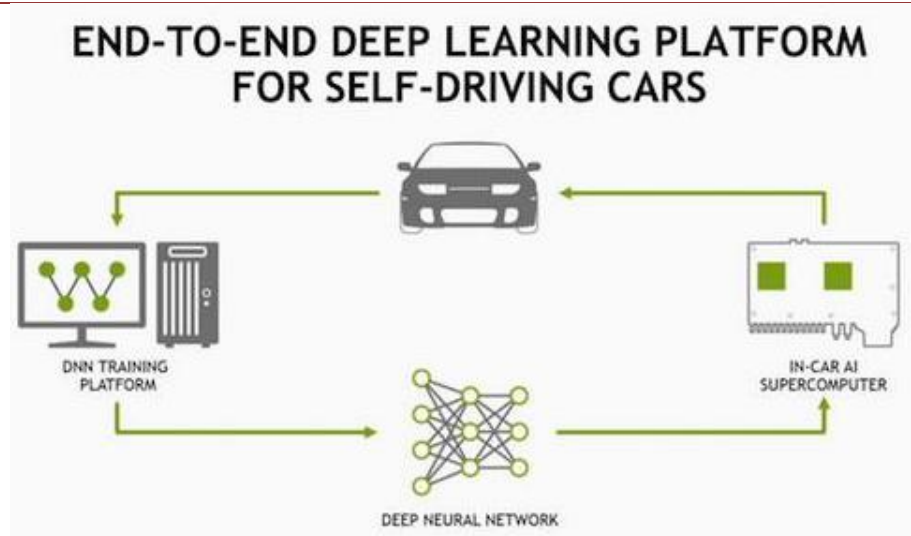
图 68: 截至 2015 年, 美国已经有 16 个州启动无人驾驶立法



资料来源: ncs1, 招商证券

“无人汽车大脑”人工智能的智能程度决定了无人驾驶的可靠性: 由于无人驾驶完全交由汽车的内置程序负责, 因此人工智能就是无人汽车的大脑, 而测距仪、雷达、传感器、GPS 等设备都是人工智能的“眼睛”。人工智能的智能程度直接决定了无人驾驶汽车在不同的路况、不同的天气、甚至一些探测设备出现故障的突发情况下能否及时做出正确的判断并灵活调整行驶策略, 最终决定了无人驾驶汽车当前最亟待突破的可靠性。NVIDIA 在 2016 年的 CES 大会上发布了“Drive PX 2”车载计算机, 以及一套与之搭配的具有学习功能的自动驾驶系统。

图 69: NVIDIA 具有学习功能的自动驾驶系统



资料来源: NVIDIA, 招商证券

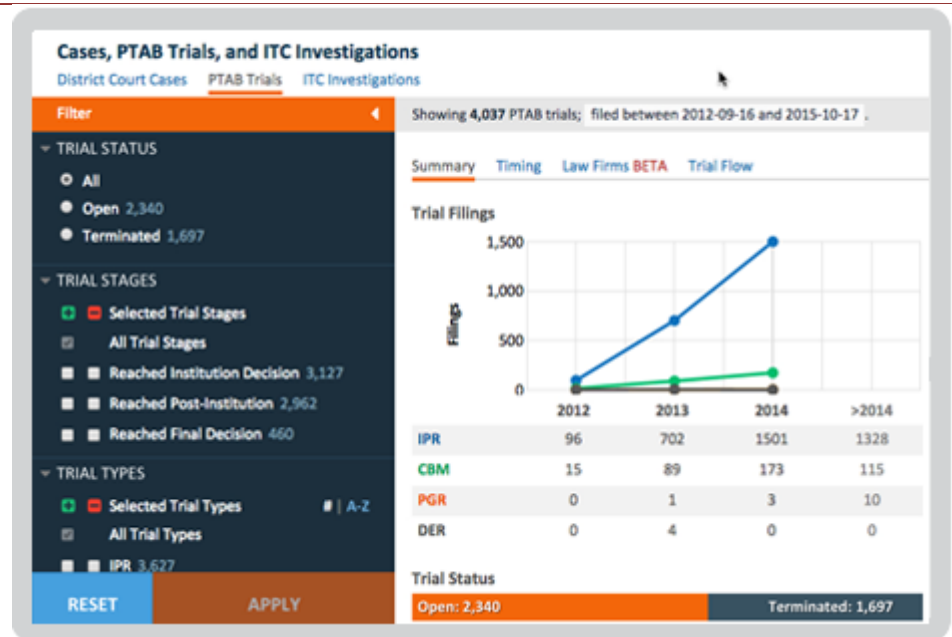
该系统的亮点在于“自我学习”, 通过让车辆自行分析路面状况, 而不是在数据库中寻找预先储存的策略实现自动驾驶, 系统背后连接着名为 NVIDIA DIGITS 的深度学习训

练平台,最终连接到 NVIDIA DRIVENET 神经网络,为车辆的自我学习和完善提供支持。并且由于它是通过判断物体的行进轨迹而不是物体本身去计算路径,因此在驾驶时受天气影响较小。

## 6、法律：案件预判与文案编纂，解放律师双手

**案件判决预测与分析，人工智能有望成为资深“助理律师”：**人工智能在案件预判与分析领域的典型应用案例是美国的知识产权诉讼研究公司 Lex Machina。Lex Machina 的抓取工具提取来自 PACER、所有 94 个联邦地区法院的独立网站、美国国际贸易委员会的电子文件信息系统以及美国专利商标局网站的文件和数据。然后通过自然语言处理技术和法律文本分类技术，将案件、诉讼摘录、组织实体、专利和诉讼结果分门别类，供用户进行审阅和搜索，从而评估原告在先前的专利诉讼是否成功，法官的具体判决特性与习惯，某个律师事务所的胜败诉记录等，进而帮助律师提高对案件的把握性，并对案件的判决进行预测和分析。

图 70: Lex Machina 协助律师进行案件分析与预判



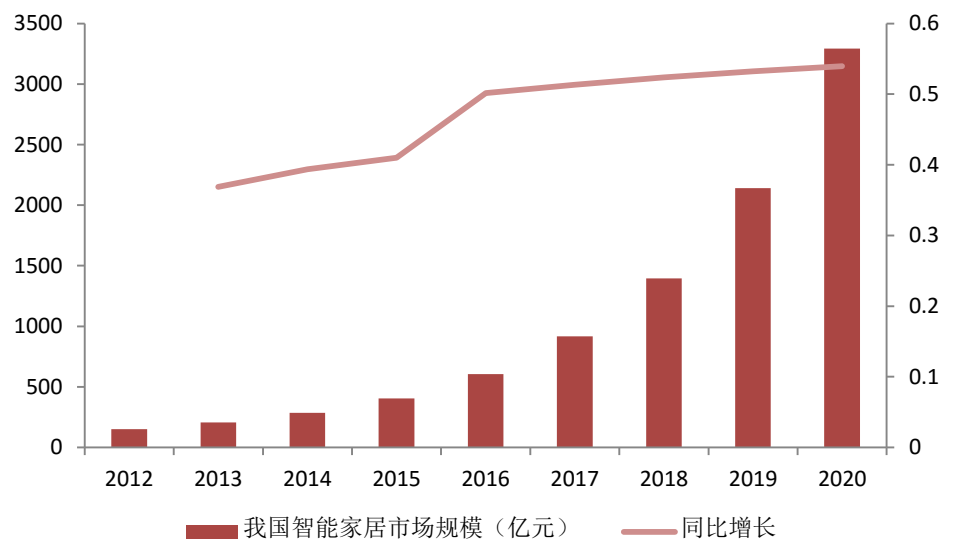
资料来源：Lex Machina、招商证券

**法律服务自动化，人工智能解放律师的双手：**法律文书往往具有较强的格式属性和一定的重复性，采用模板的方式可以一定程度上减少律师的工作量，但最原始的模板缺乏灵活性，可应用的范围也相对较窄。通过人工智能技术可以实现法律文书的自动编纂，通过机器学习可以动态调整并完善文案库和文案的编纂机制，实现法律文书模板的智能化，大幅度降低律师的工作强度，同时也有益于降低人们使用法律服务的成本。美国麻省的 Brightleaf Solutions 便是一家这样的公司，他们的语义智能引擎可以为客户提供法律合同的自动编纂，文案的准确性达到 99.99966%，可以大幅减少用户在法律服务上的开支。我们认为未来通过人工智能技术的不断完善，律师的角色有望逐渐转变为一个法律案件的指导者和策略的制定者，而相对底层的工作将全部交由机器完成。

## 7、智能家居：天花板尚远，人工智能有望成为核心

**行业天花板尚远，增速有望保持在 50%左右：**《钢铁侠》中的“Jarvis”作为智能管家，除了起到钢铁侠的小秘书的作用，还帮主人打理着日常生活，向我们展示了一个理想中的智能家居系统。虽然我们目前可能离那个无所不能的智能管家还很遥远，但智能家居对我们生活的变革确实已经开始了。根据《2012-2020 年中国智能家居市场发展趋势及投资机会分析报告》的预测，我国智能家居市场在 2016 年将达到 605.7 亿的规模，同比增长 50.15%，到 2020 年市场规模将达到 3294 亿，年均增速将保持在 50%左右，具备充足的向上延伸空间。而智能家居想达到“Jarvis”般的终极效果，必然需要引入人工智能技术，实现家居的感应式控制甚至自我学习能力。

图 71：2012-2020 年我国智能家居市场规模



资料来源：千家网、招商证券

**人工智能有望成为智能家居的核心，实现家居自我学习与控制：**按照智能家居的发展进度，大致可以分为四个阶段：手机控制、多控制结合、感应式控制、系统自我学习。当前的发展水平还处在手机控制向多控制结合的过度阶段。而从多控制结合向感应式控制甚至自我学习阶段进化时，人工智能将发挥主要功能。到今天为止，家居的实体功能已经较为全面，未来的发展重点可能在于如何使之升级改造，实现家居的自我行为及协作，因此未来人工智能在智能家居领域的应用有望成为其核心价值。人工智能对智能家居的重构可以深入到方方面面，包括：控制主机、照明系统、影音系统、环境监控、防盗监控、门窗控制、能源管理、空调系统、花草浇灌、宠物看管等等。

图 72: 智能家居产品分类

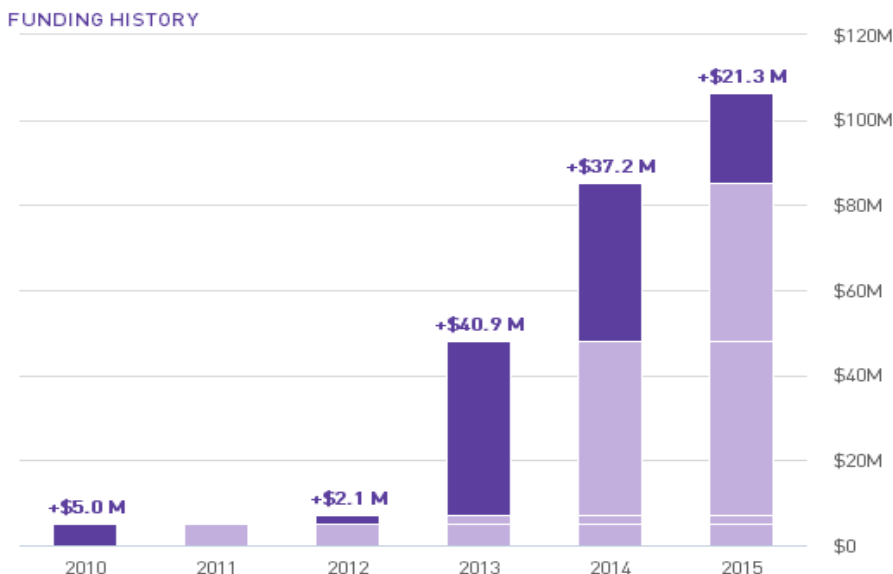


资料来源: 易观智库、招商证券

## 8、医疗：为健康诊断和药品研发插上高飞的翅膀

**健康诊断有望迎来新纪元:** 海量的病历数据和医学界的新研究成果, 单靠人工很难及时筛选并利用, 而引入人工智能技术将充分发挥这些信息的价值。例如著名的个人健康管理产品公司 Welltok 将 IBM 的 Watson 功能融入旗下产品 CafeWell Concierge APP 中, 借助 Watson 的认知计算能力理解人类语言, 实现与用户沟通的能力, 从大量数据中进行分析并为用户提供健康管理相关的答案和建议, 实现健康管理、慢病恢复训练、健康食谱等功能, 这一领域的良好前景使 Welltok 公司近年的融资额连创新高。

图 73: 智能健康管理公司 Welltok 近年融资额不断创新高



资料来源: 动脉网、招商证券

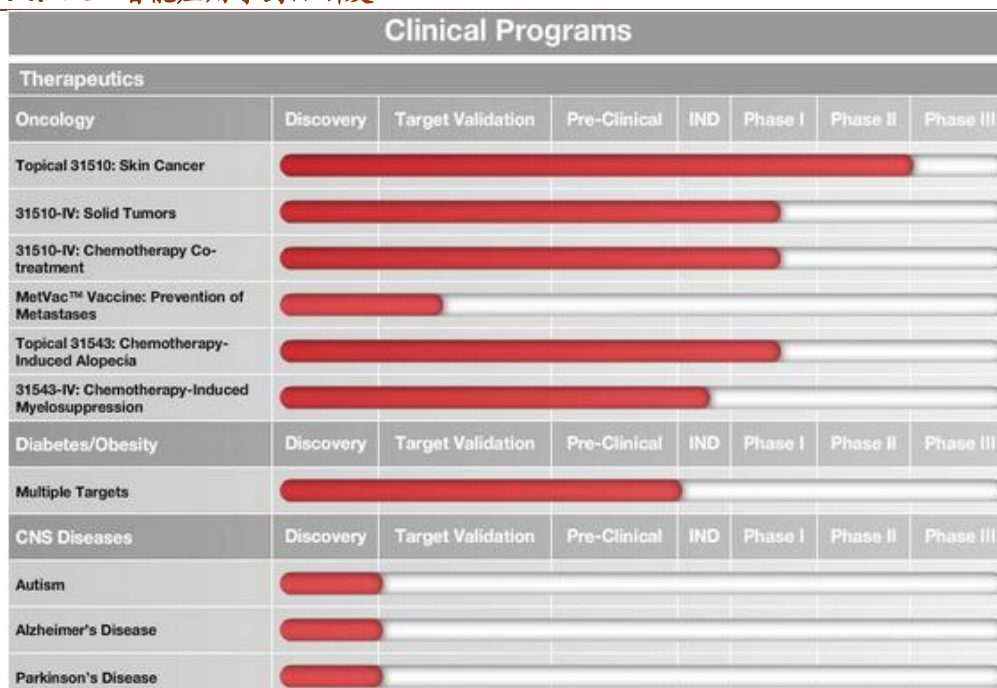
另外, 2015 年 IBM 斥资 10 亿美元收购医疗影像与临床系统提供商 Merge, 将研究如何实现 Watson 的“辨读”医学影像功能。此外, 人工智能还可以从医疗中心获得的健康数据, 通过大数据分析, 实现根据分析患者行为来制定个性化治疗方案的功能。

**人工智能将大幅缩短新药研发周期:** 一般来说, 一种新药的开发需要 10 年以上的时间,



耗资 10 亿美元左右。在传统的药物研发流程中，必须经过动物实验和 I、II、III 期临床试验。人工智能可以帮助药厂进行有效化合物以及药品副作用的筛选，在临床实验阶段也可以协助监测并分析人体对药物的反应，以加快整个药品研发的进程。例如 BergHealth 利用人工智能和大数据来探究人体分子和细胞自身防御组织，以及发病原理机制，推算人体自身分子潜在的药物化合物，利用这种技术来医治类似于糖尿病和癌症等疑难杂症，要比研究新药的时间成本与资金少一半。又如强生用 Watson 来快速分析详细的临床试验结果的科技论文，加快对不同治疗方法的对比效果研究，而如果用传统方法，则需要 3 个人花费 10 个月的时间来完成这些工作。

图 74: 人工智能应用于药品研发



资料来源：生物 360、招商证券

## 五、投资建议

我们看好人工智能开源平台百舸争流争夺 AI 生态圈会极大加速产业发展。我们看好人工智能芯片不断进化下的 AI 算法效率大幅提升。“人工智能+”时代呼啸而来，我们先期看好专用人工智能的普及，中长期看好通用人工智能的普及。

**第三次浪潮：人工智能奇点临近。**人工智能在两次高潮和低谷之后迎来第三次浪潮，各国纷纷把人工智能上升到国家战略层面，美国白宫组织四场研讨会讨论人工智能，日本提出“超智能社会”，中国发布《“互联网+”人工智能三年行动实施方案》。仿佛一夜间，各国对人工智能热情大增，我们看到人工智能领域巨头们的卡位战争已经上升到开源平台、芯片以及应用。相关一级创投金额 5 年增长了 12 倍，我们预计二级市场的机会随时来临。

**开源平台：超越 Android，AI 生态圈之争极大加速产业发展。**我们看到各大科技巨头和科研机构纷纷争夺开源人工智能技术，如 Google 的 TensorFlow、微软的 DMTK、Facebook 的 Torchnet、OpenAI、Baidu 的 Paddle 等，百舸争流千帆竞，以此来获得大量的用户需求和开发人员，建立开放共享、互利共赢的人工智能生态圈，我们认为人工智能开源平台有望成为引爆产业的导火线，从而比 Android 更深远地影响世界。

**人工智能芯片：从通用到专用，类脑计算不断进化，效率大幅提升。**我们认为人工智能芯片效率的提升和突破将带领产业实现跳跃式发展。GPU 王者 Nvidia，与其合作的组织 2 年增长 30 多倍。FPGA 兼顾性能与灵活性，与 GPU 强强对决。我们预计人工智能专用芯片在竞争中有望实现超越摩尔定律的发展速度，硬件的春天即将来临。

**应用：“人工智能+”时代呼啸而来，改变 IT 的命运。**我们先期看好专用人工智能的普及，中长期看好通用人工智能的普及。人工智能在未来几乎会给各行各业带来升级重构，包括安防、制造业、金融、交通、教育、法律、医疗等行业，AI 的触角伸向每个行业和每个人，从而改变 IT 未来的走向。

**投资建议：**建议从两个角度选择 A 股相关标的：1、数据能力；2、应用能力。重点推荐：“人工智能+安防”：**东方网力**；“人工智能+金融”：**同花顺**；“人工智能+医疗”：**思创医惠**；“人工智能+底层操作系统”：**中科创达**；“人工智能+大数据”：**神州泰岳**。关注：科大讯飞、佳都科技、汉王科技。

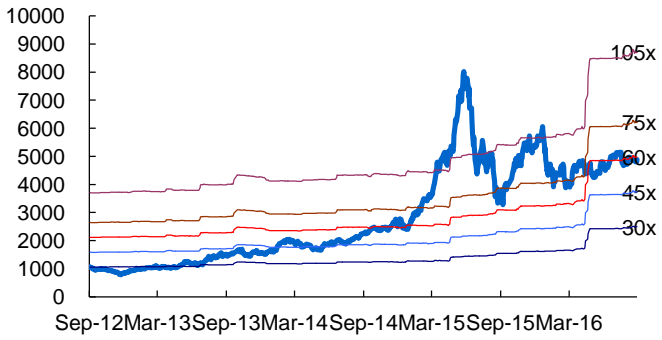
**风险提示：**政策支持力度低于预期；核心技术发展遭遇瓶颈。

表 6：重点公司主要财务指标

	股价	15EPS	16EPS	17EPS	15PE	16PE	PB	评级
东方网力	24.63	0.77	0.48	0.65	51	38	10	强烈推荐-A
同花顺	67.85	1.78	2.14	2.82	32	24	19	强烈推荐-A
思创医惠	30.01	0.34	0.52	0.70	58	43	9	强烈推荐-A
中科创达	56.17	1.17	0.34	0.56	166	101	18	强烈推荐-A
神州泰岳	11.41	0.18	0.25	0.33	45	35	4	强烈推荐-A
科大讯飞	30.52	0.33	0.40	0.55	76	55	6	强烈推荐-A
佳都科技	9.13	0.34	0.19	0.27	48	34	5	审慎推荐-A
汉王科技	25.67	0.02	0.09	0.12	275	220	7	审慎推荐-A

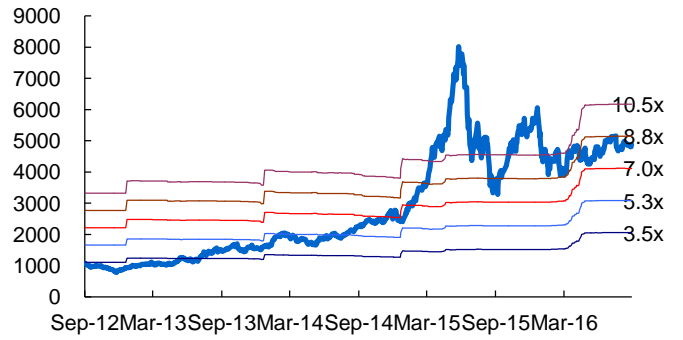
资料来源：公司数据，招商证券

图 75: 计算机行业历史PEBand



资料来源: 贝格数据、招商证券

图 76: 计算机行业历史PBBand



资料来源: 贝格数据、招商证券

参考报告:

- 1、《IT 宿命系列之: 人工智能专题深度(一) AI: 重构 IT 未来》2016/04
- 2、《“奇点”临近, AI 千亿盛宴将启——“互联网+”人工智能三年行动实施方案点评》2016/05
- 3、《人工智能跟踪: AI 的二级推进器: “云化+开源”》2016/06

## 分析师承诺

负责本研究报告的每一位证券分析师，在此申明，本报告清晰、准确地反映了分析师本人的研究观点。本人薪酬的任何部分过去不曾与、现在不与、未来也将不会与本报告中的具体推荐或观点直接或间接相关。

**刘泽晶：**2014/15 年新财富计算机行业团队第三、第五名，2014 年水晶球团队第三名。中央财经大学硕士毕业，6 年从业经验。

**周楷宁：**招商证券计算机行业高级分析师，2012、2013 年新财富第一团队成员，北京大学硕士，武汉大学学士，4 年证券从业经验。

**黄斐玉：**招商证券计算机行业分析师，北京航空航天大学硕士。具有摩托罗拉、联想集团等 IT 公司产品研发，技术预研等领域多年工作经验。

**徐文杰：**招商证券计算机行业分析师，北京航空航天大学硕士。曾供职于 IBM，VMware 等公司从事大数据、云计算领域的工作。

**宋兴未：**招商证券计算机行业分析师，美国波士顿大学计算机工程系硕士，上海交通大学电子工程系学士。2016 年加入招商证券研究所。

实习生焦欢、江鸿震、尹冠乔对本报告有贡献。

## 投资评级定义

### 公司短期评级

以报告日起 6 个月内，公司股价相对同期市场基准（沪深 300 指数）的表现为标准：

- 强烈推荐：公司股价涨幅超基准指数 20%以上
- 审慎推荐：公司股价涨幅超基准指数 5-20%之间
- 中性：公司股价变动幅度相对基准指数介于±5%之间
- 回避：公司股价表现弱于基准指数 5%以上

### 公司长期评级

- A：公司长期竞争力高于行业平均水平
- B：公司长期竞争力与行业平均水平一致
- C：公司长期竞争力低于行业平均水平

### 行业投资评级

以报告日起 6 个月内，行业指数相对于同期市场基准（沪深 300 指数）的表现为标准：

- 推荐：行业基本面向好，行业指数将跑赢基准指数
- 中性：行业基本面稳定，行业指数跟随基准指数
- 回避：行业基本面向淡，行业指数将跑输基准指数

## 重要声明

本报告由招商证券股份有限公司（以下简称“本公司”）编制。本公司具有中国证监会许可的证券投资咨询业务资格。本报告基于合法取得的信息，但本公司对这些信息的准确性和完整性不作任何保证。本报告所包含的分析基于各种假设，不同假设可能导致分析结果出现重大不同。报告中的内容和意见仅供参考，并不构成对所述证券买卖的出价，在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议。除法律或规则规定必须承担的责任外，本公司及其雇员不对使用本报告及其内容所引发的任何直接或间接损失负任何责任。本公司或关联机构可能会持有报告中所提到的公司所发行的证券头寸并进行交易，还可能为这些公司提供或争取提供投资银行业务服务。客户应当考虑到本公司可能存在可能影响本报告客观性的利益冲突。

本报告版权归本公司所有。本公司保留所有权利。未经本公司事先书面许可，任何机构和个人不得以任何形式翻版、复制、引用或转载，否则，本公司将保留随时追究其法律责任的权利。