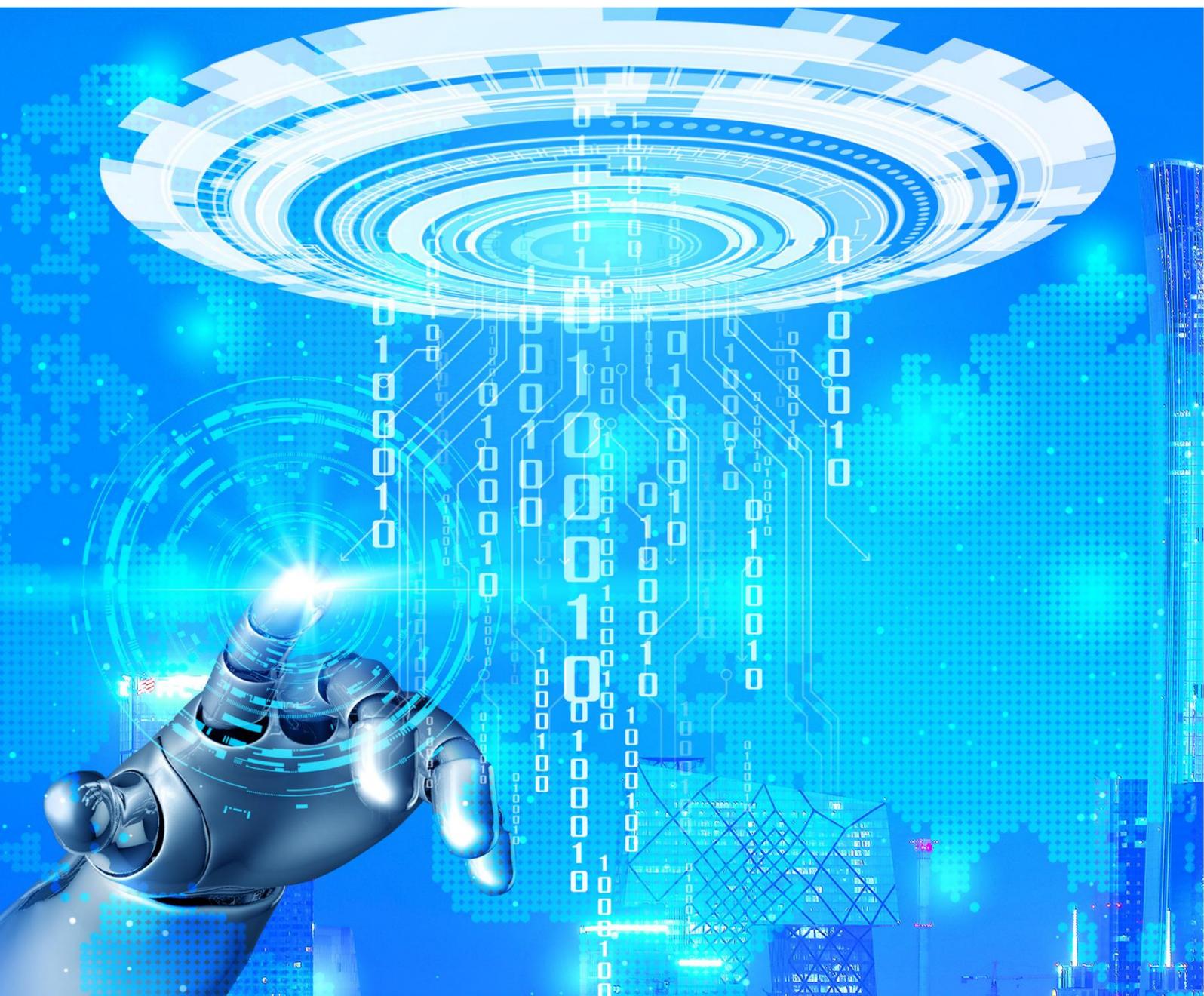


人工智能之机器学习

Research Report of Machine Learning

2020年 第1期



清华大学人工智能研究院
清华-中国工程院知识智能联合研究中心

2020年1月

1. 概述篇	1
1.1 机器学习的概念	1
1.2 机器学习的发展历史	1
2. 技术篇	2
2.1 机器学习算法分类	2
2.2 机器学习的经典代表算法	2
2.3 生成对抗网络及对抗机器学习	3
2.4 自动机器学习	4
2.5 可解释性机器学习	6
2.6 在线学习	6
2.7 BERT	7
2.8 卷积与图卷积	8
2.9 隐私保护	10
3. 深度学习篇	11
3.1 卷积神经网络	12
3.2 AutoEncoder	12
3.3 循环神经网络 RNN	13
3.4 网络表示学习与图神经网络 GNN	13
3.5 增强学习	14
3.6 生成对抗网络	14
3.7 老虎机	15
3.8 图神经网络	15
3.9 深度学习近期重要进展	16
4. 论文解读篇	18
5. 人才篇	21
5.1 学者情况概览	21
5.2 代表性学者简介	23
5.3 NeurIPS 十年高引学者	26

CONTENTS

6. 应用篇	30
7. 趋势篇	35
8. 资源篇	36
8.1 开源代码	36
8.2 预训练	37
8.3 课程	37
8.4 数据集	38
8.5 机器学习知识树	39
参考文献	41

图目录

图 1-1 机器学习发展历程	1
图 2-1 机器学习分类	2
图 2-2 GAN 发展脉络	3
图 2-3 AutoML 基本过程	5
图 2-4 ATMSeer 自动机器学习定制化工具的用户友好型交互界面	5
图 2-5 Transformer 的网络架构	8
图 2-6 图卷积示意图	9
图 3-1 深度学习模型最近若干年的重要进展	11
图 3-2 卷积神经网络的重要进展	12
图 3-3 Auto-Encoder 的重要进展	12
图 3-4 循环神经网络 RNN 的重要进展	13
图 3-5 网络表示学习与图神经网络的重要进展	13
图 3-6 增强学习的重要进展	14
图 3-7 生成对抗网络的重要进展	14
图 3-8 老虎机的重要进展	15
图 5-1 机器学习领域全球学者分布	21
图 5-2 机器学习领域学者 h-index 分布	22
图 5-3 机器学习领域中国学者分布	22
图 6-1 自动驾驶目标识别、运动预测	31
图 7-1 机器学习技术趋势	35

表目录

表 4-1 ICML 近 10 年 best paper	18
表 4-2 NeurIPS 近 10 年 best paper	19
表 5-1 机器学习领域中国与各国合作论文情况	23
表 5-2 NeurIPS 高引学者 TOP100	26
表 8-1 机器学习三级知识树	错误!未定义书签。

1 概述篇

1.1 机器学习的概念

机器学习已经成为了当今的热门话题,但是从机器学习这个概念诞生到机器学习技术的普遍应用经过了漫长的过程。在机器学习发展的历史长河中,众多优秀的学者为推动机器学习的发展做出了巨大的贡献。

从 1642 年 Pascal 发明的手摇式计算机,到 1949 年 Donald Hebb 提出的赫布理论——解释学习过程中大脑神经元所发生的变化,都蕴含着机器学习思想的萌芽。事实上,1950 年图灵在关于图灵测试的文章中就已提及机器学习的概念。到了 1952 年,IBM 的亚瑟·塞缪尔 (Arthur Samuel, 被誉为“机器学习之父”)设计了一款可以学习的西洋跳棋程序。塞缪尔和这个程序进行多场对弈后发现,随着时间的推移,程序的棋艺变得越来越好^[1]。塞缪尔用这个程序推翻了以往“机器无法超越人类,不能像人一样写代码和学习”这一传统认识。并在 1956 年正式提出了“机器学习”这一概念。

对机器学习的认识可以从多个方面进行,有着“全球机器学习教父”之称的 Tom Mitchell 则将机器学习定义为:对于某类任务 T 和性能度量 P,如果计算机程序在 T 上以 P 衡量的性能随着经验 E 而自我完善,就称这个计算机程序从经验 E 学习。

普遍认为,机器学习 (Machine Learning, 常简称为 ML) 的处理系统和算法是主要通过找出数据里隐藏的模式进而做出预测的识别模式,它是人工智能 (Artificial Intelligence, 常简称为 AI) 的一个重要子领域。

1.2 机器学习的发展历史

从机器学习发展的过程上来说,其发展的时间轴如下所示:

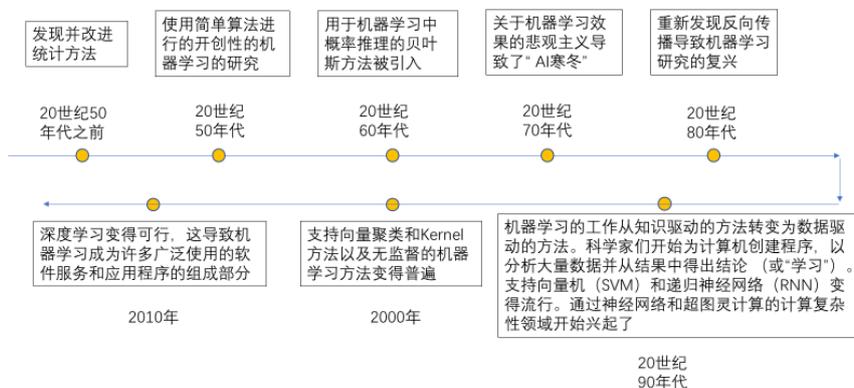


图 1-1 机器学习发展历程

2 技术篇

2.1 机器学习算法分类

机器学习算法可以按照不同的标准来进行分类。比如按函数 $f(x, \theta)$ 的不同，机器学习算法可以分为线性模型和非线性模型；按照学习准则的不同，机器学习算法也可以分为统计方法和非统计方法。但一般来说，我们会按照训练样本提供的信息以及反馈方式的不同，将机器学习算法分为监督学习、无监督学习和强化学习^[2]。

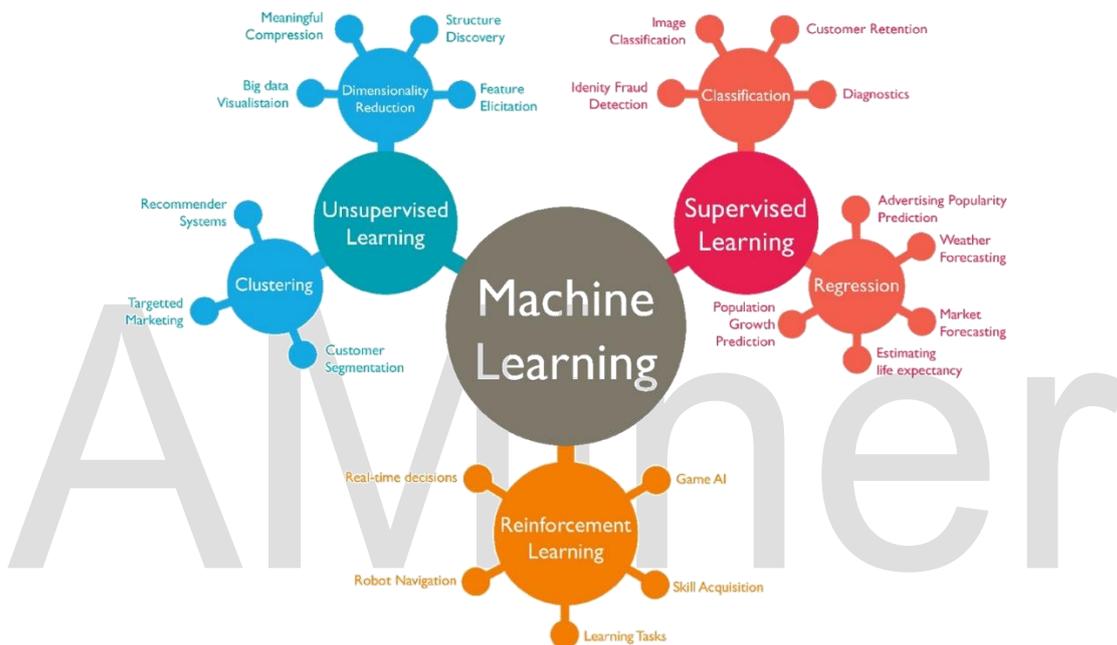


图 2-1 机器学习分类

2.2 机器学习的经典代表算法

1980 年机器学习作为一支独立的力量登上了历史舞台。在这之后的 10 年里出现了一些重要的方法和理论，典型的代表是：分类与回归树（CART, 1984）、反向传播算法（1986）、卷积神经网络（1989）。从 1990 到 2012 年，机器学习逐渐走向成熟和应用，在这 20 多年里机器学习的理论和方法得到了完善和充实，可谓是百花齐放的年代。代表性的重要成果有：支持向量机（SVM, 1995）、AdaBoost 算法（1997）、循环神经网络和 LSTM（1997）、流形学习（2000）、随机森林（2001）。机器学习代表算法包括^[3]：

- 线性回归
- 分类与回归树（CART）

- 随机森林 (Random Forest)
- 逻辑回归
- 朴素贝叶斯 (Naive Bayesian)
- k 最近邻 (kNN)
- AdaBoost
- K-均值算法 (K-Means)
- 支持向量机 (SVM)
- 神经网络 ANN (Artificial Neural Network)

2.3 生成对抗网络及对抗机器学习

2.3.1 生成对抗网络

生成对抗网络 (Generative Adversarial Networks, GAN) 是用于无监督学习的机器学习模型，由 Ian Goodfellow 等人在 2014 年提出，由神经网络构成判别器和生成器构成，通过一种互相竞争的机制组成的一种学习框架，GAN 在深度学习领域掀起了一场革命。传统的生成模型最早要追溯到 80 年代的 RBM，以及后来逐渐使用深度神经网络进行包装的 AutoEncoder，然后就是现在称得上最火的生成模型 GAN^[4]。

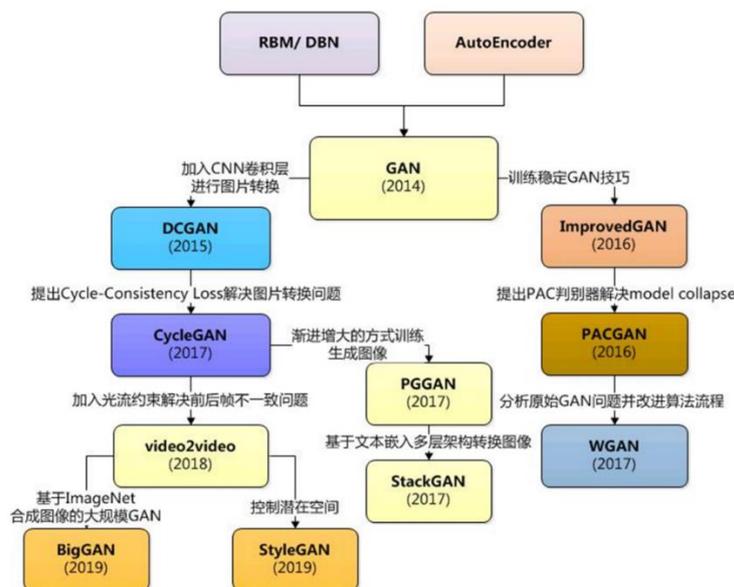


图 2-2 GAN 发展脉络

2.3.2 对抗机器学习

对抗机器学习是一个机器学习与计算机安全的交叉领域。对抗机器学习旨在给恶意环境下的机器学习技术提供安全保障。由于机器学习技术一般研究的是同一个或较为稳定的数据分布，当部署到现实中的时候，由于恶意用户的存在，这种假设并不一定成立。比如研究人员发现，一些精心设计的对抗样本（adversarial example）可以使机器学习模型失败输出正确的结果^[5]。针对模型的攻击问题，我们主要分为两大类，就是从训练阶段和推理（inference）阶段来进行讨论^[6]。

- 训练阶段的攻击

训练阶段的恶意攻击（Training in Adversarial Settings），主要的目的就是针对模型的参数进行微小的扰动，从而让模型的性能和预期产生偏差。这样的行为主要是通过数据投毒来完成的^[7]。

- 推理阶段的攻击（Inference in Adversarial Settings）

当训练完成一个模型之后，这个模型就可以看做一个 BOX，那么这个盒子中，对于我们如果是透明的话，我们就把它当成是“白盒”模型，如果这个盒子中，我们什么都看不见了，我们就把它当成“黑盒”模型。（我们在这个部分不讨论灰盒模型）那么针对白盒和黑盒的进攻手段自然是不同的，但是最终的目的都是希望能对模型的最终结果产生破坏，与预期脱离。其影响力以及攻击的构造粒度也是有所不同的。

2.4 自动机器学习

2.4.1 AutoML

自动机器学习（AutoML）旨在通过让一些通用步骤（如数据预处理、模型选择和调整超参数）自动化，来简化机器学习中生成模型的过程。AutoML 是指尽量不通过人来设定超参数，而是使用某种学习机制，来调节这些超参数。这些学习机制包括传统的贝叶斯优化，多臂老虎机（multi-armed bandit），进化算法，还有比较新的强化学习。自动机器学习不光包括大家熟知的算法选择，超参数优化，和神经网络架构搜索，还覆盖机器学习工作流的每一步。自动机器学习的用处就在于此，它帮助研究人员和从业者，自动构建机器学习管道，将多个步骤及其对应的多个选项集成为工作流，以期快速找到针对给定问题的高性能机器学习模型^[8]。AutoML 的基本过程如下图所示：虚框是配置空间，包括特征、超参数和架构；左边训练数据进入，上面的优化器和它相连，定义的测度发现最佳配置，最后出来的是模型；测试数据在模型中运行，实现预测的目的。

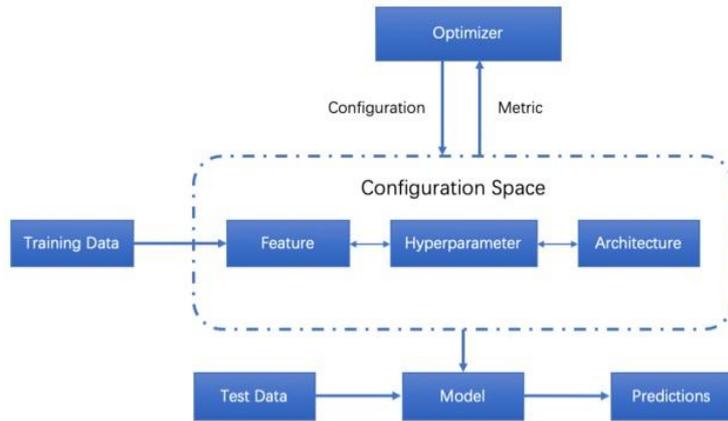


图 2-3 AutoML 基本过程

2.4.2 ATMSeer

近日，在 ACM CHI 计算系统中人的因素会议上，麻省理工学院，香港科技大学和浙江大学的研究人员共同研发出一种工具，将 AutoML 方法的分析和控制权给到用户手中。该工具名为 ATMSeer，它将 AutoML 系统、数据集和有关用户任务的一些信息作为输入，然后在用户友好型的界面内实现可视化搜索过程，界面中还能提供更多关于模型性能的信息。

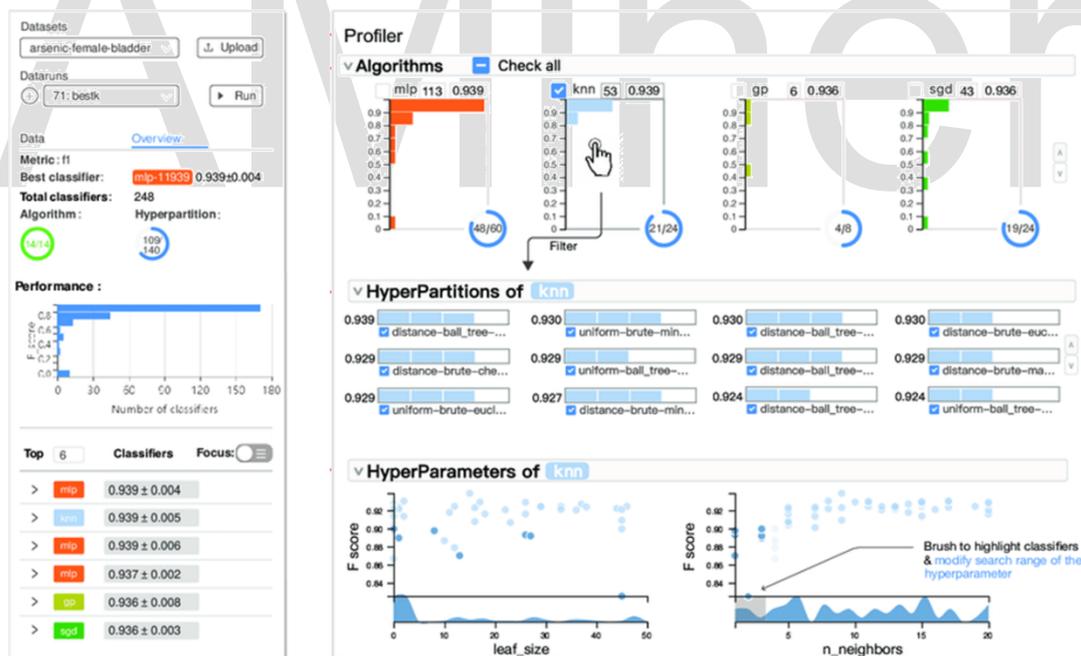


图 2-4 ATMSeer 自动机器学习定制化工具的用户友好型交互界面

上图是 ATMSeer 生成的一个用户友好界面，显示有关所选模型性能的深入信息，以及可调整的算法和参数的选项^[9]。对没有 AutoML 经验的机器学习专家的案例研究表明，让用户掌握控制权确实有助于提高 AutoML 应用的性能和效率。对生物学、金融等不同科学领域的 13 位研究生的研究也表明，确定用户对 AutoML 的搜索的自定义关键有三点：搜索的算

法数量、系统运行时间以及查找表现最好的模型。研究人员表示，这些信息可用来为用户量身定制系统^[10]。

2.5 可解释性机器学习

可解释性是指人类能够理解决策原因的程度。机器学习模型的可解释性越高，人们就越容易理解为什么做出某些决定或预测。模型可解释性指对模型内部机制的理解以及对模型结果的理解。其重要性体现在：建模阶段，辅助开发人员理解模型，进行模型的对比选择，必要时优化调整模型；在投入运行阶段，向业务方解释模型的内部机制，对模型结果进行解释。比如基金推荐模型，需要解释：为何为这个用户推荐某支基金。

机器学习流程步骤：收集数据、清洗数据、训练模型、基于验证或测试错误或其他评价指标选择最好的模型。第一步，选择比较小的错误率和比较高的准确率的高精度的模型。第二步，面临准确率和模型复杂度之间的权衡，但一个模型越复杂就越难以解释。一个简单的线性回归非常好解释，因为它只考虑了自变量与因变量之间的线性相关关系，但是也正因为如此，它无法处理更复杂的关系，模型在测试集上的预测精度也更有可能会比较低。而深度神经网络处于另一个极端，因为它们能够在多个层次进行抽象推断，所以它们可以处理因变量与自变量之间非常复杂的关系，并且达到非常高的精度。但是这种复杂性也使模型成为黑箱，我们无法获知所有产生模型预测结果的这些特征之间的关系，所以我们只能用准确率、错误率这样的评价标准来代替，来评估模型的可信性。事实上，每个分类问题的机器学习流程中都应该包括模型理解和模型解释^[11]。

2.6 在线学习

传统的机器学习算法是批量模式的，假设所有的训练数据预先给定，通过最小化定义在所有训练数据上的经验误差得到分类器。这种学习方法在小规模规模上取得了巨大成功，但当数据规模大时，其计算复杂度高、响应慢，无法用于实时性要求高的应用。与批量学习不同，在线学习假设训练数据持续到来，通常利用一个训练样本更新当前的模型，大大降低了学习算法的空间复杂度和时间复杂度，实时性强。在大数据时代，大数据高速增长的特点为机器学习带来了严峻的挑战，在线学习可以有效地解决该问题，引起了学术界和工业界的广泛关注。早期在线学习应用于线性分类器产生了著名的感知器算法，当数据线性可分时，感知器算法收敛并能够找到最优的分类面。经过几十年的发展，在线学习已经形成了一套完备的理论，既可以学习线性函数，也可以学习非线性函数，既能够用于数据可分的情况，也能够处理数据不可分的情况。下面我们给出一个在线学习形式化的定义及其学习目标。

根据学习器在学习过程中观测信息的不同，在线学习还可以再进一步分为：完全信息下的在线学习和赌博机在线学习。前者假设学习器可以观测到完整的损失函数，而后者假设学习器只能观测到损失函数在当前决策上的数值，即，以在线分类为例，如果学习器可以观测到训练样本，该问题就属于完全信息下的在线学习，因为基于训练样本就可以定义完整的分类误差函数；如果学习器只能观测到分类误差而看不到训练样本，该问题就属于赌博机在线学习。由于观测信息的不同，针对这两种设定的学习算法也存在较大差异，其应用场景也不同^[12]。

2.7 BERT

BERT 的全称是 Bidirectional Encoder Representation from Transformers，即双向 Transformer 的 Encoder。可以说是近年来自残差网络最优突破性的一项技术了。BERT 主要特点以下几点：

1) 使用了 Transformer 作为算法的主要框架，Transformer 能更彻底的捕捉语句中的双向关系^[13]；

2) 使用了 Mask Language Model (MLM)^[14]和 Next Sentence Prediction (NSP) 的多任务训练目标；

3) 使用更强大的机器训练更大规模的数据，使 BERT 的结果达到了全新的高度，并且 Google 开源了 BERT 模型，用户可以直接使用 BERT 作为 Word2Vec 的转换矩阵并高效的将其应用到自己的任务中。

BERT 的本质上是通过在海量的语料的基础上运行自监督学习方法为单词学习一个好的特征表示，所谓自监督学习是指在没有人工标注的数据上运行的监督学习。在以后特定的 NLP 任务中，我们可以直接使用 BERT 的特征表示作为该任务的词嵌入特征。所以 BERT 提供的是一个供其它任务迁移学习的模型，该模型可以根据任务微调或者固定之后作为特征提取器。BERT 的源码和模型已经在 Github 上开源，简体中文和多语言模型也已开源。

BERT 的网络架构使用的是“*Attention is all you need*”中提出的多层 Transformer 结构，其最大的特点是抛弃了传统的 RNN 和 CNN，通过 Attention 机制将任意位置的两个单词的距离转换成 1，有效的解决了 NLP 中棘手的长期依赖问题。Transformer 的网络架构如下图所示，Transformer 是一个 encoder-decoder 的结构，由若干个编码器和解码器堆叠形成。下图的左侧部分为编码器，由 Multi-Head Attention 和一个全连接组成，用于将输入语料转化成特征向量。右侧部分是解码器，其输入为编码器的输出以及已经预测的结果，由 Masked Multi-Head Attention, Multi-Head Attention 以及一个全连接组成，用于输出最后结果的条件概率。

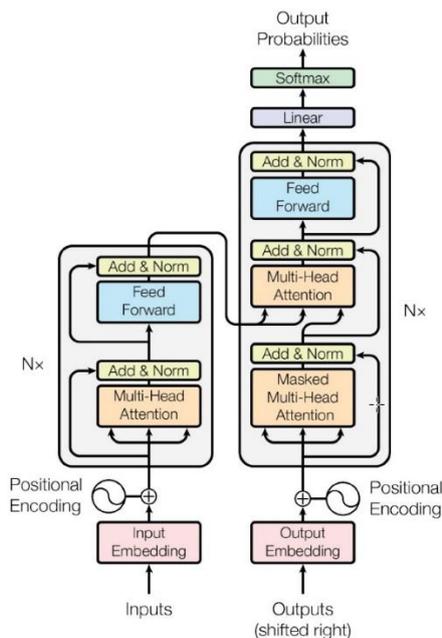


图 2-5 Transformer 的网络架构

2.8 卷积与图卷积

2.8.1 卷积

两个函数的卷积，本质上就是先将一个函数翻转，然后进行滑动叠加。在连续情况下，叠加指的是对两个函数的乘积求积分，在离散情况下就是加权求和，为简单起见就统一称为叠加。教科书上一般定义函数 f, g 的卷积 $f * g(n)$ 如下：

连续形式：

$$(f * g)(n) = \int_{-\infty}^{\infty} f(\tau)g(n - \tau)d\tau$$

离散形式：

$$(f * g)(n) = \sum_{\tau=-\infty}^{\infty} f(\tau)g(n - \tau)$$

从计算的方式上对公式进行的解释为：先对 g 函数进行翻转，相当于在数轴上把 g 函数从右边褶到左边去，也就是卷积的“卷”的由来。然后再把 g 函数平移到 n ，在这个位置对两个函数的对应点相乘，然后相加，这个过程是卷积的“积”的过程。整体看来是这么个过程：

翻转→滑动→叠加→滑动→叠加→滑动→叠加

多次滑动得到的一系列叠增加值，构成了卷积函数。卷积的“卷”，指的是函数的翻转，从 $g(t)$ 变成 $g(-t)$ 的这个过程；同时，“卷”还有滑动的意味在里面。如果把卷积翻译为“褶积”，那么这个“褶”字就只有翻转的含义了。卷积的“积”，指的是积分/加权求和。对卷积的意义理解如下：

1) 从“积”的过程可以看到，我们得到的叠增加值，是个全局的概念。以信号分析为例，卷积的结果是不仅跟当前时刻输入信号的响应值有关，也跟过去所有时刻输入信号的响应都有关系，考虑了对过去的所有输入的效果的累积。在图像处理的中，卷积处理的结果，其实就是把每个像素周边的，甚至是整个图像的像素都考虑进来，对当前像素进行某种加权处理。所以说，“积”是全局概念，或者说是一种“混合”，把两个函数在时间或者空间上进行混合。

2) 进行“卷”（翻转）的目的其实是施加一种约束，它指定了在“积”的时候以什么为参照。在信号分析的场景，它指定了在哪个特定时间点的前后进行“积”，在空间分析的场景，它指定了在哪个位置的周边进行累积处理^[15]。

2.8.2 图卷积

要理解图卷积网络的核心操作图卷积，可以类比卷积在 CNN 的地位。如下图所示，数字图像是一个二维的离散信号，对数字图像做卷积操作其实就是利用卷积核（卷积模板）在图像上滑动，将图像点上的像素灰度值与对应的卷积核上的数值相乘，然后将所有相乘后的值相加作为卷积核中间像素对应的图像上像素的灰度值，并最终滑动完所有图像的过程。

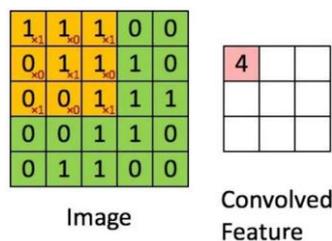


图 2-6 图卷积示意图

现实中更多重要的数据集都是用图的形式存储的，例如社交网络信息、知识图谱、蛋白质网络、万维网等等。这些图网络的形式并不像图像，是排列整齐的矩阵形式，而是非结构化的信息，那有没有类似图像领域的卷积一样，有一个通用的范式来进行图特征的抽取呢？这就是图卷积在图卷积网络中的意义。对于大多数图模型，有一种类似通式的存在，这些模型统称图卷积网络。因此可以说，图卷积是处理非结构化数据的大利器，随着这方面研究的逐步深入，人类对知识领域的处理必将不再局限于结构化数据，会有更多的目光转向这一存在范围更加广泛，涵盖意义更为丰富的知识领域^[16]。

2.9 隐私保护

2019 年，清华大学人工智能研究院院长张钹院士、唐杰教授、李涓子教授等人联合发起“AI TIME” science debate，希望用辩论的形式，探讨人工智能和人类未来之间的矛盾，探索人工智能领域的未来。AI TIME 是一群关注人工智能发展，并有思想情怀的青年人创办的圈子。AI TIME 旨在发扬科学思辨精神，邀请各界人士对人工智能理论、算法、场景、应用的本质问题进行探索，加强思想碰撞，打造成为北京乃至全国知识分享的聚集地。

在“AI TIME 4”及“AI TIME in U 之走进北邮”活动中，受邀的专家对隐私保护以及机器学习与隐私保护的联系展开了深入的探讨，相关论述观点请关注学术头条获取。

AMiner

3 深度学习篇

深度学习是近 10 年机器学习领域发展最快的一个分支，由于其重要性，三位教授（Geoffrey Hinton、Yann Lecun、Yoshua Bengio）因此同获图灵奖。深度学习模型的发展可以追溯到 1958 年的感知机（Perceptron）。1943 年神经网络就已经出现雏形（源自 NeuroScience），1958 年研究认知的心理学家 Frank 发明了感知机，当时掀起一股热潮。后来 Marvin Minsky（人工智能大师）和 Seymour Papert 发现感知机的缺陷：不能处理异或回路等非线性问题，以及当时存在计算能力不足以处理大型神经网络的问题。于是整个神经网络的研究进入停滞期。

最近 30 年来取得快速发展。总体来说，主要有 4 条发展脉络：

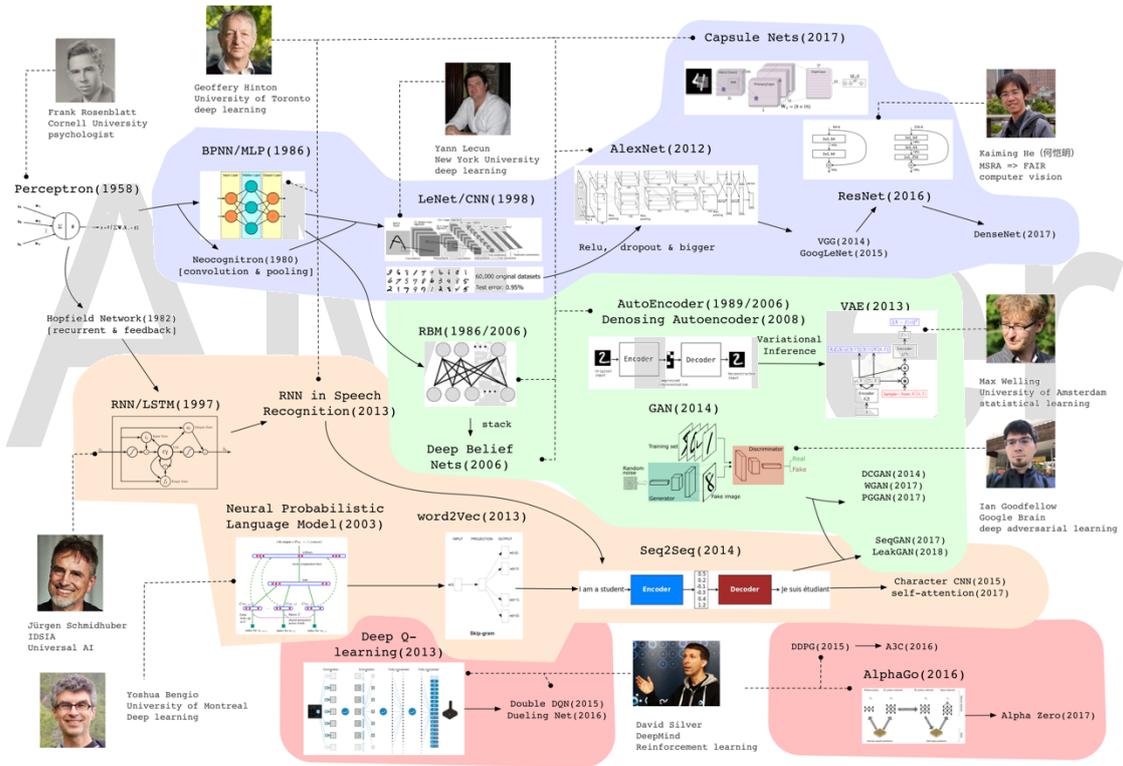


图 3-1 深度学习模型最近若干年的重要进展

以下各小节是对深度学习不同方面发展情况的梳理：

3.1 卷积神经网络

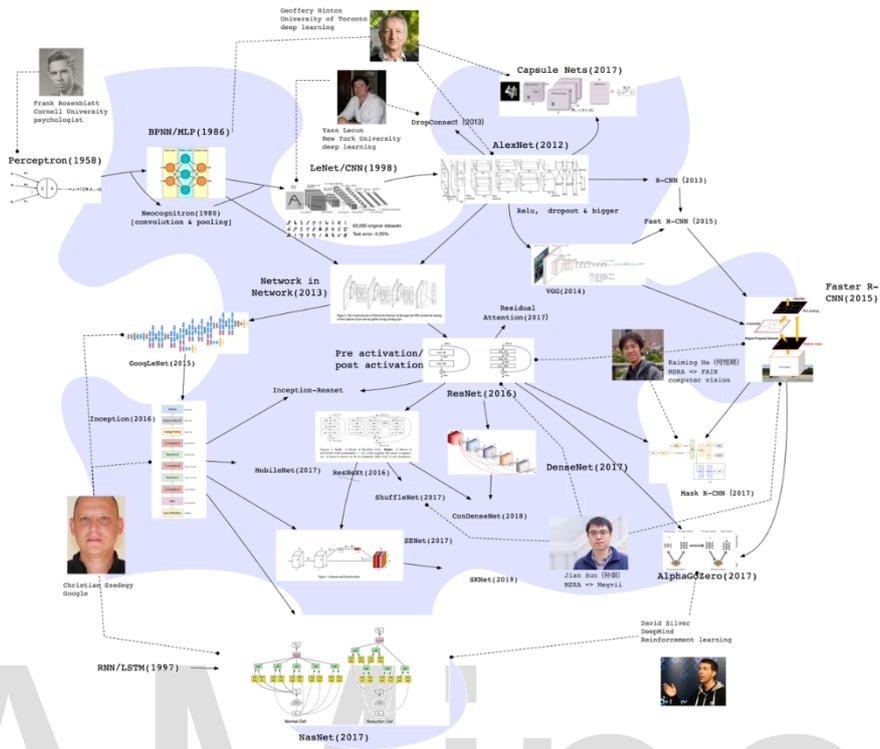


图 3-2 卷积神经网络的重要进展

3.2 AutoEncoder

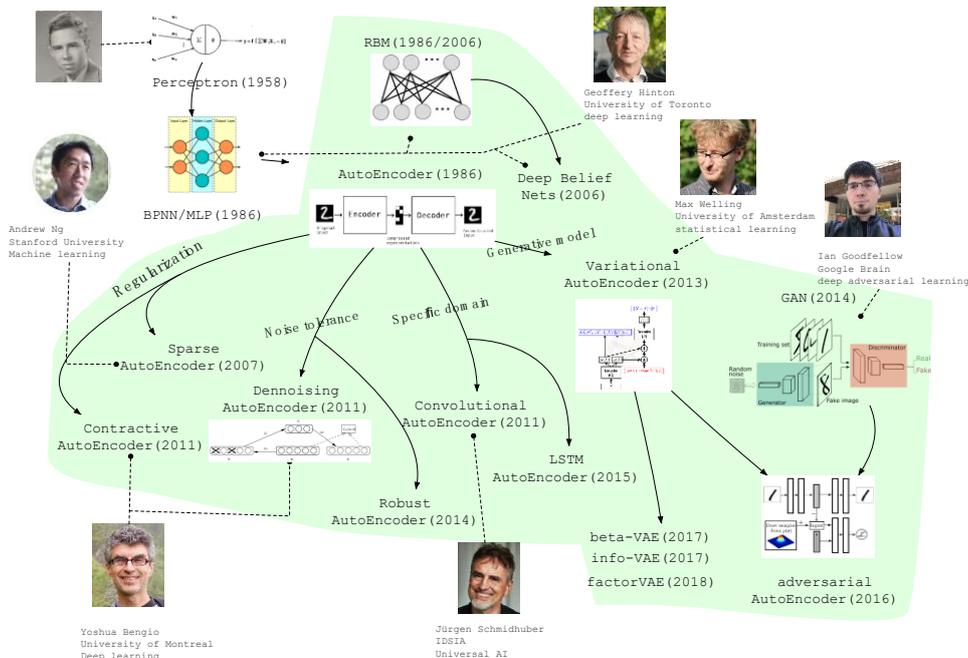


图 3-3 Auto-Encoder 的重要进展



3.3 循环神经网络 RNN

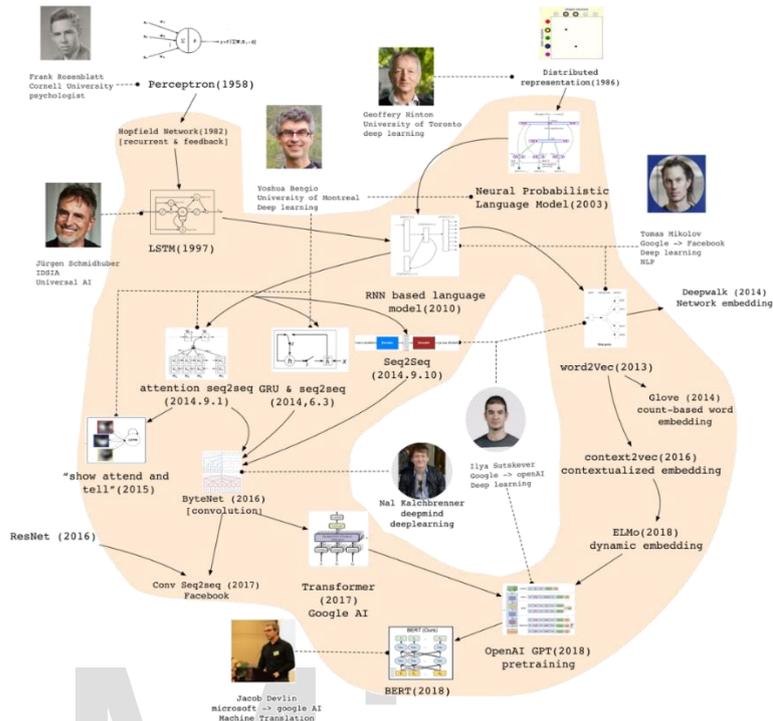


图 3-4 循环神经网络 RNN 的重要进展

3.4 网络表示学习与图神经网络 GNN

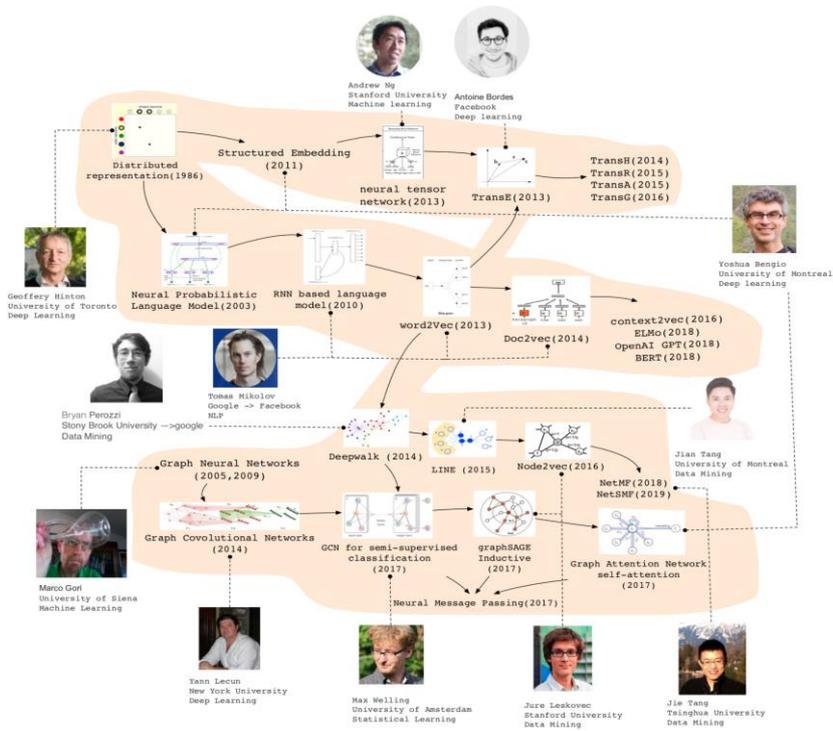


图 3-5 网络表示学习与图神经网络的重要进展

3.5 增强学习

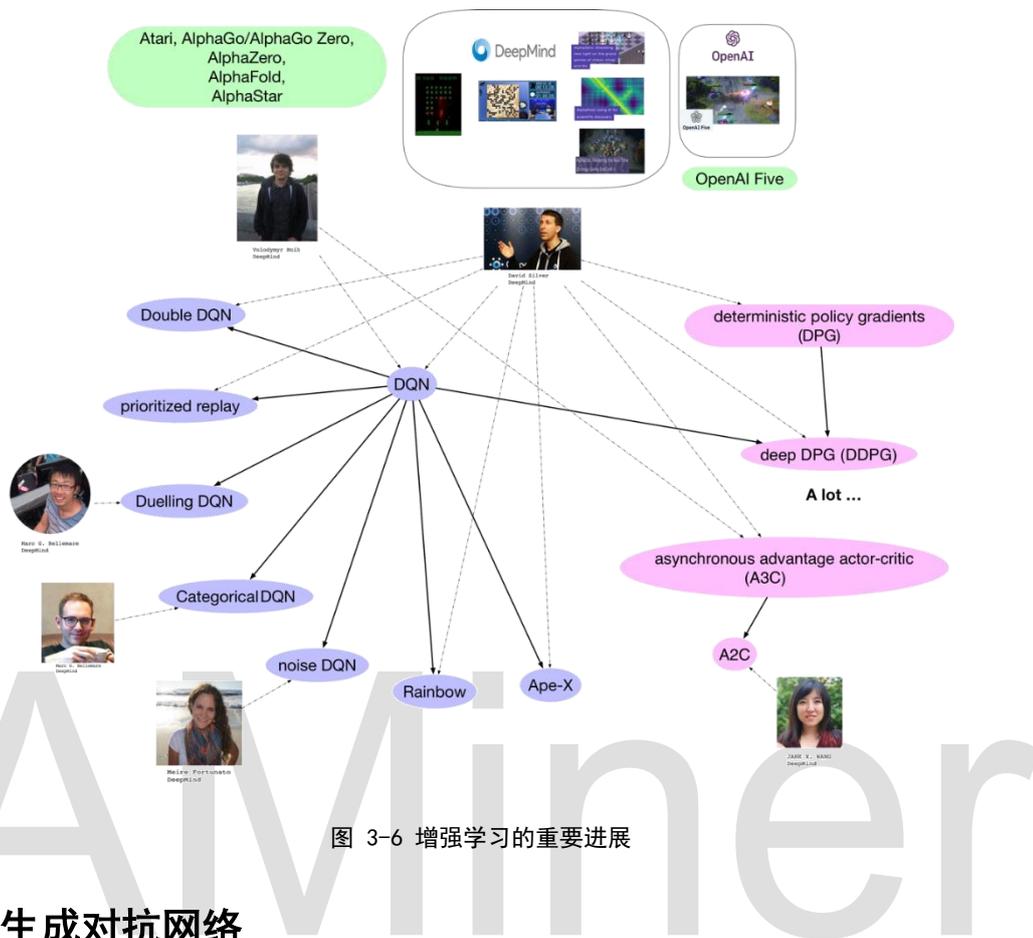


图 3-6 增强学习的重要进展

3.6 生成对抗网络

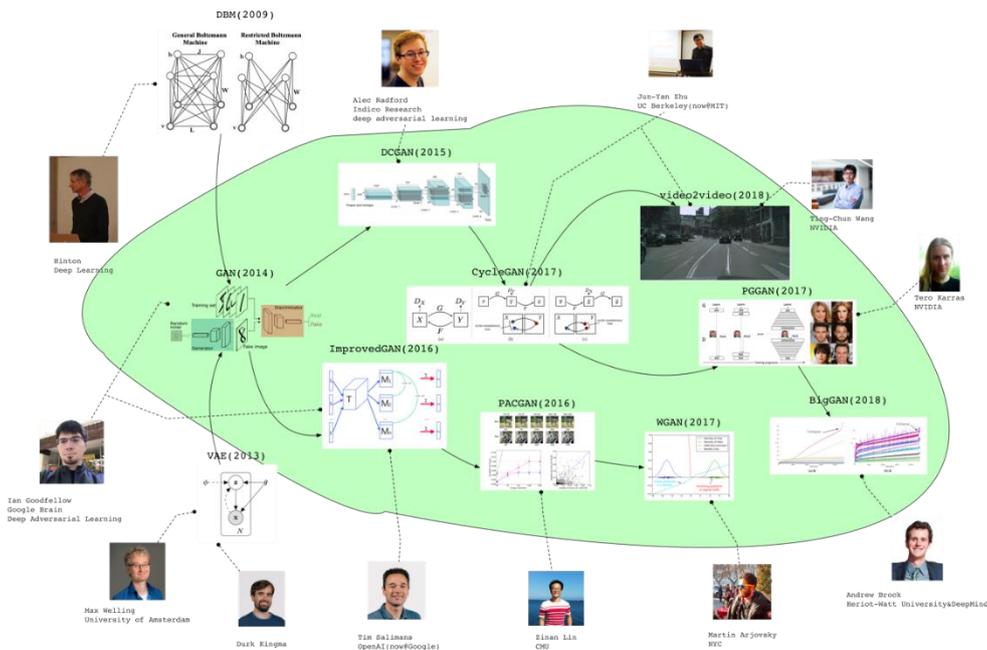


图 3-7 生成对抗网络的重要进展

3.7 老虎机

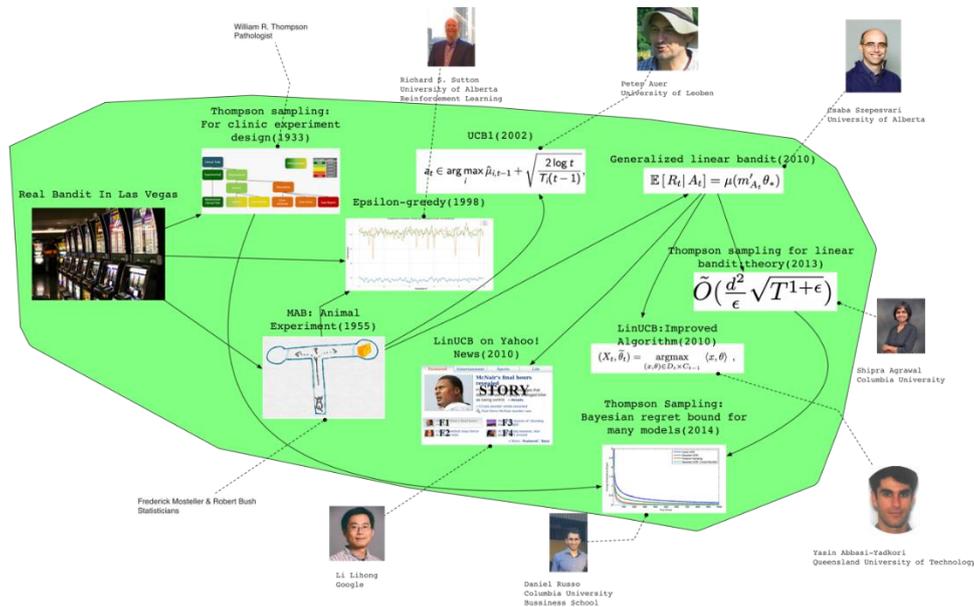


图 3-8 老虎机的重要进展

3.8 图神经网络

图神经网络的概念首先由 Gori 等人于 2005 年提出，并由 Scarselli 等人进一步阐明。这些早期的研究以迭代的方式通过循环神经网络传播邻近信息来学习目标节点的表示，直到达到稳定的固定点。该过程所需计算量庞大，而近来也有许多研究致力于解决这个难题。一般情况下，图神经网络代表的是所有用于图数据的深度学习方法。

受到卷积网络在计算机视觉领域所获巨大成功的激励，近来出现了很多为图数据重新定义卷积概念的方法。这些方法属于图卷积网络（GCN）的范畴。Bruna 等人于 2013 年提出了关于图卷积网络的第一项重要研究，他们基于谱图论（spectral graph theory）开发了一种图卷积的变体。自此，基于谱的图卷积网络不断改进、拓展、进阶。由于谱方法通常同时处理整个图，并且难以并行或扩展到大图上，基于空间的图卷积网络开始快速发展。这些方法通过聚集近邻节点的信息，直接在图结构上执行卷积。结合采样策略，计算可以在一个批量的节点而不是整个图中执行，这种做法有望提高效率。除了图卷积网络，近几年还开发出了很多替代的图神经网络。这些方法包括图注意力网络（GAT）、图自编码器、图生成网络以及图时空网络。

Battaglia 等人将图网络定位为从关系数据中学习的构建块，并在统一的框架下回顾了部分图神经网络。然而，他们整体的框架是高度抽象的，失去了每种方法在原论文中的见解。Lee 等人对图注意力模型（一种图神经网络）进行了部分调查。最近，Zhang 等人提出了一

项关于图深度学习的最新调查，却忽略了对图生成网络和图时空网络的研究。总之，现有的研究没有一个对图神经网络进行全面的回顾，只覆盖了部分图卷积神经网络且检查的研究有限，因此遗漏了图神经网络替代方法的最新进展，如图生成网络和图时空网络^[17]。

3.9 深度学习近期重要进展

在过去几年中，深度学习改变了整个人工智能的发展。深度学习技术已经开始在医疗保健，金融，人力资源，零售，地震检测和自动驾驶汽车等领域的应用程序中出现。至于现有的成果表现也一直在稳步提高。本小节将介绍深度学习近期的一些重要进展。

3.9.1 2018 年三大进展

- BERT 模型

BERT 的全称是 Bidirectional Encoder Representation from Transformers，是基于深度双向 Transformer 的预训练模型，能用所有层的上下文语境训练深度双向表征。自 Google 在 2018 年公布 BERT 在 11 项 nlp 任务中的卓越表现后，BERT 就成为 NLP 领域大火的模型。关于 BERT 的详细介绍请参见 2.9 节的内容。

- 视频到视频合成 (Video-to-Video Synthesis)

我们通常习惯由图形引擎创建的模拟器和视频游戏进行环境交互。虽然令人印象深刻，但经典方法的成本很高，因为必须精心指定场景几何、材料、照明和其他参数。一个很好的问题是：是否可以使用例如深度学习技术自动构建这些环境。NVIDIA 的研究人员解决了这个问题。他们的目标是在源视频和输出视频之间提供映射功能，精确描绘输入内容。作者将其建模为分布匹配问题，其目标是使自动创建视频的条件分布尽可能接近实际视频的条件分布。为实现这一目标，他们建立了一个基于生成对抗网络 (GAN) 的模型。在 GAN 框架内的关键思想是，生成器试图产生真实的合成数据，使得鉴别器无法区分真实数据和合成数据。他们定义了一个时空学习目标，旨在实现暂时连贯的视频^[18]。

- 图网络 (Graph Network)

DeepMind 联合谷歌大脑、MIT 等机构 27 位作者发表重磅论文“*Relational inductive biases, deep learning, and graph networks*”，提出“图网络” (Graph network)，将端到端学习与归纳推理相结合，有望解决深度学习无法进行关系推理的问题。作者认为组合泛化是人工智能实现与人类相似能力的首要任务，而结构化表示和计算是实现这一目标的关键，实现这个目标的关键是结构化的表示数据和计算。该论文讨论了图网络如何支持关系推理和组合泛化，为更复杂的、可解释的和灵活的推理模式奠定基础^[19]。

3.9.2 2019 年三大进展

- XLNet 模型

XLNet 是 CMU 与谷歌大脑提出的全新 NLP 模型，在 20 个任务上超过了 BERT 的表现，并在 18 个任务上取得了当前最佳效果，包括机器问答、自然语言推断、情感分析和文档排序。关于 XLNet 及其与 BERT 关系的详细介绍请参见 2.9 节的内容。

- MoCo

何恺明在其工作“*Momentum Contrast for Unsupervised Visual Representation Learning*”中提出了动量对比度 (MoCo) 用于无监督的视觉表示学习。从作为字典查找的对比学习的角度来看，作者构建了一个带有队列和移动平均编码器的动态字典。这样就可以实时构建大型且一致的词典，从而促进对比性的无监督学习。MoCo 在 ImageNet 分类的通用线性协议下提供了竞争性的结果。更重要的是，MoCo 学习到的表示将转移到下游任务。MoCo 可以胜过在 PASCAL VOC, COCO 和其他数据集上进行监督的预训练对等任务中的检测/细分任务，有时会大大超过它。这表明在许多视觉任务中，无监督和有监督的表征学习之间的鸿沟已被大大消除^[20]。

- DL System 2

Yoshua Bengio 在 NeuIPS 2019 上的报告“*FROM SYSTEM 1 DEEP LEARNING TO SYSTEM 2 DEEP LEARNING*”讨论了深度学习发展的方向，引起了广泛关注，清华大学的唐杰教授对其进行了深度解读，请关注学术头条公众号获取相关信息。

4 论文解读篇

ICML 和 NeurIPS 是机器学习领域的两个顶级学术会议。

- ICML

ICML, 全称 International Conference on Machine Learning, 是由国际机器学习学会 (IMLS) 主办的年度机器学习国际顶级会议。

- NeurIPS

NeurIPS, 全称 Conference and Workshop on Neural Information Processing Systems, 是机器学习和计算神经科学的国际顶级会议, 由 NIPS 基金会主办。

我们选取了 ICML 与 NeurIPS 近 10 年来的最佳论文奖, 如表 4-1、表 4-2 所列举, 并在完整版报告对这些获奖论文进行了解读。

表 4-1 ICML 近 10 年 best paper

ICML (International Conference on Machine Learning)		
年份	论文标题	作者
2019	Challenging Common Assumptions in the Unsupervised Learning of Disentangled Representations	Francesco Locatello, Stefan Bauer, Mario Lucic, Gunnar Rätsch, Sylvain Gelly, Bernhard Schölkopf, Olivier Bachem
	Rates of Convergence for Sparse Variational Gaussian Process Regression	David R. Burt, Carl E. Rasmussen, Mark van der Wilk
2018	Delayed Impact of Fair Machine Learning	Lydia T. Liu, University of California Berkeley; et al.
	Obfuscated Gradients Give a False Sense of Security: Circumventing Defenses to Adversarial Examples	Anish Athalye, Massachusetts Institute of Technology; et al.
2017	Understanding Black-box Predictions via Influence Functions	Pang Wei Koh & Percy Liang, Stanford University
2016	Ensuring Rapid Mixing and Low Bias for Asynchronous Gibbs Sampling	Christopher De Sa, Stanford University; et al.
	Pixel Recurrent Neural Networks	Aaron Van den Oord, Google; et al.
	Dueling Network Architectures for Deep Reinforcement Learning	Ziyu Wang, Google; et al.
2015	A Nearly-Linear Time Framework for Graph-Structured Sparsity	Chinmay Hegde, Massachusetts Institute of Technology; et al.
	Optimal and Adaptive Algorithms for Online Boosting	Alina Beygelzimer, Yahoo! Research; et al.
2014	Understanding the Limiting Factors of Topic Modeling via Posterior Contraction Analysis	Jian Tang, Peking University; et al.
2013	Vanishing Component Analysis	Roi Livni, The Hebrew University of Jerusalem; et al.
	Fast Semidifferential-based Submodular Function Optimization	Rishabh Iyer, University of Washington; et al.
2012	Bayesian Posterior Sampling via Stochastic Gradient Fisher Scoring	Sungjin Ahn, University of California Irvine; et al.



ICML (International Conference on Machine Learning)		
2011	Computational Rationalization: The Inverse Equilibrium Problem	Kevin Waugh, Carnegie Mellon University; et al.
2010	Hilbert Space Embeddings of Hidden Markov Models	Le Song, Carnegie Mellon University; et al.
2009	Structure preserving embedding	Blake Shaw, Tony Jebara, Columbia University

表 4-2 NeurIPS 近 10 年 best paper

NeurIPS (Neural Information Processing Systems)		
年份	论文标题	作者
2018	Non-delusional Q-learning and Value-iteration	Tyler Lu, Dale Schuurmans, Craig Boutilier
	Optimal Algorithms for Non-Smooth Distributed Optimization in Networks	Kevin Scaman, Francis Bach, Sebastien Bubeck, Laurent Massoulié, Yin Tat Lee
	Nearly Tight Sample Complexity Bounds for Learning Mixtures of Gaussians via Sample Compression Schemes	Hassan Ashtiani, Shai Ben-David, Ick Harvey, Christopher Liaw, Abbas Mehrabian, Yaniv Plan
	Neural Ordinary Differential Equations	Tian Qi Chen, Yulia Rubanova, Jesse Bettencourt, David Duvenaud
2017	Safe and Nested Subgame Solving for Imperfect-Information Games	Noam Brown, Tuomas Sandholm
	Variance-based Regularization with Convex Objectives	Hongseok Namkoong, John Duchi
	A Linear-Time Kernel Goodness-of-Fit Test	Wittawat Jitkittum, Wenkai Xu, Zoltan Szabo, Kenji Fukumizu, Arthur Gretton
2016	Value Iteration Networks	Aviv Tamar, Yi Wu, Garrett Thomas, Sergey Levine, Pieter Abbeel
	Matrix Completion has No Spurious Local Minimum	Rong Ge, Jason Lee, Tengyu Ma
	Interactive musical improvisation with Magenta	Adam Roberts, Jesse Engel, Curtis Hawthorne, Ian Simon, Elliot Waite, SAGEEV OORE, Natasha Jaques, Cinjon Resnick, Douglas Eck
2015	Competitive Distribution Estimation: Why is Good-Turing Good	Alon Orlitsky, Ananda Theertha Suresh
	Fast Convergence of Regularized Learning in Games	Vasilis Syrgkanis, Alekh Agarwal, Haipeng Luo, Robert Schapire
2014	Asymmetric LSH (ALSH) for sublinear time Maximum Inner Product Search (MIPS)	Anshumali Shrivastava, Ping Li
	A* Sampling	Chris J. Maddison, Daniel Tarlow, Tom Minka
2013	A Memory Frontier for Complex Synapses	Subhaneil Lahiri, Surya Ganguli
	Submodular Optimization with Submodular Cover and Submodular Knapsack Constraints	Rishabh Iyer, Jeff Bilmes
	Scalable Influence Estimation in Continuous-Time Diffusion Networks	Nan Du, Le Song, Manuel Gomez-Rodriguez, Hongyuan Zha
2012	No voodoo here! Learning discrete graphical models via inverse covariance estimation	Po-Ling Loh, Martin Wainwright
	Discriminative Learning of Sum-Product Networks	Robert Gens, Pedro Domingos
	Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials	Philipp Krähenbühl, Vladlen Koltun

NeurIPS (Neural Information Processing Systems)		
2011	Priors Over Recurrent Continuous Time Processes	Ardavan Saeedi, Alexandre Bouchard-Côte
	Fast and Accurate K-means for Large Datasets	Michael Shindler, Alex Wong, Adam Meyerson
2010	Construction of dependent dirichlet Processes based on Poisson Processes	Dahua Lin, Eric Grimson, John Fisher
	A Theory of Multiclass Boosting	Indraneel Mukherje, Robert E Schapire
2009	An LP View of the M-Best MAP Problem	Menachem Fromer, Amir Globerson
	Fast Subtree Kernels on Graphs	Nino Shervashidze, Karsten Borgwardt

AMiner

5 人才篇

本篇通过 AMiner 大数据平台挖掘机器学习领域顶级学术会议 ICML、NeurIPS 近 10 年的论文，提取论文中所有学者信息，从中选出 h-index 排名最靠前的 2000 位领域活跃学者，分析了学者的分布等情况，介绍了部分该领域国内外知名度较高的活跃学者。

5.1 学者情况概览

- 全球人才分布

学者地图用于描述特定领域学者的分布情况，对于进行学者调查、分析各地区竞争力现状尤为重要，下图为机器学习领域全球学者分布情况：



图 5-1 机器学习领域全球学者分布

地图根据学者当前就职机构地理位置进行绘制，其中颜色越深表示学者越集中。从该地图可以看出，美国的人才数量遥遥领先且主要分布在其东西海岸；欧洲中西部也有较多的人才分布；亚洲的人才主要分布于我国东部及日韩地区；其他诸如非洲、南美洲等地区的学者非常稀少；机器学习领域的人才分布与各地区的科技、经济实力情况大体一致。此外，在性别比例方面，机器学习领域中男性学者占比 89.8%，女性学者占比 10.2%，男性学者占比远高于女性学者。

- h-index 分布

机器学习学者的 h-index 分布如下图所示，大部分学者的 h-index 都在 30 以上，其中 h-index 小于 30 的人数最多，有 591 人，占比 29.1%。

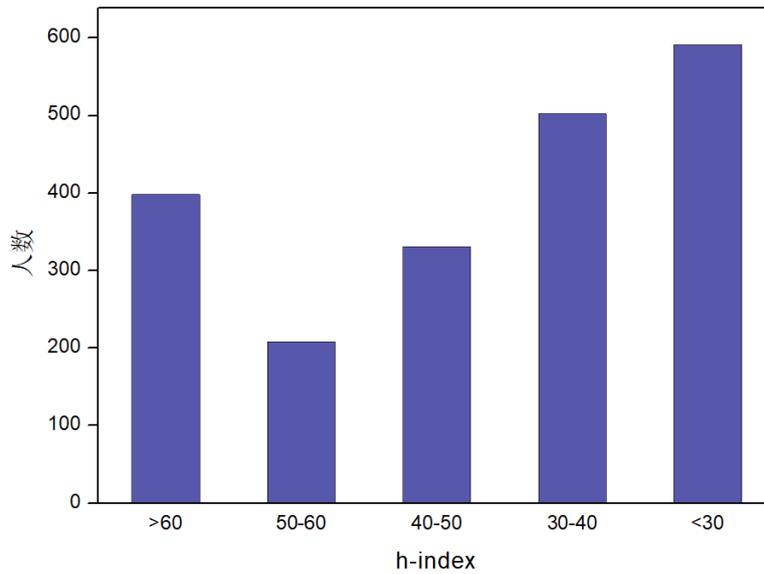


图 5-2 机器学习领域学者 h-index 分布

● 中国人才分布

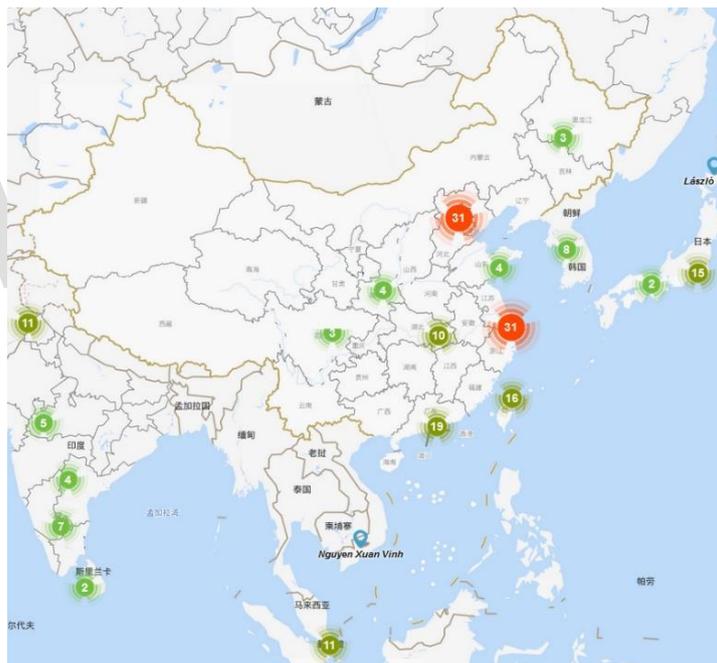


图 5-3 机器学习领域中国学者分布

我国专家学者在机器学习领域的分布如上图所示。通过上图我们可以发现，京津地区在本领域的人才数量最多，其次是长三角和珠三角地区，相比之下，内陆地区的人才较为匮乏，这种分布与区位因素和经济水平情况不无关系。同时，通过观察中国周边国家的学者数量情况，特别是与日韩、东南亚等亚洲国家相比，中国在机器学习领域学者数量较多。

中国与其他国家在机器学习的合作情况可以根据 AMiner 数据平台分析得到，通过统计论文中作者的单位信息，将作者映射到各个国家中，进而统计中国与各国之间合作论文的数量，并按照合作论文发表数量从高到低进行了排序，如下表所示。

表 5-1 机器学习领域中国与各国合作论文情况

序号	合作国家	论文数	引用数	平均引用数	总的学者数
1	中国-美国	511	26694	52	819
2	中国-英国	44	1398	32	73
3	中国-新加坡	36	1189	33	56
4	中国-澳大利亚	31	744	24	42
5	中国-印度	22	1123	51	19
6	中国-德国	17	419	25	39
7	中国-瑞士	11	233	21	22
8	中国-荷兰	6	93	16	10
9	中国-巴基斯坦	4	82	21	3
10	中国-以色列	3	23	8	6

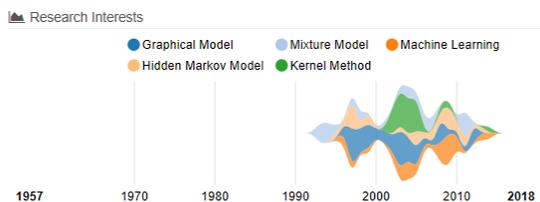
从上表数据可以看出，中美合作的论文数、引用数、平均引用数、学者数遥遥领先，表明中美间在机器学习领域合作之密切；从地域角度看，中国与欧洲的合作非常广泛，前 10 名合作关系里中欧合作共占 4 席；中国与印度合作的论文数虽然不是最多，但是平均引用数依然位列第二，说明在合作质量上中印合作也达到了较高的水平。

5.2 代表性学者简介

综合 h-index 以及领域知名度与活跃度，我们收集整理国内外机器学习领域的高水平学者，其中，国际代表性学者如：Michael I. Jordan、Yann Lecun、Geoffrey E. Hinton、Yoshua Bengio、Andrew Y Ng、Jurgen Schmidhuber、Fei-Fei Li、Daphne Koller、John D. Lafferty、Peter L. Bartlett、Michael Collins、Ian Goodfellow、David Sliver、Zoubin Ghahramani、David J.C. MacKay、Christopher Bishop、Tony Jebara、Max Welling 等；国内代表性学者如：张钹、周志华、李航、朱军、颜水成、杨强、唐杰、刘铁岩、王海峰、何晓飞、戴文渊、黄高、王立威、张长水、孙剑、林宙辰等。下面我们将对国内外机器学习领域代表性学者进行简要介绍，排名不分先后。此外，限于报告篇幅，我们对所有学者不能逐一罗列，如有疏漏，还请与 AMiner 编者联系，或者登录 <https://www.aminer.cn/> 获取更多资料。

5.2.1 国际代表性学者

● Michael I. Jordan



Michael I. Jordan，美国三院（美国国家科学院、美国国家工程院、美国艺术与科学院）院士，机器学习泰斗，被誉为人工智能领域的“根目录”之一，伯克利大学机器学习实验室 AMP Lab 联合主任，IEEE Fellow，ACM Fellow。Michael I. Jordan 是美国科学家，加州大学伯克利分校电子工程系、计算机科学和统计系杰出教授，机器学习、统计学和人工智能研究员。他是机器学习领域的领军人物之一，并且在 2016 年被 *Semantic Scholar* 称为世界上最有影响力的计算机科学家。同年也被 *AMiner* 评为机器学习最有影响力学者。

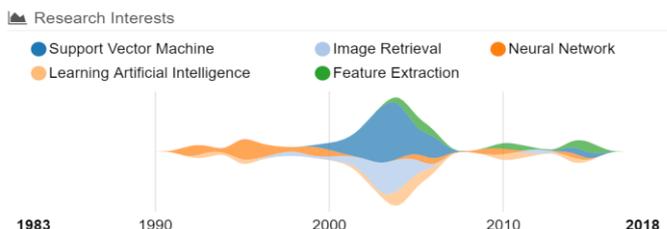
他于 1985 年获得加利福尼亚大学圣地亚哥分校博士学位。自 1988 年至 1998 年，Michael I. Jordan 任麻省理工学院教授，他的研究方向包括了计算学、统计学、认知科学以及生物科学。近年来，他的研究兴趣集中在贝叶斯非参数分析、概率图模型、谱方法、核方法、分布式计算系统、自然语言处理、信号处理和统计遗传学等问题的应用上。深度学习领域的权威 Yoshua Bengio，贝叶斯学习领域权威 Zoubin Ghahramani 及前百度首席科学家吴恩达等人都是其门下学生。

他曾获得众多奖项，在 2016 年获得 IJCAI 研究卓越奖 (IJCAI Research Excellence Award)，2015 年获得了 David E. Rumelhart 奖，2009 年获得了 ACM/AAAI Allen Newell 奖，2004 年获得 ICML 最佳学生论文奖。同时，他是 AAAI、ACM、ASA、CSS、IEEE、IMS、ISBA 和 SIAM 成员。

（其他国际代表性学者介绍请见完整版报告）

5.2.2 国内代表性学者

● 张钹



张钹，中国科学院院士，清华大学计算机科学与技术系教授，清华大学人工智能研究院院长。

张钹于 1958 年毕业于清华大学自动控制系，是国家第一批自动控制专业的毕业生。1995 年他当选为中国科学院院士。

他早期从事自动控制理论与系统研究，1979 年开始计算机科学与技术研究。从事人工智能理论、人工神经网络、遗传算法、分形和小波等理论研究；以及把上述理论应用于模式识别、知识工程、智能机器人与智能控制等领域的应用技术研究。

他针对人工智能问题求解计算复杂性、指数爆炸的主要困难，提出了问题分层求解的商空间理论，解决了不同粒度空间的描述、它们之间相互转换、复杂性分析等理论问题。在此基础上提出统计启发式搜索算法，基于拓扑的空间规划方法和关系矩阵的规划算法，对克服计算量的指数爆炸很有成效。还提出了研究不确定性处理、定性推理、模糊分析、证据合成等新原理。指导并参加建成了陆地自主车、图像与视频检索等实验平台。

张钹和同期同事成了国内最早接触到人工智能的研究者，并成为我国在这方面的首批专家。

在学术研究上的主要贡献是提出问题分层求解的商空间理论，通过代数的方法，系统地解决了不同层次求解空间的问题表达、复杂性分析、不同层次空间之间信息、算子及推理机制等的相互转换关系。在上述理论基础上，他进一步提出了统计启发式搜索算法，基于拓扑的空间规划方法以及基于关系矩阵的时间规划算法等，极大降低了计算复杂性，具有重要的应用价值。其专著《问题求解理论及应用》全面总结了他在人工智能理论研究中的成果，其英文版于 1992 年由 Elsevier Science Publishers B.V.(Nortn-Holland)出版，中文版获国家教委颁发的高校出版社优秀学术专著特等奖。澳大利亚专家 Ronald Walts 在计算机杂志 *The Australian Computer Journal* (1995) 对《问题求解理论及应用》(英文版)的评论为“这是一部在重要研究领域的优秀著作”。美国学者 Harold S.Stone 认为，张钹等在统计启发式搜索等方面的工作，是“最近几年中国学者作出的很有意义的贡献”，“将新一代计算技术的前沿向前推进了”。

他在国内外共发表论文 100 多篇，中英文专著有《问题求解理论及应用》(中英版)以及《人工神经网络理论及应用》等。

他于 1994 年当选为俄罗斯自然科学院外籍院士；1995 年当选为中国科学院院士；2011 年德国汉堡大学授予自然科学名誉博士；2015 年 1 月 31 日张钹获得 2014 CCF 终身成就奖。

他的社会任职有：智能技术与系统国家重点实验室主任、校学位委员会副主任、信息科学与技术学院学术委员会主任；中国自动化学会机器人专业委员会副主任及智能控制专业委

员会副主任；《计算机学报》副主编；国家高技术“863”计划智能机器人主题专家组成员；河南科技大学兼职院士；计算机学术委员会主任。

（其他国内代表性学者介绍请见完整版报告）

5.3 NeurIPS 十年高引学者

本月，NeurIPS 2019 在加拿大温哥华正式拉开帷幕。作为机器学习领域最重要的顶会，NeurIPS 一直有着很强的影响力和排名，被认为是神经计算方面最好的会议之一。随着近几年深度学习的崛起，NeurIPS 不仅成为了学术界的新星，也引起了工业界的高度关注，注册人数从数年前的几百人跃升到今年的近万人。根据 AMiner 数据平台的统计分析，NeurIPS 的 H5 指数为 149，10H 值为 34641，在人工智能方向会议中排名第二。通过对 NeurIPS 近十年接收论文引用量的统计分析，我们评选出了 NeurIPS 高引学者 TOP100 榜单。

表 5-2 NeurIPS 高引学者 TOP100

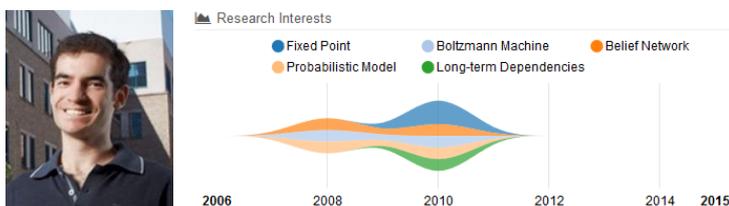
序号	姓名	单位	总引用量
1	Ilya Sutskever	OpenAI	67457
2	Geoffrey E. Hinton	多伦多大学	47482
3	Alex Krizhevsky	Google	44218
4	Yoshua Bengio	蒙特利尔大学	18714
5	Greg Corrado	Google	17218
6	Jeffrey Dean	Google	17218
7	Kai Chen	Google	16139
8	Tomas Mikolov	Facebook	15166
9	Ian Goodfellow	Apple	13480
10	Kaiming He	Facebook AI	12605
11	Jian Sun	旷视科技	12601
12	Aaron Courville	蒙特利尔大学	12407
13	Ross B. Girshick	Facebook AI	11495
14	Shaoqing Ren	Momenta	11093
15	Mehdi Mirza	DeepMind	10713
16	Jean Pouget-Abadie	Google	10618
17	Bing Xu	布兰迪斯大学	10618
18	David Warde-Farley	DeepMind	10618
19	Sherjil Ozair	蒙特利尔大学	10618
20	Quoc V. Le	Google	10180
21	Oriol Vinyals	DeepMind	9663
22	Andrew Y. Ng	斯坦福大学	6632
23	Andrew Zisserman	牛津大学	5570
24	Koray Kavukcuoglu	DeepMind	5223
25	Karen Simonyan	DeepMind	4883

26	Ruslan Salakhutdinov	苹果 AI	4456
27	Antonio Torralba	麻省理工学院	4249
28	Ryan P. Adams	普林斯顿大学	4116
29	Francis Bach	法国国家信息与自动化研究所	3350
30	Xi Chen	纽约大学	3257
31	Honglak Lee	Google Brain	3248
32	David M. Blei	哥伦比亚大学	3191
33	Lukasz Kaiser	Google AI	3146
34	Jason Yosinski	Uber 人工智能实验室	3042
35	Hugo Larochelle	Google Brain	2992
36	Jeff Clune	Uber 人工智能实验室	2989
37	Tong Zhang	香港大学	2889
38	Marc' Aurelio Ranzato	Facebook AI	2844
39	Hod Lipson	哥伦比亚大学	2802
40	Pieter Abbeel	加州大学伯克利分校	2761
41	Jasper Snoek	Google Brain	2734
42	Jakob Uszkoreit	Google	2710
43	Wojciech Zaremba	OpenAI	2669
44	Yann LeCun	Facebook	2616
45	Tim Salimans	OpenAI	2566
46	Pradeep D. Ravikumar	卡内基梅隆大学	2541
47	Jason Weston	Facebook	2539
48	Rob Fergus	纽约大学	2441
49	Michael I. Jordan	加州大学伯克利分校	2427
50	Samy Bengio	Google	2354
51	Richard Socher	Salesforce	2332
52	Eric Poe Xing	卡内基梅隆大学	2303
53	Christopher D. Manning	斯坦福大学	2248
54	Martin J. Wainwright	加州大学伯克利分校	2176
55	Aude Oliva	麻省理工人工智能实验室	2113
56	Jurgen Schmidhuber	慕尼黑工业大学	2082
57	Jianxiong Xiao	普林斯顿大学计算机科学系	2079
58	Rajat Monga	Google	2052
59	Matthieu Devin	Google	2052
60	Mark Z. Mao	Google	2052
61	Andrew Senior	纽约州立大学帕切斯学院	2052
62	Paul A. Tucker	北卡罗莱纳州立大学	2052
63	Ke Yang	耶鲁大学	2052
64	Bolei Zhou	香港中文大学	1914
65	Agata Lapedriza	西班牙巴塞罗那自治大学	1914
66	Richard Zemel	多伦多大学	1897
67	Nathan Srebro	丰田工业大学芝加哥分校	1848
68	Vladlen Koltun	Intel Labs	1829

69	Inderjit S. Dhillon	德克萨斯大学奥斯汀分校	1826
70	Razvan Pascanu	加拿大蒙特利尔大学	1811
71	Alexander J Smola	亚马逊	1754
72	Benjamin Recht	加州大学伯克利分校	1719
73	Antoine Bordes	Facebook	1689
74	Nicolas Usunier	Facebook	1631
75	Joshua B. Tenenbaum	麻省理工学院	1628
76	Kyunghyun Cho	纽约大学	1626
77	Hao Su	加州大学圣地亚哥分校	1589
78	Jure Leskovec	斯坦福大学	1569
79	Faruk Ahmed	加州大学伯克利分校	1554
80	Vincent Dumoulin	Google	1554
81	Jifeng Dai	商汤科技	1508
82	Philipp Krahenbuhl	斯坦福大学	1503
83	Trevor Darrell	加州大学伯克利分校	1499
84	John Schulman	OpenAI	1488
85	Yan Duan	加拿大蒙特利尔算法研究所	1470
86	Chong Wang	微软研究院	1439
87	Volodymyr Mnih	DeepMind	1428
88	Song Han	麻省理工学院	1427
89	William J. Dally	斯坦福大学	1427
90	Prateek Jain	微软研究院	1413
91	Stephen J. Wright	威斯康星大学	1412
92	Sewoong Oh	韩国首尔延世大学	1409
93	Xiaoou Tang	香港中文大学	1405
94	Dit-Yan Yeung	香港科技大学	1400
95	John Wright	哥伦比亚大学	1374
96	Arvind Ganesh	伊利诺伊大学香槟分校	1374
97	Shankar Rao	伊利诺伊大学香槟分校	1374
98	Yigang Peng	商汤科技	1374
99	Yi Ma	加州大学伯克利分校	1374
100	Rein Houthoofd	乐元素	1359

下面是本榜单 TOP10 学者的简介，其中 Geoffrey E. Hinton、Yoshua Bengio、Ian Goodfellow 已在上节中进行了介绍，本节不再重复。

● Ilya Sutskever



Ilya Sutskever, 目前担任 OpenAI 的首席科学家。他对深度学习领域做出了几项重大贡献。他是卷积神经网络 AlexNet 的联合创始人。他与 Oriol Vinyals 和 Quoc Le 一起发明了从序列到序列的学习方法。Sutskever 也是 AlphaGo 和 TensorFlow 的共同发明人。

他获得了多伦多大学计算机科学系的计算机科学学士学位和博士学位。2012 年毕业后, Sutskever 在斯坦福大学跟随 Andrew Ng 做了两个月的博士后。然后他回到多伦多大学, 加入了欣顿的新研究公司 DNNResearch, 这是欣顿研究小组的一个分支。四个月后谷歌收购了 DNNResearch, 并聘请 Sutskever 为谷歌 Brain 的研究科学家。在谷歌大脑中, Sutskever 与 Oriol Vinyals 和 Quoc Le 一起创建了序列到序列的学习算法。2015 年, Sutskever 被麻省理工学院技术评论评为 35 岁以下的 35 名创新者。2015 年底, 他离开谷歌, 成为新成立的 OpenAI 研究所的所长。Sutskever 是 NVIDIA NTECH 2018 和 AI Frontiers Conference 2018 的主题演讲者

(其他 TOP10 学者介绍请见完整版报告)

AMiner

6 应用篇

机器学习是人工智能研究的核心内容，它的应用已遍及人工智能的各个分支，随着机器学习能力的增强和技术的发展，其应用前景也十分广泛，近年来我们看到机器学习与金融、自动驾驶、健康和医疗、制造、零售等行业产生更为紧密的融合，并开始实现大规模的商业应用。本篇将介绍机器学习在各行业中的应用，此外，机器学习算法的应用场景以及在企业中的应用将在完整版报告中进一步介绍。

6.1 金融行业应用

- 欺诈检测

使用机器学习进行欺诈检测时，先收集历史数据并将数据分割成三个不同的部分，然后用训练集对机器学习模型进行训练，以预测欺诈概率。最后建立模型，预测数据集中的欺诈或异常情况。与传统检测相比，这种欺诈检测方法所用的时间更少。由于目前机器学习的应用量还很小，仍然处于成长期，所以它会在几年内进一步发展，从而检测出复杂的欺诈行为。

- 股票市场预测

当今，股票市场俨然已成为大家关注的热点，但是，如果不了解股票运作方式和当前趋势，要想击败市场则非常困难。随着机器学习的使用，股票预测变得相当简单。这些机器学习算法会利用公司的历史数据，如资产负债表、损益表等，对它们进行分析，并找出关系到公司未来发展的有意义的迹象。

- 财资部（Treasury）/客户关系管理（CRM）/现货交易（Spot Transactions）

客户关系管理（CRM）在小额银行业务中占有十分突出的地位，但在银行内部的财资空间却没什么作用。因为财资部有自己的产品群，如外汇、期权、掉期交易（Swaps）、远期交易（Forwards）以及更为重要的现货交易（Spots）。线上交易需要结合这些产品的复杂程度、客户风险、市场与经济行为以及信用记录信息，这对银行来说几乎是一个遥远的梦想。

- 聊天机器人/私人财务助理

聊天机器人可以担当财务顾问，成为个人财务指南，可以跟踪开支，提供从财产投资到新车消费方面的建议。财务机器人还可以把复杂的金融术语转换成通俗易懂的语言，更易于沟通。一家名为 Kasisto 的公司的聊天机器人就能处理各种客户请求，如客户通知、转账、支票存款、查询、常见问题解答与搜索、内容分发渠道、客户支持、优惠提醒等^[21]。

6.2 自动驾驶

将汽车内外传感器的数据进行融合，借此评估驾驶员情况、进行驾驶场景分类，都要用到机器学习。自动驾驶汽车的设计制造面临着诸多挑战，如今，各大公司已经广泛采用机器学习寻找相应的解决方案。汽车中的 ECU（电子控制单元）已经整合了传感器数据处理，如何充分利用机器学习完成新的任务，变得至关重要。潜在的应用包括将汽车内外传感器的数据进行融合，借此评估驾驶员情况、进行驾驶场景分类。这些传感器包括像激光雷达，雷达，摄像头或者是物联网。

车载信息娱乐系统所运行的应用，能从传感器数据融合系统中获取数据。举个例子，如果系统察觉驾驶员发生状况，有能力把车开到医院。基于机器学习的应用，还包括对驾驶员的语言和手势识别以及语言翻译。相关的算法被分类为非监督和监督算法。它们两者的区别在于学习的方式。

在自动驾驶汽车上，机器学习算法的一个主要任务是持续渲染周围的环境，以及预测可能发生的变化。这些任务可以分为四个子任务：目标检测、目标识别或分类、目标定位、运动预测。

机器学习算法可以简单地分为 4 类：决策矩阵算法、聚类算法、模式识别算法和回归算法。可以利用一类机器学习算法来完成两个以上的子任务。例如，回归算法能够用于物体定位和目标识别或者是运动预测^[22]。

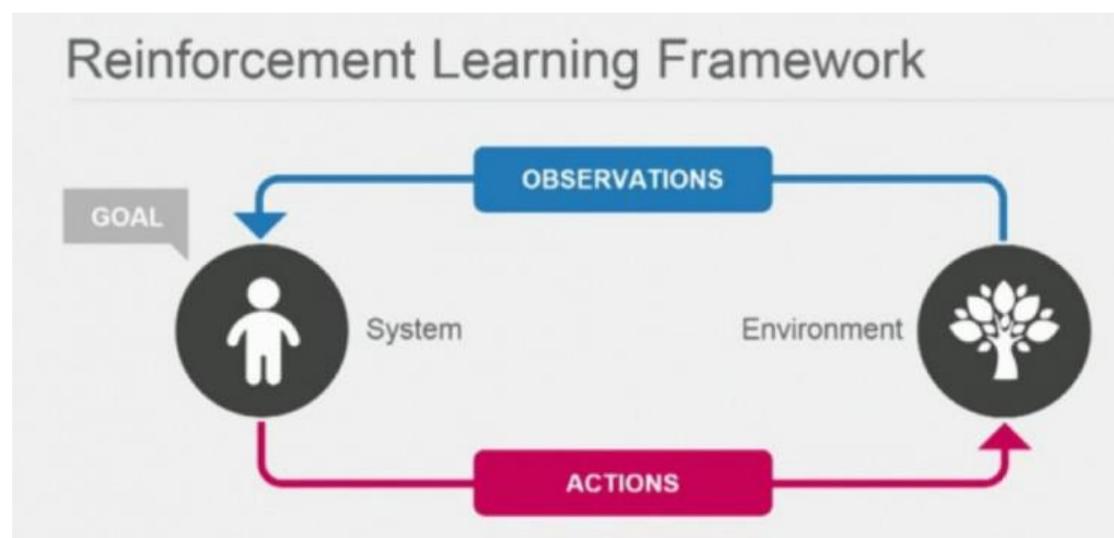


图 6-1 自动驾驶目标识别、运动预测

6.3 健康和医疗

为了更好地了解人工智能和机器学习正如何改变医疗保健行业,我们通过一些具体案例,并且这些案例可以有力证明这些前沿技术的实用价值。

- 判断发展中国家的结核病情况

识别图像中的模式 (Pattern) 是现有人工智能系统中最强有力的一点,研究人员现在正在训练人工智能检查胸部 x 光片,识别结核病。这项技术可以为缺乏放射学家的结核病流行地区带来有效的筛查和评估手段。

- 一种治疗退伍军人创伤后应激障碍 (PTSD) 的人工智能

退伍军人创伤后成长计划与 IBM Watson 合作使用人工智能和分析技术,以确保更多患有创伤后应激障碍的退伍军人能够完成心理治疗。使用这些技术,他们的完成率从不到 10% 上升到 73%。根据退伍军人事务部的统计,80% 的患有创伤后应激障碍的退伍军人在确诊后一年内完成治疗计划,然后康复。在 300 万阿富汗和伊拉克战争老兵中,大约五分之一患有创伤后应激障碍。

- 检测脑出血

以色列医疗技术公司 MedyMatch 和 IBM Watson Health 正在使用人工智能,通过检测颅内出血,帮助医院急诊室的医生更有效地治疗中风和头部外伤患者。AI 系统使用临床洞察力 (clinical insight)、深度学习、患者数据和机器视觉来自动标记潜在的脑出血,以供医生检查。

- 优化管理工作流程并消除等待时间

行政和助理工作是 AI 起作用的主要领域。埃森哲表示,语音到文本转录等省时的 workflow 功能有可能替代为医疗专业人员订购测试和处方以及在图表中写笔记等任务--任何涉及非患者护理的任务。这相当于节省了 17% 的医生工作时间和 51% 的注册护士工作时间。

- 检测阿尔茨海默病

现在,人工智能机器人只需要不到一分钟的时间,就可以根据语音模式和声音来诊断阿尔茨海默氏病,准确率达到 82%,而且这种准确率还在不断提高。人工智能系统可以处理单词之间的停顿长度、处理代词优于专有名词的任何偏好、处理过于简单的描述以及语音频率和幅度的变化。所有这些因素对于人类听众来说都很难高精度地记录和检测,但人工智能系统却能够进行客观和可量化的分析。

- 癌症诊断

用于检测和诊断癌症的传统方法包括计算机断层扫描（CT）、磁共振成像（MRI）、超声和 X 射线。不幸的是，许多癌症无法通过这些技术得到足够准确的诊断，从而可靠地挽救生命。微阵列基因图谱的分析是一种替代方法，但这项技术需要计算很多小时，除非这项技术可以使用 AI 替换。现在已经被证明，斯坦福大学的人工智能诊断算法与由 21 名经委员会认证的皮肤科医生的团队一样有效地从图像中检测潜在的皮肤癌。Startup Enlitic 正在使用深度学习来检测 CT 图像中的肺癌结节，其算法比作为一个团队工作的专家胸科医生的准确率高 50%。

● 机器人辅助手术

在价值潜力方面，机器人辅助手术是人工智能辅助方向的佼佼者。AI-enabled 机器人技术可以通过集成实时操作矩阵、来自实际手术医生的数据以及来自手术前病历的信息来提高和指导手术器械的精度。事实上，埃森哲报告说，人工智能机器人技术带来的进步缩短了 21% 的停留时间^[23]。

6.4 零售业

IDC 副总裁 Ivano Ortis 最近分享了他的观点“人工智能将把分析带到一个新的水平，并将成为零售创新的基础，这已经得到了全球半数零售商的认可。人工智能可以实现规模化、自动化和前所未有的精度，当适用于超细微客户细分和上下文交互的时候，可推动客户体验”。

鉴于人工智能和机器学习的能力，很容易看到人工智能和机器学习是如何成为零售商强大的工具。现在，计算机可以读取、倾听和了解数据，从数据中进行学习，立即且准确地推荐下一个最佳动作，而不需要明确的编程。这对那些希望能够准确预测需求、预期客户行为、优化和个性化客户体验的零售商来说是一个福音。

零售业已经开始感受到人工智能和机器学习的影响了，例如：

- 零售商正在通过机器学习结合物联网技术来预测需求，优化商店业务并减轻员工负担。
- 基于店内摄像头检测提供个性化的广告，承担店员部分的半手动的、通过在平板电脑或者触屏终端设备查看客户的消费记录。
- 零售商可以监控排队结账的等候时间，以了解个别店面的流量和商店销售效率，然后进行分类和调整店面布局来实现购物篮、满意度和销售的最大化。
- 系统现在可以通过把计划调整为按需活动，来识别和预测客户行为，改善员工生产效率。
- 摄像头系统可以在店内员工之前检测易腐产品的新鲜状态。

- 实体店正在实现很多操作任务的自动化，例如设置货架定价，确定产品分类和混合，优化促销等。
- 店内应用可以显示客户在特定通道停留了多长时间，根据个人消费记录和偏好数据，提供有针对性的优惠和建议（通过他/她的移动设备）。

机器学习可以帮助减轻推动利用可用数据所需的分析任务。当部署了一个全公司范围的、实施的分析平台时，这将成为所有公司职能优化决策所依赖的事实来源^[24]。

6.5 制造业

与自动驾驶汽车一样，随着物联网的发展，制造业企业可以从安置在生产线各环节的传感器收集大量的生产数据。

然而，这些数据并没有被充分利用。随着从复杂系统收集到众多参数的数据，数据分析变成了一项艰巨的任务。机器学习在制造业中的最大应用将是异常检测。

据统计，到 2030 年，全球的淡水需求预计将超过供应近 40%。为协助各企业实现净零水循环使用的目标，美国水处理公司 Ecolab（艺康集团）正通过包括 Azure 和 Dynamics CRM Online 在内的微软云平台帮助全球企业实现可持续运营。

与全球各地数以千计传感器相连的云平台能收集实时用水数据，并通过机器学习和商业智能分析全球各地的生产用水运营解决方案，不仅提高效率，还能降低水、能源消耗及运营成本。

尽管在这个领域之前已经进行过一些分析尝试，未来将会有更多机器学习通过监督学习和建模来预测风险和失败。

此外，机器学习也将推动工业自动化的实现，通过观察生产线和数据流来学习，并能够精确优化生产过程，降低生产成本，加快生产周期，从而节省人工分析数据的时间成本和资金成本^[25]。



7 趋势篇

领域技术分析系统 (<http://trend.aminer.cn>) 可以基于 AMiner 超过 2 亿篇论文的数据进行深入挖掘, 对技术趋势、国际趋势、机构趋势及学者趋势等方面进行分析。机器学习是一门多领域交叉学科, 设计概率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科。在下面的各种趋势分析中, 研究基础是期刊/会议 *Journal of Machine Learning Research*、*Machine Learning*、*International Conference on Machine Learning*、*Conference and Workshop on Neural Information Processing Systems* 近 20 年的论文。

● 技术趋势

技术趋势分析如下图所示。技术趋势分析描述了技术的出现、变迁和消亡的全过程, 可以帮助研究人员理解领域的研究历史和现状, 快速识别研究的前沿热点问题。图中每条色带表示一个话题, 其宽度表示该术语在当年的热度, 与当年该话题的论文数量呈正相关, 每一年份中按照其热度由高到低进行排序。通过技术趋势分析可以发现当前该领域的热点研究话题 Top10 是: Neural Network、Machine Learning、Deep Neural Networks、Deep Learning、Support Vector Machine、Reinforcement Learning、Feature Selection、Deci Tree、Data Mining、Artificial Neural Network。

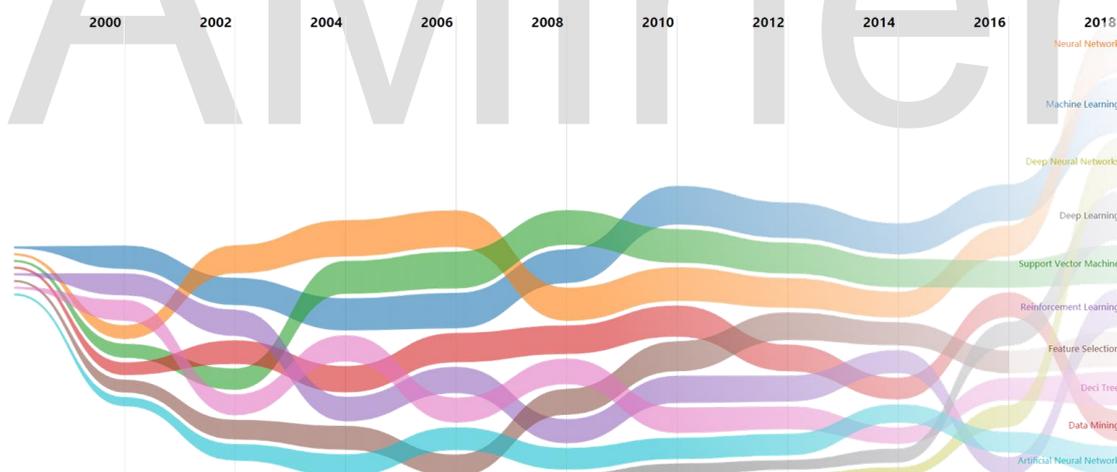


图 7-1 机器学习技术趋势

根据技术趋势分析我们可以发现, 该领域当前最热门的话题是 Neural Network, 从全局热度来看, Neural Network 一直保持着较高的话题热度, 2002-2006 年期间保持着最高的热度并于 2018 年重登榜首。

(国际趋势、机构趋势及学者趋势等方面分析请见完整版报告)

8 资源篇

本篇中我们搜集整理了机器学习领域的若干资源,希望能对读者朋友更好地了解机器学习有所帮助,同时也欢迎大家补充。

8.1 开源代码

1) 深度学习框架 Pytorch

<https://github.com/pytorch/pytorch>

2) 深度学习框架 Tensorflow

<https://github.com/tensorflow/tensorflow>

3) 机器学习库 scikit-learn

<https://github.com/scikit-learn/scikit-learn>

4) 自然语言处理时下最流行的 Transformer 库

<https://github.com/huggingface/transformers>

5) 自然语言处理工业级别库 spaCy

<https://github.com/explosion/spaCy>

6) 自然语言处理 Toolkit NLTK

<https://github.com/nltk/nltk>

7) 自然语言处理经典模型库 Gensim

<https://github.com/RaRe-Technologies/gensim>

8) 图计算库 NetworkX

<https://github.com/networkx/networkx>

9) XGBoost

<https://github.com/dmlc/xgboost>

10) 深度学习 FastAI

<https://github.com/fastai/fastai>



8.2 预训练

1) Glove

<https://nlp.stanford.edu/projects/glove/>

2) FastText

<https://fasttext.cc/>

3) MUSE

<https://github.com/facebookresearch/MUSE>

4) ELMo

<https://allennlp.org/elmo>

5) BERT

<https://github.com/google-research/bert>

6) XLNet

<https://github.com/zihangdai/xlnet>

7) XLM

<https://github.com/facebookresearch/XLM>

8) OpenAI-GPT2

<https://github.com/openai/gpt-2-output-dataset>

9) ResNet, VGG

<https://keras.io/applications/>

10) YOLOv2

<https://github.com/experiencor/keras-yolo2>

8.3 课程

1) stanford cs231n ML

<http://cs231n.stanford.edu/>

2) FastAI ML

<https://www.fast.ai/2018/09/26/ml-launch/>

3) OpenAI RL

<https://gym.openai.com/>

4) stanford cs234 RL

<https://web.stanford.edu/class/cs234/index.html>

5) CMU 10701 ML

https://www.cs.cmu.edu/~lwehbe/10701_S19/

6) CMU 11747 NN4NLP

<http://www.phontron.com/class/nn4nlp2019/description.html>

7) Coursera ML

<https://www.coursera.org/learn/machine-learning>

8) edX

<https://www.edx.org/course/machine-learning>

9) udacity

<https://www.udacity.com/course/intro-to-machine-learning-nanodegree--nd229>

10) deeplearning.ai

<https://www.coursera.org/specializations/deep-learning>

8.4 数据集

1) [CV] ImageNet

<http://www.image-net.org/>

2) [CV] CoCo

<http://cocodataset.org/>

3) [CV] PASCAL VOC

<http://host.robots.ox.ac.uk/pascal/VOC/index.html>

4) MNIST



<http://yann.lecun.com/exdb/mnist/>

5) [NLP] GLUE

<https://gluebenchmark.com/>

6) [NLP] XNLI

<https://github.com/facebookresearch/XNLI>

7) [Recommendation] MovieLens

<https://grouplens.org/datasets/movielens/>

8) [NLP] WikiText

<https://blog.einstein.ai/the-wikitext-long-term-dependency-language-modeling-dataset/>

9) [NLP] SQuAD

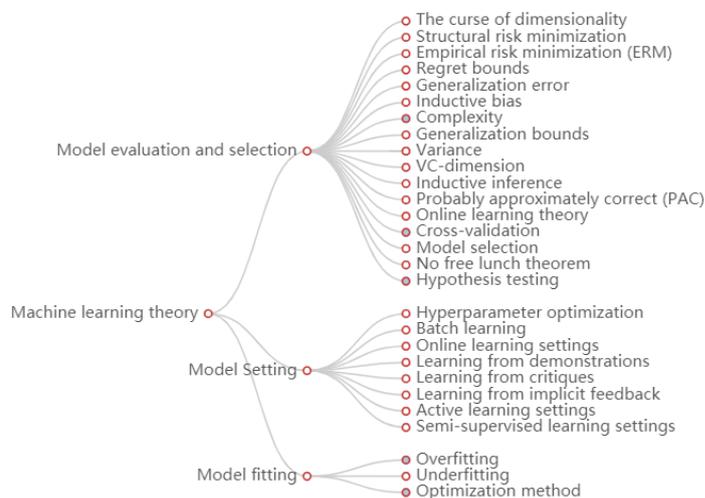
<https://rajpurkar.github.io/SQuAD-explorer/>

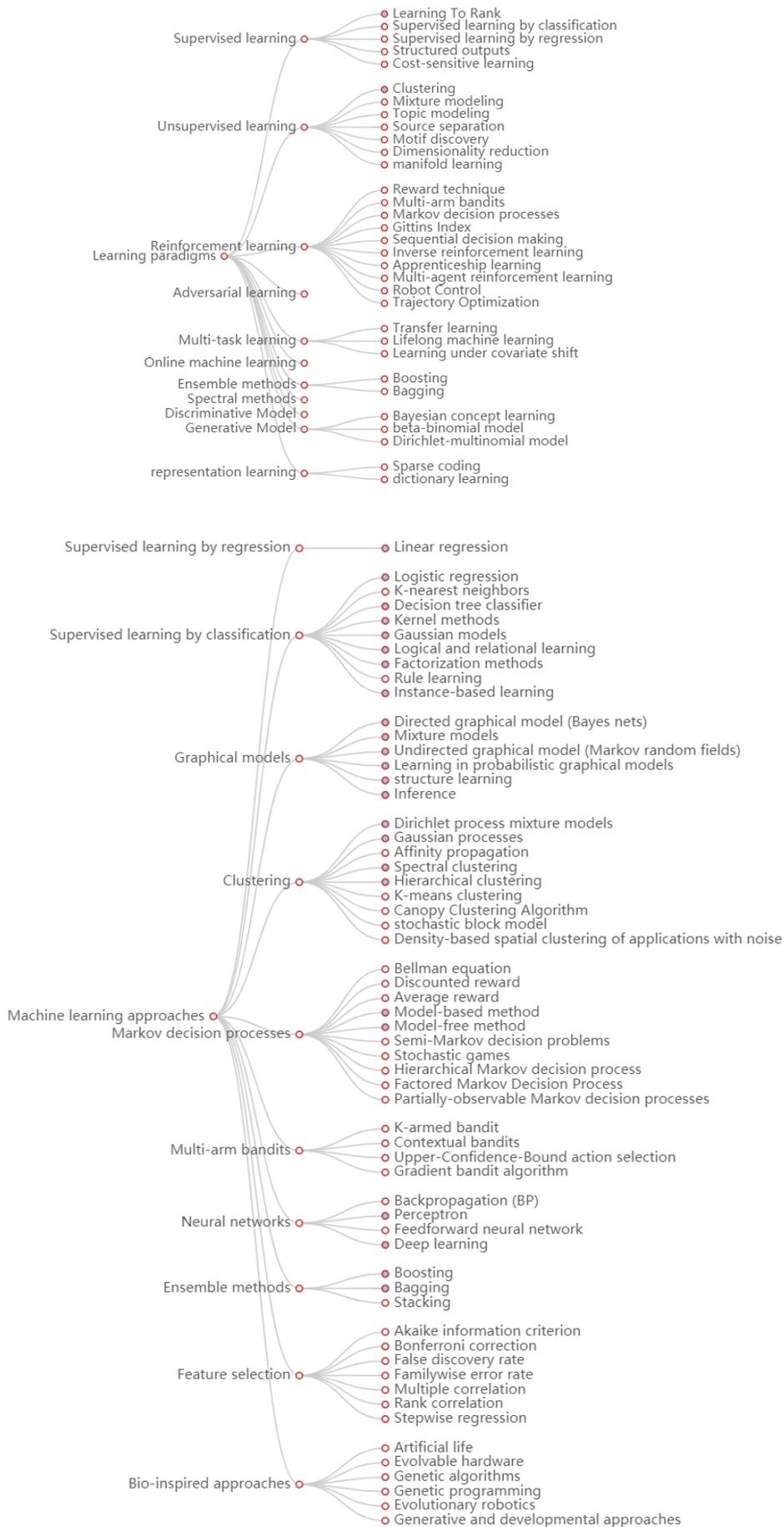
10) [CV] CIFAR

<https://www.cs.toronto.edu/~kriz/cifar.html>

8.5 9 机器学习知识树

本报告分析了近年来机器学习领域的高水平学术论文，挖掘出了相关关键词，结合知识图谱技术整理了机器学习八级知识树，详细数据可联系 <https://www.aminer.cn/data> 下载原始数据，鉴于自动分析技术和论文采集的局限性，图谱还可以进一步完善，欢迎读者批评指正，我们会根据根据读者的反馈定期更新。Machine learning 的二级分类包括 Machine learning theory、Learning paradigms 以及 Machine learning approaches，它们的知识树展示如下：





参考文献

- [1] Samuel A L. (1959). Some studies in machine learning using the game of checkers[J]. *IBM Journal of research and development*, 3(3): 210-229.
- [2] Machine Learning Types. <https://towardsdatascience.com/machine-learning-types-2-c1291d4f04b1>.
- [3] 史上最全机器学习经典算法详解. <https://zhuanlan.zhihu.com/p/95135478>.
- [4] 一文看懂 GAN 演进图谱. <https://www.infoq.cn/article/GCGIboPIfTpBe9dEqF3m>.
- [5] 对抗机器学习. <https://www.jiqizhixin.com/graph/technologies/a1490c68-7868-4b6a-a775-097d458c64c1>.
- [6] 深入浅出对抗性机器学习. <https://zhuanlan.zhihu.com/p/52561355>.
- [7] Kearns, M., & Li, M. (1993). Learning in the presence of malicious errors. *SIAM Journal on Computing*, 22(4): 807-837.
- [8] 什么是自动机器学习, 它有哪些用处. <https://www.boxuegu.com/news/1359.html>.
- [9] Wang, Q., Ming, Y., Jin, Z., Shen, Q., Liu, D., Smith, M. J., ... & Qu, H. (2019). Atmseer: Increasing transparency and controllability in automated machine learning. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 681.
- [10] Cracking open the black box of automated machine learning. <https://techxplore.com/news/2019-06-black-automated-machine.html>.
- [11] 机器学习模型可解释性的详尽介绍. <https://www.jiqizhixin.com/articles/2019-10-30-9>.
- [12] 人工智能之机器学习篇——在线学习. <https://baijiahao.baidu.com/s?id=159433714663599109&wfr=spider&for=pc>.
- [13] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*. 5998-6008.
- [14] Taylor, W. L. (1953). Cloze procedure: A new tool for measuring readability. *Journalism Bulletin*, 30(4):415-433.
- [15] 如何通俗易懂地解释卷积. <https://www.zhihu.com/question/22298352/answer/637156871>.
- [16] 一文读懂图卷积 GCN. https://mp.weixin.qq.com/s/Rd-MBAgq_i-PsaopzNmACQ.
- [17] 图神经网络. <https://www.jiqizhixin.com/graph/technologies/c39cf57b-df95-4c9e-9a8a-0d8ea330d625>.
- [18] 2018 年深度学习的主要进步. https://blog.csdn.net/weixin_42137700/article/details/85754604.
- [19] 图网络模型原理详解. https://blog.csdn.net/weixin_40871455/article/details/86515934.
- [20] Momentum Contrast for Unsupervised Visual Representation Learning 无监督胜有监督, 刷新检测分割任务. https://blog.csdn.net/weixin_43876801/article/details/103148773.

- [21] 机器学习是在金融领域的应用，离不开大数据. <https://www.jianshu.com/p/0f3b638a4a74>.
- [22] 机器学习算法在自动驾驶领域的应用大盘点. https://blog.csdn.net/weixin_34050005/article/details/90434845.
- [23] 人工智能与机器学习技术在医疗保健行业中的应用. <https://zhuanlan.zhihu.com/p/34918987>.
- [24] 当人工智能和机器学习遇到零售. <http://www.cniteyes.com/archives/32001>.
- [25] 2017 年机器学习将在这四大行业得到全面应用. <https://baijiahao.baidu.com/s?id=1567230473301588&wfr=spider&for=pc>.

AMiner

版权声明

AMiner 研究报告版权为 AMiner 团队独家所有，拥有唯一著作权。AMiner 咨询产品是 AMiner 团队的研究与统计成果，其性质是供用户内部参考的资料。

AMiner 研究报告提供给订阅用户使用，仅限于用户内部使用。未获得 AMiner 团队授权，任何人和单位不得以任何方式在任何媒体上（包括互联网）公开发布、复制，且不得以任何方式将研究报告的内容提供给其他单位或个人使用。如引用、刊发，需注明出处为“AMiner.org”，且不得对本报告进行有悖原意的删节与修改。

AMiner 研究报告是基于 AMiner 团队及其研究员认可的研究资料，所有资料源自 AMiner 后台程序对大数据的自动分析得到，本研究报告仅作为参考，AMiner 团队不保证所分析得到的准确性和完整性，也不承担任何投资者因使用本产品与服务而产生的任何责任。

顾问：朱军、唐杰

编辑：景晨、刘佳、邵洲、殷达、赵杨奥

数据：赵慧军



获取200页完整纸质版



关注学术头条