## 人工智能的进程

### 编程

提前设置好交互逻辑，利用计算机强大的计算能力快速重复的执行，从而将人类从简单枯燥的重复劳动中解脱出来。但是较高智能的任务无法通过简单的交互程序设计逻辑规则，比如：人脸识别，聊天机器人，自动驾驶等

### 人工智能

让机器获得像人类一样的智能机制，最早出现在1956年的达特茅斯会议上

三个阶段：

推理期：师徒通过总结归纳逻辑规律来实现，但是显式的规则过于简单，难以表达复杂的抽象的规则。

机器学习（知识库 + 推理）：1970年，通过庞大的专家系统来模拟人类专家的智能水平。难点是：复杂抽象的概念很难用代码实现，比如图片识别。因此而产生了机器学习（神经网络包含在其中）

深度神经网络：深度学习特指基于深层神经网络实现的模型或算法在计算机视觉、自然语言处理、机器人等领域取得了重大突破。

### 机器学习

#### 有监督学习

样本集包含样本x和样本标签y，算法模型需要学习到映射𝑓𝜃: 𝒙 → 𝒚，，训练时优化𝑓𝜃: 𝒙 到真实标签y之间的误差来优化回归参数θ，使得下一次预测更精准。

常见的有监督学习有：线性回归，逻辑回归，支持向量机，随机森林等

#### 无监督学习

无法收集到样本标签，因此只能由算法自己去发现数据的模态。其中有一类自监督学习(Self-supervised Learning)： 将自身作为监督信号，即𝑓𝜃: 𝒙 → x。

常见的无监督学习有：K-means算法，自编码器，生成对抗网络等

#### 强化学习

也称增强学习，指算法与环境交互来学习解决问题的策略，并没有明确的“正确”信号作为监督，而是在与环境交互的过程中获得环境反馈的奖励信号，不过不能作为计算动作与“正确动作”之间的误差来优化算法。

常见强化学习算法有：DQN, PPO等

### 神经网络

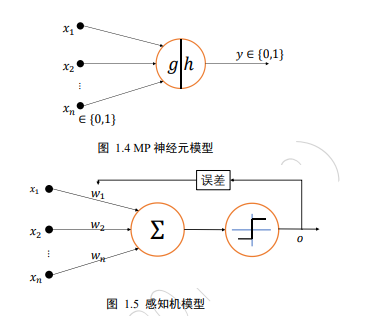
是一种通过神经网络从数据中学习的算法。受限于计算能力和数据量，早起的神经网络一般在1-4层，表达能力有限。

#### 发展史

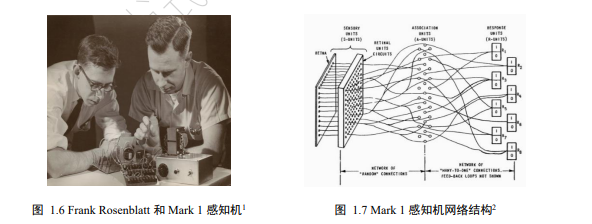
浅层神经网络

1943年，心理学家Warren McCulloch 和逻辑学家 Walter Pitts 根据生物神经元（Neuron）结构，提出最早的神经元模型，成为MP神经元模型。该模型的输出f(x) = h(g(x))，其中g(x) = ∑𝑖 𝑥𝑖 , 𝑥𝑖 ∈ {0,1}。模型通过g（x）的值来完成输出值的预测，如果g(x) >= 0 输出1， 如果g(x) < 0， 输出为0

可以看出MP神经元模型并没有学习能力，只能以固定逻辑完成判断。



1958年，美国心理学家Frank Rosenblatt 提出了第一个可以自动学习权重的神经元模型，称为感知机（Perceptron）如图1.5所示，输出值o 和真实值y知己恩的误差用于调整神经元的权重参数{ , , … , }。感知机随后基于“Mark 1 感知机”硬件实现，如图 1.6 图 1.7 所示，输入为 400 个单元的图像传感器，输出为 8 个节点端子，可以成功识别一些英文 字母。我们一般认为 1943 年~1969 年为人工智能发展的第一次兴盛期。



1969年，美国科学家Marvin Minsky 等人出版的《Perceptions》一书中指出了感知机等线性模型的主要缺陷，即无法处理简单的异或XOR等线性不可分问题。这直接导致以感知机为代表的神经网络研究陷入了低谷期。一般认为1969-1982年为人工智能发展的第一次寒冬。

反向传播算法：Backpropagation 简称BP算法：至今仍是现代深度学习的核心理论基础。

1982年John Hopfild的循环链接的Hopfield的提出，开启了1982年-1995年的第二次人工智能复兴大潮。期间相继提出了卷积神经网络，循环神经网络，反向传播算法等算法模型。

1989年，Yann LeCun等人将BP算法应用在手写数字图片识别上，取得了巨大成功，这套系统成功商用在邮政编码识别，银行支票识别等系统上

1997年，应用最广泛的循环神经网络的变种之一 LSTM 被Jürgen Schmidhuber提出，同年双向循环神经网络也被提出

遗憾的是，神经网络的研究随着以支持向量机（Support Vector Machine，简称SVM）为代表的传统机器学习算法兴起而逐渐进入低谷，成为人工智能的第二次寒冬。

支持向量机拥有严格的理论基础，需要的样本量较少，同时也具有良好的泛化能力，

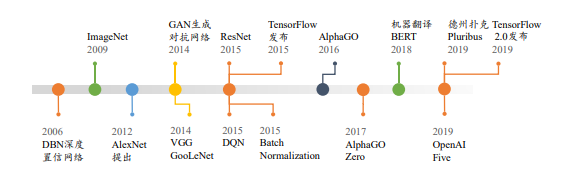
神经网络理论基础欠缺，可解释性查，很难训练深层网络，性能一般

### 深度学习

2006年，Geoffrey Hinton等人发现通过逐层预训练的方式可以较好地训练多层神经我网络，并且在MNIST手写数字图片数据集上取得了优于SVM的错误率。开启了第3次人工智能分复兴。在论文中，Geoffrey Hinton 首次提出了 Deep Learning 的概念，这也是(深层)神 经网络被叫做深度学习的由来。2011年，Xavier Glorot提出了线性正六单元（Rectified Linear Unit, RelU）激活函数，这是现在使用最广泛的激活函数之一。2012年, Alex Krizhevshy提出了8层的神经网络 AlexNet ,，它采用了ReLU激活函数，并使用Dropout技术防止过拟合，同时抛弃了逐层预训练的方式，直接在2块GTX580 CPU 上训练网络。AlexNet 在 ILSVRC-2012图片识别比赛中获得了第一名，比第二名在Top-5错误率上低了惊人的10.9%

自AlexNet 模型提出后， 各种各样的算法模型相继被发表，其中有VGG系列，GooLeNet, ResNet系列， DenseNet系列等等，其中，ResNet 系列网络实现简单，效果显著，很快将网络的层数提升至数百层，甚至上千层，同时保持性能不变甚至更好。

除了有监督学习领域取得了惊人的成果，在无监督学习和强化学习领域也取得了巨大 的成绩。2014 年，Ian Goodfellow 提出了生成对抗网络，通过对抗训练的方式学习样本的 真实分布，从而生成逼近度较高的图片。此后，大量的生成对抗网络模型被提出，最新的 图片生成效果已经达到了肉眼难辨真伪的逼真度。2016 年，DeepMind 公司应用深度神经 网络到强化学习领域，提出了 DQN 算法，在 Atari 游戏平台中的 49 个游戏取得了人类相 当甚至超越人类的水平；在围棋领域，DeepMind 提出的 AlphaGo 和 AlphaGo Zero 智能程 序相继打败人类顶级围棋专家李世石、柯洁等；在多智能体协作的 Dota2 游戏平台， OpenAI 开发的 OpenAI Five 智能程序在受限游戏环境中打败了 TI8 冠军 OG 队，展现出了 大量专业级的高层智能的操作。图 1.9 列出了 2006 年~2019 年之间重大的时间节点。

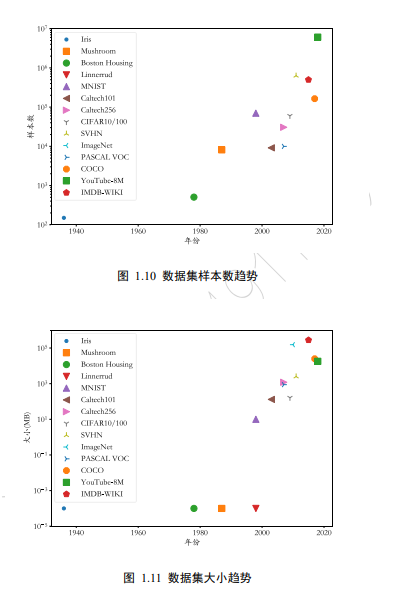


### 深度学习与其他算法比较

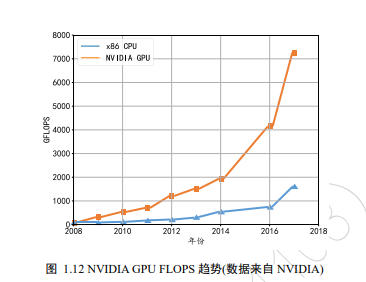
### 相互关系

### 深度学习的特点

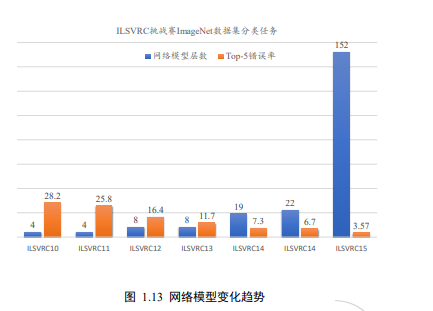
数据集大能有效防止过拟合，但是成本高，清洗过程会带入手动采集的主观偏差和随机误差。因此研究数据量需求 较少的算法模型是非常有用的一个方向。



计算能力



网络规模



### 深度学习的应用

#### 计算机视觉

图片识别：图片识别(Image Classification) 是常见的分类问题。神经网络的输入为图片数据，输出 值为当前样本属于每个类别的概率，通常选取概率值最大的类别作为样本的预测类别。图 片识别是最早成功应用深度学习的任务之一，经典的网络模型有 VGG 系列、Inception 系 列、ResNet 系列等。

目标检测：目标检测(Object Detection) 是指通过算法自动检测出图片中常见物体的大致位置，通 常用边界框(Bounding box)表示，并分类出边界框中物体的类别信息，如图 1.15 所示。常 见的目标检测算法有 RCNN，Fast RCNN，Faster RCNN，Mask RCNN，SSD,YOLO 系列 等。

语义分割：语义分割(Semantic Segmentation) 是通过算法自动分割并识别出图片中的内容，可以 将语义分割理解为每个像素点的分类问题，分析每个像素点属于物体的类别，如图 1.16 所 示。常见的语义分割模型有 FCN，U-net，SegNet，DeepLab 系列等。

视频理解：视频理解(Video Understanding) 随着深度学习在 2D 图片的相关任务上取得较好的效 果，具有时间维度信息的 3D 视频理解任务受到越来越多的关注。常见的视频理解任务有 视频分类，行为检测，视频主体抽取等。常用的模型有 C3D，TSN，DOVF，TS\_LSTM 等。

图片生成：图片生成(Image Generation) 通过学习真实图片的分布，并从学习到的分布中采样而获 得逼真度较高的生成图片。目前主要的生成模型有 VAE 系列，GAN 系列等。其中 GAN 系 列算法近年来取得了巨大的进展，最新 GAN 模型产生的图片样本达到了肉眼难辨真伪的 效果，如图 1.17 为 GAN 模型的生成图片。

除了上述应用，深度学习还在其他方向上取得了不俗的效果，比如艺术风格迁移(图 1.18)，超分辨率，图片去燥/去雾，灰度图片着色等等一系列非常实用酷炫的任务，限于篇 幅，不再敖述。



#### 自然语言处理

机器翻译：机器翻译(Machine Translation) 过去的机器翻译算法通常是基于统计机器翻译模型，这 也是 2016 年前 Google 翻译系统采用的技术。2016 年 11 月，Google 基于 Seq2Seq 模型上 线了 Google 神经机器翻译系统(GNMT)，首次实现了源语言到目标语言的直译技术，在多 项任务上实现了 50~90%的效果提升。常用的机器翻译模型有 Seq2Seq，BERT，GPT， GPT-2 等，其中 OpenAI 提出的 GPT-2 模型参数量高达 15 亿个，甚至发布之初以技术安全 考虑为由拒绝开源 GPT-2 模型。

聊天机器人：聊天机器人(Chatbot) 聊天机器人也是自然语言处理的一项主流任务，通过机器自动与 人类对话，对于人类的简单诉求提供满意的自动回复，提高客户的服务效率和服务质量。 常应用在咨询系统、娱乐系统，智能家具等中。

#### 强化学习

虚拟游戏 ：相对于真实环境，虚拟游戏平台既可以训练、测试强化学习算法，有可以避 免无关干扰，同时也能将实验代价降到最低。目前常用的虚拟游戏平台有 OpenAI Gym， OpenAI Universe，OpenAI Roboschool，DeepMind OpenSpiel，MuJoCo 等，常用的强化学 习算法有 DQN，A3C，A2C，PPO 等。在围棋领域，DeepMind AlaphGo 程序已经超越人 类围棋专家；在 Dota2 和星际争霸游戏上，OpenAI 和 DeepMind 开发的智能程序也在限制 规则下战胜了职业队伍。

机器人(Robotics) ：在真实环境中，机器人的控制也取得了一定的进展。如 UC Berkeley 在机器人的 Imitation Learning，Meta Learning，Few-shot Learning 等方向取得了不少进展。 美国波士顿动力公司在人工智能应用中取得喜人的成就，其制造的机器人在复杂地形行 走，多智能体协作等任务上表现良好(图 1.19)。

自动驾驶(Autonomous Driving) ：被认为是强化学习短期内能技术落地的一个应用方 向，很多公司投入大量资源在自动驾驶上，如百度、Uber，Google 无人车等，其中百度的 无人巴士“阿波龙”已经在北京、雄安、武汉等地展开试运营，图 1.20 为百度的自动驾驶 汽车。

