# 深度学习历史

## 时势造英雄

### 什么是深度学习？

深度学习是一种层次化的机器学习方法，它模仿人脑神经网络的结构和功能，通过多层非线性变换来对数据进行建模和识别。深度学习算法可以自动学习特征，从而能够处理各种类型的数据，比如图像、音频、自然语言等，实现各种任务，比如图像分类、语音识别、自然语言处理等。深度学习已经在许多领域得到广泛的应用，包括计算机视觉、自然语言处理NLP、语音识别、推荐系统、医疗、金融等。比如，在计算机视觉领域，深度学习已经实现了人类水平的图像分类、目标检测、图像生成等任务；在自然语言处理领域，深度学习已经成为文本分类、情感分析、机器翻译等任务的主要方法之一；在医疗领域，深度学习已经被用于疾病诊断、药物发现等任务。深度学习的应用前景非常广阔，将会对未来的科技和社会发展产生重要影响。

### 深度学习的历史简述

1980年，基于传统的感知器结构，深度学习创始人，加拿大多伦多大学教授杰弗里·辛顿（G. Hinton）采用多个隐含层的深度结构来代替代感知器的单层结构，多层感知器模型（Multi\_Layer Perceptron）是其中最具代表性的，而且多层感知器也是最早的深度学习网络模型。1974年，Paul Werbos提出采用反向传播法来训练一般的人工神经网络，随后，该算法进一步被杰弗里·辛顿、燕·勒存（Y. LeCun）等人应用于训练具有深度结构的神经网络。反向传播法根据神经网络输出层的计算误差来调整网络的权值，直到计算误差收敛为止。但是，反向传播法训练具有多隐含层的深度网络的网络参数的学习性能并不好，因为具有多隐含层的深度网络的网络参数的训练问题是一个非凸问题，基于梯度下降的反向传播法很容易在训练网络参数时收敛于局部极小值。此外，反向传播法训练网络参数还存在很多实际问题，比如需要大量的标签样本来训练网络的权值，多隐含层的神经网络权值的训练速度很慢，权值的修正随着反向传播层数的增加逐渐削弱等。

面对采用反向传播法来训练具有多隐含层的深度网络的网络参数时存在的缺陷，一部分研究人员开始探索通过改变感知器的结构来改善网络学习的性能，由此产生了很多著名的单隐含层的浅层学习模型，如SVM、logistic regression、Maximum entropy model和朴素贝叶斯模型等。浅层学习模型能够有效地解决简单或者具有复杂条件限制的问题，但受限于只含一个隐含层，所以浅层学习模型特征构造的能力有限，不能有效处理包含复杂特征的问题[5]。为了同时解决具有多隐含层的深度网络在参数训练时存在的缺陷和浅层网络特征构造能力有限的问题，一些研究人员开始尝试采用新的参数训练方法来训练多隐含层的深度网络。

1984年,日本学者福岛邦彦提出了卷积神经网络的原始模型神经感知机（Neocognitron）。1998年，燕·勒存（Y. LeCun）提出了深度学习常用模型之一卷积神经网络（Convoluted Neural Network， CNN）。2006年，杰弗里·辛顿（G. Hinton）提出了深度学习的概念，随后与其团队在文章《A fast Learning Algorithm for Deep Belief Nets》中提出了深度学习模型之一，深度信念网络，并给出了一种高效的半监督算法：逐层贪心算法，来训练深度信念网络的参数，打破了长期以来深度网络难以训练的僵局。从此，深度学习的大门打开，在各大政府、高校和企业中掀起了研究深度学习的大浪潮。2009年，Yoshua Bengio提出了深度学习另一常用模型：堆叠自动编码器（Stacked Auto-Encoder,SAE），采用自动编码器来代替深度信念网络的基本单元：限制玻尔兹曼机，来构造深度网络。

约翰·霍普菲尔德(John Hopfield)在1982提出了Hopfield网络，是最早的递归神经网络（Recurrent Neural Network，RNN）。因Hopfield网络实现困难，没有合适的应用场景，86年后逐渐被前向神经网络取代。1990年，出现了Elman&Jordan SRN两种新的RNN网络，同样因为没有合适的应用场景，很快淡出了研究人员视线。Dalle Molle人工智能研究所的主任Jurgen Schmidhuber在论文《THE VANISHING GRADIENT PROBLEM DURING recurrent neural networks and problem solutions》提出了LSTM，促进了循环神经网络的发展，特别是在深度学习广泛应用的今天，RNN（LSTM）在自然语言处理领域，如机器翻译、情感分析、智能对话等，取得了令人惊异的成绩。

深度学习自2006年产生之后就受到科研机构、工业界的高度关注。最初，深度学习的应用主要是在图像和语音领域。从2011年开始，谷歌研究院和微软研究院的研究人员先后将深度学习应用到语音识别，使识别错误率下降了20%-30%。2012年，杰弗里·辛顿的学生IIya Sutskever和Alex Krizhevsky在图片分类比赛ImageNet中，使用深度学习打败了Google团队，深度学习的应用，使得图片识别错误率下降了14%。2012年6月，谷歌首席架构师Jeff Dean和斯坦福大学教授AndrewNg主导著名的GoogleBrain项目，采用16万个CPU来构建一个深层神经网络，并将其应用于图像和语音的识别，最终大获成功。2014年，Facebook基于深度学习技术的DeepFace项目实现了人脸识别准确率达到97%以上。2016年，谷歌公司发布了基于深度学习开发的AlphaGo，并在围棋比赛中战胜国际顶尖围棋高手李世石。2017年，AlphaGo升级版AlphaGo Zero采用强化学习算法，以100:0的比分轻而易举地打败之前的AlphaGo。此外，深度学习在搜索领域也获得广泛关注。如今，深度学习已经在图像、语音、自然语言处理、CTR预估、大数据特征提取等方面获得广泛的应用。

随着深度学习技术的不断进步以及计算机硬件水平的提高，深度学习在许多领域中开始得到广泛应用。总的来说，深度学习技术已经成为当前最热门的研究方向之一，其在图像识别、自然语言处理、医疗、金融等领域都展现出了巨大的优势，并不断推动相关领域的进步和创新。未来，深度学习技术的应用前景十分广阔，备受关注和期待。

### 深度学习的优势

1. 对大规模数据的处理能力：深度学习模型可以利用多层神经网络进行端到端的学习，从而对海量数据进行高效处理。相比传统机器学习算法，深度学习能够更好地挖掘数据中的潜在关系和模式，使得模型的准确性和可靠性更高。
2. 自适应能力：深度学习模型可以自动调整自己的权重和参数，从而不需要手工进行特征工程或者选择数据集，简化了模型构建的流程。这种自适应能力也意味着深度学习模型具有更好的泛化能力，能够适应新的数据并保持良好的性能。
3. 可以处理多类型数据：深度学习可以应用于处理各种类型的数据，包括文本、图像、音频等。例如，在自然语言处理领域，深度学习可以通过RNN、LSTM和Transformer等模型来进行文本分类、情感分析、机器翻译等任务。在计算机视觉领域，深度学习可以利用卷积神经网络（CNN）来进行图像分类、目标检测和图像分割等任务。
4. 高效性：深度学习具有高度并行的计算能力，能够利用GPU等现代硬件设备来加速训练过程，提高模型的效率。此外，深度学习在推理过程中也能够利用硬件设备进行加速，从而实现快速的响应。
5. 可扩展性：深度学习可以通过增加层数和节点数来增强模型的表达能力，从而可以对更复杂的问题进行建模。例如，在自然语言处理领域，通过使用Transformer模型可以处理更长的文本序列；在计算机视觉领域，通过增加网络深度和宽度可以提高模型的性能，同时也带来了更多的计算和存储需求。

### 深度学习的应用

1. 计算机视觉：深度学习模型在图像分类方面的应用已经非常成熟，能够识别出图像中的不同物体和场景。在目标检测方面，深度学习技术可以实现自动化的物体检测和跟踪，用于智能监控和自动驾驶等领域。在图像分割方面，深度学习可以将图像分成不同的区域，用于医学影像分析和无人机图像处理等应用。
2. 自然语言处理（NLP）：深度学习技术在文本分类方面的应用包括垃圾邮件过滤、新闻分类和情感分析等。在命名实体识别方面，深度学习可以自动地识别出文本中的人名、地名和组织名等实体。在机器翻译方面，深度学习可以将源语言的句子翻译成目标语言的句子，在多语种交流和跨国企业合作等领域有广泛应用。
3. 语音识别：深度学习技术在自动语音识别领域被广泛应用，能够将语音转换为文本。在说话人识别方面，深度学习可以识别说话人的身份和情感状态等信息。在语音合成方面，深度学习可以生成自然流畅的语音，用于智能客服和虚拟助手等应用。
4. 推荐系统：深度学习技术在推荐系统中被用于基于内容的推荐和协同过滤推荐。基于内容的推荐根据用户的历史行为和兴趣，推荐相关的商品或服务。协同过滤推荐则根据用户与其他用户的相似性，为用户推荐其他用户喜欢的商品或服务。
5. 自动驾驶技术：深度学习技术在自动驾驶汽车中被用于对道路标记、障碍物和其他车辆的识别。通过深度学习模型的训练，自动驾驶汽车可以实现智能导航、自动避障和自动驾驶等功能，提高行车安全和驾驶效率。
6. 医疗保健：深度学习技术在医学影像分析方面的应用包括肺癌筛查、脑部疾病检测和乳腺癌诊断等。在疾病诊断方面，深度学习可以自动分析患者的症状和体征，提供精准的诊断和治疗建议。在药物研发方面，深度学习可以帮助科学家快速地筛选出具有潜力的化合物，加速新药开发的过程。