

## 1.2 深度学习的发展历程

深度学习基本上是深层神经网络的一个代名词，最早神经网络数学模型源自“A logical calculus of the ideas immanent in nervous activity”。神经活动内在思想的逻辑演算。

早期的 McCulloch-Pitts Neuron 结构使用了简单的线性加权和的方式来模拟这个变化。随后，Frank Rosenblatt 教授在 1958 年提出了感知机模型（perceptron），这是首个可以根据样例数据来学习特征权重的模型。但是感知机模型只能解决线性可分问题无法解决异或问题（1969 年）。

到 20 世纪 80 年代末，第二波神经网络研究因为分布式知识表达和神经网络反向传播算法兴起。

分布式知识表达大大加强了模型的表达能力，解决了类似异或问题等线性不可分问题。

反向传播算法大幅降低了训练神经网络所需要的时间。1991 年 LSTM 模型诞生可以对较长的序列进行建模，比如一句话或者一段文章。它是解决很多自然语言处理、机器翻译、语言识别、时序预测等问题最有效的方法。现在用的多的是卷积神经网络和循环神经网络。

在神经网络发展的同时，传统的机器学习算法也有了突破，比如 1998 年末的支持向量机，已经可以使得错误率低于 0.8%超越了神经网络，因为当时训练深层神经网络需要大量数据。

但随着计算机性能的进一步提高，已经云计算、GPU 的出现，2012 年深度学习系统 AlexNet 赢得了冠军，从此 Deep Learning 作为深度神经网络的代名词被大家所熟知。