

前馈神经网络

机器学习我们已经知道可以分为两大流派：

1. 频率派，这个流派的方法叫做统计学习，根据具体问题有下面的算法：
 1. 正则化，L1, L2 等
 2. 核化，如核支撑向量机
 3. 集成化，AdaBoost, RandomForest
 4. 层次化，神经网络，神经网络有各种不同的模型，有代表性的有：
 1. 多层感知机
 2. Autoencoder
 3. CNN
 4. RNN

这几种模型又叫做深度神经网络。

2. 贝叶斯派，这个流派的方法叫概率图模型，根据图特点分为：
 1. 有向图-贝叶斯网络，加入层次化后有深度有向网络，包括
 1. Sigmoid Belief Network
 2. Variational Autoencoder
 3. GAN
 2. 无向图-马尔可夫网络，加入层次化后有深度玻尔兹曼机。
 3. 混合，加入层次化后有深度信念网络

这几个加入层次化后的模型叫做深度生成网络。

从广义来说，深度学习包括深度生成网络和深度神经网络。

From PLA to DL

- 1958, PLA
- 1969, PLA 不能解决 XOR 等非线性数据
- 1981, MLP, 多层感知机的出现解决了上面的问题
- 1986, BP 算法应用在 MLP 上, RNN
- 1989, CNN, Universal Approximation Theorem, 但是于此同时, 由于深度和宽度的相对效率不知道, 并且无法解决 BP 算法的梯度消失问题
- 1993, 1995, SVM + kernel, AdaBoost, RandomForest, 这些算法的发展, DL 逐渐没落
- 1997, LSTM
- 2006, 基于 RBM 的深度信念网络和深度自编码
- 2009, GPU的发展
- 2011, 在语音方面的应用
- 2012, ImageNet
- 2013, VAE
- 2014, GAN

- 2016, AlphaGo
- 2018, GNN

DL 不是一个新的东西，其近年来的大发展主要原因如下：

1. 数据量变大
2. 分布式计算的发展
3. 硬件算力的发展

非线性问题

对于非线性的问题，有三种方法：

1. 非线性转换，将低维空间转换到高维空间（Cover 定理），从而变为一个线性问题。
2. 核方法，由于非线性转换是变换为高维空间，因此可能导致维度灾难，并且可能很难得到这个变换函数，核方法不直接寻找这个转换，而是寻找一个内积。
3. 神经网络方法，将复合运算变为基本的线性运算的组合。