

Feedforward Neural Network 03 Non-Linear Problem

Chen Gong

12 November 2019

实际上在 1958 年就已经成功的提出了 Perceptron Linear Analysis (PLA), 标志着人工智能的正式诞生。但是, Minsky 在 1969 年提出 PLA 无法解决非线性分类问题, 让人工智能陷入了 10 年的低谷。后来的发展, 人们开始寻找到越来越多的, 解决非线性分类问题的方法。于是, 我们提出了三种解决非线性问题的方法。

1 Non-Transformation

这实际上就是一种明转换, 将向量从 input space 转换到 feature space, 可以写做 $\phi = \mathcal{X} \mapsto \mathcal{Z}$ 。在 Conver's theory 中提出, 高维空间比低维空间更加容易线性可分。很显然对于一个异或问题 (XOR) 来说, 我们将 $x = (x_1, x_2) \xrightarrow{\phi} z = (x_1, x_2, (x_1 - x_2)^2)$

$$\begin{array}{ccc|ccc} 0 & 1 & \longrightarrow & 1 & & 0 & 1 & 1 & \longrightarrow & 1 \\ 1 & 0 & \longrightarrow & 1 & \xrightarrow{\phi} & 1 & 0 & 1 & \longrightarrow & 1 \\ 1 & 1 & \longrightarrow & 0 & & 1 & 1 & 0 & \longrightarrow & 0 \\ 0 & 0 & \longrightarrow & 0 & & 0 & 0 & 0 & \longrightarrow & 0 \end{array} \quad (1)$$

很显然在三维空间中, 进行空间映射后, 就会变得比较容易进行线性划分了。可以自己画图来进行验证, 这里不再作图。

2 Kernel Method

这实际上是一种暗转的思路, 也就是令 $K(x, x') = \langle \phi(x), \phi(x') \rangle$, 在这个核函数中实际上隐藏了一个 ϕ , 而 $x, x' \in \mathcal{X}$ 。

3 Neural Network

神经网络算法实际上就是一个 Multit-Layer Perceptron (MLP), 有时也会被称为, Feedforward Neural Network (FNN), 所以大家在其他书上见到这几种描述都不要感到意外。我们以 XOR (位运算) 为例吧。在我们的逻辑运算中, 大致有四种运算方法。

$$\begin{array}{cccc} XOR & OR & AND & NOT \\ \oplus & \vee & \wedge & \neg \end{array} \quad (2)$$

而后三种运算为基础运算，因为异或运算实际上是可以由后三种运算组成，也就是 $x_1 \oplus x_2 = (\neg x_1 \wedge x_2) \vee (x_1 \wedge \neg x_2)$ 。实际上就是先做两个与运算，然后做一个或运算。把一个线性不可分的东西来分层实现，将特征空间进行了分解而已。然后，在分层运算中插入了激活函数，来达到非线性映射的效果。这部分内容，比较的简单，而且网上也有大量的资料，此处就不再做过多的阐述。

实际上神经网络就是一个有向无环图。所以，某种意义上说可以引入概率图的模型，当然这就是后话了。