Deep Neural Network (DNN)

Alex Tse
University College London

Deep Neural Networks

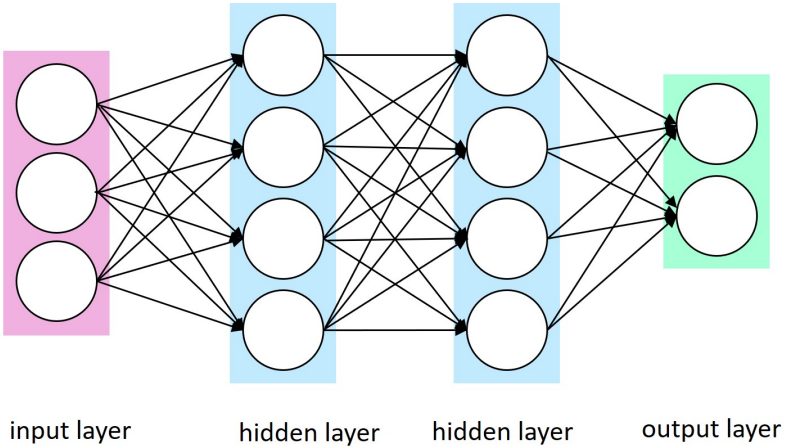


Figure: An example of deep neural networks. Source: [1]

Network Architecture

DNNs to approximate continuous functions in $\mathcal{C}(\mathbb{R}^d, \mathbb{R}^e)$.

Let $h^{(l)}$ denote the l^{th} layer with n_l neurons.

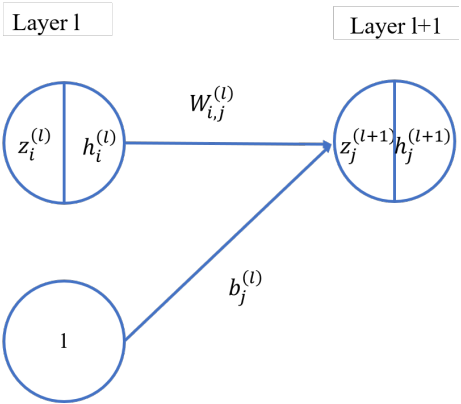
- Input layer ($h^{(0)} : \mathbb{R}^d \to \mathbb{R}^{n_0}$ with $n_0 = d$): $h^{(0)}(x) = x, \forall x \in \mathbb{R}^d$.
- Hidden layer ($h^{(l)} : \mathbb{R}^d \to \mathbb{R}^{n_l}$): $\forall l \in \{1, 2, \cdots, L-1\}$.
- Output layer ($h^{(L)} : \mathbb{R}^d \to \mathbb{R}^e$).

$(h^{(l)})_{l=1}^L$ is defined in the following recursive way that for any $x \in \mathbb{R}^d$,

$$\begin{aligned} z^{(l+1)}(x) &= h^{(l)}(x)W^{(l)} + b^{(l)}, \\ h^{(l+1)}(x) &= \sigma_{l+1}(z^{(l+1)}(x)), \end{aligned}$$

where $W^{(l)}$ is a $n_l \times n_{l+1}$ matrix, $b^{(l)}$ is a n_{l+1} dimensional row vector and σ_l is the activation function on the l^{th} layer. Here $\theta := (W^{(l)}, b^{(l)})_{l=1}^{L-1}$ is the parameters of DNNs, which can be trained from data.

Neural Network Building Block



Formulas

For any $l \in \{1, \cdots, (L-1)\}$,

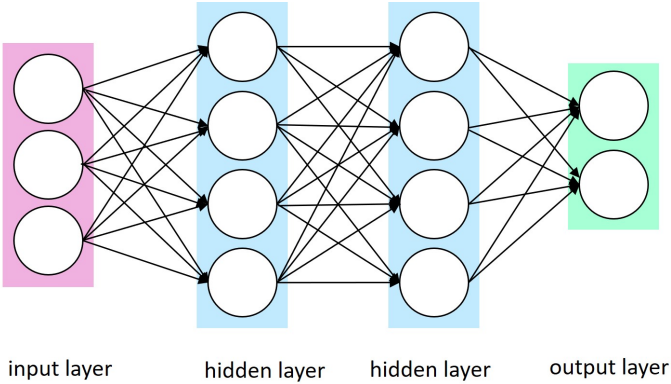
$$z_j^{(l+1)} = \sum_{i=1}^{n_l} W_{i,j}^{(l)} h_i^{(l)} + b_j^{(l)},$$

$$h_i^{(l+1)} = \sigma_{l+1}(z_i^{(l+1)}),$$

where $i \in \{1, \cdots, n_l\}$, $j \in \{1, \cdots, n_{l+1}\}$, $W_{i,j}^{(l)}$ is the weight from incoming node i to output node j on the layer l.

Questions

Can you write down the mathematical formulation of the DNN example shown in the following figure? What is the corresponding number of model parameters?



Thanks for your attention!

References I

Hao Ni, Xin Dong, Jinsong Zheng, and Guangxi Yu.
An Introduction to Machine Learning in Quantitative Finance (Chinese version).
Tsinghua University Press, 2021.