

深度学习

作者: Calvin

QQ: 179209347

Mail: 179209347@qq.com

介绍

笔记简介:

- 面向对象: 深度学习初学者
- 依赖课程: **线性代数, 统计概率**, 优化理论, 图论, 离散数学, 微积分, 信息论

知乎专栏:

<https://zhuanlan.zhihu.com/p/693738275>

Github & Gitee 地址:

https://github.com/mymagicpower/AIAS/tree/main/deep_learning

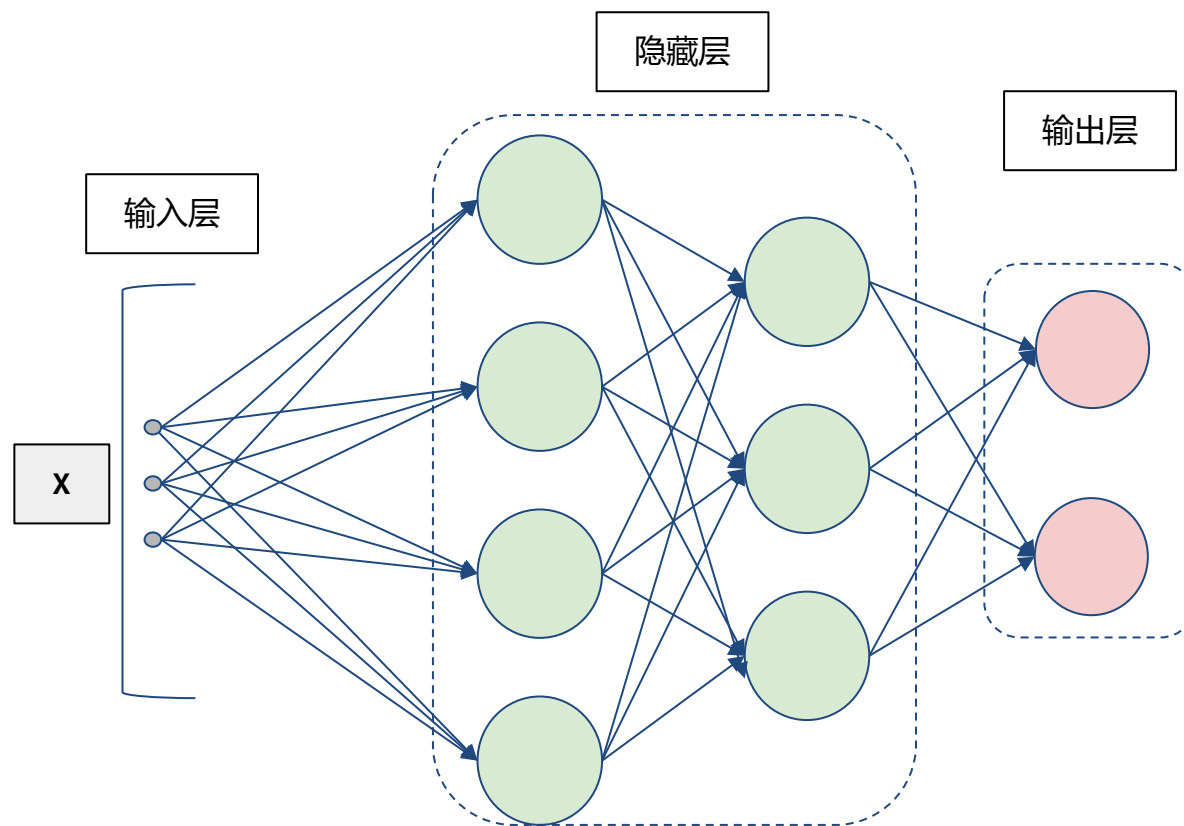
https://gitee.com/mymagicpower/AIAS/tree/main/deep_learning

* 版权声明:

- 仅限用于个人学习
- 禁止用于任何商业用途

前馈神经网络

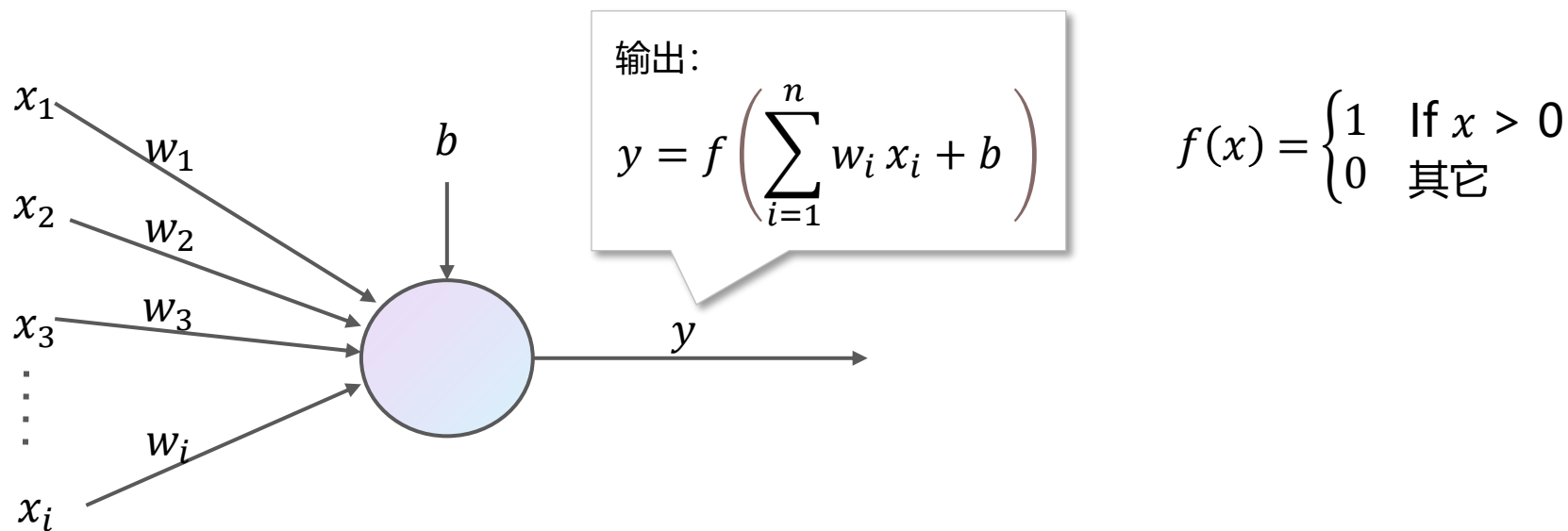
在前馈神经网络中，各神经元分别属于不同的层。整个网络中无反馈，数据从输入层向输出层单向传播，可用一个有向无环图表示。



感知机

感知机是一种最简单形式的前馈神经网络，由美国学者Frank Rosenblatt在1957年提出。感知机接收多个输入信号，并根据每个输入信号的权重进行加权求和，然后将这个加权和加上偏置输入到一个激活函数中，最终输出一个二元的结果（通常是0或1）。

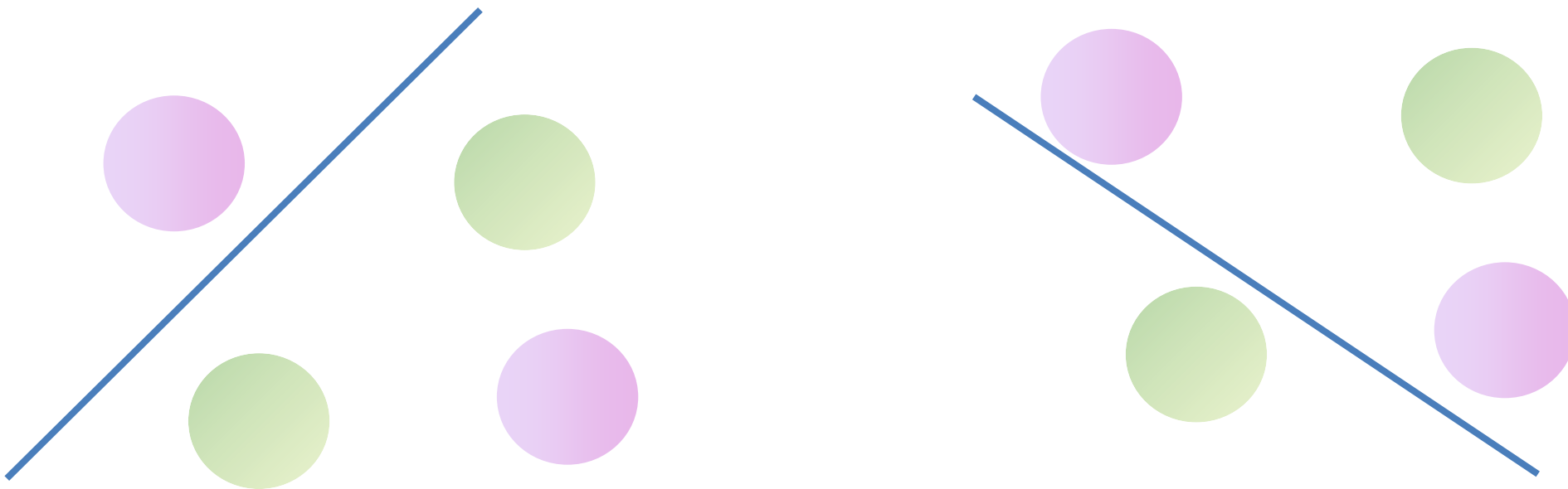
与M-P神经元模型需要人为确定参数不同，感知机**能够通过训练自动更新参数**。其训练方式为有监督学习，即根据感知机的预测结果与真实标签之间的误差来更新权重，直到达到一定的准确度或收敛条件为止。



信号 $x_1, x_2, x_3, \dots, x_i, \dots, x_n$ 输入神经元时，乘以各自的权重 w_i 后求和，加上偏置 b ，再由激活函数 $f()$ 处理后得到输出 y 。

感知机

尽管感知机在简单的线性可分问题上表现良好，最终输出一个二元的结果（通常是0或1）。但是对于非线性可分问题，感知机无法进行有效的学习。



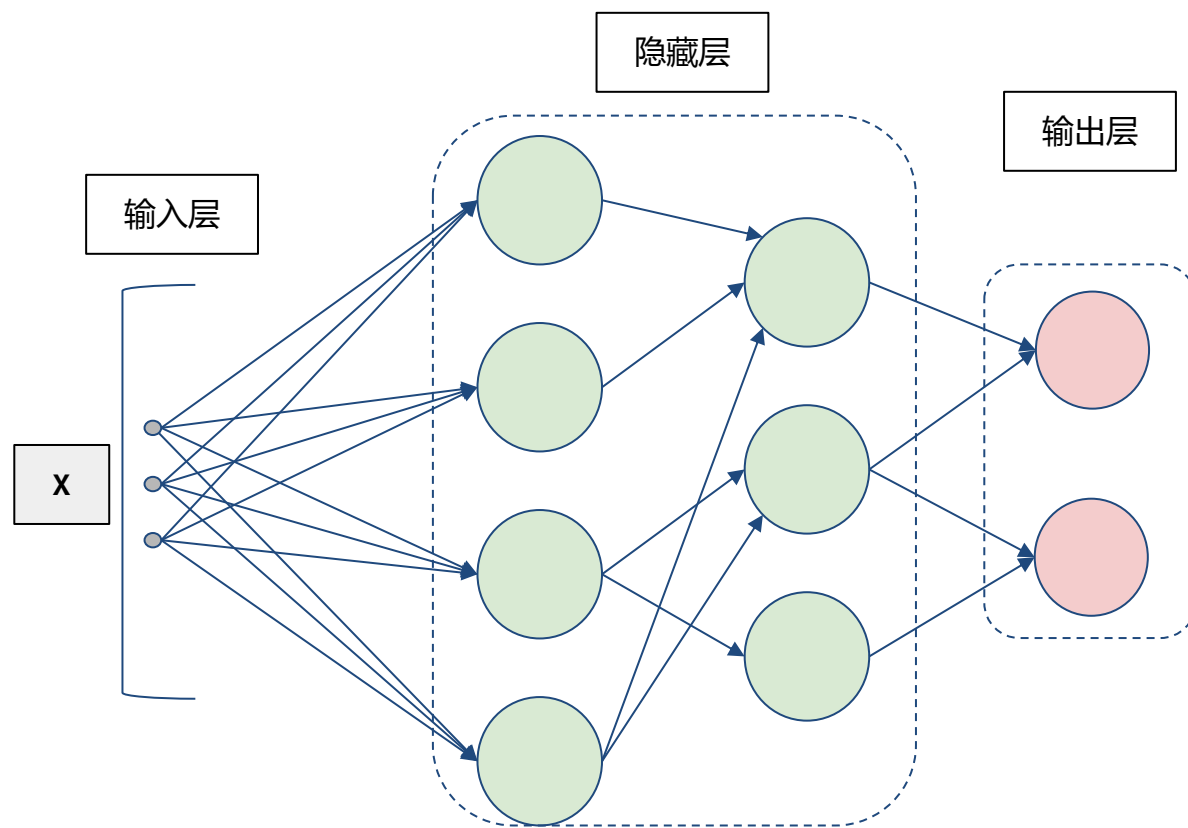
XOR问题 (Minsky & Papert, 1969)
感知机不能拟合 XOR 函数，感知机是一个二分类模型，它只能线性分割。

多层感知机模型

多层感知机（Multilayer Perceptron, MLP）是一种深层前馈神经网络，由一个输入层、一个或多个隐藏层和一个输出层组成。每一层都由多个神经元（或称为节点）组成，相邻层之间的神经元之间有连接，并且每个连接都有一个权重。

MLP模型的工作原理如下：

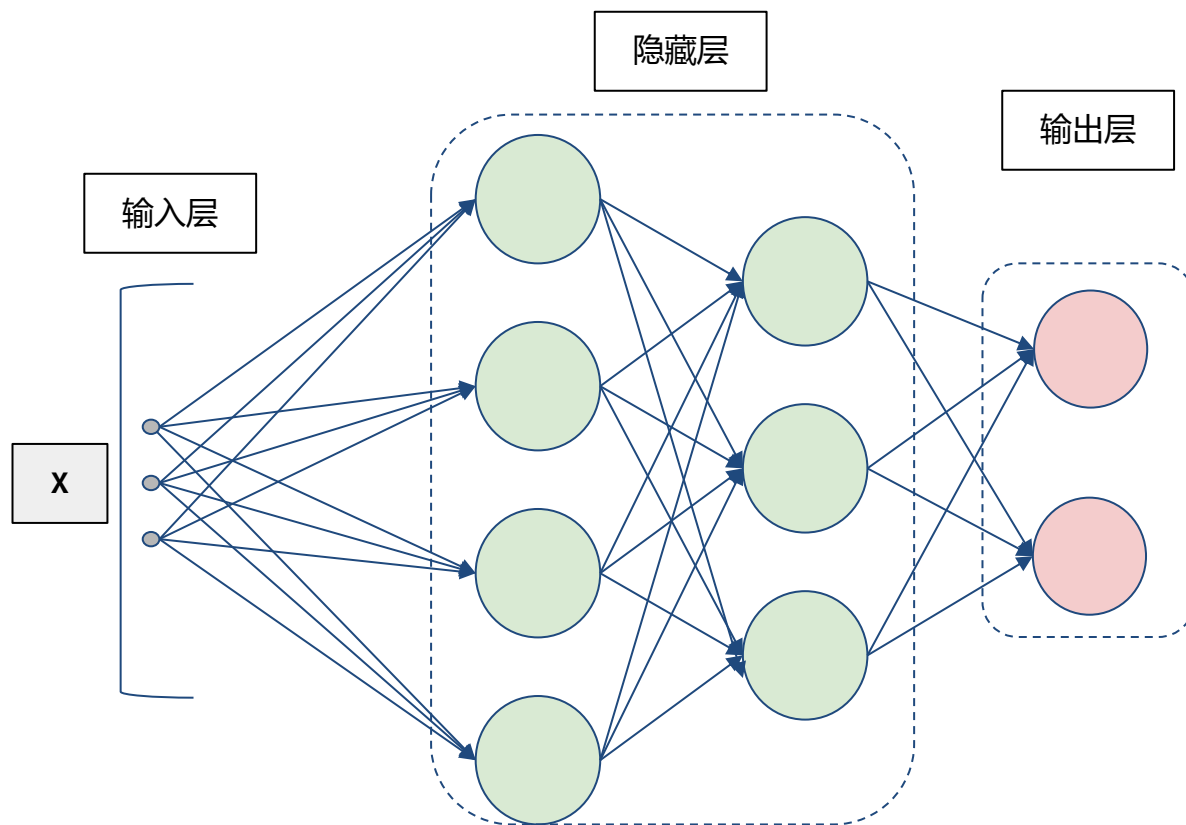
- 输入层：接收输入数据，并将数据传递给下一层。
- 隐藏层：对输入数据进行加权求和，并应用激活函数（如ReLU、Sigmoid、Tanh等）来引入非线性特性。隐藏层的作用是学习输入数据的复杂特征。
- 输出层：接收来自隐藏层的数据，进行加权求和并应用激活函数，输出最终的预测结果。



多层感知器模型

全连接神经网络

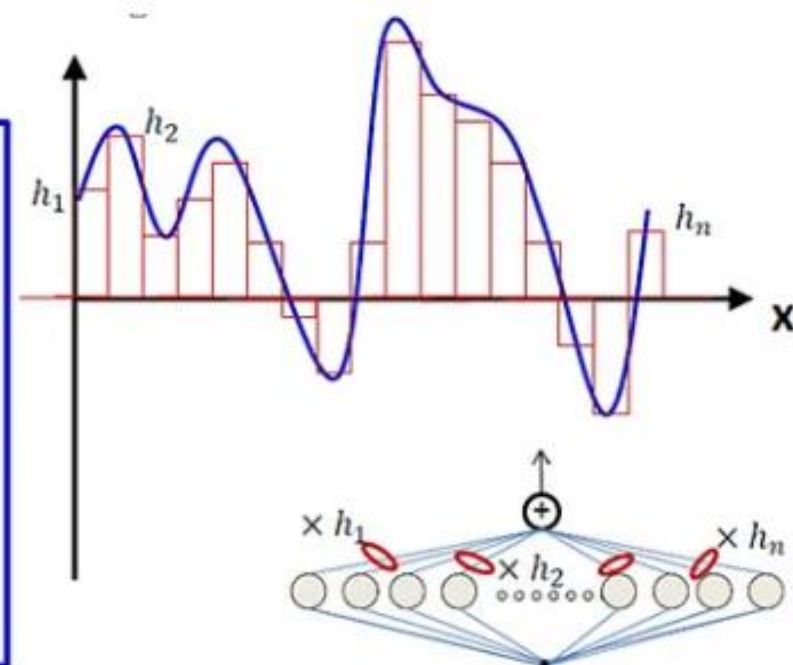
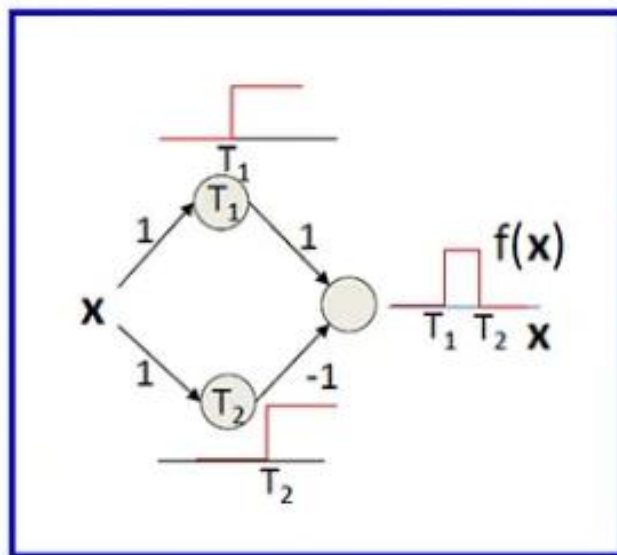
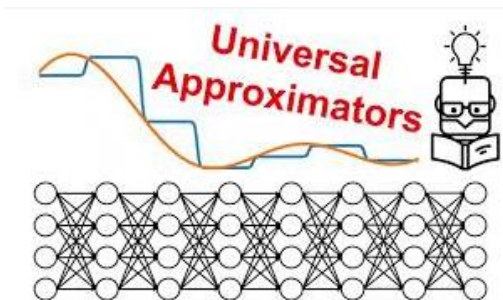
全连接神经网络（Fully Connected Neural Network）是一种最基本的神经网络结构，也被称为多层感知机（Multilayer Perceptron, MLP）。在全连接神经网络中，神经元之间的每一对相邻层之间的神经元都是相互连接的，即每个神经元都与前一层的所有神经元相连。



全连接神经网络（也是前馈神经网络）
图中每一层的每个结点与后一层的所有结点相连接

通用近似定理

通用近似定理 (Universal approximation theorem) 是指神经网络具有足够的神经元 (隐藏层节点) 时, 可以以**任意精度逼近任何连续函数**。这个定理是神经网络理论中的重要结果之一, 它表明具有足够多参数的神经网络可以在理论上逼近任何复杂的函数关系。也就是说神经网络可以作为一个“万能”函数来使用。





Thank

You