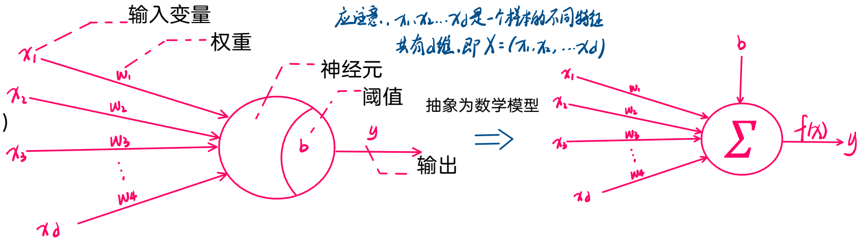
神经网络

神经元模型

- 基本单位：神经元
- 灵感来源
  - 生物神经网络
    - 当神经元兴奋时
    - 向相连的神经元发送化学物质
    - 从而改变这些神经元内的电位
  - 激活
    - 如果某个神经元的电位超过一个阈值
    - 它就会被激活，即兴奋起来
- 对比决策树
  - 决策树模型：通过if-then方式描述非常复杂的规则集合
  - 神经网络模型
    - 可以通过权重调整的方式对任意复杂的函数
    - 以任意精度拟合
- 图示（MP神经元模型）

应注意，$x_1, x_2 \dots x_d$ 是一个样本的不同特征，共有d维，即 $X=(x_1, x_2, \dots x_d)$



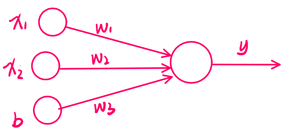
- 公式表达
  - d 样本的维度
  - ω 权重（个数与维度相等）
  - $f(x)$
    - 激活函数或激励函数，如符号函数sign
    - $sign(x)=\begin{cases}1, & x \geq 0 \\ 0, & x<0\end{cases}$ （模拟神经元状态，要么激活，要么没激活）
  - 输出结果：样本与权重相乘，再与阈值相加，通过激活函数得到

一个感知器

- 描述
  - 神经网络之所以能够以任意精度拟合任意复杂的连续函数，一个重要的原因是感知器可以非常容易地实现逻辑与，或，非运算
- 与运算
  - 逻辑表

x_1	x_2	y
0	0	0
0	1	0
1	0	0
1	1	1

  - 感知器



其中b的偏置项，为了优化方便另b=1并赋予1个权重

  - 赋值
    - 假定 $w_1=15$、$w_2=10$，$w_3=-20$，激活函数选用sign函数
    - 则 $y=sign(w_1x_1+w_2x_2+w_3b)=sign(15x_1+10x_2-20)$ 此时就实现了逻辑与（if $y>0$: $y=1$; else: $y=0$）
- 或运算
  - 逻辑表

x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	1

  - 赋值
    - 假定 $w_1=15$、$w_2=50$，$w_3=-10$
    - 则 $y=sign(w_1x_1+w_2x_2+w_3b)=sign(15x_1+50x_2-10)$
- 非运算
  - 逻辑表

x_1	y
0	1
1	0

  - 赋值
    - 假定 $w_1=-15$，$w_2=0$，$w_3=10$
    - 则 $y=sign(w_1x_1+w_2x_2+w_3b)=sign(-15x_1-10)$

感知器模型可以非常容易地解决线性问题，而对于非线性问题，如最简单的异或问题，则无法采用一个感知器模型将其完全分离。

若采用两个感知器模型，则每个感知器就只学习一个线性分界线，然后通过逻辑与运算，取两部分的并集，就能完美结局异或问题。

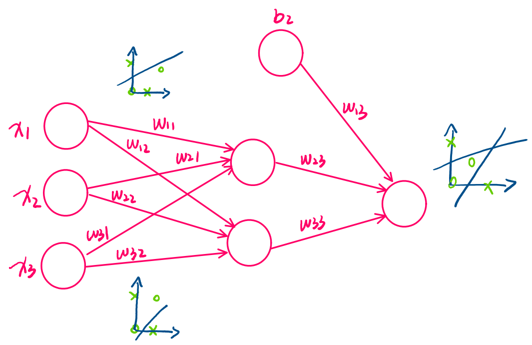
神经网络模型沿用了这种思想，采用增加感知器的方式在原始空间解决非线性分类。

神经网络的网络结构

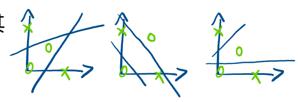
若采用两个感知器模型，则每个感知器就只学习一个线性分界线，然后通过逻辑与运算，取两部分的并集，就能完美结局异或问题。
神经网络模型沿用了这种思想，采用增加感知器的方式在原始空间解决非线性分类。

两个感知器

异或问题：当一条分界线不足以分类时，可以添加多条分界线

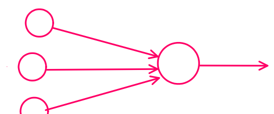
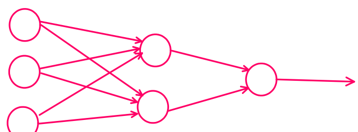
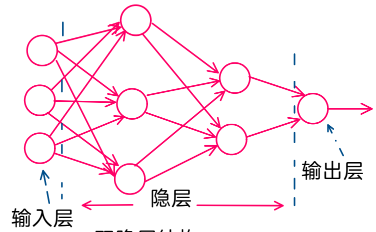


应注意，由于感知器模型得到的分界线并不唯一，因此有感知器组成的神经网络模型其学习结果也不是唯一的。因此上述求与运算的两个分界线可能有多种形式。

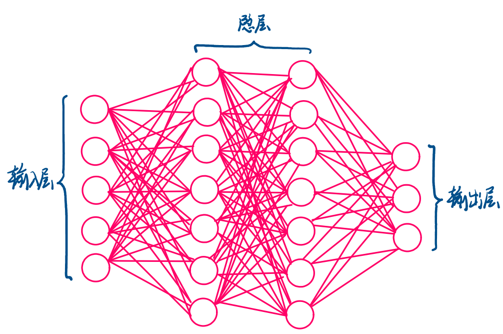


神经网络

多个感知器模型构成了神经网络模型，神经网络中随着感知器的增加，其模型表达能力也会增加

网络结构	分类区域	可理解的问题
 无隐层结构	线性超平面 n维线性空间中，维度为n-1的子空间	线性可分问题，逻辑与或非运算
 单隐层结构	开的凸区或闭的凹区 针对欧氏空间	简单的非线性问题，如异或问题
 双隐层结构	任意形状	任意复杂的非线性问题

多层前馈网络



第一层：输入层，没有激活函数，只是将输入信号传递给隐层的神经元

最后一层
- 输出层，有激活函数
- 负责将中间层的结果加权
- 实现不同的逻辑运算并得到最后的预测结果

中间层
- 隐层，有激活函数
- 隐层可以有多层，每层中可以有多个神经元
- 隐层中第一层的每一个神经元都与前一层（输入层）的所有神经元构成一个感知器模型（即逻辑回归模型）
- 隐层中第二层的每一个神经元都与前一层（隐层的第一层）的所有神经元构成一个感知器模型

- 多层前馈网络
  - 中间层
    - 隐层可以有多层，每层中可以有多个神经元
    - 隐层中第一层的每一个神经元都与前一层（输入层）的所有神经元构成一个感知器模型（即逻辑回归模型）
    - 隐层中第二层的每一个神经元都与前一层（隐层的第一层）的所有神经元构成一个感知器模型
  - 连接方式
    - 前馈网络是指每层的神经元与下一层的神经元以全连接的形式互联
    - 同层的神经元之间不存在连接
    - 跨层的神经元之间不存在跨层连接
  - 参数
    - 位置
      - 输入层与隐层之间有权重
      - 隐层与隐层之间有权重
      - 隐层与输出层之间有权重
    - 假设
      - 输入数据的维度为m
      - 采用双隐层网络
        - 第一个隐层采用n个隐层节点（即n个神经元）
        - 第二个隐层采用t个隐层节点
      - 输出层为s个神经元
    - 权重个数
      - $sum(w)=(m+1)n+(n+1)t+(t+1)s$ （加1是因为每个神经元都有一个偏置项b）
      - 在上图中，权重个数为$sum(w)=(5+1)\times 7+(7+1)\times 7+(7+1)\times 3=122$，即图中需要训练的权重数为122