



神经网络算法

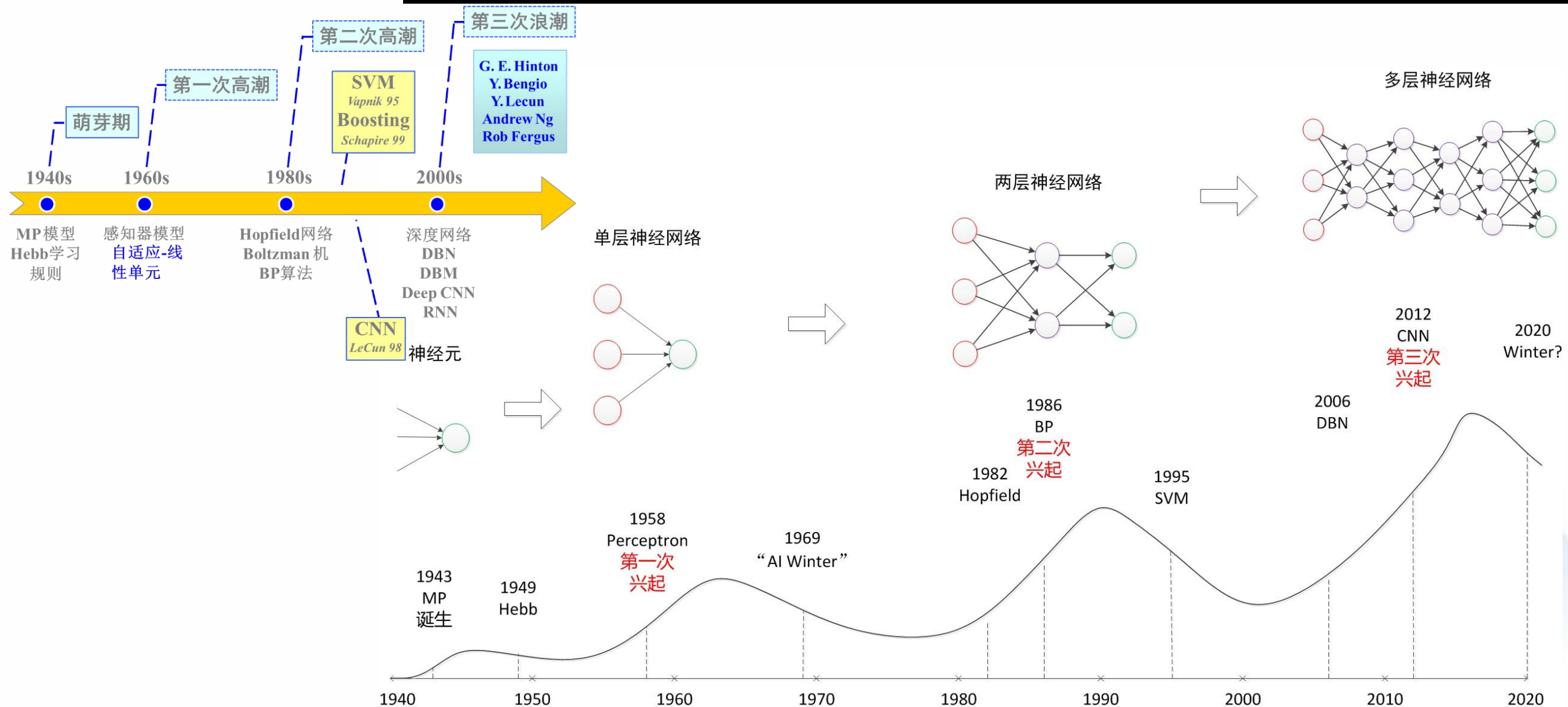
Artificial Neural Network, ANN

肖 玲



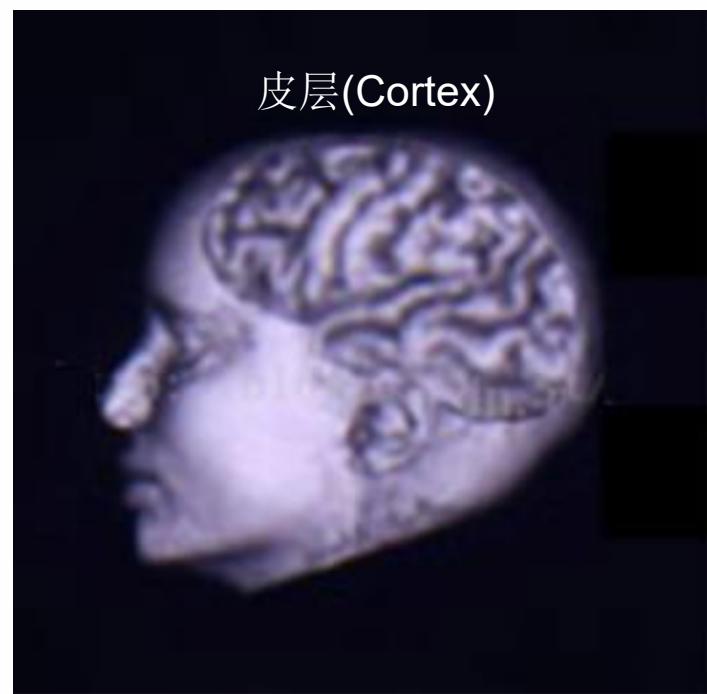


神经网络的“三起三落”





- 神经网络是一种模拟人脑的神经网络以期能够实现类人工智能的机器学习技术。



皮层(Cortex)



大脑有两层: 灰色的外层和白色的内层

<http://t/zzwu>

1 大脑半球像半个核桃

2 大脑皮层由灰质白质组成
图 大脑的外形和切片形状

表1人和几种动物的神经细胞的数目

动物种类	神经细胞的数目 (数量级)
蜗牛	10,000 ($=10^4$)
蜜蜂	100,000 ($=10^5$)
蜂雀	10,000,000 ($=10^7$)
老鼠	100,000,000 ($=10^8$)
人类	86,000,000,000 ($=10^{11}$)
大象	100,000,000,000 ($=10^{11}$)

灰色层只有几毫米厚, 其中紧密地压缩着几十亿个被称作**neuron** (神经细胞、神经元) 的微小细胞。白色层在皮层灰质的下面, 占据了皮层的大部分空间, 是由神经细胞相互之间的无数连接线组成。

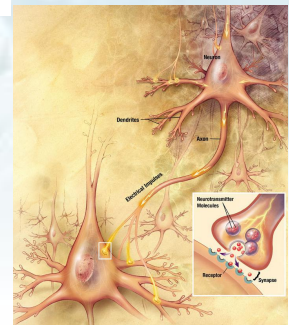
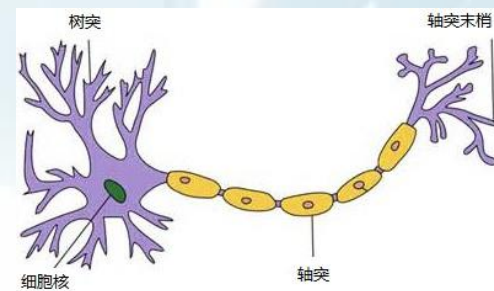
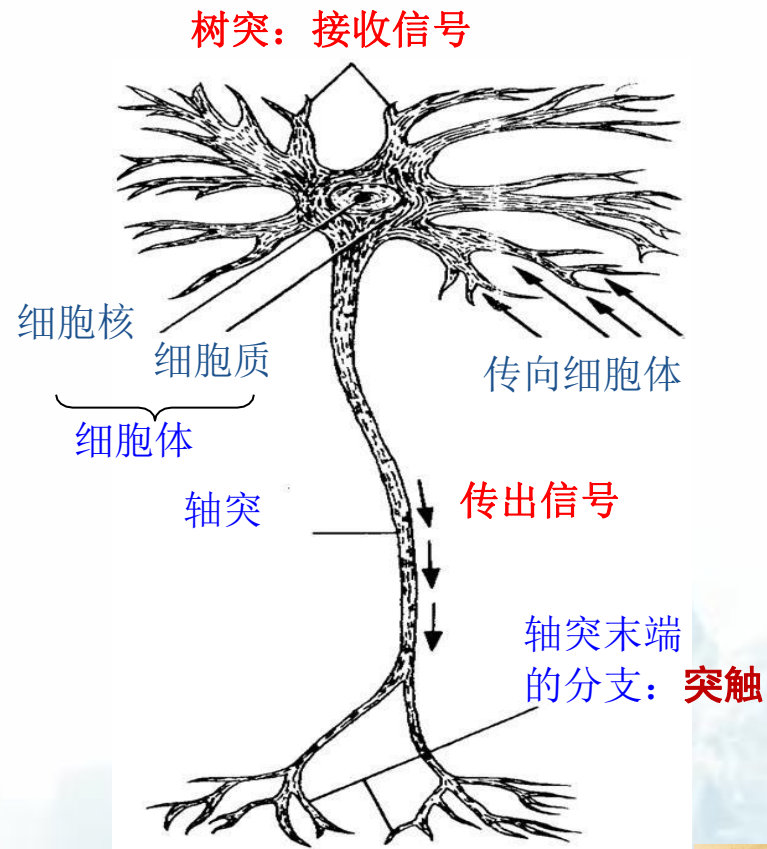
生物神经元

□生物神经元

- ①神经元间通过突触两两相连
- ②树突接收来自多个神经元的信号
- ③轴突根据树突传递过来的综合信号的强弱是否超过某一阈值来决定是否将该信号传递给下一个神经元

□生物神经元的启示

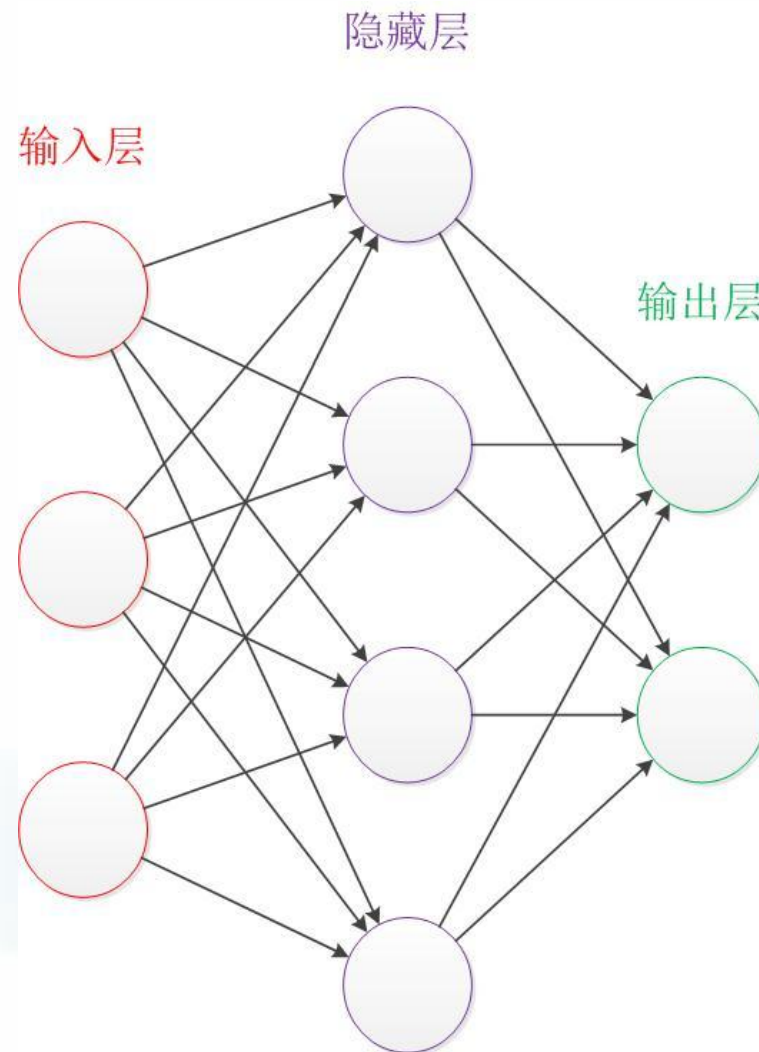
- ①每个神经元都是一个**多输入单输出**的信号处理单元
- ②神经元具有**阈值特性**





经典的神经网络结构

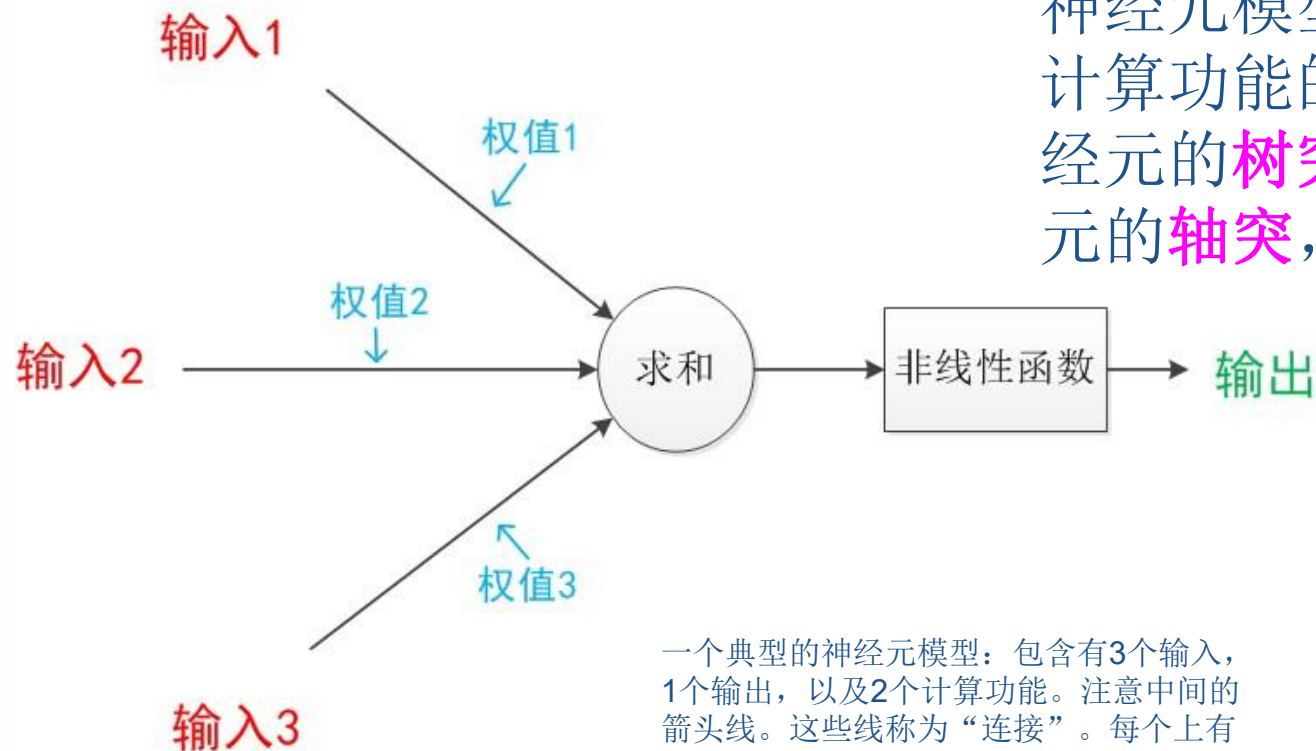
1. 设计一个神经网络时，输入层与输出层的节点数往往是固定的，中间层则可以自由指定；
2. 神经网络结构图中的拓扑与箭头代表着预测过程时数据的流向，跟训练时的数据流有一定的区别；
3. 结构图里的关键不是圆圈（代表“神经元”），而是**连接线**（代表“神经元”之间的连接）。每个连接线对应一个不同的权重（其值称为**权值**），这是需要训练得到的。





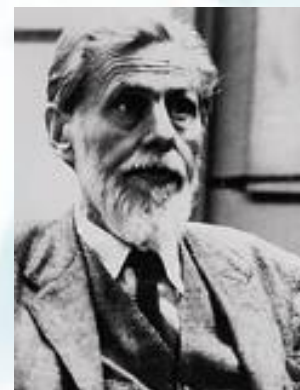
神经元模型MP

- 1943年，心理学家McCulloch和数学家Pitts参考了生物神经元的结构，发表了抽象的神经元模型MP。



一个典型的神经元模型：包含有3个输入，1个输出，以及2个计算功能。注意中间的箭头线。这些线称为“连接”。每个上有一个“权值”。连接是神经元中最重要的东西。每一个连接上都有一个权重。

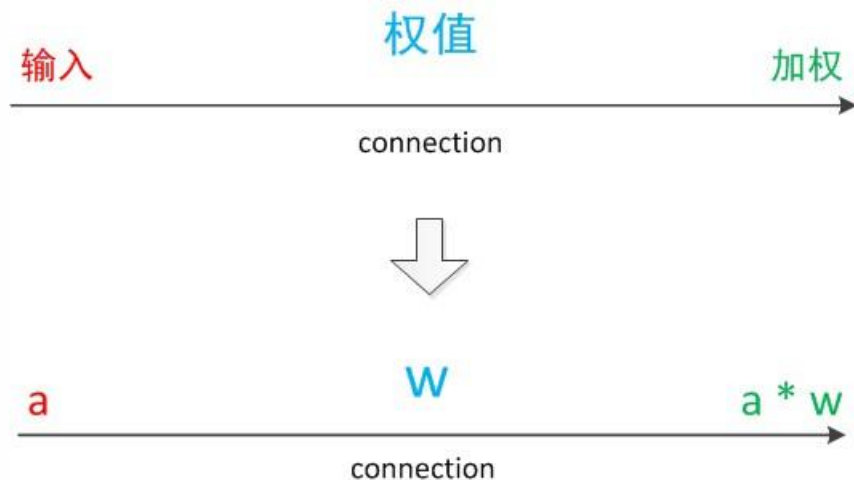
神经元模型是一个包含输入，输出与计算功能的模型。**输入**可以类比为神经元的**树突**，而**输出**可以类比为神经元的**轴突**，计算则可以类比为细胞核。



Warren McCulloch和 Walter Pitts

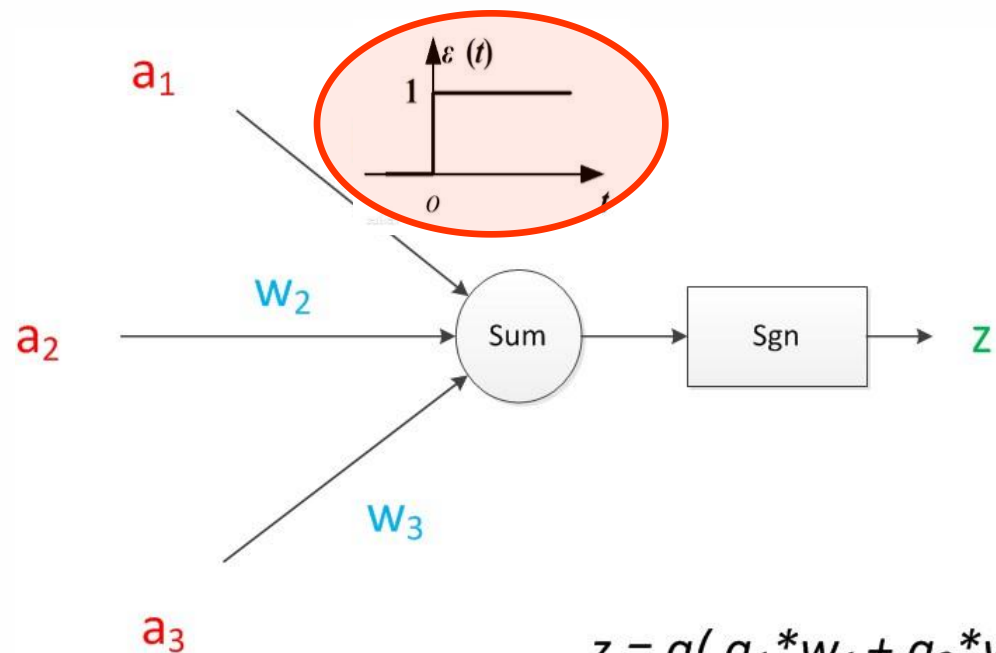


神经元模型MP



我们使用 a 来表示输入，用 w 来表示权值。一个表示连接的有向箭头可以这样理解：在初端，传递的信号大小仍然是 a ，端中间有加权参数 w ，经过这个加权后的信号会变成 $a*w$ ，因此在连接的末端，信号的大小就变成了 $a*w$ 。而在神经元模型里，每个有向箭头表示的是值的加权传递。

如果我们将神经元图中的所有变量用符号表示，并且写出输出的计算公式的话，就是右图。可见 z 是在输入和权值的线性加权和叠加了一个函数 g 的值。在MP模型里，函数 g 是sgn函数，也就是取符号函数。这个函数当输入大于0时，输出1，否则输出0。

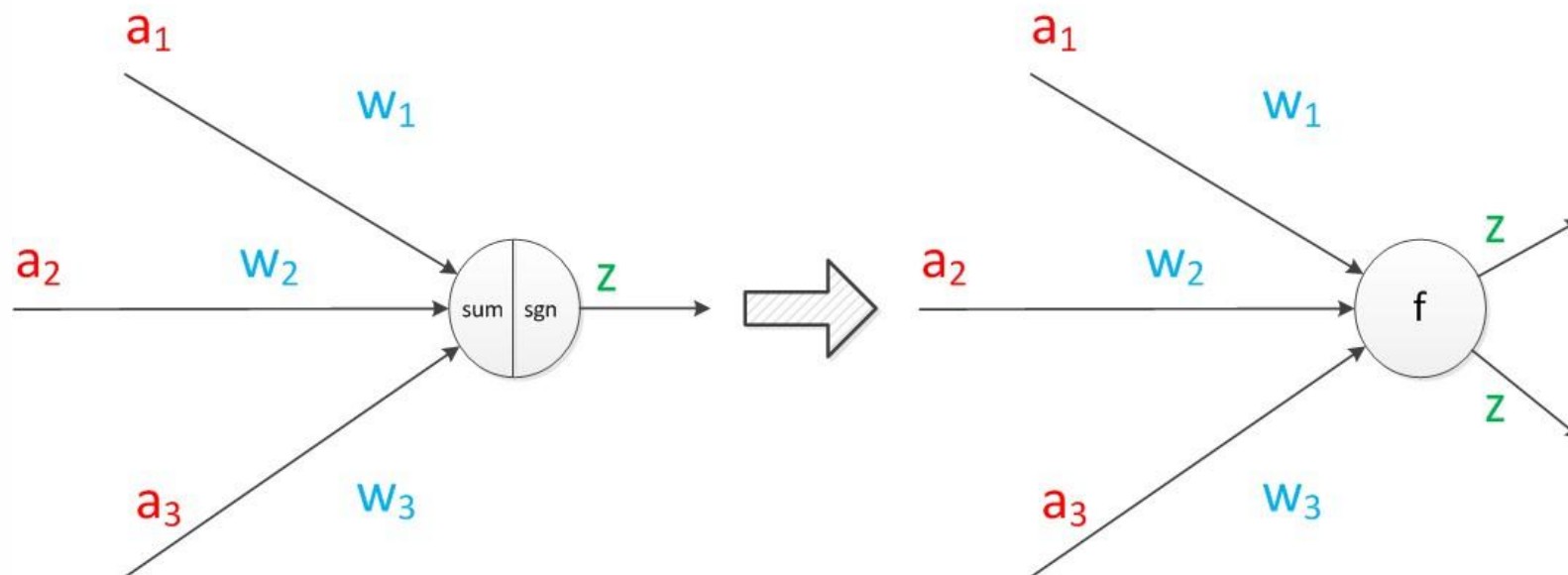


$$z = g(a_1 * w_1 + a_2 * w_2 + a_3 * w_3)$$



神经元模型MP

- 当我们用“神经元”组成网络以后，描述网络中的某个“神经元”时，我们更多地会用“单元”（unit）来指代。



下面对神经元模型的图进行一些扩展。首先将sum函数与sgn函数合并到一个圆圈里，代表神经元的内部计算。其次，把输入 a 与输出 z 写到连接线的左上方，便于后面画复杂的网络。最后说明，一个神经元可以引出多个代表输出的有向箭头，但值都是一样的。

神经元可以看作一个计算与存储单元。计算是神经元对其输入进行计算功能。存储是神经元会暂存计算结果，并传递到下一层。

神经元模型的使用可以这样理解：

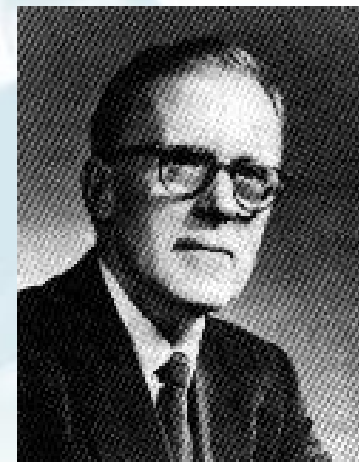
我们有一个数据，称之为样本。样本有四个属性，其中三个属性已知，一个属性未知。我们需要做的就是通过三个已知属性预测未知属性。

具体办法就是使用神经元的公式进行计算。三个已知属性的值是 a_1 ， a_2 ， a_3 ，未知属性的值是 z 。 z 可以通过公式计算出来。

这里，已知的属性称之为特征，未知的属性称之为目标。假设特征与目标之间确实是线性关系，并且我们已经得到表示这个关系的权值 w_1 ， w_2 ， w_3 。那么，我们就可以通过神经元模型预测新样本的目标。



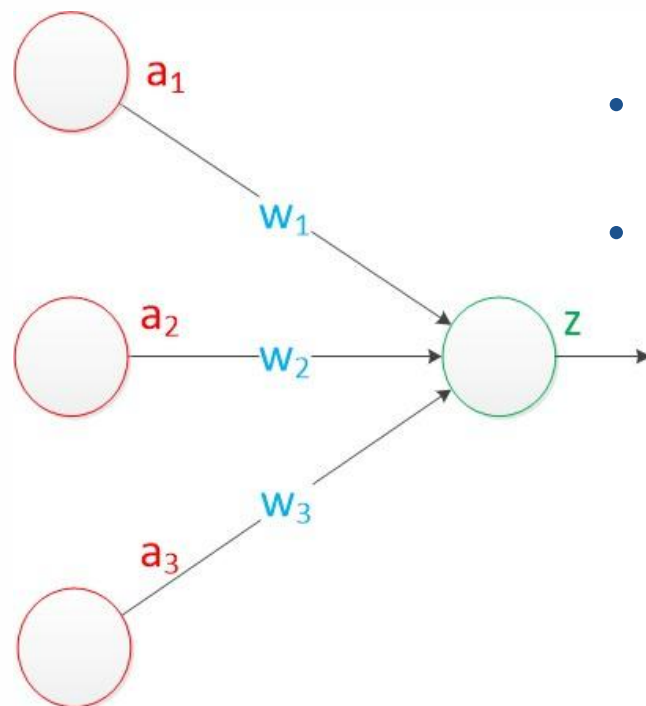
- 1943年发布的MP模型，虽然简单，但已经建立了神经网络大厦的地基。但是，MP模型中，权重的值都是预先设置的，因此不能学习。
- 1949年心理学家Hebb提出了Hebb学习率，认为人脑神经细胞的突触（也就是连接）上的强度上可以变化的。于是计算科学家们开始考虑用调整权值的方法来让机器学习。这为后面的学习算法奠定了基础。
- 尽管神经元模型与Hebb学习律都已诞生，但限于当时的计算机能力，直到接近10年后，第一个真正意义的神经网络才诞生。





单层神经网络（感知器）

1958年，计算科学家Rosenblatt提出了由两层神经元组成的神经网络——“感知器”（Perceptron）



- 在“感知器”中，有两个层次。分别是输入层和输出层。
- 输入层里的“**输入单元**”只负责传输数据，不做计算。
- 输出层里的“**输出单元**”则需要对前面一层的输入进行计算。

在原来MP模型的“输入”位置添加神经元节点，标志其为“输入单元”。其余不变，从本图开始，我们将权值 w_1 , w_2 , w_3 写到“连接线”的中间。

