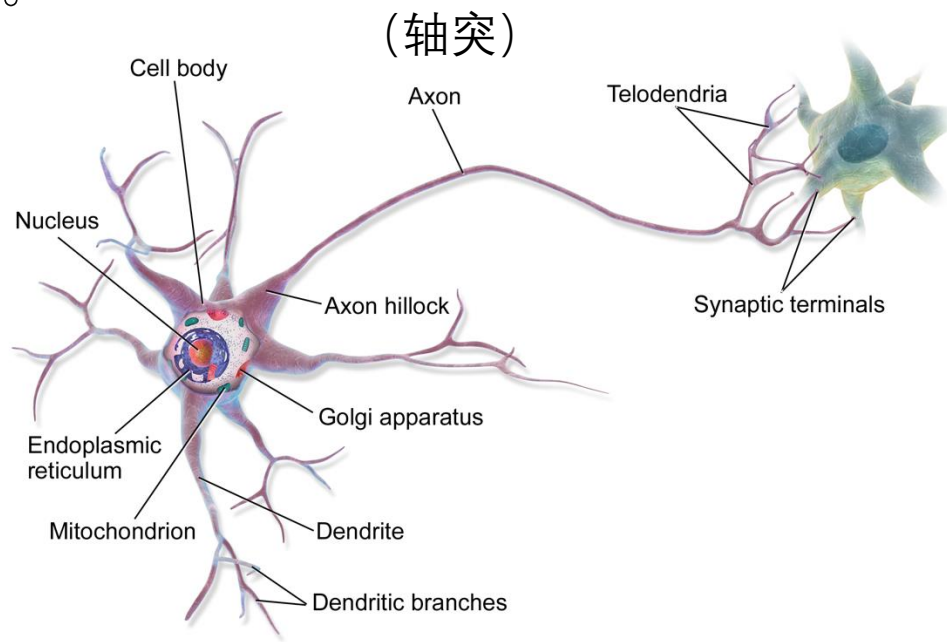


第5章 人工神经网络

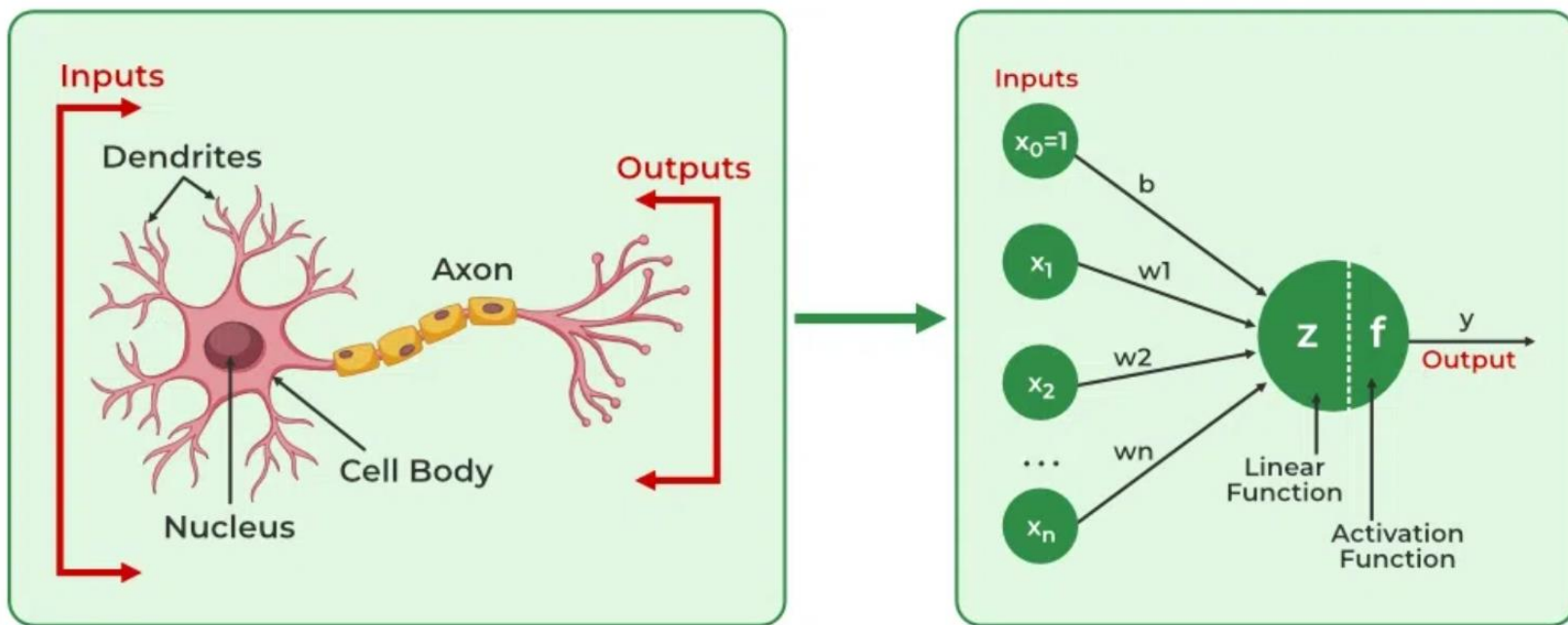
神经网络

- 生物学中的**神经网络**指的是由神经元（即神经细胞）组成的复杂网络系统，这些神经元通过突触相互连接并传递电信号和化学信号，以实现信息处理和传递。
- **神经元 (Neurons)**：生物神经元是神经系统的基本单元，负责接收、处理和传递信息。
- **突触 (Synapses)**：突触是神经元之间的连接点，通过化学或电信号传递信息。突触可以是兴奋性或抑制性的，决定了信号传递的方式和强度。



人工神经网络

- 生物神经网络是人工神经网络（ANN）发展的灵感来源。
- 人工神经网络（Artificial Neural Networks, ANN）是一种模仿生物神经网络（例如人脑）的信息处理系统。它由许多相互连接的“神经元”组成，每个神经元接收输入信号，并通过加权求和、激活函数处理，产生输出信号。



人工神经网络发展史

年份	事件	重要性
1958	感知器（Perceptron）提出	开启神经网络研究的先河
1986	反向传播算法（Backpropagation）普及	使得多层神经网络的训练成为可能
1998	LeNet-5 提出	在手写数字识别上取得成功
2006	深度信念网络（Deep Belief Networks, DBN）提出	引领深度学习的兴起
2012	AlexNet 在 ImageNet 上取得突破	推动深度学习在图像识别领域的应用
2014	VGG 和 Inception 模型提出	进一步提升图像识别性能
2015	ResNet 提出	解决了深层网络的退化问题
2017	Transformer 模型提出	变革了自然语言处理领域
2018	BERT 提出	实现了双向语言表示的重大突破
2020	GPT-3 发布	推动语言生成模型的发展

人工神经网络： 神经元

- 神经网络的核心组件是**神经元**， 每一个网络由**大量的神经元**构成
- 在成人大脑中， 神经元的总数在800亿到1000亿之间。

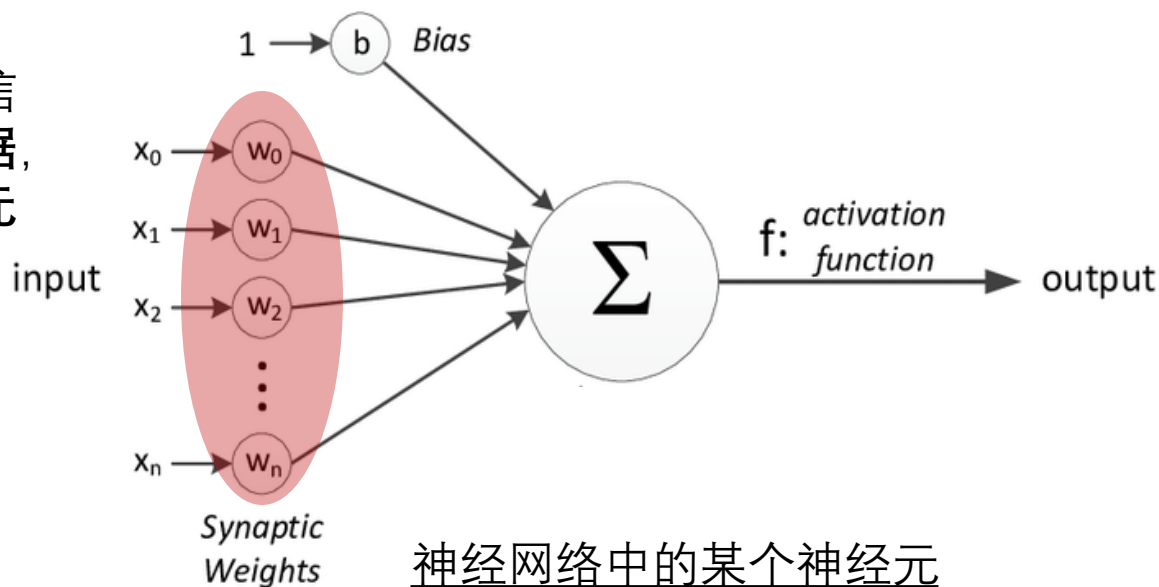
	神经网络模型	参数数量（百万）
1	LeNet-5	0.06
2	AlexNet	60
3	VGG-16	138
4	Inception-v1 (GoogLeNet)	6
5	ResNet-50	25
6	BERT-base	110
7	BERT-large	340
8	GPT-3	1750
9	EfficientNet-B0	5.3
10	EfficientNet-B7	66

更多的参数意味着网络可以学习和表示更复杂的数据模式，但也需要更多的数据和计算资源来训练。

人工神经网络：连接和权重

- 神经元之间通过连接相互作用，每个连接有一个对应的权重。
- 权重决定了输入信号对神经元输出的影响力大小。权重是神经网络中需要学习和调整的参数。

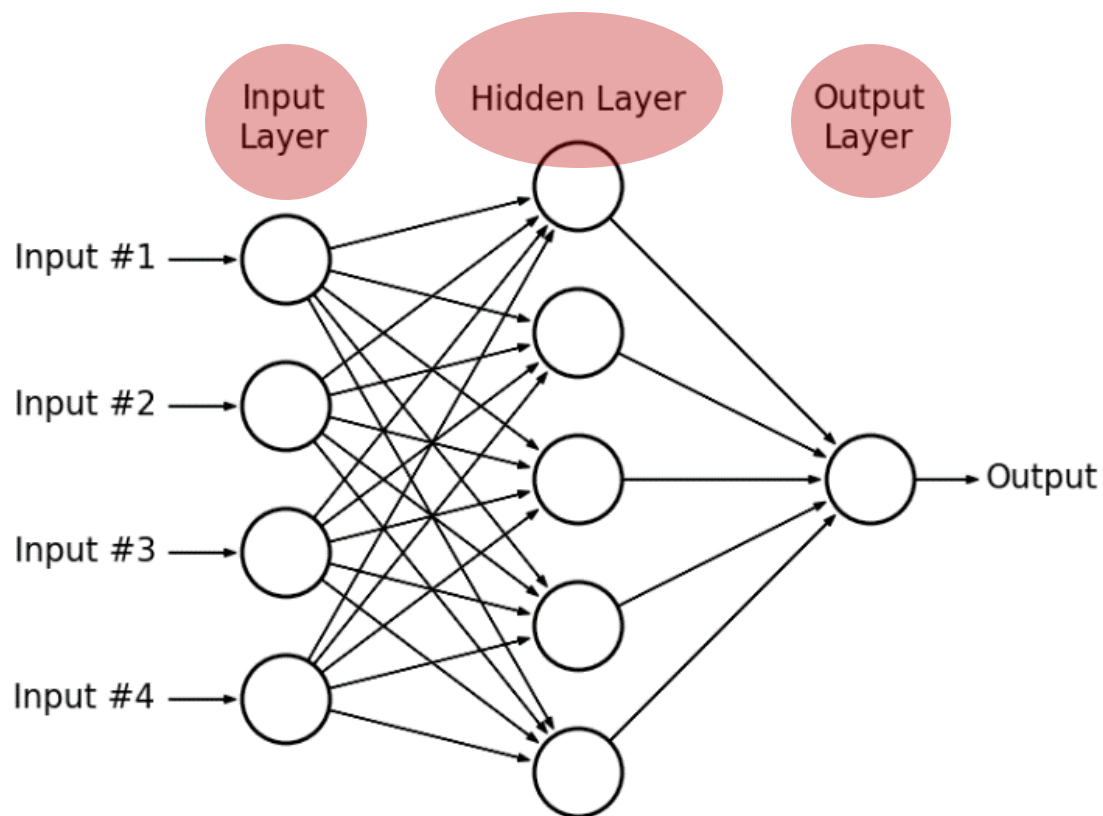
某个神经元的输入信号可以来自**原始数据**，也可以是**其它神经元的输出**



$$Output = f(w_0x_0 + w_1x_1 + w_2x_2 + \dots + w_nx_n + b)$$

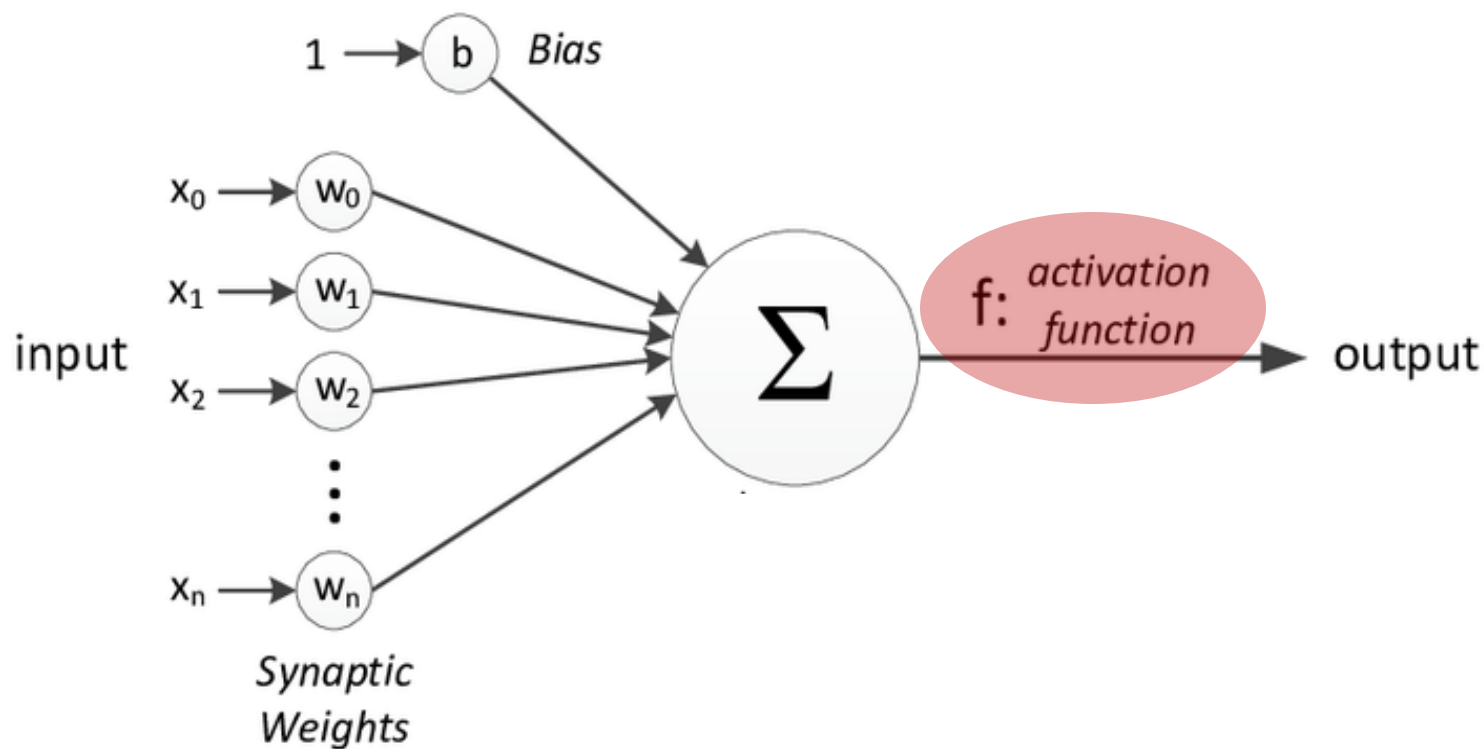
人工神经网络：层

- 神经网络通常分为三种层：输入层、隐藏层和输出层。
- **输入层**：接收外部数据作为输入信号。
- **隐藏层**：位于输入层和输出层之间，负责中间处理和特征提取。
- **输出层**：产生最终结果或预测值。



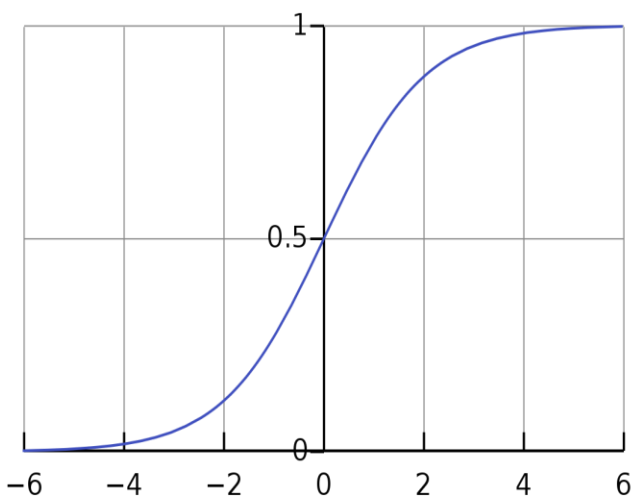
人工神经网络：激活函数

- 激活函数是非线性函数，用于将神经元的加权和转换为输出信号。
- 引入非线性。



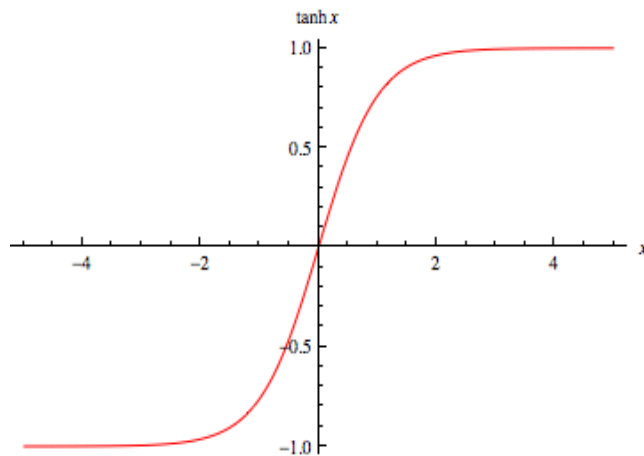
人工神经网络：激活函数

- 常用的激活函数 (Sigmoid, Tanh, ReLU)



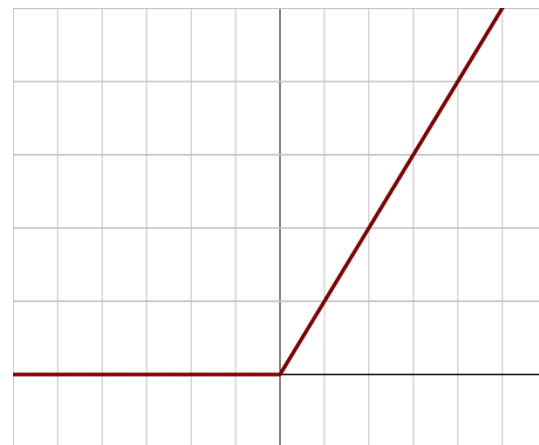
Sigmoid函数

$$f(x) = \frac{1}{1 + e^{-x}}$$



tanh函数

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$



ReLU函数

$$f(x) = \max(0, x)$$

$$f'(z) = f(z)(1 - f(z))$$

$$f'(z) = 1 - f^2(z)$$

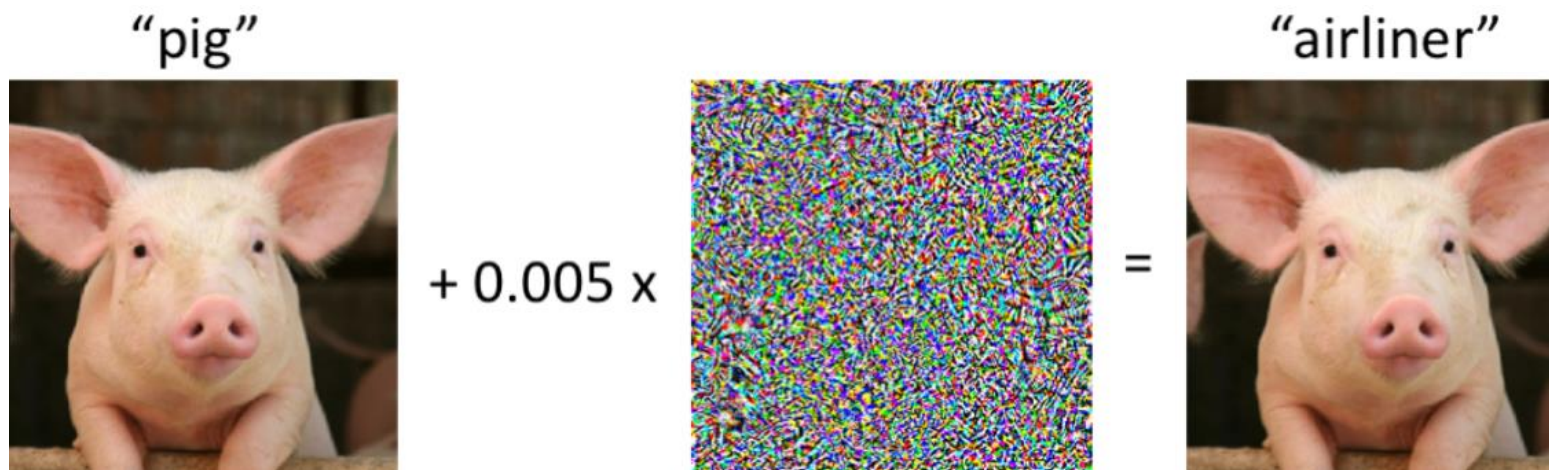
人工神经网络的优点

- **并行处理**：神经网络中的神经元可以同时处理多个输入信号，这使得神经网络具有并行处理的能力。这种并行性使得神经网络在处理大量数据时具有优势。
- **分布式表示与学习**：神经网络的知识以分布式的方式存储在连接权重中。这种表示方式使得神经网络具有较强的容错能力，即使部分神经元受损，仍然可以保持一定程度的性能。
- **非线性映射**：神经网络可以表示和处理非线性问题，这使得它们在许多复杂任务上具有优势。非线性激活函数使神经网络能够学习并表示复杂的非线性关系。

神经网络的缺点

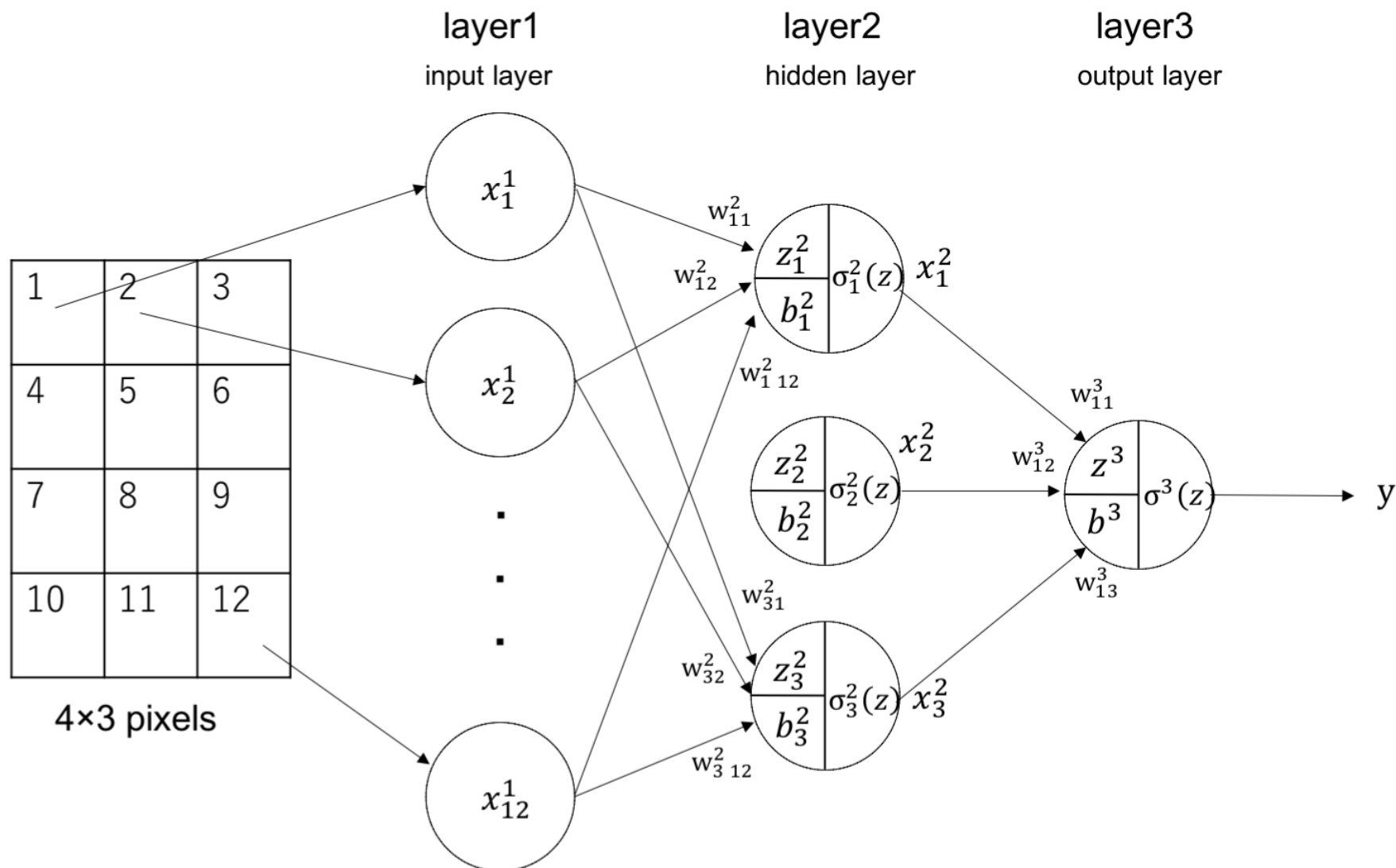
- 神经网络是一个“黑箱” $\Leftrightarrow$ 神经网络的可解释性差。

原因：参数多（数以百万，十亿计），非线性变换多，导致难以追踪和解释神经网络的决策过程。



对抗攻击：通过对输入数据进行微小的、人类不易察觉的修改，使模型产生错误的预测或分类结果。

人工神经网络：前向传播



人工神经网络：前向传播

1. 计算隐藏层的加权输入和激活值：

- 隐藏层加权输入 $z^{(1)}$ ：

$$z^{(1)} = W^{(1)} \cdot x + b^{(1)}$$

- 隐藏层激活值 $a^{(1)}$ ：

$$a^{(1)} = \sigma(z^{(1)})$$

其中 σ 是Sigmoid激活函数，定义为 $\sigma(z) = \frac{1}{1+e^{-z}}$ 。

2. 计算输出层的加权输入和激活值：

- 输出层加权输入 $z^{(2)}$ ：

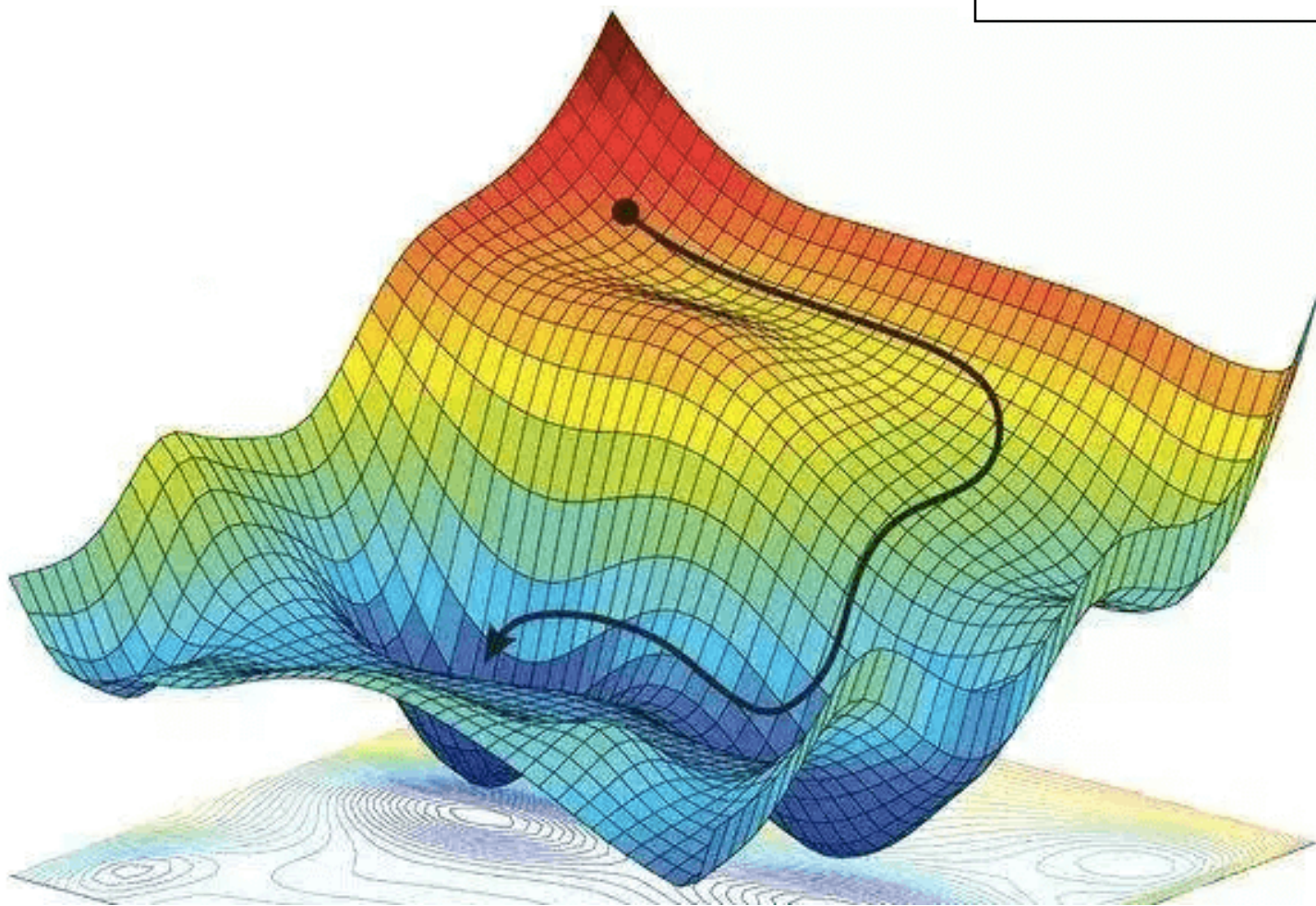
$$z^{(2)} = W^{(2)} \cdot a^{(1)} + b^{(2)}$$

- 输出层激活值 $a^{(2)}$ ：

$$a^{(2)} = \sigma(z^{(2)})$$

梯度下降

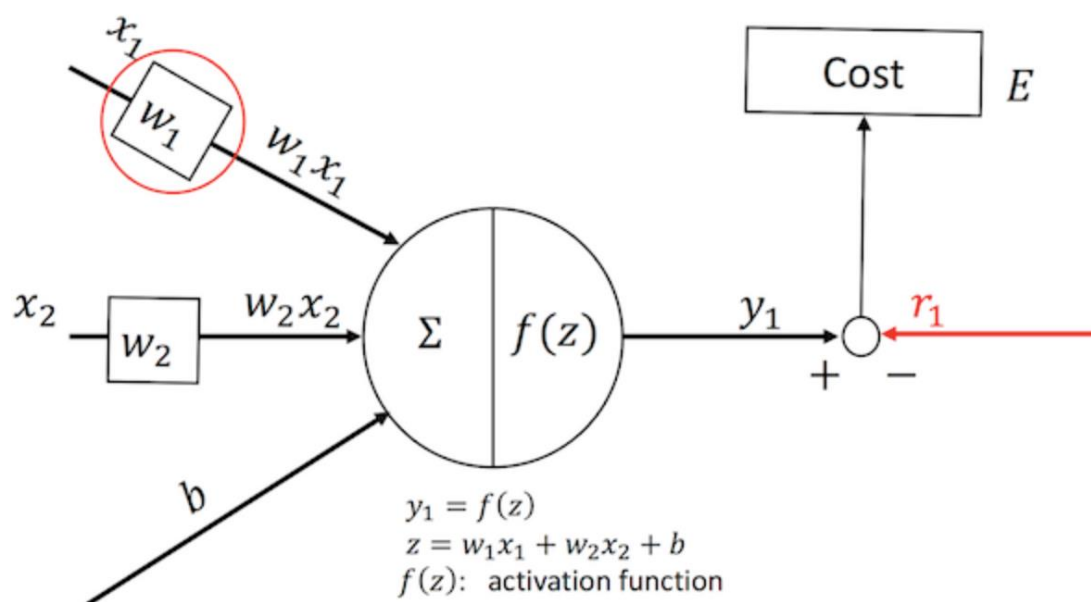
$$w_{t+1} \leftarrow w_t - \gamma \frac{\partial E}{\partial w}$$



人工神经网络：反向传播

- E 是损失函数

$$w_{t+1} \leftarrow w_t - \gamma \frac{\partial E}{\partial w}$$



$$z = w_1x_1 + w_2x_2 + b$$

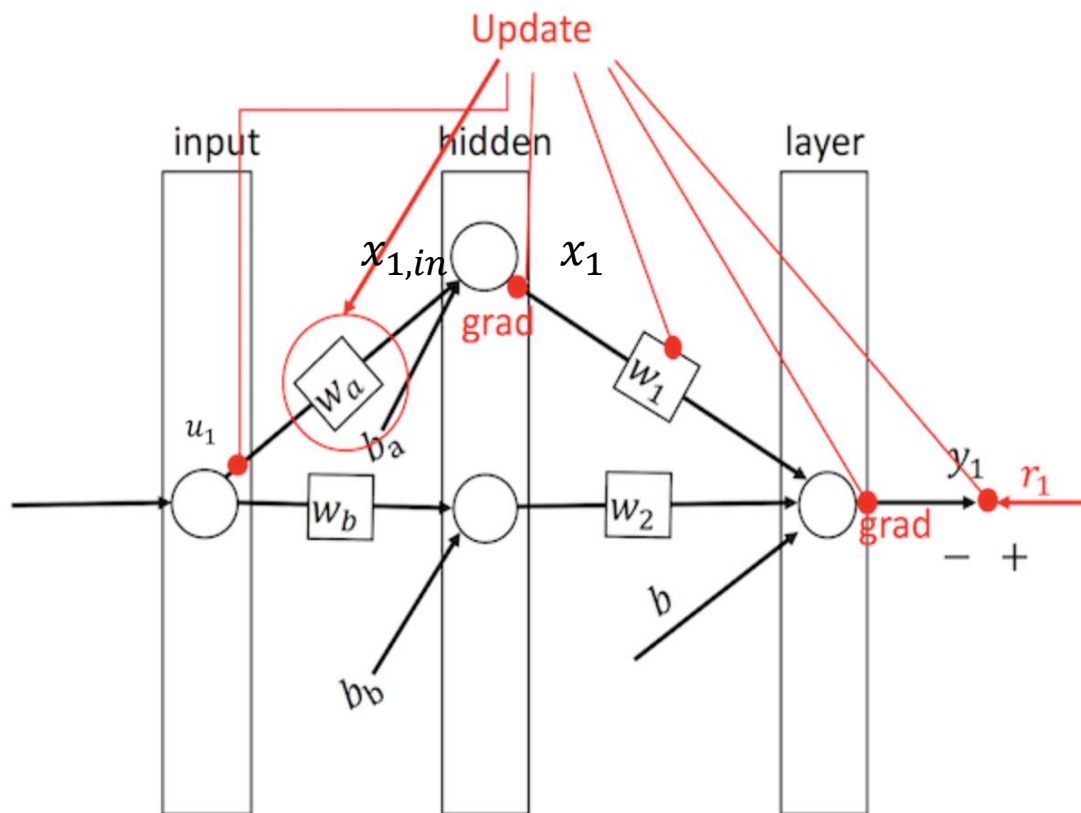
$$y_1 = f(z)$$

损失函数: $E = (y_1 - r_1)^2$

$$\frac{\partial E}{\partial w_1} = \frac{\partial E}{\partial y_1} \frac{\partial y_1}{\partial z} \frac{\partial z}{\partial w_1}$$

人工神经网络：反向传播

$$w_{t+1} \leftarrow w_t - \gamma \frac{\partial E}{\partial w}$$



$$x_{1,in} = w_a u_1 + b_a$$

$$x_1 = f(x_{1,in})$$

$$z_1 = w_1 x_1 + w_2 x_2 + b$$

$$y_1 = f(z_1)$$

$$E = (y_1 - r_1)^2$$

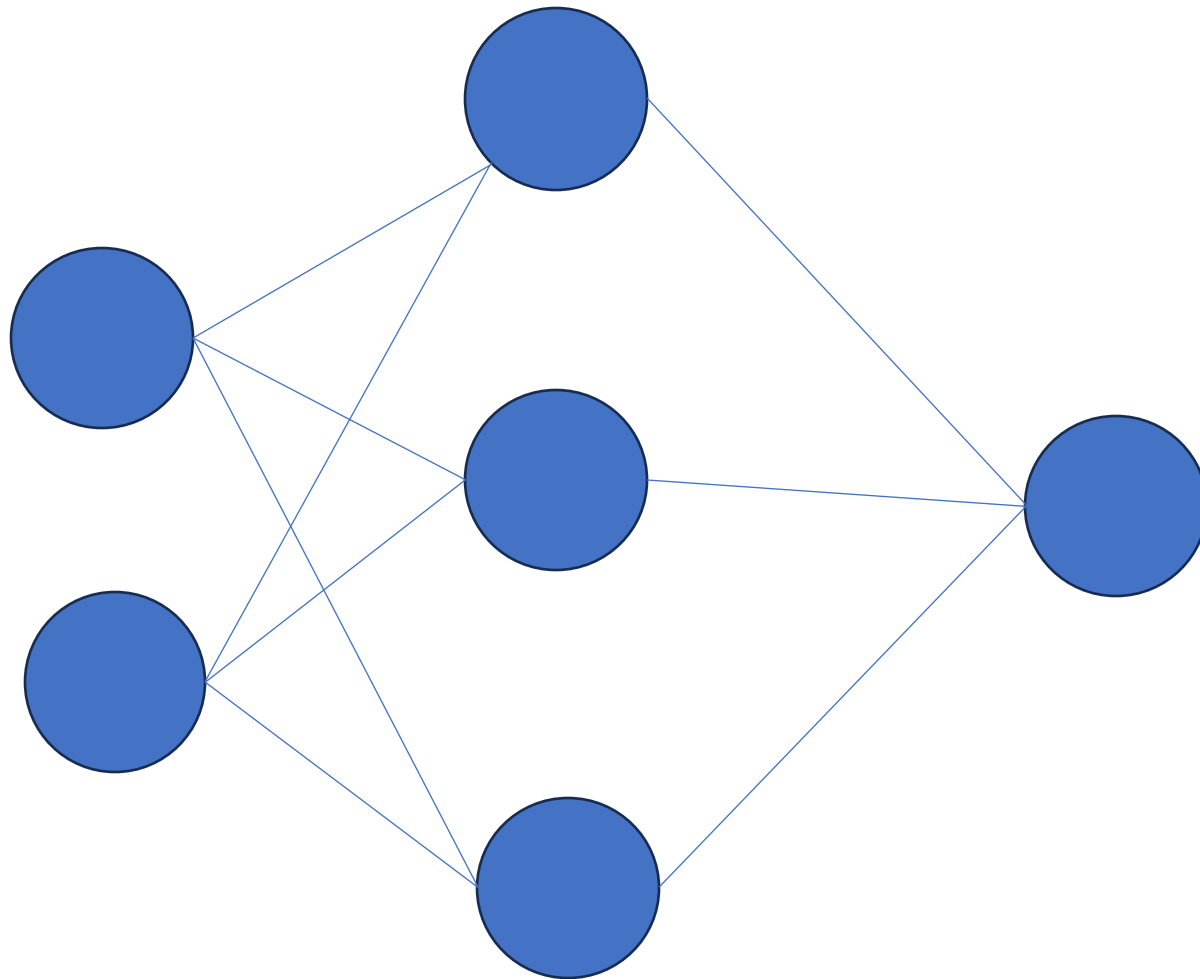
$$\frac{\partial E}{\partial w_a} = \frac{\partial E}{\partial y_1} \frac{\partial y_1}{\partial z_1} \frac{\partial z_1}{\partial x_1} \frac{\partial x_1}{\partial x_{1,in}} \frac{\partial x_{1,in}}{\partial w_a}$$

人工神经网络：训练

1. 初始化 权重和偏置
2. 前向传播(forward propagation)
3. 计算1个批次的损失 (Loss function/cost function)
4. 反向传播 (back propagation)
5. 参数更新 (包括权重, 偏置)
6. 送下一个批次的数据前向传播并重复2到5

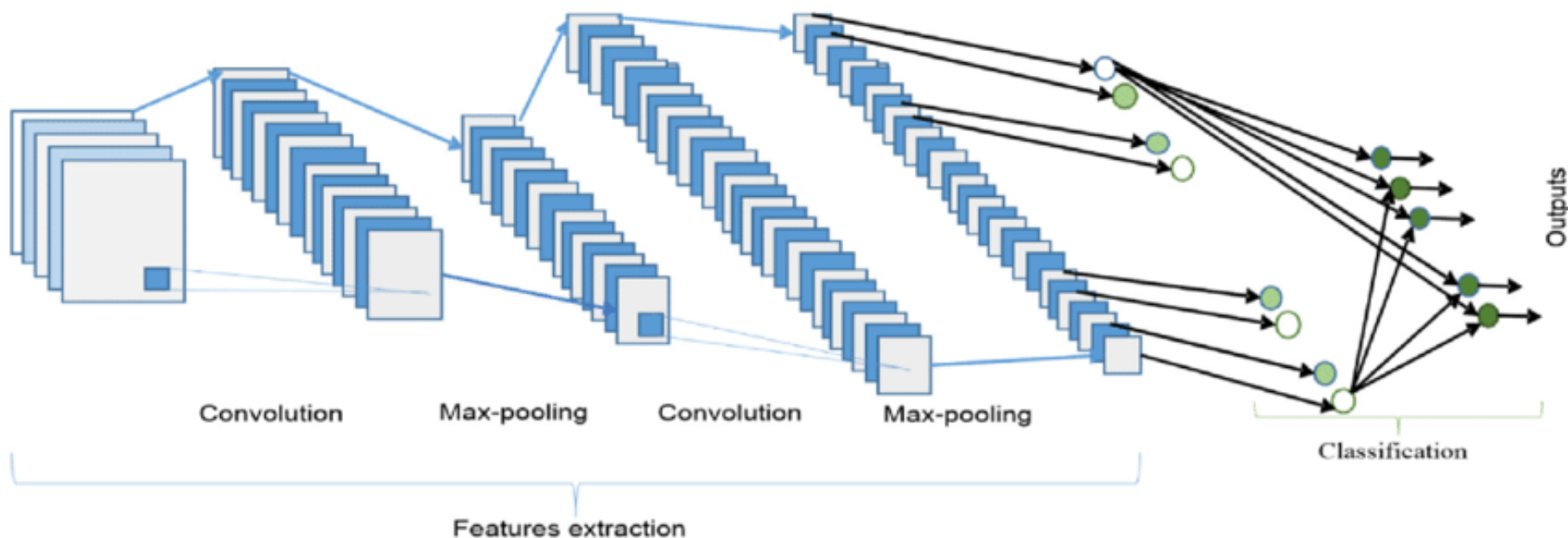
思考：

这个神经网络中有多少个参数？



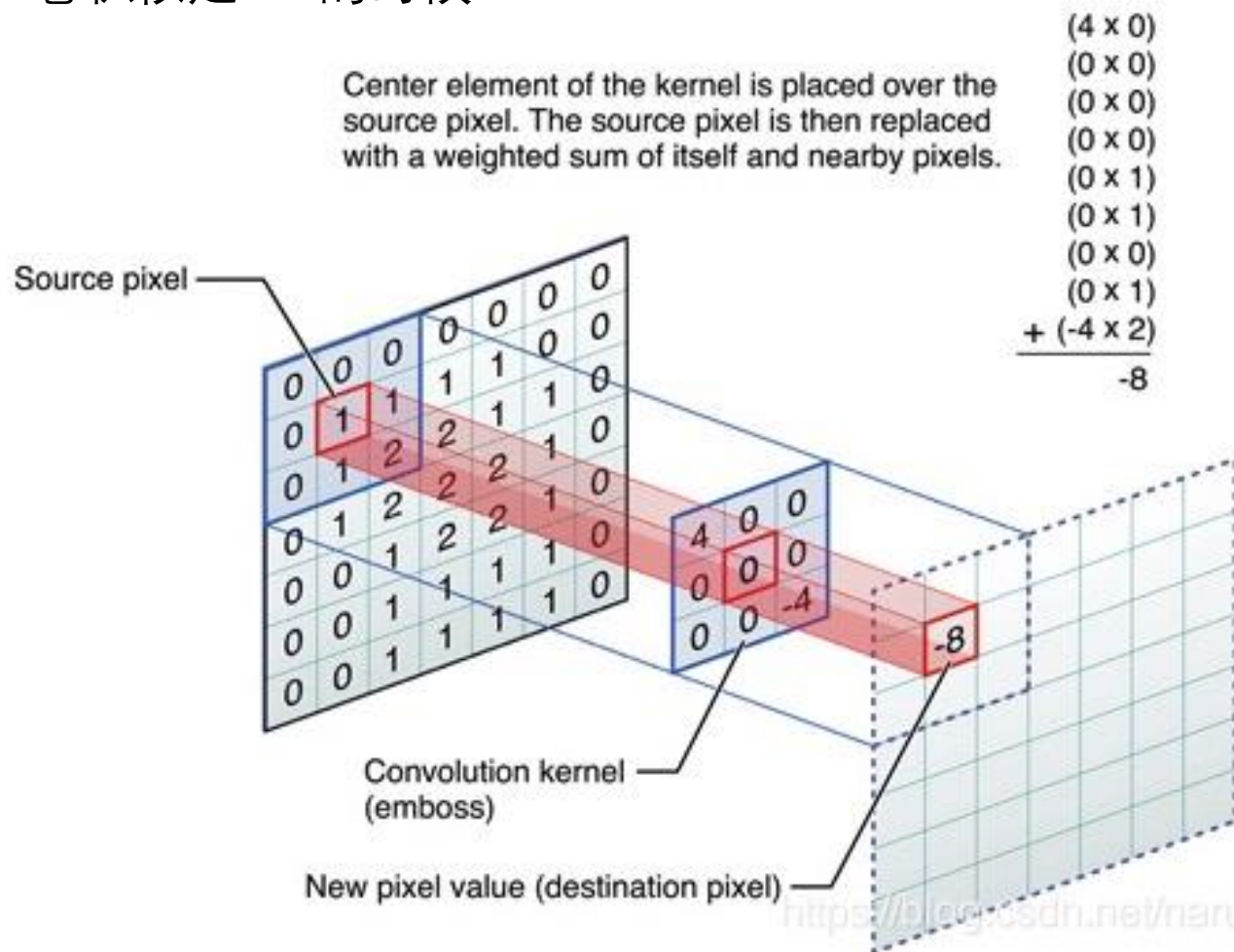
卷积神经网络 (CNN)

- 每个神经元上也不再是一个数值，而是一副“图像”（矩阵）
- 每一个连接上不再有权重，而是一个卷积核
- 由“加权和”变为“卷积和”



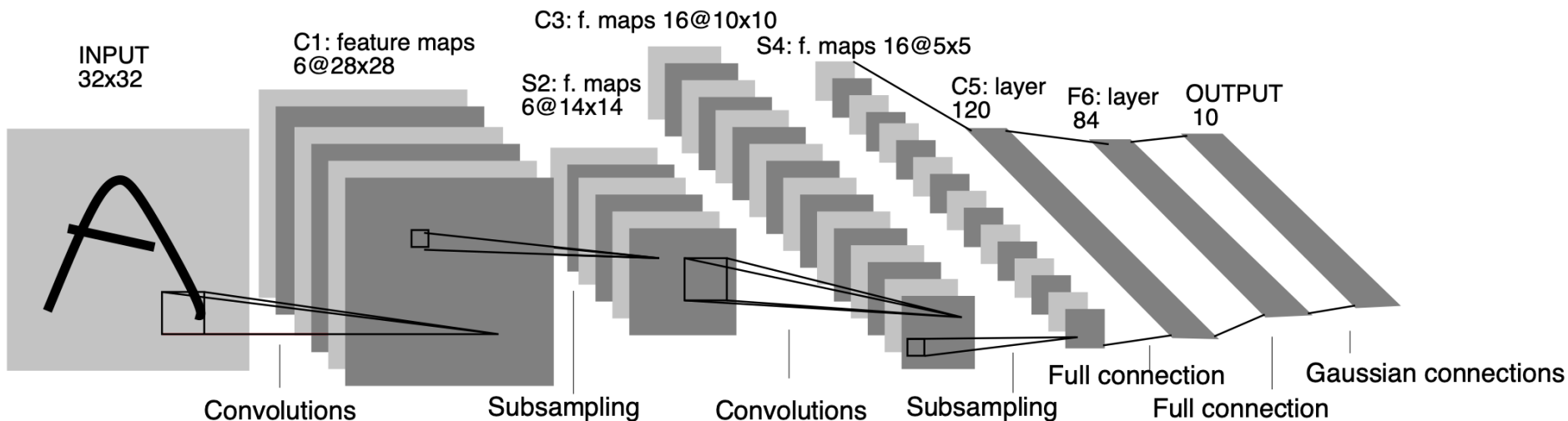
卷积神经网络

- 对图像做卷积
- 卷积核是3*3的时候



卷积网络的发展

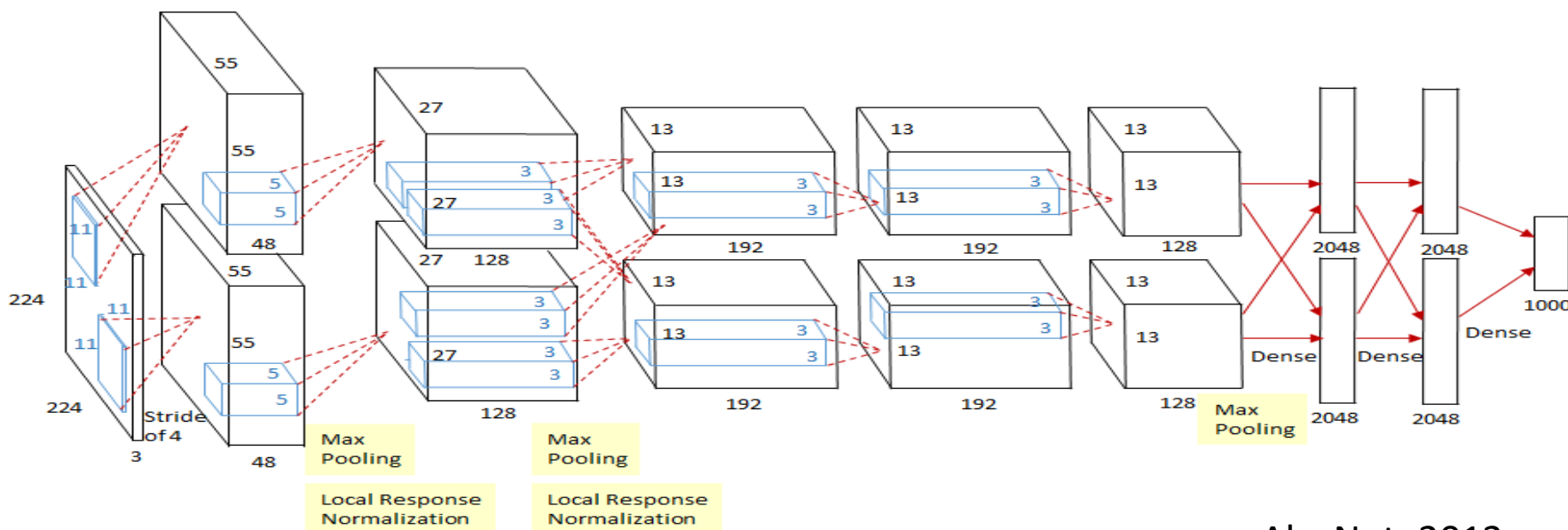
- LeNet是一种早期的卷积神经网络（Convolutional Neural Network, CNN），由Yann LeCun于1998年提出。它是一种用于手写数字识别的神经网络结构。尽管LeNet在当今的深度学习应用中可能被认为是相对简单的模型，但它对卷积神经网络领域的发展产生了重要影响，也为后续更复杂的网络结构奠定了基础。



LeNet, 1998

卷积网络的发展

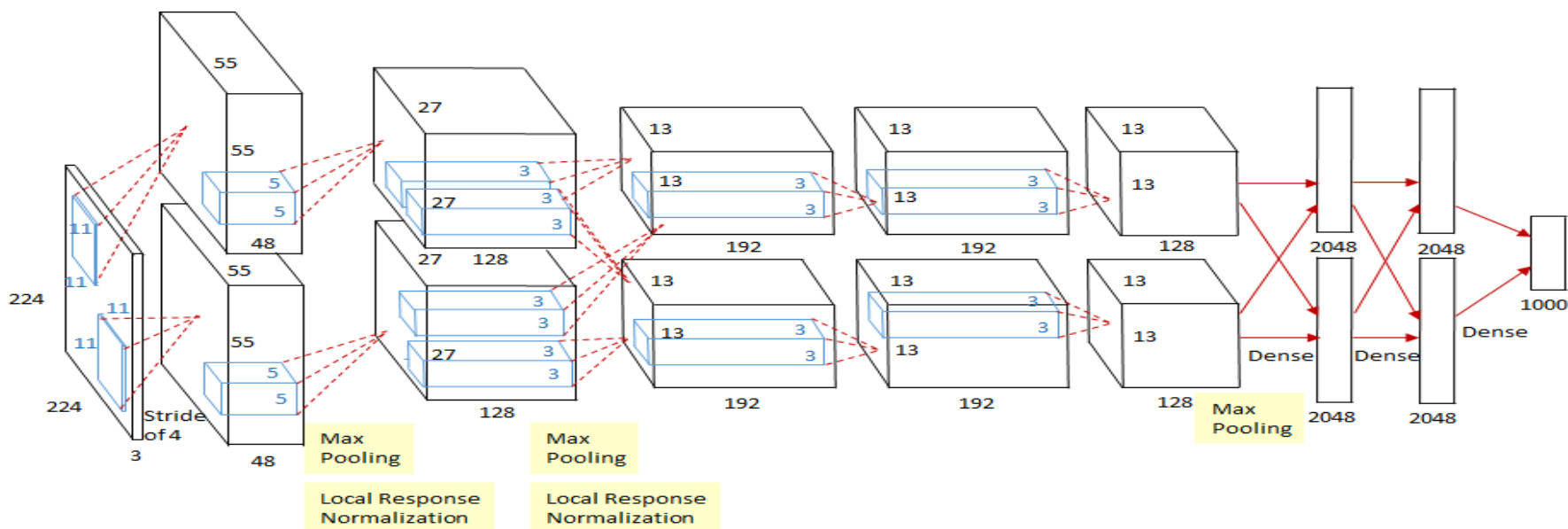
AlexNet是一种深度卷积神经网络（Convolutional Neural Network, CNN），由Alex Krizhevsky、Ilya Sutskever和Geoffrey Hinton在2012年提出。它在当年的ImageNet Large Scale Visual Recognition Challenge（ILSVRC）比赛中取得了突破性的成绩，大幅度提高了图像分类任务的准确率，从而引发了深度学习领域的快速发展。



AlexNet, 2012

卷积网络的发展

AlexNet, 2012



AlexNet参数数量总数达到60965224个
而LeNet只有接近60000个（网络的规模大了1000倍多）

为什么产生了大规模神经网络？

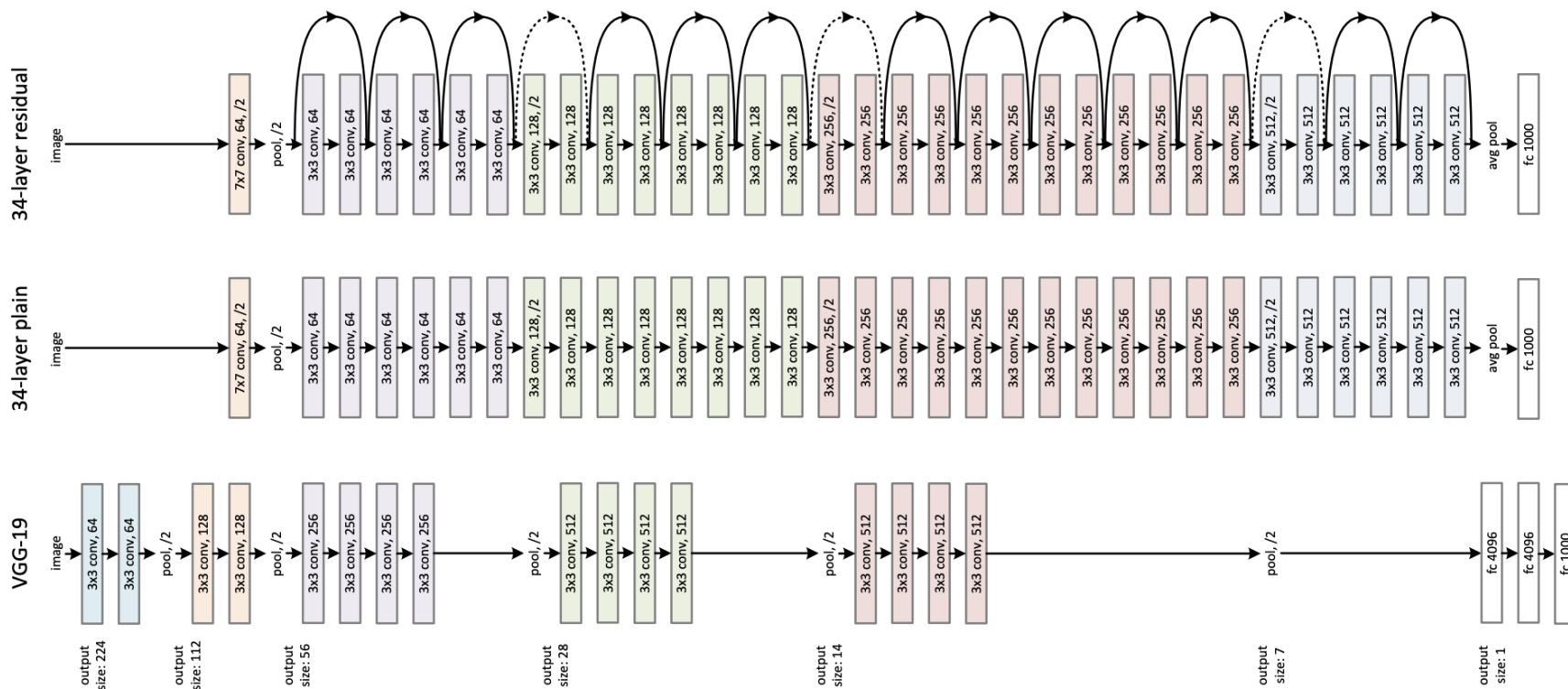
- 计算机硬件的发展
- 与日俱增的数据量

卷积网络的发展



2014年：Simonyan和Zisserman提出了VGGNet，这是一种具有更深层次的卷积神经网络，主要特点是使用 3×3 小卷积核。VGGNet在ILSVRC2014比赛中取得了优异成绩

卷积网络的发展



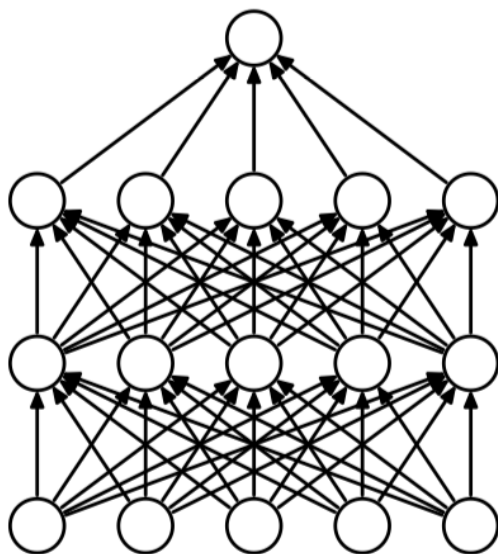
2015年：何恺明等人提出了残差网络（ResNet），它引入了残差连接，使得卷积神经网络可以达到更深的层次（如152层甚至更深），同时提高了性能。ResNet在ILSVRC2015比赛中获得了冠军。

一些常用的训练技巧

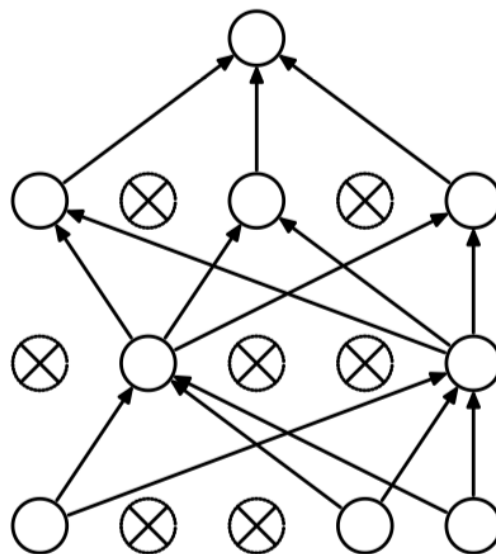
- **Dropout**

在训练过程中随机地将一些神经元的输出置为零，从而使得网络不能过于依赖某些特定的神经元，从而提高模型的泛化能力。

这种操作导致每次更新权重时网络的结构都不同，实际上相当于在训练过程中组合了多个不同的神经网络。



(a) Standard Neural Net



(b) After applying dropout.

一些常用的训练技巧

- **Finetune（微调）**

"finetune"（微调）指的是在预训练模型的基础上，对模型的某些部分进行进一步训练，以便适应特定任务或数据集。这种方法通常用于迁移学习中。

1. **语音识别**：使用预训练的语音识别模型，在特定语言或方言数据上进行微调，提高识别准确性。
2. **文本分类**：利用预训练的语言模型（如BERT、GPT-3），在特定领域的数据上进行微调，以实现情感分析、垃圾邮件检测等任务。
3. **人脸识别**：通过在大型人脸数据集上预训练的模型，在特定的数据集上微调，以提高识别特定群体或环境中的人脸准确性。

一些常用的训练技巧

- **Data augmentation（数据增强）**
 - 数据增强（Data Augmentation）是一种用于扩展训练数据集的技术，通过生成多样化的训练数据来提高模型的泛化能力，从而减少过拟合
 - 平移，翻转，旋转，缩放，颜色变换，添加噪声，遮挡等

