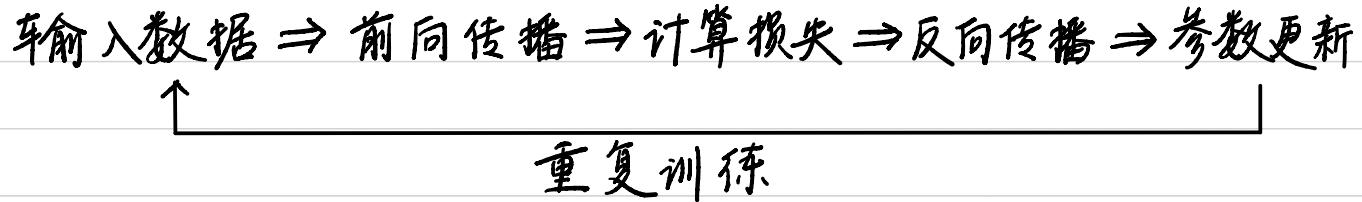


神经网络训练

整个过程是：神经网络训练不断调整参数，让模型预测更接近真实答案



① 前向传播 (Forward Propagation)

每一层执行 $z = Wx + b \quad a = f(z)$

W 是权重， b 是偏置， f 是激活函数 \rightarrow (ReLU 或 Sigmoid 等)

该层得到模型的预测值 \hat{y}

② 计算损失 (Loss)

比较模型预测值 \hat{y} 和真实标签 y ，看看预测错了多少
常见的损失函数

\rightarrow 回归任务：均方误差 (MSE)

$$L = \frac{1}{N} \sum_i (\hat{y}_i - y_i)^2$$

\rightarrow 分类任务：交叉熵 (Cross Entropy)

$$L = - \sum_j y_j \log \hat{y}_j$$

③ 反向传播 (Back Propagation)

计算每个参数对损失的影响，也就是计算梯度

用输出层开始，使用链式法则反向传递误差
每一层都根据上一层的误差更新自己的梯度

④ 参数更新 (Update Parameter)

利用反向传播得到的梯度，更新参数，使损失函数 $L(\theta)$ 逐渐减小

参数： $\theta = \{W, b\}$ ，学习率： η ，梯度 $\nabla_{\theta} L(\theta)$

更新方式： $\theta \leftarrow \theta - \eta \nabla_{\theta} L(\theta)$