大量的（输入特征，标签）对构成数据集，将数据集放入设计好的神经网络中，网络优化参数得到模型。

**MP模型**：一系列输入，乘以权值，求和，再经过非线性函数的处理得到输出

本层每个神经元与前一层每个神经元之间均有连接，我们称之为**全连接网络**

**损失函数**：表达预测值和标准答案之间的差距 （差的平方求和取均值就可以）

**梯度下降**：**目的**是寻找一组参数w和b，使得损失函数最小

**梯度**：函数对各参数求偏导后的向量，函数梯度下降方向是函数减小方向

**梯度下降法**：沿损失函数梯度下降的方向，寻找损失函数的最小值，得到最优参数的方法

**学习率：**梯度下降的速度。

**反向传播**：逐层求损失函数对各参数的偏导，更新参数