# 线性模型的局限性：

线性模型假设特征与输出呈线性关系，但现实中可能存在复杂非线性关系，如前一天观看人数与隔天观看人数可能先正相关后负相关，线性模型无法拟合这种曲线，这种限制被称为模型偏差。

# 分段线性曲线的表示：

分段线性曲线可通过常数项加上多个特定函数（如 “蓝方” 函数）组合而成，“蓝方” 函数具有先水平、再斜坡、再水平的特性，通过调整其参数可组合出各种分段线性曲线，且足够多的分段线性曲线可逼近任意连续曲线。

# Sigmoid 函数的作用：

可用 Sigmoid 函数逼近 “蓝方” 函数，其表达式为 y = c \* Sigmoid (b + w\*x1)，当 x1 趋近于无穷大时，y 趋近于 c；当 x1 负向趋近于无穷大时，y 趋近于 0，通过调整 w、b、c 可改变其斜率、位置和高度。

# 模型的扩展与向量表示：

模型可扩展到多个特征，用 j 表示特征编号，通过矩阵和向量相乘简化运算，输入特征向量 x 乘以权重矩阵 w 加上偏置向量 b 得到 r，r 通过 Sigmoid 函数得到 a，a 与 c 的转置相乘再加上 b 得到输出 y。

# 优化过程与梯度下降：

未知参数 θ 包含 w、b、c 等，优化目标是找到使损失函数 L (θ) 最小的 θ，采用梯度下降法，先随机初始化 θ0，计算梯度 g，然后迭代更新 θ，实际中会将数据分成批次计算梯度。

# 激活函数的选择：

除 Sigmoid 外，ReLU（Rectified Linear Unit）也是常用激活函数，表达式为 c \* max (0, b + w\*x1)，两个 ReLU 叠加可形成 Hard Sigmoid，激活函数统称 Activation Function。

# 神经网络与深度学习：

Sigmoid 或 ReLU 等函数被称为神经元（Neuron），多个神经元组成神经网络（Neural Network），多层隐藏层的神经网络称为深度学习（Deep Learning），其命名是为了重振神经网络技术，避免 “Neural Network” 之前的负面印象。

# 实验与过拟合问题：

实验表明增加 ReLU 数量和网络层数可降低训练集损失，但可能导致过拟合（训练集表现好但测试集表现差），如 4 层网络在训练集损失更低但测试集损失更高，选择模型时应关注其在未知数据上的表现。