

第三次作业

1. 多层感知机网络-BP 算法

考虑某分类任务中，有参数如下的多层感知机结构。其中隐藏层使用 **Sigmoid** 激活函数，Loss 函数使用输出层向量 $\mathbf{p} = (p_1, p_2)^T$ 与 one-hot 标签向量 $\mathbf{y} = (y_1, y_2)^T$ 的二元交叉熵损失函数，即

$$\text{Loss} = -\frac{1}{2} \sum_{i=1}^2 y_i \ln(p_i) + (1 - y_i) \ln(1 - p_i)$$

使用 SGD 随机梯度下降法更新参数，每批只使用一个样本。当前迭代使用的样本是 $\{0.2, 0.5, \text{货车}\}$ ，即输入 $\{\mathbf{x}_1, \mathbf{x}_2\}^T$ 为 $\{0.2, 0.5\}^T$ ，标签 $(y_1, y_2)^T$ 为 $\{0, 1\}^T$ 。

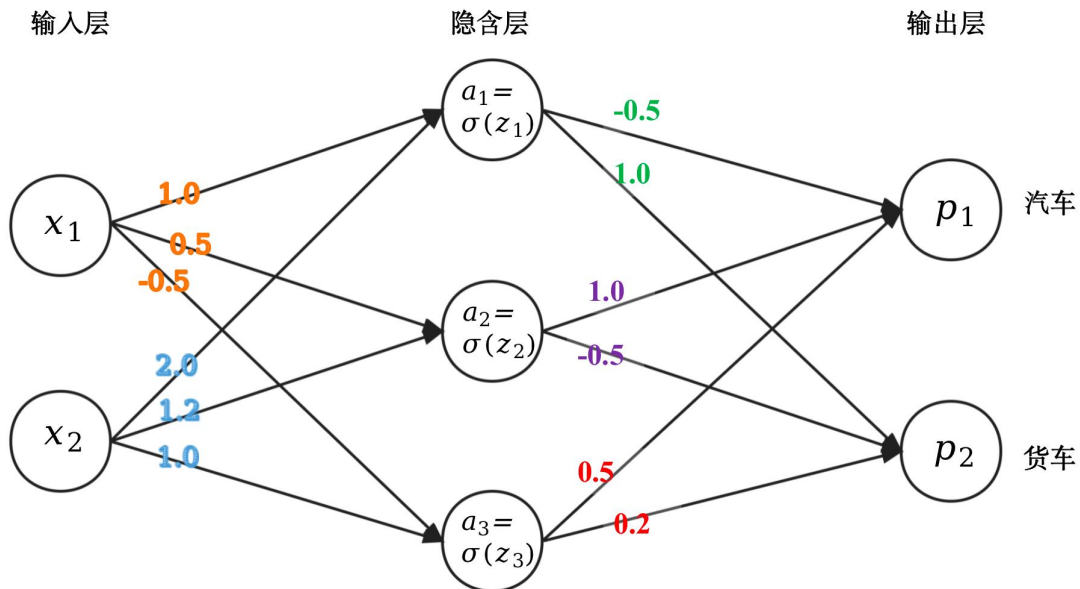


图 1. MLP 网络架构图

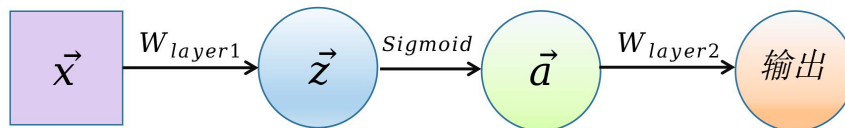


图 2. MLP 网络流程图

$$W_{Layer1} = \begin{bmatrix} 1, & 2 \\ 0.5, & 1.2 \\ -0.5, & 1.0 \end{bmatrix}, W_{Layer2} = \begin{bmatrix} -0.5, & 1.0, & 0.5 \\ 1.0, & -0.5, & 0.2 \end{bmatrix}$$

(矩阵形式的网络模型参数)

(1) 计算前向传播过程，给出各个神经元（包括隐藏层和输出层）和损失函数的计算结果；

(2) 设学习率为 0.1，计算一次参数更新后的 W_{Layer1} 和 W_{Layer2} 。

(适当保留有效位数进行计算即可)。