

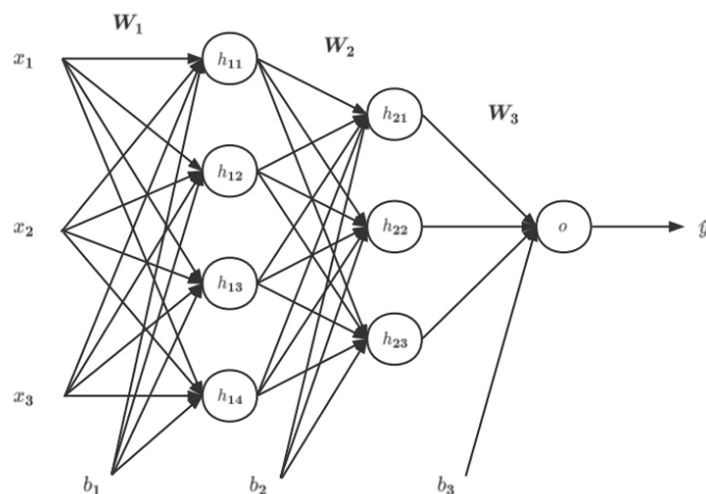
人工智能原理-作业5

Author: 夏弘宇 2023011004

T1

问题描述

- 考虑含两组隐层单元的三层前馈神经网络：



$$W_1 = \begin{bmatrix} 0.1 & 0.1 & 0.3 \\ 0.2 & 0.3 & 0.2 \\ 0.1 & 0.2 & 0.1 \\ 0.3 & 0.1 & 0.1 \end{bmatrix}, \quad W_2 = W_1^T, \quad W_3 = [0.2 \quad 0.2 \quad 0.4]$$

$$[b_1, b_2, b_3] = [0.5, 0.5, 0.2]$$

各参数的初始值如上所示。定义该网络的隐藏层单元的激活函数为 $h = \cos(z)$ ，输出单元为 Logistic 函数。

T1-1

- Q: 当输入为 $(x_1, x_2, x_3) = (0.05, 0.10, 0.05)$ 时，计算该神经网络输出 y 的值。请写明必要的计算过程。
- A:

输入向量

$$x = \begin{bmatrix} 0.05 \\ 0.10 \\ 0.05 \end{bmatrix}$$

第一隐藏层计算

- 权重矩阵 W_1 和 偏置 $b_1 = 0.5$:

$$W_1 = \begin{bmatrix} 0.1 & 0.1 & 0.3 \\ 0.2 & 0.3 & 0.2 \\ 0.1 & 0.2 & 0.1 \\ 0.3 & 0.1 & 0.1 \end{bmatrix}, \quad b_1 = 0.5$$

- 线性组合 $z_1 = W_1 \cdot x + b_1$:

$$\begin{aligned} z_1 &= \begin{bmatrix} 0.1 \cdot 0.05 + 0.1 \cdot 0.10 + 0.3 \cdot 0.05 \\ 0.2 \cdot 0.05 + 0.3 \cdot 0.10 + 0.2 \cdot 0.05 \\ 0.1 \cdot 0.05 + 0.2 \cdot 0.10 + 0.1 \cdot 0.05 \\ 0.3 \cdot 0.05 + 0.1 \cdot 0.10 + 0.1 \cdot 0.05 \end{bmatrix} + 0.5 \\ &= \begin{bmatrix} 0.005 + 0.010 + 0.015 \\ 0.010 + 0.030 + 0.010 \\ 0.005 + 0.020 + 0.005 \\ 0.015 + 0.010 + 0.005 \end{bmatrix} + 0.5 \\ &= \begin{bmatrix} 0.03 \\ 0.05 \\ 0.03 \\ 0.03 \end{bmatrix} + 0.5 = \begin{bmatrix} 0.53 \\ 0.55 \\ 0.53 \\ 0.53 \end{bmatrix} \end{aligned}$$

- 激活函数 $a_1 = \cos(z_1)$:

$$a_1 = \begin{bmatrix} \cos(0.53) \\ \cos(0.55) \\ \cos(0.53) \\ \cos(0.53) \end{bmatrix} \approx \begin{bmatrix} 0.862 \\ 0.853 \\ 0.862 \\ 0.862 \end{bmatrix}$$

第二隐藏层计算

- 权重矩阵 $W_2 = W_1^T$ 和 偏置 $b_2 = 0.5$:

$$W_2 = \begin{bmatrix} 0.1 & 0.2 & 0.1 & 0.3 \\ 0.1 & 0.3 & 0.2 & 0.1 \\ 0.3 & 0.2 & 0.1 & 0.1 \end{bmatrix}, \quad b_2 = 0.5$$

- **线性组合** $z_2 = W_2 \cdot a_1 + b_2$:

$$\begin{aligned}
 z_2 &= \begin{bmatrix} 0.1 \cdot 0.862 + 0.2 \cdot 0.853 + 0.1 \cdot 0.862 + 0.3 \cdot 0.862 \\ 0.1 \cdot 0.862 + 0.3 \cdot 0.853 + 0.2 \cdot 0.862 + 0.1 \cdot 0.862 \\ 0.3 \cdot 0.862 + 0.2 \cdot 0.853 + 0.1 \cdot 0.862 + 0.1 \cdot 0.862 \end{bmatrix} + 0.5 \\
 &= \begin{bmatrix} 0.0862 + 0.1706 + 0.0862 + 0.2586 \\ 0.0862 + 0.2559 + 0.1724 + 0.0862 \\ 0.2586 + 0.1706 + 0.0862 + 0.0862 \end{bmatrix} + 0.5 \\
 &= \begin{bmatrix} 0.6016 \\ 0.6007 \\ 0.6016 \end{bmatrix} + 0.5 = \begin{bmatrix} 1.102 \\ 1.101 \\ 1.102 \end{bmatrix}
 \end{aligned}$$

- **激活函数** $a_2 = \cos(z_2)$:

$$a_2 = \begin{bmatrix} \cos(1.102) \\ \cos(1.101) \\ \cos(1.102) \end{bmatrix} \approx \begin{bmatrix} 0.452 \\ 0.453 \\ 0.452 \end{bmatrix}$$

输出层计算

- **权重矩阵** W_3 和 **偏置** $b_3 = 0.2$:

$$W_3 = \begin{bmatrix} 0.2 & 0.2 & 0.4 \end{bmatrix}, \quad b_3 = 0.2$$

- **线性组合** $z_3 = W_3 \cdot a_2 + b_3$:

$$\begin{aligned}
 z_3 &= 0.2 \cdot 0.452 + 0.2 \cdot 0.453 + 0.4 \cdot 0.452 + 0.2 \\
 &= 0.0904 + 0.0906 + 0.1808 + 0.2 \\
 &= 0.562
 \end{aligned}$$

- **Logistic 函数** $\sigma(z_3) = \frac{1}{1+e^{-z_3}}$:

$$\sigma(0.562) = \frac{1}{1 + e^{-0.562}} \approx \frac{1}{1 + 0.570} \approx 0.637$$

最终输出

$$\hat{y} \approx 0.637$$

T1-2

- Q: 在T1-1的基础上, 若 $y=0.95$, 采用最小化均方误差作为优化准则, 请根据BP算法计算参数 W_3 的梯度。

- A:

1. 输出层误差计算

均方误差损失函数：（问了王子安助教，答曰系数不要紧，故按照自己之前学的乘了二分之一）

$$E = \frac{1}{2}(\hat{y} - y)^2$$

Logistic函数导数：

$$\sigma'(z_3) = \sigma(z_3)(1 - \sigma(z_3)) = \hat{y}(1 - \hat{y})$$

输出层误差项：

$$\begin{aligned}\delta_3 &= \frac{\partial E}{\partial z_3} = (\hat{y} - y) \cdot \sigma'(z_3) = (0.637 - 0.95) \times 0.637 \times (1 - 0.637) \\ &= (-0.313) \times 0.637 \times 0.363 \approx -0.0724\end{aligned}$$

2. W_3 梯度计算

$$\begin{aligned}\frac{\partial E}{\partial W_3} &= \frac{\partial E}{\partial z_3} \cdot \frac{\partial z_3}{\partial W_3} = \delta_3 \cdot a_2^T = -0.0724 \times [0.452 \quad 0.453 \quad 0.452] \\ &\approx [-0.0327 \quad -0.0328 \quad -0.0327]\end{aligned}$$

3. 梯度结果

W_3 的梯度为：

$$\nabla W_3 = [-0.0327 \quad -0.0328 \quad -0.0327]$$

T1-3

- Q: 在T1-2的基础上，若采用梯度下降更新参数，且学习率设置为0.1，写出更新后的参数 W_3 。
- A:

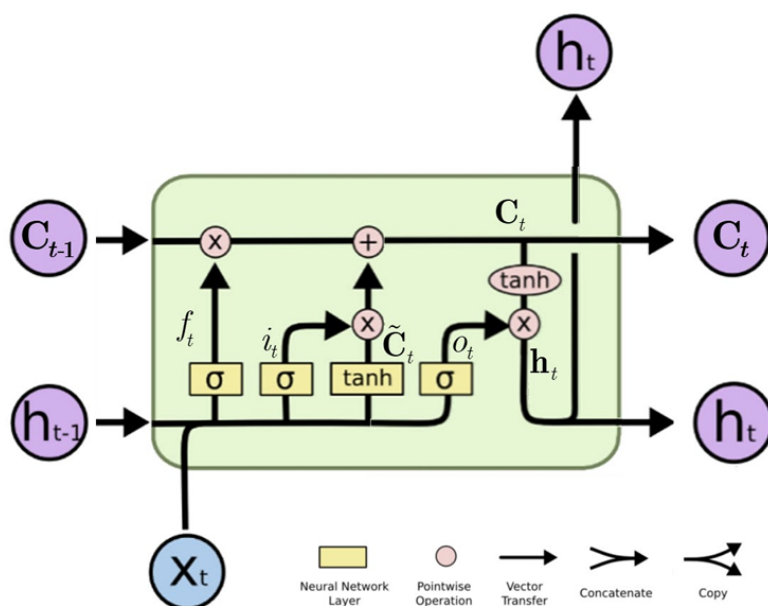
$$\begin{aligned}W_3^{new} &= W_3 - \eta \cdot \nabla W_3 \\ w_{3,1}^{new} &= 0.2 - 0.1 \times (-0.0327) \approx 0.2033 \\ w_{3,2}^{new} &= 0.2 - 0.1 \times (-0.0328) \approx 0.2033 \\ w_{3,3}^{new} &= 0.4 - 0.1 \times (-0.0327) \approx 0.4033 \\ W_3^{new} &= [0.2033 \quad 0.2033 \quad 0.4033]\end{aligned}$$

T4

问题描述

尝试对单个时间步的LSTM网络进行权值更新。已知输入值 x_t 、上一时刻的输出状态 h_{t-1} 、记忆状态 C_{t-1} 、初始参数向量 θ_0 （包括输入门 i_t 、遗忘门 f_t 、输出门 o_t 、候选记忆 \tilde{C}_t 等门控单元的权重矩阵 W ）和理论输出值 d 分别为：

- 输入值 $x_t = [1]$
- 上一时刻的输出状态 $h_{t-1} = [0]$ ，记忆状态 $C_{t-1} = [0]$
- 初始参数 $W_f = [0.5, 0.5]$ ， $W_i = [0.4, 0.4]$ ， $W_o = [0.5, 0.5]$ ， $W_C = [0.4, 0.4]$ ，不考虑偏置项
- 理论输出值 $d = 0.6$



T4-1

Q: 标准LSTM单元如图所示，其中 σ 代表Sigmoid函数，请你根据如下公式进行前向传播，计算输出值 h_t 和记忆状态 C_t 。

$$f_t = \sigma(\mathbf{W}_f \times [\mathbf{h}_{t-1}, \mathbf{x}_t])$$

$$i_t = \sigma(\mathbf{W}_i \times [\mathbf{h}_{t-1}, \mathbf{x}_t])$$

$$\tilde{C}_t = \tanh(\mathbf{W}_C \times [\mathbf{h}_{t-1}, \mathbf{x}_t])$$

$$o_t = \sigma(\mathbf{W}_o \times [\mathbf{h}_{t-1}, \mathbf{x}_t])$$

$$\mathbf{C}_t = f_t \times \mathbf{C}_{t-1} + i_t \times \tilde{C}_t$$

$$\mathbf{h}_t = o_t \times \tanh(\mathbf{C}_t)$$

A:

$$\begin{aligned} [h_{t-1}, x_t] &= [0, 1] \\ f_t &= \sigma(0.5 \times 0 + 0.5 \times 1) = 0.6225 \\ i_t &= \sigma(0.4 \times 0 + 0.4 \times 1) = 0.5987 \\ \tilde{C}_t &= \tanh(0.4 \times 0 + 0.4 \times 1) = 0.3799 \\ C_t &= 0.6225 \times 0 + 0.5987 \times 0.3799 = 0.2274 \\ o_t &= \sigma(0.5 \times 0 + 0.5 \times 1) = 0.6225 \\ h_t &= 0.6225 \times \tanh(0.2274) = 0.1393 \end{aligned}$$

\$\$

\$\$

T4-2

Q: 假设损失函数为 $L = \frac{1}{2}(y - d)^2$, 学习率为 $\alpha = 0.1$, 其中 y 代表当前时间步的网络输出, 即 h_t , 请你进行反向传播更新 W_o 的权重。

A:

$$\begin{aligned} \frac{\partial L}{\partial h_t} &= h_t - d = 0.1393 - 0.6000 = -0.4607 \\ \frac{\partial h_t}{\partial o_t} &= \tanh(C_t) = 0.2237 \\ \frac{\partial o_t}{\partial z_o} &= o_t(1 - o_t) = 0.2350 \\ \frac{\partial L}{\partial W_o} &= \begin{bmatrix} 0 \\ -0.4607 \times 0.2237 \times 0.2350 \times 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -0.0242 \end{bmatrix} \\ W_o^{\text{new}} &= \begin{bmatrix} 0.5000 \\ 0.5000 \end{bmatrix} - 0.1 \times \begin{bmatrix} 0 \\ -0.0242 \end{bmatrix} = \begin{bmatrix} 0.5000 \\ 0.5024 \end{bmatrix} \end{aligned}$$