

**گزارش تمرین سوم برنامه نویسی**

**هوش محاسباتی**

**دکتر عباد زاده**

**پویا ترابی**

**۹۴۳۱۰۳۱**

$$\begin{aligned}\frac{d \text{cost}}{dW} &= \frac{d \text{cost}}{dy} \frac{dy}{dW} = \frac{1}{2} \times 2 \times 1 \times (y - y_0) \times \left( X \times S(WX + b) \times (1 - S(WX + b)) \right) \\ &= (y - y_0) \cdot S(WX + b) \cdot (1 - S(WX + b)) \cdot \underbrace{\begin{bmatrix} x_0 \\ x_1 \end{bmatrix}}_X = \begin{bmatrix} \frac{\partial \text{cost}}{\partial w_0} \\ \frac{\partial \text{cost}}{\partial w_1} \end{bmatrix}\end{aligned}$$

$$\frac{d \text{cost}}{db} = \frac{d \text{cost}}{dy} \frac{dy}{db} = (y - y_0) \times (1 \times S(WX + b) \times (1 - S(WX + b)))$$

$$\begin{aligned}\frac{d \text{cost}}{dW} &= \frac{d \text{cost}}{dy} \frac{dy}{dZ} \times \left[ \begin{array}{c} \frac{dz_0}{dW} \\ \frac{dz_1}{dW} = 0 \end{array} \right] \times \left( \frac{dS(WX + b_0)}{d(WX + b_0)} \times \frac{d(WX + b_0)}{dW} \right) \frac{dz_0}{dW} \quad (5) \\ &= (y - y_0) S(ZU + b_2) (1 - S(ZU + b_2)) U^T \begin{bmatrix} x_0 & x_1 \\ 0 & 0 \end{bmatrix} \cdot S(WX + b_0) (1 - S(WX + b_0))\end{aligned}$$

$$\frac{d \text{cost}}{dV} = \dots = (y - y_0) S(ZU + b_2) (1 - S(ZU + b_2)) U^T \begin{bmatrix} 0 & 0 \\ x_0 & x_1 \end{bmatrix} \cdot S(VX + b_1) (1 - S(VX + b_1))$$

$$\frac{d \text{cost}}{dU} = \frac{d \text{cost}}{dy} \frac{dy}{dU} = (y - y_0) S(ZU + b_2) (1 - S(ZU + b_2)) \begin{bmatrix} z_0 \\ z_1 \end{bmatrix}$$

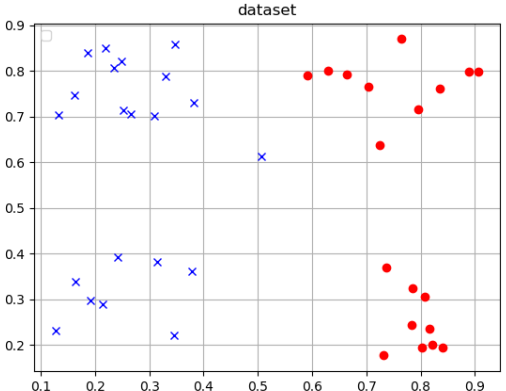
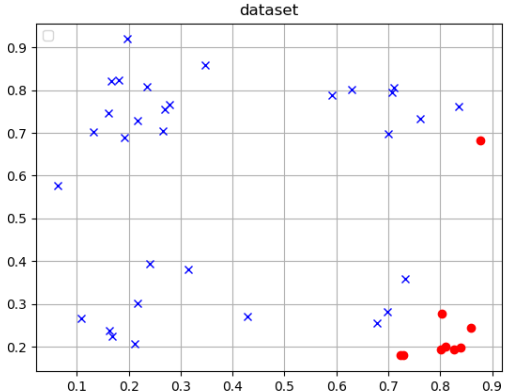
$$\frac{d \text{cost}}{db_2} = (y - y_0) S(ZU + b_2) (1 - S(ZU + b_2))$$

$$\frac{d \text{cost}}{db_0} = (y - y_0) S(ZU + b_2) (1 - S(ZU + b_2)) U^T \begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot S(WX + b_0) (1 - S(WX + b_0))$$

$$\frac{d \text{cost}}{db_1} = (y - y_0) S(ZU + b_2) (1 - S(ZU + b_2)) U^T \begin{bmatrix} 0 \\ 1 \end{bmatrix} \cdot S(VX + b_1) (1 - S(VX + b_1))$$

سوال ۴- برای دو اجرای مختلف موارد زیر برای شبکه یک لایه ای آورده شده است:

- مقدار تابع  $cost$
- نمودار  $scatter$  برای خروجی شبکه در داده های تست
- دقت شبکه برای داده های تست

آزمایش ۱	آزمایش ۲
	
$cost = 19.843236$	$cost = 19.810309$
Accuracy = 45%	Accuracy = 57.5%

خروجی تابع  $cost$  برای epoch های متوالی یک اجرا در زیر آورده شده است:

network cost in epoch 0: 38.287470

network cost in epoch 100: 34.844854

network cost in epoch 200: 28.260500

network cost in epoch 300: 23.170122

network cost in epoch 400: 22.270451

network cost in epoch 500: 21.596994

network cost in epoch 600: 20.782403

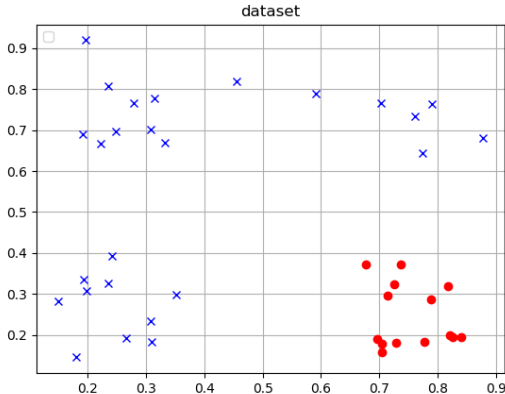
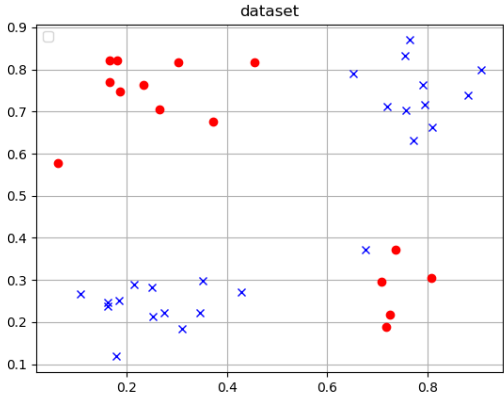
network cost in epoch 700: 20.058202

network cost in epoch 800: 19.809255

network cost in epoch 900: 19.809255

.....

network cost in epoch 9900: 19.809255

آزمایش ۱	آزمایش ۲
	
cost = 15.084494	cost = 0.011897
Accuracy = 75%	Accuracy = 97.5%

خروجی تابع cost برای epoch های متوالی یک اجرا در زیر آورده شده است:

network cost in epoch 0: 35.175152

network cost in epoch 100: 29.694122

...

network cost in epoch 500: 19.499090

network cost in epoch 600: 18.954199

...

network cost in epoch 1500: 15.210069

...

network cost in epoch 2200: 9.634957

...

network cost in epoch 3000: 1.441947

...

network cost in epoch 6100: 0.121783

...

network cost in epoch 9400: 0.064711

network cost in epoch 9500: 0.064115

network cost in epoch 9600: 0.063552

network cost in epoch 9700: 0.063021

network cost in epoch 9800: 0.062521

network cost in epoch 9900: 0.062050

## تحلیل نتایج

- در شبکه یک لایه ای، مدل در epoch ۱۰ اول در یک بهینه محلی گیر می کرد و بعد از آن بهبودی در وزن ها ایجاد نمی شد. دقت شبکه در اجراهای مختلف کمتر از ۶۰٪ بدست آمد. مقدار هزینه هم معمولا 19.xx بدست می آمد.  
احتمالا فضای جستجوی بردار وزن ها شامل کمینه های محلی زیاد با مقدار 19.xx است و شبکه خیلی سریع به این مقادیر هم گرا می شود.  
می توانم نتیجه بگیرم از شبکه یک لایه انتظار دقت بیشتر از ۶۰٪ نباید داشت.
- در شبکه دو لایه ای، مدل در تمام epoch ها همواره در حال بهبود بود. دقت بسیار بهتر از شبکه یک لایه ای بود (97% vs. 60%).  
با توجه به اینکه دقت مناسب است، می توانم نتیجه بگیرم شبکه دو لایه ای مدل جوابگویی برای پیش بینی این مجموعه داده است.