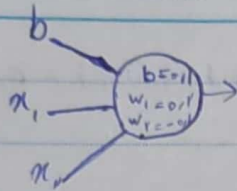## Problem 1

$$\hat{y} = w_1 x_1 + w_2 x_2 + b$$
$$w = w + \eta \, e \, x$$
$$b = b + \eta e$$

Function : ReIU $b = 0,1$ $w_1 = 0,2$ $w_2 = -0,1$

iteration 1: $\hat{y} = 0,2(0) - 0,1(0) + 0,1 = 0,1 \xrightarrow{\text{Relu}} \hat{y} = 0,1$

Input: $x_1 = 0$ $x_2 = 0$ $e = (0 - 0,1) = -0,1$

$y = 0$ update weights:

$b = 0,1 + 0,05(0 - 0,1) = 0,095$

$w_1 = 0,2 + 0,05 (0 - 0,1) x_0 = 0,2$

$w_2 = -0,1 + 0,05 (0 - 0,1) x_0 = -0,1$

Input: $x_1 = 0$ $x_2 = 1$

$y = 0$ $\hat{y} = 0,2(0) - 0,1(1) + 0,095 = -0,005 \xrightarrow{\text{ReIU}} \hat{y} = 0$

$b = 0,95 + 0,05(0) = 0,095$ $e = 0 - 0 = 0$

$w_1 = 0,2 + 0,05 (0)^0 = 0,2$ چون خطا e صفر است، وزن‌ها تغییری

$w_2 = -0,1 + 0,05 (0) x_1 = -0,1$ نمی‌کند.عوض

Input: $x_1 = 1$, $x_2 = 0$

$y = 0$ $\hat{y} = 0,2(1) - 0,1(0) + 0,095 = 0,295 \xrightarrow{\text{Relu}} \hat{y} = 0,295$

$b = 0,095 + 0,05 (-0,295) = 0,081$ $e = 0 - 0,295 = -0,295$

$w_1 = 0,2 + 0,05 (-0,295) x_1 = 0,187$

$w_2 = -0,1 + 0,05 (-0,295) x_0 = -0,1$

Input: $x_1 = 1$ $x_2 = 1$

$y = 1$ $\hat{y} = 0,187(1) - 0,1(1) + 0,081 = 0,147 \xrightarrow{\text{Relu}} \hat{y} = 0,147$

$b = 0,081 + (0,05)(0,853) = 0,123$ $e = 1 - 0,147 = 0,853$

$w_1 = 0,187 + (0,05)(0,853) x_1 = 0,227$

$w_2 = -0,1 + (0,05)(0,853) x_1 = -0,058$

iteration 2: $\hat{y} = 0,227(0) + (-0,058)(0) + 0,123 = 0,123 \xrightarrow{\text{Relu}} \hat{y} = 0,123$

$x_1 = 0$ $x_2 = 0$ $e = 0 - 0,123 = 0,123$

$w_1 = 0,227 + 0,05(-0,123)(0) = 0,227$

$w_2 = -0,058 + 0,05(-0,123)(0) = -0,058$

$b = 0,123 + 0,05(-0,123) = 0,117$

Input $x_1 = 0$, $x_2 = 1$ $\Rightarrow$ $\hat{y} = 0,224(0) + -0,049(1) + 0,114 = 0,084 \rightarrow$ Relu $0,084$

$w_1 = 0,224 + 0,05(-0,084) x_0 = 0,224$ $e = 0 - 0,084 = -0,084$

$w_2 = 0,049 + 0,05(-0,084)(1) = -0,041$

$b = 0,114 + 0,05(-0,084) = 0,114$

Input $x_1 = 1$ $x_2 = 0$ $\Rightarrow$ $\hat{y} = 0,224(1) + (-0,041)(0) + 0,114 = 0,340 \rightarrow$ Relu $0,340$

$w_1 = 0,224 + 0,05(-0,340)(1) = 0,210$ $e = 0 - 0,340 = -0,340$

$w_2 = -0,041 + 0,05(-0,340)(0) = -0,041$

$b = 0,114 + 0,05(-0,340) = 0,094$

Input $x_1 = 1$ $x_2 = 1$ $\Rightarrow$ $\hat{y} = 0,210(1) + (-0,041)(1) + 0,094 = 0,210 - 0,041 + 0,094 = 0,263$

$w_1 = 0,210 + 0,05(0,0448)(1) = \boxed{0,258}$ w_1 $0,263$ Relu

$w_2 = -0,041 + 0,05(0,0448)(1) = \boxed{0,014}$ w_2 $e = 1 - 0,263 = 0,0448$

$b = 0,094 + 0,05(0,0448) = \boxed{0,144}$ b

Problem 1 , function : step function

$b = 0,1$ $w_1 = 0,2$ $w_2 = -0,1$ Learning Rate = 0,05

Iteration 1

Input $x_1 = 0$ $x_2 = 0$ ($d = -1$) $\Rightarrow 0,2(0) + (-0,1)(0) + 0,1 = 0,1 \xrightarrow{\text{step func}}$ $\hat{y} = 1$ Eror = $-1 - 1 = -2$

$w_1 = 0,2 + 0,05(-2)(0) = 0,2$

$w_2 = -0,1 + 0,05(-2)(0) = -0,1$

$b = 0,1 + 0,05(-2) = 0$

Input $x_1 = 0$ $x_2 = 1$ $d = -1 \Rightarrow 0,2(0) + (-0,1)(1) + 0 = -0,1 \xrightarrow{\text{step func}}$ $\hat{y} = -1$ Eror = $-1 - (-1) = 0$

$w_1 = 0,2 + 0,05(0)(0) = 0,2$

$w_2 = -0,1 + 0,05(0)(1) = -0,1$

$b = 0 + 0,05(0) = 0$

x_1	x_2	d
0	0	-1
0	1	-1
1	0	-1
1	1	1

Input $x_1=1$ $x_2=0$ $d=-1$ $\Rightarrow 0,2(1) + (-0,1)(0) + 0 = 0,2 \overset{step func}{\Longrightarrow} \hat{y}=1$ $Eror = -1-1 = -2$

$w_1 = 0,2 + 0,05 (-2)(1) = 0,1$

$w_2 = -0,1 + 0,05 (-2)(0) = -0,1$

$b = 0 + 0,05 (-2) = -0,1$

Input $x_1=1$ $x_2=1$ $d=1$ $\Rightarrow 0,1(1) + (-0,1)(1) + (-0,1) = -0,1 \overset{step func}{\Longrightarrow} \hat{y}=-1$ $Eror\ 1-(-1)$
$= 2$

$w_1 = 0,1 + 0,05 (2)(1) = 0,2$

$w_2 = -0,1 + 0,05 (2)(1) = 0$

$b = -0,1 + 0,05 (2) = 0$

Iteration 2

Input $x_1=0$ $x_2=0$ $d=-1$ $\Rightarrow 0,2(0) + 0(0) + 0 = 0 \longrightarrow \hat{y}=1$ $Eror = -1-1 = -2$

$w_1 = 0,2 + 0,05 (-2)(0) = 0,2$

$w_2 = 0 + 0,05 (-2)(0) = 0$

$b = 0 + 0,05 (-2) = -0,1$

Input $x_1=0$ $x_2=1$ $d=-1$ $\Rightarrow 0,2(0) + 0(1) + (-0,1) \overset{=-0,1}{\longrightarrow} \hat{y}=-1$ $Eror = -1-(-1) = 0$

جون خط مغرشده پارامترها تغیر نمی کند

Input $x_1=1$ $x_2=0$ $d=-1$ $\Rightarrow 0,2(1) + 0(0) + (-0,1) = 0,1 \rightarrow \hat{y}=1$ $Eror = -1-1 = -2$

$w_1 = 0,2 + 0,05 (-2)(1) = 0,1$

$w_2 = 0 + 0,05 (-2)(0) = 0$

$b = 0 + 0,05 (-2) = -0,2$

Input $x_1=1$ $x_2=1$ $d=+1$ $\Rightarrow 0,1(1) + 0(1) + (-0,2) = -0,1 \Rightarrow \hat{y}=-1$ $Eror = 1-(-1) = 2$

$w_1 = 0,1(1) + 0,05 (2)(1) = 0,2$ ضابطه w_1

$w_2 = (0 + 0,05(2)(1) = 0,1$ ضابطه w_2

$b = -0,2 + 0,05(2)(1) = -0,1$ ضابطه bias

Problem 2

برای ساده‌تر کردن مدل بدون کاهش دقت مدلمان می‌بایست روی معماری

طراحی کنیم تعداد پارامترهای آن کمتر شود. برای این مورد می‌توانیم از چند روش استفاده کنیم:

۱- تعداد نورون‌های لایه مخفی را کم کنیم. هرچه تعداد نورون‌ها را کم کنیم پارامترها نیز کم می‌شود، اما از آن طرف

ممکن است دقت مدل نیز کاهش یابد. اما با توجه به این که مسئله ما مسئله ساده‌ای است می‌توان

تعداد لایه‌های مخفی را از ۳۲ تا ۶۴ کاهش دهیم.

۲- کاهش تعداد لایه‌ها: با کاهش لایه‌ها تعداد پارامترها کم می‌شود. برای مسئله ما یک لایه مخفی کافی است، زیرا

تعداد و پیچیدگی داده‌های ما در دیتاست نیز زیادی ندارند.

۳- کاهش تعداد ورودی‌ها: می‌توانیم از تکنیک کاهش ابعاد استفاده کنیم مانند (PCA) که ویژگی‌های اصلی تصویر

را حفظ می‌کنیم. می‌توانیم با استفاده از این روش تعداد لایه ورودی را از ۷۸۴ به ۲۰۰ کاهش دهیم.

۴- انتخاب بهتر تابع فعالساز و بهتران برای کاهش پیچیدگی از توابع خطی مانند $ReLU$ و $LeakyReLU$

استفاده کنیم.

معماری پیشنهادی:

لایه ورودی: مثلا ۲۰۰ یا ۲۰۰ ویژگی بجای ۷۸۴ | تابع فعالساز: $ReLU$ و $LeakyReLU$

لایه مخفی: یک لایه مخفی با ۳۲ یا ۶۴ نورون

لایه خروجی: ماژول خروجی برای دسته‌بندی اعداد ۰ تا ۹

problem 3

$x_1 = 0,25$ $x_2 = 0,9$ $y(d) = 0,5$ $b = 1$ learning Rate = 1

act fun: sigmoid , $w_{13} = 0,1$, $w_{14} = 0,4$, $w_{23} = 0,8$, $w_{24} = 0,7$, $w_{35} = 0,3$, $w_{45} = 0,1$

current output : $H_3 = 0,1(0,25) + 0,8(0,9) + 1 = 1,745 \xrightarrow{\text{sigmoid}} \sigma(H_3) = 0,8527$

$H_4 = 0,4(0,25) + 0,7(0,9) + 1 = 1,73 \xrightarrow{\text{sigmoid}} 0,8491$

$y_5 = 0,3(0,8527) + 0,9(0,8491) + 1 = 2,0154V \xrightarrow{\text{sigmoid}} \sigma(y_5) = \boxed{0,8823}$ (خروجی شبکه در این مرحله ی اول)

$\underbrace{y_5(1-y_5)}_{\text{مشتق تابع سیگموید}} \quad y-y_5 \quad \underbrace{y_5(1-y_5)}$

$\text{Error}(y_5) = (y-y_5) \times \sigma'(y_5) = (0,5 - 0,8823) \times 0,8823(1-0,8823) = -0,0397$

Back propagte Error to H_3 , H_4 $\longrightarrow \sigma_3 = \sigma_5 \times w_{35} \times \sigma'(h_3)$

$H_3: \sigma'(h_3) = y_3(1-y_3) = 0,8527 \times (1-0,8527) = 0,1257$ $\text{Error} = -0,0397 \times 0,3 \times 0,1257 =$
$= -0,0015$

$H_4: \sigma'(H_4) = y_4(1-y_4) = 0,8491 \times (1-0,8491) = 0,1274 \Rightarrow \text{Error} = -0,0397 \times 0,9 \times 0,1274 = -0,0047$

Update the weights: $w_{ij} \leftarrow w_{ij} + \eta \times \sigma_j \times h_i$

$w_{35} = 0,3 + 1(-0,0397) \times 0,8527 = 0,266$
$w_{45} = 0,9 + 1(-0,0397) \times 0,8491 = 0,844$

$w_{13} = 0,1 + 1(-0,0015) \times 0,25 = 0,099$
$w_{14} = 0,8 + 1(-0,0015) \times 0,9 = 0,798$
$w_{23} = 0,4 + 1(-0,0047) \times 0,25 = 0,398$
$w_{24} = 0,7 + 1(-0,0047) \times 0,9 = 0,995$

End of epoch 1

SABA

Epoch 2 : Calculate the weighted sum..

$H_r = 0,099 \times 0,٢٥ + 0,٧٩٨ \times 0,9 + 1 = 1,٧٨٢ \xrightarrow{\text{sigmoid}} y_r = 6(h_r) = 0,٨٥٢$

$H_f = 0,٣٩٨ \times 0,٢٥ + 0,٨٩٨ \times 0,9 + 1 = 1,٤٧٨ \xrightarrow{\text{sigmoid}} y_r = 6(h_f) = 0,٨٤٢$

$O_٥ = 0,٢٤٤ \times 0,٨٥٢ + 0,٨٤٤ \times 0,٨٢٤ + 1 = 1,٩٥٤ \xrightarrow{\text{sigmoid}} y_٥ = 6(h_٥) = 0,٨٧٤$

Calculate the Error $O_٥$:

$\acute{6}(O_٥) = 0,٨٧٤ \times (1 - 0,٨٧٤) = 0,١٠٨ \qquad \text{Error} \Rightarrow (0,٥ - 0,٨٧٤) \times 0,١٠٨ = -0,٠٤٠٨$

Backpropagate the Error to H_r & H_f

$\delta_r = -0,٤٠٨ \times 0,٢٤٤ \times 0,١٢٥ = -0,٠٠١٢$

$\delta_f = -0,٤٠٨ \times 0,٨٤٤ \times 0,١٣٢٤ = -0,٠٠٤٤$

Update the weights

$W_{r٥} : 0,٢٤٤ + 1(-0,٠٤٠٨)(0,٨٥٢) = 0,٢٤١$

$W_{f٥} : 0,٨٤٤ + 1(-0,٠٤٠٨)(0,٨٤٢) = 0,٨٤١$

$W_{١r} : 0,099 + 1(-0,٠٠١٢)(0,٢٥) = 0,09٨$

$W_{١f} : 0,٧٩٨ + 1(-0,٠٠١٢)(0,9) = 0,٧٩٧$

$W_{٢r} : 0,٣٩٨ + 1(-0,٠٠٤٨)(0,٢٥) = 0,٣٩٧$

$W_{٢f} : 0,٨٩٨ + 1(-0,٠٠٤٨)(0,9) = 0,٨٩١$