

11:00

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

مطلوبه
desired

1

12:00

$$w_1 = w_2 = 1$$

13:00

$$b = 1$$

برای محاسبه

14:00

input (1, 1) (مرحله اول)

$$y = 1 + 1 + 1 = 3$$

15:00

$$w_1 = 1 + 1(-1 - 3) = -3$$

16:00

$$w_2 = 1 + 1(-1 - 3) = -3$$

$$b = -3$$

17:00

input (1, -1) (مرحله دوم)

$$y = -3 + 3 - 3 = -3$$

$$w_1 = -0.05 + 0.1(1 + 0.05)_{p-1} = 0.05V \approx 0.05$$

$$b w_1 = -0.05 + (-1)(1 + 0.05)_{p-1} = -0.15V \approx -0.15$$

$$b = w_1 = -0.05$$

input(-1, 1)

(p-1)

$$y = -0.05 + 0.15 + 0.05 = 0.15$$

~~$$w_1 = 0.05 + 0.1(1 + 0.05)_{p-1} = 0.15V \approx 0.15$$~~

~~$$w_2 = 0.15 + 0.1(1 + 0.05)_{p-1} = 0.25V \approx 0.25$$~~

$$w_1 = 0.05 + 0.1(1 - 0.05)_{p-1} = -0.05V \approx -0.05$$

$$w_2 = 0.15 + 0.1(1 - 0.05)_{p-1} = 0.25V \approx 0.25$$

$$b = 0.05 + 0.1(1 - 0.05) = 0.15V \approx 0.15$$

input(-1, -1) : (p-1)

~~$$y = -0.05_{p-1} + 0.15_{p-1} + 0.15 = -0.05$$~~

$$w_1 = -0.05 + 0.1(1 + 0.05)_{p-1} = 0.15V \approx 0.15$$

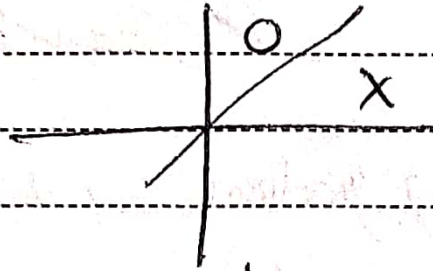
$$w_2 = 0.15 + 0.1(1 + 0.05)_{p-1} = 0.25V \approx 0.25$$

$$b = 0.15 + 0.1(1 + 0.05) = 0.25V \approx 0.25$$

قسمت اول

⑤ در توابع فعال‌سازی خطی قابلیت‌های مدل یک نورون را به صورت خطی

حداکثر



در توابع فعال‌سازی غیر خطی قدرت پیگیری دارند و در مورد ضیف زیلای از انواع Data

می‌توانند پاسخ‌دهی مناسب داشته باشند مثل \tanh / $ReLU$ / sigmoid

همچنین الگوریتم BP در این توابع هست که در توابع خطی نیست

(۲) قسمت دوم

$bias = random, weight = 0$

الف) ~~این روش را می توانیم به~~

این از مشکلات $Vanishing\ gradient$ می باشد مقدار بردارهای می تواند صفر باشد

که باعث مشکلات هنگام $learn$ و آپدیت کردن پارامترهای می شود

می شود به این می شود همگرا می بسیار کند می شود

$bias = 0, weight = random$

ب)

به طور کلی ممکن است باز هم مشکل داشته باشیم اما نسبت به حالت قبل باز هم

سریعتر خواهیم بود

(۲) قسمت سوم

به طور کلی به این روش و آپدیت کردن شبکه های عصبی می توانیم به این حالت

تیم می توانیم داشته باشیم و MLP می توانیم به این روش از قبل می توانیم به این

به طور کلی:

one layer perceptron < Adaline < G1 < madaline < MLP

② نکته چهارم

به طور کلی استفاده از مشتق دوم می تواند دیدگاهی مانند سرعت نزدیک شدن ما به $global\ minimum$ را بدهد (از آنجایی که مشتق اول حاصله را می دهد)

با دانستن سرعت می توانیم هنگامی که نزدیک به $global\ minimum$ هستیم سرعت

حرکت کنیم و هنگامی که دور از $global\ minimum$ هستیم سرعت خود را کم کنیم

البته از تعابیر آن این است که ممکن است ما به نقطه $local\ minimum$ برسیم

و در آن نقطه متوقف شویم