

Nand	x_1	x_2	bias
①	-1	1	1
②	+1	-1	1
③	-1	+1	1
④	+1	-1	1

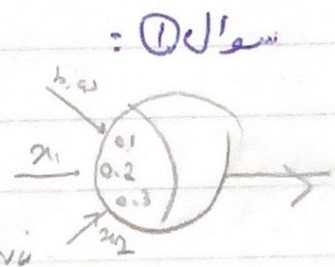
$$\text{bias} = 0.1$$

$$\alpha = 0.1$$

$$y = b + \sum x_i w_i$$

$$b(\text{new}) = b(\text{old}) + \eta(d - y)$$

$$w_i(\text{new}) = w_i(\text{old}) + \eta(d - y)x_i$$



① initialize random numbers for inputs: $x_1 = 0.2$, $x_2 = 0.3$ (مثال/شماره)

$$b = 0.1$$

* First round: input = $(1, 1)$: $y = 0.1 + (-0.2) + (0.3) = 0.4$

$$b(\text{new}) = 0.1 + 0.1(-1 - 0.4) = 0.04 \rightarrow 0.06$$

$$w_1 = -0.2 + 0.1(-0.6) = -0.26 \rightarrow 0.06 \text{ change}$$

$$w_2 = -0.3 + 0.1(-0.6) = -0.36 \rightarrow 0.06 \text{ change}$$

* Second round: $(-1, +1)$: $y = 0.04 - 0.26(-1) - 0.36(1) = -0.06$

$$b = 0.04 + 0.1(1 - (-0.06)) = 0.15 \rightarrow (0.11 \text{ change})$$

$$w_1 = -0.26 + 0.1(1.06) = -0.4 \rightarrow (0.14 \text{ change})$$

$$w_2 = -0.36 + 0.1(1.06)x_1 = -0.25 \rightarrow (0.11 \text{ change})$$

* Third round: $(+1, -1)$: $y = 0.15 + -0.4(+1) - 0.25(-1) = 0$

$$b = 0.15 + 0.1(1 - 0) = 0.16$$

$$w_1 = -0.4 + 0.1(1) = -0.30$$

$$w_2 = -0.25 + 0.1(-1) = -0.35$$

$$-0.1$$

$$4 \text{ Fourth round: } (-1, -1) \Rightarrow y = 0.16 + \underset{\substack{\downarrow \\ +1}}{(-1 \times -0.3)} + \underset{+0.35}{(-0.35 \times -1)} = 0.81$$

$$b_{\text{new}} = 0.16 + \underset{0.19}{0.1 \left(\overset{0.019}{1 - 0.81} \right)} \approx 0.18 \rightarrow \text{not change}$$

$$w_1 = -0.3 + \underset{0.019 \times -1}{0.1(0.19)} = -0.32 \rightarrow \text{not change}$$

$$w_2 = -0.35 + \underset{1}{-0.019} = -0.37 \rightarrow \text{not change}$$

الف) فعال ساز غیر خطی، خطاف فعال ساز خطی، خردی بی نوزن تناسب (Q2)

باورهای سنت - در تیم بلا مدل غیر خطی در مدل یارگیری به نوزن اجماره ی صورت استفاده

اساس خطا یا تفاوت مدل آموزش شده در تیم خطی نباشد. داده های غیر خطی را مدل سازی

ج) خطی

$$6 \left((y = w_0 + w_1x_1 + \dots + w_nx_n + b) \right)$$

الف) ① Random bias
② Weights

ب) ② Zero bias
Random weights

حالت ① همان دوره فرمول ارائه می کند، در صورتی که وزن ها صفر باشد خردی نوزن همواره

بی مقدار ثابت است. در تیم برای آن به علت ثابت بودن 0 است. در حالت

حالت ② همگامی وزن ها $(w_{new} = w_{old} + \eta \times \Delta w)$ به علت همگامی نوزن برای آن وزن ها تغییر می کند

در تیم مدل به اصطلاح تنها یک feature را آموزش می دهد و هیچ شیئی در مورد خود ندارد

حالت ③: در حالتی که $bias = 0$ و وزن ها زنده باشد، $y = w_0 + w_1x_1 + \dots + w_nx_n$ مقدارهای اولیه

است و با توجه به الگوریتم Gradient Descent، هر چه خطا وزن ها $update bias$

می شود اما در نهایت کار وزن ها $bias$ به مقدار مطلوب خواهد رسید.

در تیم تلاش می کنند برای هر آن مدل را

مدل‌ها: Adeline , Perceptron , Madeline , MLP

(6)

مدل‌های Perceptron و Adeline برای دسته‌بندی داده‌های خطی مورد استفاده قرار می‌گیرند. در ساختار ساده‌تری دارند. به صورت جزئی برای داده‌ای که به صورت خطی تفکیک پذیر باشند در مدل با بهترین است اما برای ساختارهای غیر خطی به کار قرار می‌گیرد. MLP و Madeline انتخاب‌های بهتری هستند و به کار خبره‌ای دلیلی در ساختار پیچیده‌تر در موارد غیر خطی تصمیم‌گیری با آن‌ها در بهتری دارند.

(7)

$$\Delta w = -H^{-1} \frac{dE}{dw}, \quad H = \frac{\partial^2 E}{\partial w^2}$$

همان لحظه در محاسبه است، در این روش برای آپدیت کردن وزن‌ها و Δw استفاده شده

از تزارایان درجه دوم (هستند درجه دوم) در تزارایان E استفاده می‌شود

از آن جایی که از مشتق درجه دو استفاده می‌شود، سرعت و اصطلاح همگرایی سریع‌تر به اتفاق

می‌افتد. همگرایی نسبت به دقت بالاتر کی انجام می‌پذیرد

لا دارد. اما پیچیدگی محاسباتی زیادی دارد و از همایان این دست است. محاسبه است

نام، تعداد $nodes$ زیاد می‌کنند