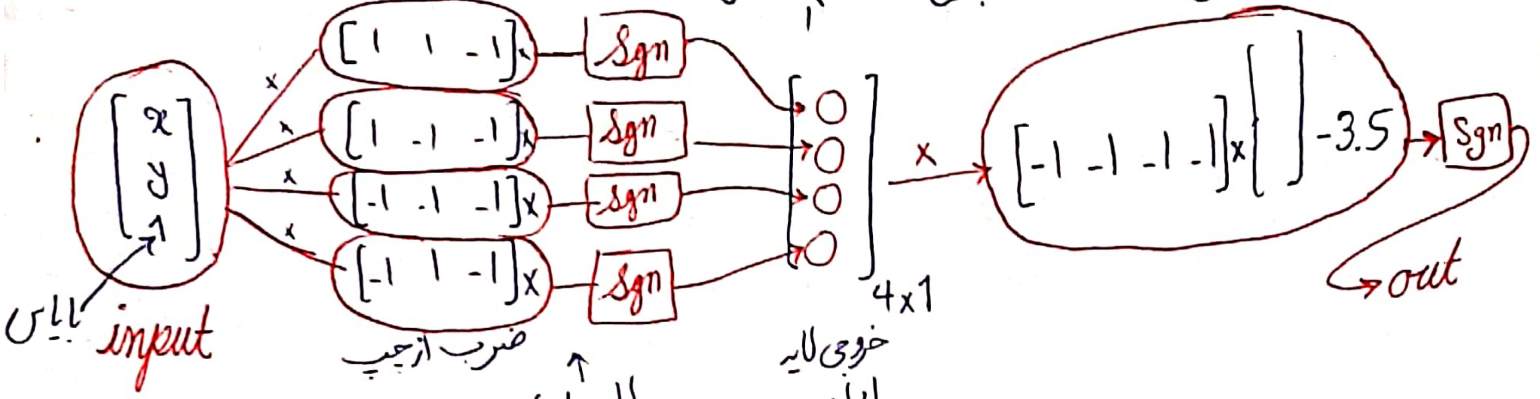
نهال مظاهری 98102346

مقدمه‌ای بر یادگیری ماشین - تمرین سری 3

1 شبکه عصبی با دو لایه مخفی

هدف تشکیل مسئله طبقه بندی $|x|+|y|-1 \gtrless 0$ است . بردار ورودی را augmented می‌کنیم



2 مشتق برداری

$$\frac{\partial f_1}{\partial x}\Big|_x = \begin{bmatrix} \frac{1}{\pi}\times\sin(\pi x_2) & x_1\cos(\pi x_2) \\ e^{x_1-1}x_2^2 & 2e^{x_1-1}x_2 \\ x_2 & x_1 \end{bmatrix}\Bigg|_x = \begin{bmatrix} 0 & 1 \\ 4 & 4 \\ 2 & 1 \end{bmatrix}$$

$$f_1\left(\begin{bmatrix}1\\2\end{bmatrix}\right) = \begin{bmatrix}0\\4\\2\end{bmatrix}, \quad \frac{\partial f_2}{\partial x}\Big|_{f_1(x)} = \begin{bmatrix} 1 & 1 & 1 \\ 2x_1 & 2x_2 & 2x_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 8 & 4 \end{bmatrix}$$

$$\frac{\partial f}{\partial x} = \frac{\partial f_2}{\partial x}\Big|_{f(x)} \times \frac{\partial f_1}{\partial x}\Big|_x = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 8 & 4 \end{bmatrix} \times \begin{bmatrix} 0 & 1 \\ 4 & 4 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 6 \\ 40 & 36 \end{bmatrix}$$

3 الگوریتم back-propagation

$$y = f(z_2) = f(W_2^T y_1) = f(W_2^T f(z_1)) = f(W_2^T f(W_1 x))$$

$$\frac{\partial y}{\partial W_2} = \frac{\partial y}{\partial z_2}\Big|_{z_2} \times \frac{\partial z_2}{\partial W_2}\Big|_{W_2, y_1} = f'(z_2) \times y_1^T = f'(W_2^T f(W_1 x)) \times (f(W_1 x))^T$$

$$\frac{\partial y}{\partial W_1} = \frac{\partial y}{\partial z_2} \times \frac{\partial z_2}{\partial y_1} \times \frac{\partial y_1}{\partial z_1} \times \frac{\partial z_1}{\partial W_1} = f'(z_2) \times W^T \times f'(z_1) \times x$$

4 کرنل گوسی

$$K(x,x') = \sum_{n=0}^{\infty} \phi_n(x)\phi_n(x') = \sum_{n=0}^{\infty} \left(\frac{1}{\sqrt{n!}}\right)^2 (xx')^n e^{-\frac{x^2}{2}-\frac{x'^2}{2}} = \exp\left(-\frac{x^2+x'^2}{2}\right) \sum_{n=0}^{\infty} (xx')^n \times \frac{1}{n!}$$

$$= \exp\left(-\frac{x^2+x'^2}{2}\right) \times \underbrace{\sum_{n=0}^{\infty} \frac{(xx')^n}{n!}}_{\text{بسط } \exp(xx')} = \exp\left(-\frac{x^2+x'^2}{2} + \frac{2xx'}{2}\right) = \exp\left(-\frac{(x-x')^2}{2}\right) \checkmark$$

5 مدل SVM

$$\phi(x_1) = (1, 0, 0) \qquad \phi(x_2) = (1, 2, 2)$$

الف) ~~معادله مرز تصمیم گیری به شکل $w_0 + \begin{bmatrix}0\\1\\1\end{bmatrix} z$ که $-\infty < z < \infty$ است می باشد~~

$$w^T \begin{bmatrix}1\\0\\0\end{bmatrix} + w_0 \leq -1$$

$$w^T \begin{bmatrix}1\\2\\2\end{bmatrix} + w_0 \geq 1$$

مرز تصمیم گیری عمود بر پاره خط واصل $\phi(x_1)$ و $\phi(x_2)$ خواهد بود چرا که لیبل متفاوت دارند پس w موازی با $\phi(x_1) - \phi(x_2) = \begin{bmatrix}0\\1\\1\end{bmatrix}$ خواهد بود

ب) چون بردار بهینه برداری است که دو نقطه را بیشترین فاصله از هم جدا می کند پس با خط واصل این دو نقطه موازی است و مارجین برابر نصف فاصله بین دو نقطه خواهد بود

$$\text{margin} = \frac{\|\vec{\phi(x_1)} - \vec{\phi(x_2)}\|}{2} = \frac{2\sqrt{2}}{2} = \sqrt{2}$$

ج)

$$\|w\| = \frac{\sqrt{2}}{2} \Rightarrow w = \left[0, \frac{1}{2}, \frac{1}{2}\right]^T$$

د)

$$\left.\begin{array}{l} [0, \tfrac{1}{2}, \tfrac{1}{2}] \times \begin{bmatrix}1\\0\\0\end{bmatrix} + w_0 = -1 \Rightarrow w_0 = -1 \\ [0, \tfrac{1}{2}, \tfrac{1}{2}] \times \begin{bmatrix}1\\2\\2\end{bmatrix} + w_0 = 1 \Rightarrow 2 + w_0 = 1 \end{array}\right\} \Rightarrow \underline{w_0 = -1}$$

ه)

$$f(x) = -1 + \left[0 \quad \tfrac{1}{2} \quad \tfrac{1}{2}\right] \times \phi(x)$$