

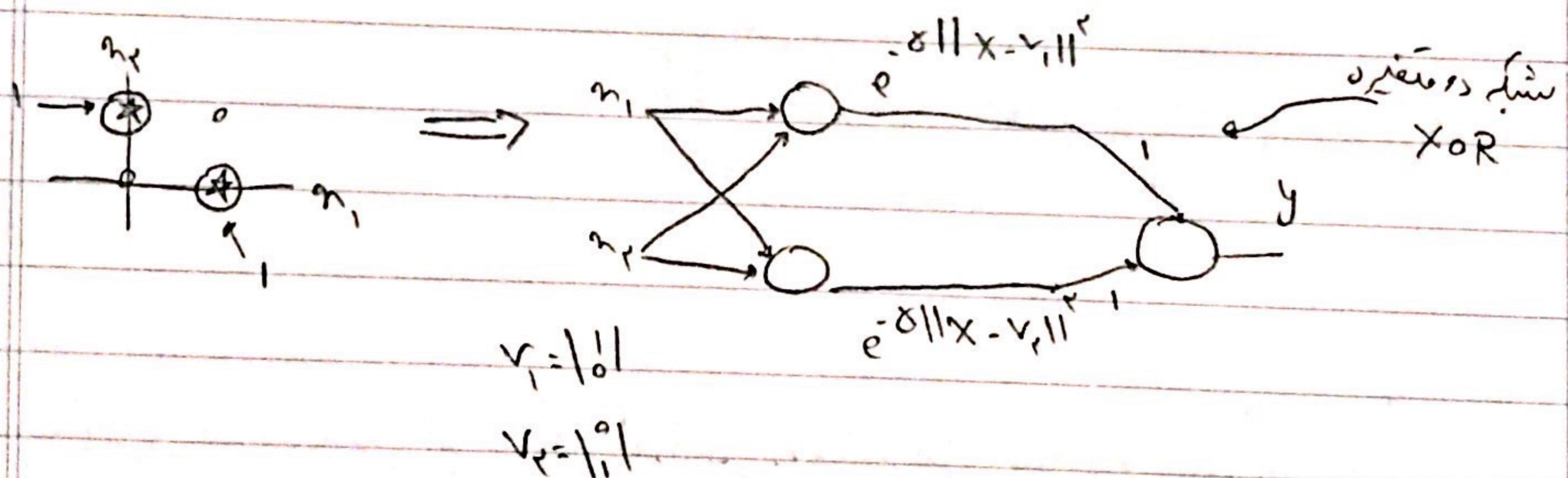
تمرین اول هریش محلسیاست

۱) خبر - با توجه به نوادره XOR فنی توانیم با یک خط صفرها را از نیکه جدا کنیم. از آنچه ای

که با یک پرسپترون تنها می‌توان داده‌ها را با یک خط از هم جدا کرد، در نتیجه فنی توانیم با

یک پرسپترون حاصل XOR را محاسبه کنیم.

$$n \oplus y = n_1 \oplus n_2 \rightarrow (n_1 \oplus n_1, \theta_{n_1} \oplus \theta_{n_2}, \dots \oplus \theta_{n_n}) = (n_1, \theta_{n_1}) \oplus (n_2, \theta_{n_2}) \oplus \dots \oplus (n_n, \theta_{n_n})$$



با توجه به شبکه دو متغیره و رابطه ۴)، برای پیاده‌سازی دو محاسبه XOR. n متغیر باشد و داده‌ها

از متغیرها XOR بگیریم و بعد دوباره دوی نتیجه آن بصین کار را تکرار کنیم. با توجه به این عملکرد

در رایه اول نیاز به $\frac{n}{2}$ XOR داریم (به ازای هر دو متغیر، یک XOR). در رایه بعدی نیاز به

$\frac{n}{2}$ XOR داریم (از نتیجه $\frac{n}{2}$ XOR قبلی دو تا دو $\frac{n}{2}$ XOR می‌گیریم). و به همین صورت ادامه

من دعیم تا به کم $\frac{n}{2}$ پرسپترون لفغان نزوجی است. در واقع تعداد رایه‌های این به شکل

نیز حسابی می شود:

تعداد رای (ها) (شامل رای اخراج)

$$\frac{n}{K} = 1 \rightarrow n = K \rightarrow K = \log^n$$

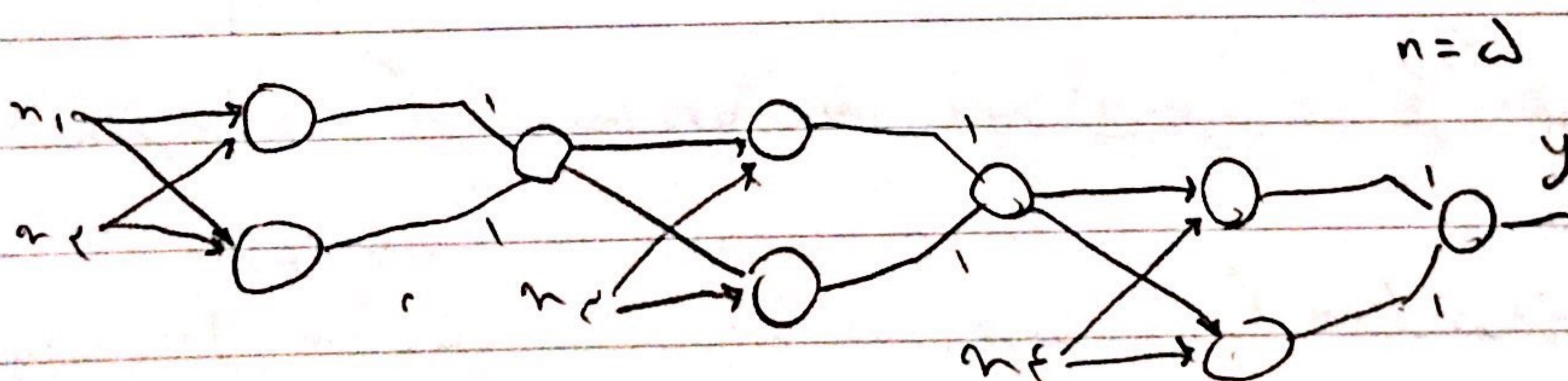
پس \log^n تا لای فراغیم داشت. از طرفی با توجه به شبکه تابع XOR و متغیرهای در پر سیسترون نیاز دارد.

در نتیجه تعداد نورون استاد در شده برابر است با:

$$n + \frac{n}{2} + \frac{n}{4} + \dots + \frac{n}{2^k} \approx 2n \rightarrow \underline{\Theta(n)}$$

۲) برای نیازهای این شبکه، ابتدا بین دو متغیر اول XOR، احاسیس کنیم، سپس نتیجه آن

با متغیر سوم XOR می تبدیل و بدین ترتیب تا متغیر آخر پیش می رویم.



در کل برای حساب هر XOR، ۳ پرسیترون داریم. در نتیجه تعداد کل نورون ها برابر است با:

$$3(n-1) + 1 = 3n - 2 \rightarrow \underline{\Theta(n)}$$

d) به فرم ۵۰۹ تابع مطابق XOR بسکل زیر است:

$$y = x_1 \oplus x_2 \oplus x_3 \oplus x_4 \oplus x_5 \oplus x_6 \oplus x_7 \oplus x_8 \oplus x_9$$

اگر تعداد فردی از متغیرها باشد خروجی است. برای پایانهای آن هر term را با

بازخ می‌بریم تعداد

پرسشی داشتیم

$$= 2^{n-1}$$
$$= 2^{(n-1)} + 2^{(n-2)} + \dots + 2^0$$

نحوه ترمیم

$$\boxed{\frac{n-1}{2}}$$

نتیجه حداقل تعداد نویسندهای برابر است با

$$a = f\left(\sum_{k=0}^L w_k^L a_k^{L-1} + b_j^L\right)$$

برای لایه پنجم اول فواد فیم داشت:

$$\begin{array}{l} w_{ik} = 0 \\ b_j^L = 0 \end{array} \rightarrow (A) \rightarrow , \Delta x_1, \Delta x_2, \Delta x_3, \Delta x_4 = z = 2$$

برای نورون های A_4, A_3, A_2, A_1

$$\begin{aligned} & \downarrow \\ & \frac{e^z - e^{-z}}{e^z + e^{-z}} \end{aligned}$$

$$(D) \rightarrow , \Delta x_4, 99 + , \Delta x_3, 99 + , \Delta x_2, 99 = 1,55 = z$$

برای نورون های V_3, V_2, V_1

$$(E) \rightarrow E_1 \rightarrow , \Delta x_1, 9 + , \Delta x_2, 9 + , \Delta x_3, 9 + , \Delta x_4, 9 = 1,18 = z$$

$$\rightarrow a = \text{swish}(z) = 1.18 \times 0.18 = 1.18 \leftarrow E_1 \quad \text{برای نورون}$$

$$\downarrow z, \text{sigmoid}(z)$$

$$E_2 \rightarrow , \Delta x_1, 9 + , \Delta x_2, 9 + , \Delta x_3, 9 = 1.50 = z$$

$$\rightarrow a = \text{swish}(1.50) = 1.50 \times 0.1 = 1.50 \leftarrow E_2 \quad \text{برای نورون}$$

$$Y \rightarrow , \Delta x_1, 1.50 + , \Delta x_2, 1.00 = 0.150 + 0.100 = 1.25 = z$$

$$\rightarrow a = \text{sigmoid}(1.25) = \boxed{0.79} \quad \text{خروجی شبکه مصنوعی}$$

$$NSE = \frac{1}{N} \sum_{i=1}^N (y - y^*)^2$$

نحوی مجموعه انتظار $y^* = 1$

(6)

$$\rightarrow \text{MLP output} = 0,89 \rightarrow MSE = \frac{1}{N} (0,89 - 1)^2 \approx 0,02$$

$f = \tanh$

$$(1) \quad z = \frac{2}{\alpha = 0,89} \quad A_1 \quad f = \tanh$$

$$(1) \quad z = \frac{1,89}{\alpha = 0,89} \quad D_1 \quad f = \text{swish}$$

$$(1) \quad z = \frac{1,11}{\alpha = 1,29} \quad E_1 \quad f = \text{sigmoid}$$

$$(1) \quad z = \frac{1,11}{\alpha = 0,89} \quad Y \quad y = 1$$

$$(1) \quad z = \frac{1,11}{\alpha = 1,29} \quad E_2 \quad y = 0,1$$

$$(1) \quad z = \frac{1,11}{\alpha = 0,9} \quad D_2$$

$$(1) \quad z = \frac{1,11}{\alpha = 0,9} \quad D_3$$

$$(1) \quad z = \frac{1,11}{\alpha = 1,09} \quad E_3$$

$$(1) \quad z = \frac{1,11}{\alpha = 0,9} \quad A_2$$

* برای بروزرسانی درون های خروجی A_2 ، باید از جلو شروع کنیم

$$w^{new} = w^{old} - y \frac{\partial E}{\partial w}$$

: (back propagation)

- ابتدا از شروع س لتیم که زیرتر بروزرسان م شود:

$$\frac{\partial E}{\partial w_{A_2 E_1}} = \frac{\partial E}{\partial E_1} \times \frac{\partial E_1}{\partial Z_1} \times \frac{\partial Z_1}{\partial w_{A_2 E_1}} \rightarrow Z_1 = D_1 \cdot w_{D_1 E_1} + A_2 \cdot w_{A_2 E_1} + D_2 \cdot w_{D_2 E_1} + D_3 \cdot w_{D_3 E_1}$$

$$\frac{\partial E}{\partial E_1} = \frac{\partial E}{\partial y} \times \frac{\partial y}{\partial Z_1} \times \frac{\partial Z_1}{\partial E_1} \rightarrow Z_1 = E_1 \cdot w_{E_1 y} + E_2 \cdot w_{E_2 y}$$

$$\frac{\partial E}{\partial y} = (y - y^*) \rightarrow y = \text{sigmoid}(Z_1)$$

$$\frac{\partial y}{\partial z_1} = \sigma(z_1)(1-\sigma(z_1)), \quad \frac{\partial z_1}{\partial E_1} = w_{E_1, y}$$

$$\rightarrow \frac{\partial E}{\partial E_1} = (y - \hat{y}) \sigma(z_1)(1-\sigma(z_1)) w_{E_1, y}$$

$$\frac{\partial E_1}{\partial z_1} \rightarrow a = \text{swish}(z_1) = z_1 \sigma(z_1) = z_1 \sigma(z_1)$$

$$\frac{\partial z_1}{\partial z_2} \rightarrow \frac{\partial a}{\partial z_2} = \sigma(z_2) + z_1 \sigma(z_1)(1-\sigma(z_1)) = \sigma(z_2)(1+z_1(1-\sigma(z_1)))$$

$$\rightarrow \frac{\partial z_1}{\partial z_2} = A_2$$

$$\frac{\partial w_{A_2 E_1}}{\partial z_2} \rightarrow \frac{\partial E}{\partial z_2} = (y - \hat{y}) \sigma(z_1)(1-\sigma(z_1)) w_{E_1, y} \sigma(z_2) \\ \times (1+z_1(1-\sigma(z_1))) A_2$$

$$= (1, \dots, 1) \times 1 \times 1 \times \dots \times 1 \times 0, 1 \times (1 + 1, \dots, 1, 0) = 1 \times 1, 1 \times 1, 1 \times 1, \dots, 1, 0$$

$$\rightarrow \frac{\partial w_{A_2 E_1}}{\partial z_2} = \underbrace{\omega - \alpha_{old} v}_{\text{new}}$$

$$\frac{\partial E}{\partial w_{A_2 D_1}} = \frac{\partial E}{\partial z_2} \times \frac{\partial z_2}{\partial D_1} \times \frac{\partial D_1}{\partial z_2} \rightarrow z_2 = A_1 w_{A_1 D_1} + A_2 w_{A_2 D_1} + A_3 w_{A_3 D_1} + A_4 w_{A_4 D_1}$$

$$\frac{\partial z_2}{\partial w_{A_2 D_1}} = A_2$$

$$\frac{\partial D_1}{\partial z_2} \rightarrow a = \tanh(z_2) \rightarrow a' = 1 - \tanh(z_2)$$

$$\frac{\partial E}{D_1} = \frac{\partial E}{\partial E_1} \times \frac{\partial E_1}{\partial z_r} \times \frac{\partial z_r}{\partial D_1} \rightarrow \frac{\partial z_r}{\partial D_1} = ^w D_{rE_1}$$

These terms remain
the same

$$\rightarrow \frac{\partial E}{D_1} = (y_j^*) (6(z_r)(1-6(z_r)) w_{E_1} y_6(z_r) (1+z_r(1-6(z_r))) w_{DE_1}$$

$w_{A_2 D_1}$
 $\times (1-\tanh^2(z_r)) A_2$

$$= 0,01V \times 1,2 \times 0,1 \Lambda^W \approx 0,000V \rightarrow w_{A_2 D_1}^{new} = 1,2 - 0,000V = 1,199V$$

$$\frac{\partial E}{\partial w_{A_2 D_2}} = \frac{\partial E}{\partial E_1} \times \frac{\partial D_2}{\partial z_r} \times \frac{\partial z_r}{\partial w_{A_2 D_2}} \rightarrow z_r = A_1 w_{A_1 D_r} + A_2 w_{A_2 D_r} + A_3 w_{A_3 D_r} + A_4 w_{A_4 D_r}$$

$$\frac{\partial E}{\partial D_r} = \frac{\partial E}{\partial E_1} \times \frac{\partial E_1}{\partial z_r} \times \frac{\partial z_r}{\partial D_r} \rightarrow \frac{\partial z_r}{\partial D_r} = ^w D_{rE_1}$$

$$\rightarrow \frac{\partial E}{\partial w_{A_2 D_r}} = 0,1 \Lambda V \times 1,2 \times 0,1 \Lambda^W \approx 0,000V \rightarrow w_{A_2 D_r}^{new} = 1,2 - 0,000V = 1,199V$$

$\frac{\partial E}{\partial w_{A_2 D_r}}$

$$\frac{\partial E}{\partial w_{A_1 D_r}} = \frac{\partial E}{\partial E_1} \times \frac{\partial D_r}{\partial z_r} \times \frac{\partial z_r}{\partial w_{A_1 D_r}} \Rightarrow z_r = A_1 w_{A_1 D_r} + A_2 w_{A_2 D_r} + A_3 w_{A_3 D_r} + A_4 w_{A_4 D_r}$$

$$\frac{\partial E}{\partial D_r} = \frac{\partial E}{\partial E_1} \times \frac{\partial E_1}{\partial z_r} \times \frac{\partial z_r}{\partial D_r} \rightarrow \frac{\partial z_r}{\partial D_r} = ^w D_{rE_1}$$

$$\frac{\partial E}{\partial w_{A_2 D_2}} = 0,01V \lambda_{A_2 D_2} e^{w_{A_2 D_2}} = 0,0001V \rightarrow V_{A_2 D_2} = 0,0001V = 0,0001$$

الگوریتم stochastical gradient descent (a) با این صورت عمل می کند که برای هر نمونه ای کوادرا شود

تابع هزینه را جدا محسوب می کند. این الگوریتم در مقایسه با batch gradient descent، سریعتر است و

دستیگیری آن نیز سریعتر است. از این اوش می توان در شبکهای باردهای زیاد استفاده کرد. در واقع

این الگوریتم جواب که می دهد بینه نیست ولی خطاکاری این سریعتری در مقایسه با batch gradient descent دارد.

batch gradient descent دفعه داده های نمونه یکجا به کار می رود از تابع هزینه ماسبه می شود.

ب عبارت دیگر پارامترها در هر epoch آبیتیت می شوند. با توجه به این روشن، الگوریتم لذی است و

زمانی که داده های زیادی را مشاهده باشیم، استفاده از این الگوریتم به صرفه نی باشد. این الگوریتم معمولاً برای

منحنی های smooth بکار ریده می شوند.

روشن min-batch gradient descent از داده ها را استفاده می کند و تابع هزینه را

از طریق آن ها محسوب می کند و پارامترها را آپدیت می کنند. به طور مثال از ۱۰۰۰ داده تنها ۱۰ داده

را در نظر می کنند.

۱) افزایش مقدار الگوریتم باعث می شود مقادیر ضرایب روابع کاوش بینند و از بیش برآورده نیز محروم شوند (کاملاً بردارهای آموزش مکن نیستند و دارای حوصله است).

البته در صورت که متغیرهای خیلی بزرگ باشند ممکن است کارایی الگوریتم پایین بباشد.

در داده های برآورده ممکن است به جای بیان هفتم انتخاب مفهوم انتخاب.

منظم سازی ۱) هنگام انتخاب استعداده می شود. در واقع این واسطه ضرایب را به مقدار ممکن

کاهش می دهد و برای باقی متغیرها و شاخصهای کاربردی اهمیت منظم سازی و ضرایب را

به صورت تکراری کاهش می دهد ولی به صفر نزدیک نمی کند. با این تعاریف نتیجه می تواند از این

روش زمانی استفاده می شود که متغیرها همیشه دارند.

۲) dropout: تیز روش های regularization می باشد و نتود یکملکر آن بین تنعل است که

برخی از نوروت ها (به بجز نوروت های لایه آخر) را حذف می کند. به صورت که به هر نوروت یک

احتمالی نسبت می دهد و بعد از آن به طور تصادفی یک نوروت از نوروت ها را حذف می کند

و در میزانی از آن ها استفاده نمی کند. در واقع dropout ذره بذره batch ترها بخشی از

نورون ها را نادیده می نماید که این کار باعث می شود در هر چند ترکیب مختلف آن نورون ها

در محاسبات استفاده شوند. تقاضه drop out باشد، عبارت دلایل نورون لغزش دستینام و همین

است.

در نهایت استفاده از روش drop out و استخراج مجموعه داده از پیش آمد

overfitting جلوگیری می کند.

(d)

X: نرخ یادگیری زیاد، Z: نرخ یادگیری بینی



از یه جایی به بعد loss نمی تواند

به مقادیر بینی برسد

loss به صورت یکنماخت

به مقادیر بینی برسد

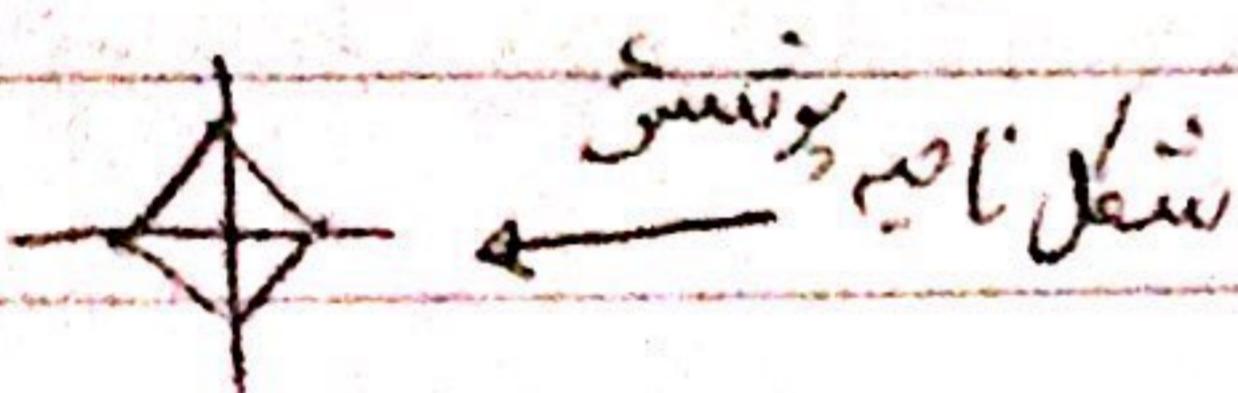
۱۰

$$d_k(x, y) = \left(\sum_{i=1}^n (x_i - y_i)^k \right)^{\frac{1}{k}} \rightarrow \underbrace{\sum_{i=1}^n (x_i - y_i)}_{\textcircled{*}}$$

حالت اول: $k=1$

رایمہ ۱) نشان دهنده فاصله منتهی یا
euclidean distance

ناحیه پوشش = نتاً روی محیط از مرکز باشد فاصله کوچک و اصرار است باشند.

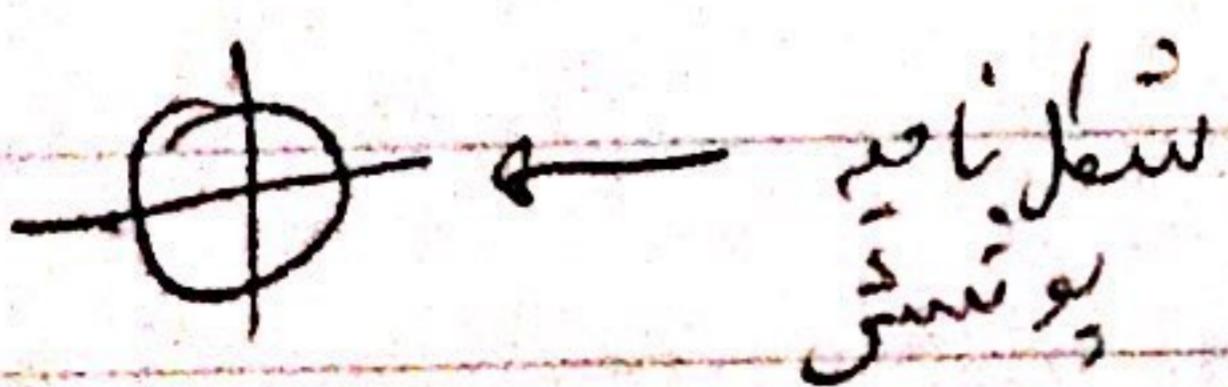


$$d_k(x, y) = \left(\sum_{i=1}^n (x_i - y_i)^k \right)^{\frac{1}{k}} \rightarrow \underbrace{\left(\sum_{i=1}^n (x_i - y_i)^2 \right)^{\frac{1}{2}}}_{\textcircled{*}\textcircled{*}}$$

حالت دوم: $k=2$

رایمہ ۲) نشان دهنده رایمہ فاصله اکلیدس یا
euclidean distance

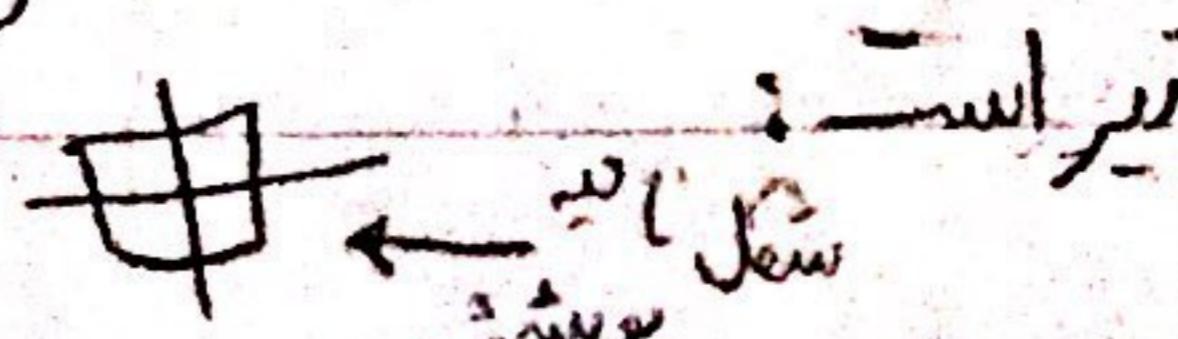
ناحیه پوشش = درایردادی به مرکز مبدأ متماثل و سطع واحد



$$\lim_{K \rightarrow +\infty} \left(\sum_{i=1}^n (x_i - y_i)^k \right)^{\frac{1}{k}} \rightarrow \underbrace{\left(\sum_{i=1}^n (x_i - y_i)^{\infty} \right)^{\frac{1}{\infty}}}_{\textcircled{*}\textcircled{*}}$$

$K = +\infty$ (6)

رایمہ ۳) نشان دهنده فاصله chebyshev یا لیپسکی که ناحیه پوشش آن بر شکل



NAHAK

$$e^{-\|x-v\|^2} \left[\begin{array}{cc} n_1 & n_2 \\ n_3 & \end{array} \right] \left. \right\}^m \xrightarrow{\text{تعداد را به دلایی}} \text{مارکوس} \xleftarrow{\text{مارکوس}} \text{اخرزس}$$

برای ماسه ۶ ابتدا پایه ماتریس ۶ را نتیل دهیم. ماتریس 6×1

است که به معنای این است که هر کدام از داده آنرا می باید مکار یعنی را حل کند.

$$G = \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} \quad \begin{cases} \text{ردارهای آموزشی} \\ \text{نوردهای راسانه} \end{cases}$$

$$W = \left[\begin{matrix} y_1 & y_2 \\ \vdots & \vdots \end{matrix} \right] \left\{ l \rightarrow Y = \left[\begin{matrix} h_1 & h_2 \\ \vdots & \vdots \end{matrix} \right] \right\}^m$$

$$L_2 \text{ مضمون سادی} \rightarrow L_2(w) := \frac{1}{2} \sum_{i=1}^m \|Y_i - G_i w\|^2 + \frac{\lambda}{2} \|w\|^2$$

فتد
لتد
متدا
متدا ر مخلوب

۱- پارامتر منظم سازی

$$\rightarrow \frac{\partial L_2(w)}{\partial w} \rightarrow \frac{\partial}{\partial w} \left(\frac{1}{2} \sum_{i=1}^m \|y_i - G_i w\|^2 + \frac{\lambda}{2} \|w\|^2 \right)$$

$$\rightarrow \frac{\partial L_2(w)}{\partial w} = \sum_{i=1}^m -G_i^T(y_i - G_i w) + \lambda w = 0 \rightarrow \frac{\partial L_2}{\partial w} = -G^T(Y - Gw) + \lambda w = 0$$

$$\rightarrow \overset{\top}{G} G w - \overset{\top}{G} y + \lambda w = 0 \rightarrow (\overset{\top}{G} G + \lambda I) w = \overset{\top}{G} y$$

$$W = (G^T G + \lambda I)^{-1} G^T$$

input

1	1	1	1
1	1	1	1
1	1	1	1
1	1	1	1

kernel

1	-1	1
-1	1	-1
1	-1	1

9

(a) با توجه به ماتریس input و کرنل، متوجه مسُویم ماتریس خروجی 2×2 است.

$$C_{11} = (1 \times 1) + (-1 \times 1) + (1 \times 1) + (-1 \times 0) \\ + (1 \times 1) + (-1 \times 1) + (1 \times 1) + (-1 \times 1) \\ \rightarrow (1 \times 1) = 1$$

ماتریس خروجی:

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

$$C_{12} = (1 \times 1) + (-1 \times 1) + (1 \times 1) + (-1 \times 1) + (1 \times 1) + (-1 \times 1) + (1 \times 1) + (-1 \times 1) \\ + (1 \times 0) = 0$$

$$C_{22} = (1 \times 1) + (-1 \times 1) + (1 \times 1) + (-1 \times 1) + (1 \times 1) + (-1 \times 0) + (1 \times 1) + (-1 \times 1) \\ + (1 \times 0) = 1$$

$$C_{21} = (1 \times 0) + (-1 \times 1) + (1 \times 1) + (-1 \times 1) + (1 \times 1) + (-1 \times 1) + (1 \times 1) + (-1 \times 1) \\ + (1 \times 1) = 1$$

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{ماتریس خروجی} \leftarrow$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{خروجی دانعی} = (b)$$

$$\Delta E = \frac{1}{r} (y - y^*) \rightarrow E = \frac{1}{r} ((r_1), (r_2), (r_3), (r_4)) = \frac{1}{r} (r + 1) = r, \text{ جاوترس}$$

$$\frac{\partial E}{\partial k_a}, \frac{\partial E}{\partial c_a}, \frac{\partial c_a}{\partial k_a} \rightarrow \frac{\partial E}{\partial c_a}, (a - y_a^*), \frac{\partial c_a}{\partial k_a} = n_a \sim \text{input}$$

$$\rightarrow \frac{\partial E}{\partial k_{11}}, \frac{\partial E}{\partial c_{11}}, \frac{\partial c_{11}}{\partial k_{11}}, \frac{\partial E}{\partial c_{12}}, \frac{\partial c_{12}}{\partial k_{11}}, \frac{\partial E}{\partial c_{21}}, \frac{\partial c_{21}}{\partial k_{11}}, \frac{\partial E}{\partial c_{22}}, \frac{\partial c_{22}}{\partial k_{11}}$$

$$\Rightarrow \frac{\partial E}{\partial k_{11}} = r \times 1 + o \times r + l \times o + o \times l = r \rightarrow k_{11}^{\text{new}} = 1 - o, 1 \times r = \boxed{-1, \sqrt{A}}$$

$$\frac{\partial E}{\partial k_{12}} = r \times r + o \times r + l \times l + o \times r = o \rightarrow k_{12}^{\text{new}} = -1 - o, 1 \times o = \boxed{-1, 0}$$

$$\frac{\partial E}{\partial k_{13}} = r \times r + o \times r + l \times r + o \times r = A \rightarrow k_{13}^{\text{new}} = 1 - o, 1 \times r = \boxed{-1, \sqrt{A}}$$

$$\frac{\partial E}{\partial k_{21}} = r \times o + o \times l + l \times l + o \times r = 1 \rightarrow k_{21}^{\text{new}} = -1 - o, 1 \times 1 = \boxed{-1, 1}$$

$$\frac{\partial E}{\partial k_{22}} = r \times l + o \times r + l \times r + o \times r = r \rightarrow k_{22}^{\text{new}} = 1 - o, 1 \times r = \boxed{-1, \sqrt{r}}$$

$$\frac{\partial E}{\partial k_{23}} = r \times r + o \times r + l \times r + o \times o = V \rightarrow k_{23}^{\text{new}} = -1 - o, V = \boxed{-1, V}$$

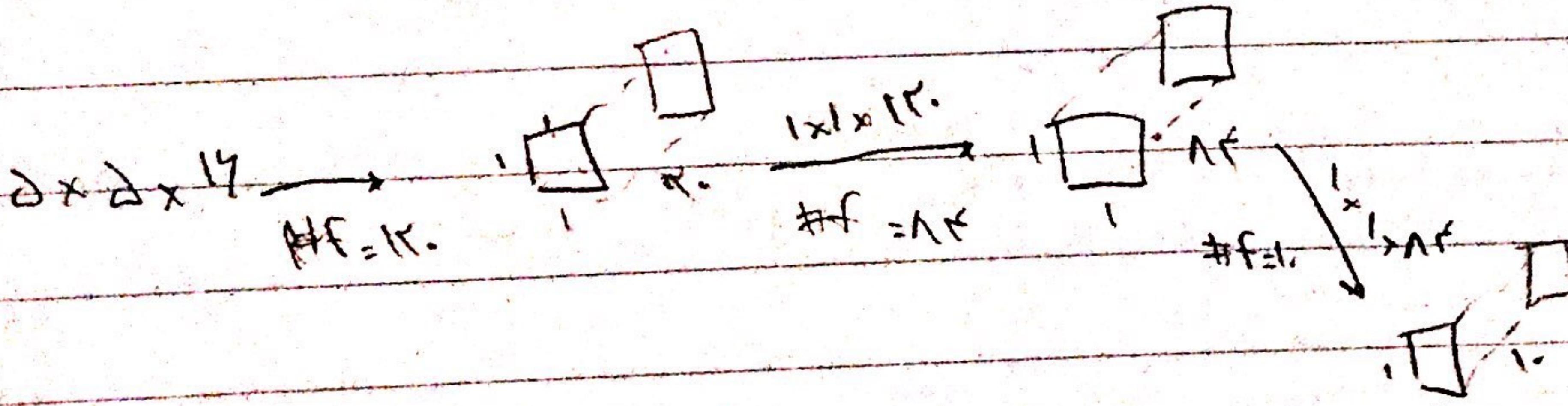
$$\frac{\partial E}{K_{31}} = (x^1 \rightarrow 0 \times 1 + 1 \times 1 + 0 \times 1) = 1 \rightarrow K_{31}^{new} = 1 - 0,1 \times 1 = \boxed{0,9}$$

$$\frac{\partial E}{K_{32}} = (x^1 \rightarrow 0 \times 1 + 1 \times 1 + 0 \times 1) = 1 \rightarrow K_{32}^{new} = 1 - 0,1 \times 1 = \boxed{0,9}$$

$$\frac{\partial E}{K_{33}} = (x^1 \rightarrow 0 \times 0 + 1 \times 1 + 0 \times 0) = 1 \rightarrow K_{33}^{new} = 1 - 0,1 \times 1 = \boxed{0,9}$$

\downarrow	\wedge	$-1,0$	\uparrow
\downarrow	\rightarrow	$-1,0$	\leftarrow
$0,9$	\rightarrow	$1,0$	\downarrow

kernel new



(a) در سکون های ۵۰۰ راتیغ دعوایی شیوه مقیس ناصله بین دخرون های دعوایی

استثنی می کند. بدین سطل که وقت دخرون شیوه هم داشتند و در دعوایی یادی برخواهند.

گرفته از راتیغ دعوایی ناصله بین دخرون، استثنی می کند را باعث تصحیح ناپذیر دخرون

پردازی میداریم باشد. با توجه به تغییرات داده شده، راتیغ دعوایی تأثیر بخوبی دارد

نرا بینهای دارد زیرا لینی دسته بندی و خوش بندی، استثنی می کند.
ناصله دخرون های

(b) ترخ پارالیس ترخ است اما در ۵۰۰ تغییف لذت های هنرها در هر رطیعه بعد

تغییرات در راتیغ دعوایی می کند داده های جدید در حل آزمیش چقدر در میزان تغییرات

وزن های تغییر انجام گنند، در اینجا این ترخ بالا در نظر گرفته عن شود و نهانی که به میزان

فروزانهای دلیری رسیدم، این ترخ کاهشی می باشد.

۲) برای مسأله دوی فزن ها روش های متعددی وجود دارد که دلیل این است مقدار دهن

از این فرن های دارای طبقه انتقال آنرا مشخص نمایند.

آنچه بگذرم (مفرغ)، اختصار مادر دهن های دو فرود حلق و همینطور

در صول آزمایش نزدیک بماند در نهایت نتوان مبتلای پویش داد.

۳) با افزایش تعداد نورون ها و نتیجه آن افزایش شبکه ایوانای شده در تشخیص خوب های بیشتر

و سردو دست آن شبکه در خوب شدن افزایشی یابد.

۴) از اینجا که تعداد نورون ها را زیاد کردیم، تعداد عملیات زیاد شد و در نهایت بیشیدن محاسبه

ازایشی می یابد.

(۵) از کامپیوتر عالی فاصله اقلیمی نورون ها از نقاط

۶) دوم نورون طای لئین فاصله را بر عبارت $\frac{1}{m}$ از انتخاب کنیم.

دوفلای نزدیکی $n = [a_1, a_2, a_3, a_4]$

نورون ها $m = [b_1, b_2, b_3, b_4]$

نامه انتیس برای استریا

$$\rightarrow (n_i, b_j) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_n - b_n)^2}$$

$$\xrightarrow{n} e(n_1, b_1) = 0,25, e(n_1, b_2) = 0,25, e(n_1, b_3) = 0,45$$

نامه انتیس

$$e(n_1, b_4) = 1, e(n_1, b_5) = 1,00, e(n_1, b_6) = 0,45$$

$$e(n_1, b_7) = 0,8, e(n_1, b_8) = 0,84, e(n_1, b_9) = 0,9$$

$$(جی نورون) b_1 = n_1 \text{ برای BMV} \leftarrow$$

$$\xrightarrow{n} e(n_2, b_1) = 0,25, e(n_2, b_2) = 0,25, e(n_2, b_3) = 0,25$$

$$e(n_2, b_4) = 0,25, e(n_2, b_5) = 0,7, e(n_2, b_6) = 1,00$$

$$e(n_2, b_7) = 1,00, e(n_2, b_8) = 0,8, e(n_2, b_9) = 0,8$$

$$(پرس نورون) b_2 = n_2 \text{ برای BMV} \leftarrow$$

$$\xrightarrow{n} e(n_3, b_1) = 0,45, e(n_3, b_2) = 0,25, e(n_3, b_3) = 0,45$$

$$e(n_3, b_4) = 0,8, e(n_3, b_5) = 0,8, e(n_3, b_6) = 0,45$$

$$e(n_3, b_7) = 0,44, e(n_3, b_8) = 0,8, e(n_3, b_9) = 0,8$$

$$(پرس نورون) b_3 = n_3 \text{ برای BMV} \leftarrow$$

$$w(i,j) = w(i,j) + \alpha * h(i,j,1) * (x \cdot w(i,j)) \quad (1)$$

$$\hookrightarrow w(1,1) = w(1,1) + 0.1 h(1,1,1)$$

$$\rightarrow h(1,1,1) = e^{-d(i,j)} = e^{0} = 1 \quad \times ([0.1, 0.5, -0.5, 0.1]) \\ \rightarrow [0.1, 0.5, -0.5, 0.1]$$

$$w(1,1) = w(1,1) + 0.1 \times ([0.1, 0.5, -0.5, 0.1, 1]) = [0.1, 0.5, -0.5, 0.1, 1] +$$

$$([0.1, 0.5, -0.5, 0.1, 1]) = [0.1, 0.5, -0.5, 0.1, 1]$$

$$w(1,2) = w(1,2) + 0.1 h(1,2,1) ([0.0, 0.5, -0.5, 0, 0])$$

$$\rightarrow h(1,2,1) = 0.1 \Delta \rightarrow w(1,2) = [0.1, 0.5, -0.5, 0.1, 1] + [0.1, 0.5, -0.5, 0.1, 1]$$

$$= [0.1, 0.5, -0.5, 0.1, 1]$$

$$w(2,1) = w(2,1) + 0.1 \times 0.2 (-0.5, 0.5, -0.5, 0, 0)$$

$$\rightarrow ([0.4, 0.4, 0.4, 0, 1]) + [0.1, 0.2, 0.1, 0, 0] = [0.5, 0.6, 0.5, 0, 1]$$

نیزی کاربری نبود. برای نمود خواهیم داشت:

$$w(1,1) = w(1,1) + \alpha_1 h(1,1) (n - w(1,1))$$

$$\rightarrow h(1,1) = \alpha_1 v$$

$$w(1,1) = \{0.9, 0.8, 0.7, 0.6\} + \alpha_1 \times 0.5 \times [2.9, -0.5, 0.5, -0.5]$$

$$= [0.9, 0.8, 0.7, 0.6]$$

$$w(2,1) = [0.9, 0.7, 0.5, 0.3] + \alpha_1 [2.9, -0.5, 0.5, -0.5]$$

$$= [0.9, 0.7, 0.5, 0.3]$$

$$w(3,1) = [0.5, 0.3, 0.1, -0.1] + \alpha_1 [0.5, 0.3, 0.1, -0.1] =$$

$$= [0.5, 0.3, 0.1, -0.1]$$

$$w(4,1) = [0.1, 0.1, 0.1, 0.1] + \alpha_1 [0.1, 0.1, 0.1, 0.1] = [0.1, 0.1, 0.1, 0.1]$$

نیز نویسنده ای.

$$w(2,2) = [0.9, 0.7, 0.5, 0.3] + \alpha_1 [0.1, 0.1, 0.1, 0.1] = [0.9, 0.7, 0.5, 0.3]$$

$$w(3,2) = [0.5, 0.3, 0.1, -0.1] + \alpha_1 [0.1, 0.1, 0.1, 0.1] =$$

$$[0.5, 0.3, 0.1, -0.1]$$

$$w(4,2) = [0.1, 0.1, 0.1, 0.1] + \alpha_1 [0.1, 0.1, 0.1, 0.1] =$$

$$[0.1, 0.1, 0.1, 0.1]$$

در روش استنادی ارتباط لایه‌ای، یادداشت نفره به نورون‌ها در هر لایه، ارتباط ورودی و

(۱)

خروجی بسته آورد و می‌شود و در واقع در این روش بررسی شود که هر کدام از سقراطی‌های ورودی

چه گام‌های تدریب خود را تأثیرگذارد. این روش دقیق‌تر است و جزئیات بیشتری از تابعیت می‌دهد.

هزینه‌آلت زیاد است.

در این نظرسنجی بررسی با استنادی از شرکهای معتبر مانند انتل و شود کدام یک از داده

های ورودی دستخوش تأثیر دارند (برخی از مثال‌های پیشنهادی (theories) این روش بررسی

پالایی دارد و در عین حال پیشنهاد آن است که به طور دقیق نمود ابتکار داشتم.

۱. سیان نور.

۱- خوشبینی دلایل معیارهای متناظر مانند silhouette و Kalinski و silhovette می‌باشد.

نقد از نورون‌ها

کیفیت خوشبینی در ۳۰m آهرزین دیده، به انعداد داده‌ها و همچنین محبوی آن پیشنهاد داده.

در صورت وجود ارتباط مناسب و خوشبینی درست انجام می‌شود.

برای پیشنهاد دسته خوشبینی می‌توان کارهای ساخته انتقالی ناصله مناسب با استنادی از تابع انتقام دار