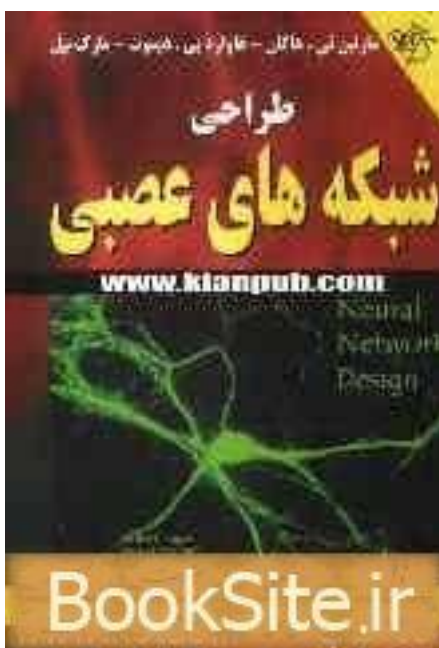




دانشگاه آزاد اسلامی واحد بافت

جزوه درس طراحی

شبکه های عصبی



تهیه کننده :

کاظم فریدی

دانشجوی کارشناسی ارشد مهندسی معماری سیستمهای کامپیوتری

با سلام

دوستان عزیز جزوه اسکن شده مربوط به درس طراحی شبکه عصبی استاد ابدالی در ترم گذشته میباشد یه باهمت یکی از دانشجویان جمع آوری و بر روی سایت جهت استفاده شما قرار گرفته است . با تشکر فریدی

«جایگاه اول» شبکه های عصبی / شبکه های عصبی حاکمان

Subject:

شبکه های عصبی از
Ali Motie Nasrabadi

Date:

تاریخچه شبکه های عصبی: یک شبکه پردازش با اتصالات زیاد

اولین انسان که تلاش کردند ساختاری شبیه مغز انسان ایجاد کنند

چون شبکه های عصبی نیاز به محاسبات داشتند تا زمان ایجاد مدارهای مجامعه در این درس تلاش ایجاد مدلی برای مغز انسان است.

در مغز انسان ^{۱۵} سلول عصبی (نرون) هر نرون معادل یک پردازنده

کوچک است. حدود ۱۰۰ میلیون سلول عصبی در مغز انسان قرار دارد.

هر کدام از این سلول ها یک یا چند ورودی و یک خروجی هستند که به صورت موازی

در کنار هم قرار دارند با ^{۱۶} لینک (پیوند) برای هر سلول عصبی وجود دارد و هر یک

ارتباط با سایر نرون ها یعنی ۱۰۰ تا ده هزار شبکه برای سایر نرون ها

دارند:

ساختار و رسی شبکه های عصبی که دستورالعمل اتصالات آن است.

ده میلیون میلیارد سیم کشی داشته باشیم که امکان پذیر نیست بر فرض

یک زمین به اندازه هم شبیه مدیریت کن که با آن نیست به همین دلیل

© MICRO

مغز انسان به آسانی امکان پذیر نیست و الگوریتم آن فقط قادر بر یادگیری
سخت

هر نرون سرعت آن $\approx$ یک میلی ثانیه است ^{یعنی} در هر ثانیه هزار
تا پنج دهه

در مقایسه با کامپیوترها انسان سرعتش کم است چون کامپیوتر محولی $\approx$ میلیون
دستور و انجام می دهد یعنی کامپیوتر محولی $\approx$ هزار برابر سریع تر است .
از خصوصیات انسان هوشمندی است که دارد .
الگوریتمی که مغز انسان را نه می دارد که ما نمی کنیم تمام کارهای ما طبق الگوریتم است
که در مغز انسان وجود دارد .

اکموزشن و یادگیری باعث ایجاد یک الگوریتم در مغز انسان شود .
بزرگترین خصوصیات شبکه عصبی همین یادگیری است یعنی یک الگوریتم
می دهیم .

بزرگترین خصوصیاتش این است که ما را در الگوریتم پیچیده ای قرار می دهد
ایجاد این که ما از درونش خبر نداریم .

Subject :

Date :

این انحراف در رد و ترلید در مغز نشی چیز نیست نه انحراف نردن باشد نردن وجود دارد در واقع لیب های ما لیب ها در می نیستند .

نکته !

« یک سری از نردن ها از بعضی ها نردن ها و زنی شیه است »

حلقه یک مسیر ارتباطی از ما هیچ با یک دست وصل است ؟ اول یک مسیر

ارتباطی از یک دست به مغز وصلی شود و بعد از مغز به وصلی شود

نکته !

مهم ترین خصوصیت مغز این آموزش است .

مغز یک یک سری نردن است نه بهم لیب دارند نه این لیب ها دقت

نه نگارند مغز را مشخص می کنند .

این ها به هم

هدف کلیه علمی معنوی : ساخت یک سری نردن ۱۱ اقبال کوا آموزش کان

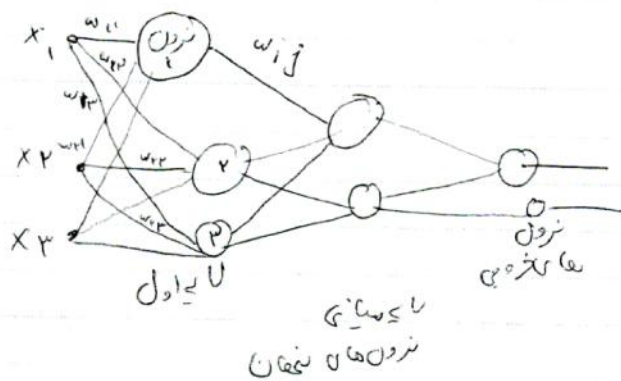
هدف از آموزش ؟ این است که این لیب ها براساس یک وزن دست

بهم وصل کنند که کار ما را انجام دهند .

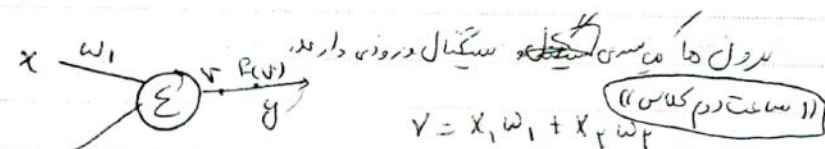
© MICRO

شبکه های عصبی

۳ لایه است



من باید وزن ها، ۱، ۲، ۳ را طوری تنظیم کنیم که هر وقت بیج اومد بیج درخوری
به ۵ و هر وقت شماره اومد شماره درخوری به ۵



$$y = x_1 w_1 + x_2 w_2$$

$$F(x_1 w_1 + x_2 w_2) = y$$

$$F(u) = y$$

بیج

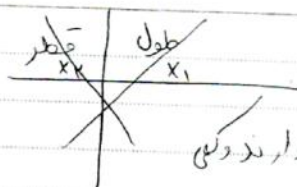
مقدار قطر	
2	2
0.2	6

شماره ها طول و قطر تانگی است

یا قطر تانگی شبکه های عصبی

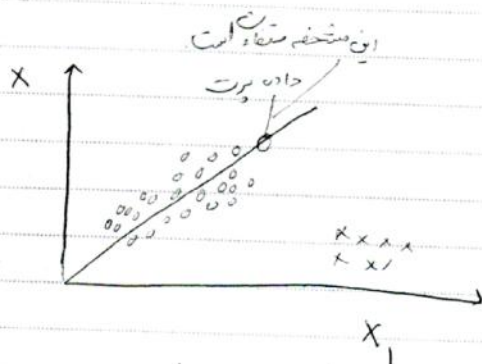
Subject :

Date :



داده های x_1 و x_2 که خصوصیات داده ها دیگر را نشان می دهند

مشاهدات هستند.



اولین انتخاب داده ها باید بتوان باشد

بعد از آن ورودی x_1 به گره های x_2 باید انتخاب کنیم به تناسب باشند

برای دسته بندی پیچیده و شماره باید از گره و مشخص استفاده کنیم

برای تشخیص ورودی ها باید از گره و مشخص گره بکار ببریم

۲- کاربرد شبکه های عصبی:

۱- دسته بندی (جداسازی پیچیده، چهره، جلد سازه، بیماری ها، حیوانات بیگانه)

۲- مدل سازی (مدل های فیزیکی برای راحتی یادگیری می کنند)
اگر مدل داشته باشیم می توانیم آن را به آسانی بنویسیم و به تناسب با سیستم
است مدل سازی

کاربرد همیشگی است.

مزیت: نیاز به پیچیدگی ها و الگوریتمی ندارد. مدل اندازه گیری فشار آب درون

محزن: در بسیاری نیاز به داده الگوریتم ندارد بلکه از روی جدولی که از گذشته سیستم

داریم می آوریم. با استفاده از این جدول ها می بینند که عصبی آموزش می بینند

و رفتار کل سیستم را شبیه عصبی یاد می گیرند.

از این مدل ها با سیستم و یک جدول برای آن تنظیم کنیم و این داده ها را به

تبدیل می کنیم. می دهیم و یک مدل برای انجام دهیم و یک مدل ایجاد می کنیم.

اگر مدل ما دقیق باشد و داده ها را می بیند مدل ما درست جواب می دهد.

در انتخاب داده ها باید دقت کنیم کلام پارامتر مفید تر و خاص تر هستند.

این یا را به ما مشخص می گوید.

استخوان خیره: کسی است همه جواب برای داند.

مثالی از کارها: تبدیل عصبی. مثل این که تبدیل عصبی برای بررسی استحکام سازه ها

۱- اول معیارها را با استفاده از مشخصات مشخص انجام می دهیم

۲- کلیه جدول ها مربوط به استحکام سازه ها

۳- بعدی روم سازه ها مختلف و بررسی می کنیم و در جدول مشخص می کنیم. در انتها

Date:

Date:

بعد میزان سیم و تحریک شود مشخص می کنیم که این محرک می کنیم کدوم مدلی از ساختمان ماست و بعد این مدل رو می سازیم ساختمان پیدا می کنیم که این مدلی گوید که ساختمان ما تحریک می شود یا خیر.

نکته! دقیقاً کنیم مدلی که پیدا می کنیم عصبی برای زلزله ۶ رشتی باشد نه بیشتر نه کمتر که تا زیر زلزله را روی ساختمان بررسی کنیم.

محرک های در این کار جا که مختلفی با ساختمان های مختلفی بررسی کنیم در روی شبکه های دیوار کنیم

خطی آن
کاربرد ۳ آدره های ما، ها : هر جایی که سیم داده داشته باشیم بخوانیم

مدل ش کنیم شبکه های عصبی این کار را انجام می دهد.

کمی شبکه های عصبی قابلیت جاگزینی مغز انسان را دارد یا خیر

نکته! برای دقیق شدن شبکه های عصبی داده ها باید دقیق و خوب باشند.

تمام داده ها باید دقیق باشند و هست داشته باشند

نکته! اگر شبکه های عصبی همه چی بر اساس داده است. شبکه های عصبی با سیم

داده منجر داده
۱- از اول داده ها را مشخص می کنیم و مشخص می کنیم این داده
بیخ است و داده ای داریم همه است. مثلاً ۱ همه جایی که منور می
۲- شبکه های عصبی جایی است که منور و غیر منور
یکی نیست منور می کنیم

شبکه ی عصبی بدون علم:

۲- داده مارا اتوماتیک خوش کرده بندی کند و دسته بندی کند

مثلا بگه ارایه طول از عرض شبکه است که دسته رارای ^{طول شون شبکه دین}

درکه . قلا بلی به اکوزن در مورد بر اساس قدر وزن جانات های به

قدرت و جاقن زن و قد بلند و لاغر مرد باشن . این دسته بندی ها بر اساس

اکموشن است .

این دسته بندی همراست چون میکن است شبکه ها عصبی از دیدگاه های گاه می کند که ما به گاه نمی کنیم هوار از یک دیدگاه گاه می کند چون بصورت رندم

است و گاه به صورت به ترتیبی گاه به گاه می شود

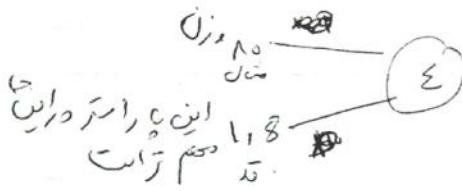
به این روش به دنبال اثبات روش شبکه عصبی بدون علم است ۱۰

مثال:

ایجاد شبکه ی عصبی برای کشتی گیر و والیبال ^{مخصوصه} همرا تفتیک شون :

داده : ترکیب قد و وزن ^(۱) ^(۲) که در شتی گیر ها وزن شته است

یا را متر وزن ۹۰ kg
قد ۱۸۰



ارزش دوتا کم است

Student:

Date:

از ابتدا ما باید همه داده ها را از ماشین کسب یعنی صلا در وزن تمام وزن ها را

بین ۱ تا ۱۰۰۰۰۰ بیاوریم و قد ها را نیز را بین ۰ تا ۱۰۰۰۰۰ بیاوریم.

ی	هدف یا Target یا خروجی	وزن	قد
۰	۰	۰-۱	۰-۱

۷ هدف از آموزش: وزن ها را چگونه ای ایجاد کند که خروجی شبکه ی عصبی نزدیک

نشود به هدف ما.

هدف ما این است که به خروجی شبکه ی عصبی ما D حد در کم می شود
مفهوم آگروگیشن این داده ها را با بد این قدر انجام بد هیم تا کم می شوند.

خطا $D-y$ است که باید کم شود به بد در درش انجام می شود تا خود ترا اصلاح

کن یعنی امری در نشود می گیم خود ترا اصلاح کن یک داده را نشان

می دهیم (اینترینسک) $intra$ $shen$ یک داده، یک پیچ و یک ماس

۲- همه داده ها را نشان می دهیم همه پیچ ها، همه صوره ها، یک $epoch$

که معلوم بماند که یک بار در گذرد باید چندین بار انجام بد هیم.

هدف از تست در شبکه های عصبی؟

همیشه صحت آموزش دانش شبکه های عصبی هدف خود کلاس است که در سطح تدریس باشد
برگشتن از این حفظ کرده و این باید از داده ها است که می بینیم که آموزش نباید از این

همان : هیچ و صفر
مثال:

اول عدد را هیچ و صفر بگیریم و چند را که آموزش می بینیم بدیم به یاد می آید
نحوه انتخاب هیچ و صفر:

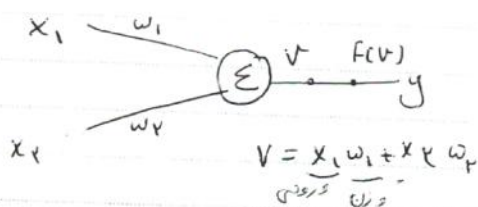
اول بر روی کلاس خود تا حدی داریم بعد تا حدی که را را بنویسیم و به صورت
تصادفی یکی را انتخاب کنیم بعد هیچ و صفر دهانی و انداز می گیریم و می بینیم
از این عدد تا باز می آید که کرده برای امتحان که داده ها آموزش
۳۰ تا می ماند که برای امتحان در آموزش یکبار نمی روند که در آموزش
نبودند اگر درست بود یعنی یا در گفته می اگر خطا داشت یعنی شبکه عصبی دارد
حفظ می کنند پس به آموزش متوقف می شود.

← خطای هموم // شبکه ی عصبی //

ADALINE ← شبکه ی عصبی داریم
perceptron ←

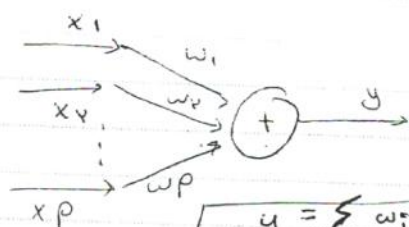
: ADALINE

شبکه ی perceptron است با تابع فعال خطی است نه بیشینه اش و یاد داده است
خطا است به چه معنی است؟



$$F(v) = v = y = x_1 w_1 + x_2 w_2$$

$p=2$ تعداد ورودی ها
طول - عرض
مصفوفه



$$y = \sum w_i x_i = X^T \cdot w$$

این عبارت را به صورت ماتریس می توان نوشت

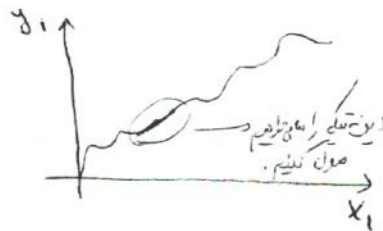
$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = x_1 w_1 + x_2 w_2$$

ورودی ها وزن ها

$$X^T \Rightarrow X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

که هر راسی آدراسی

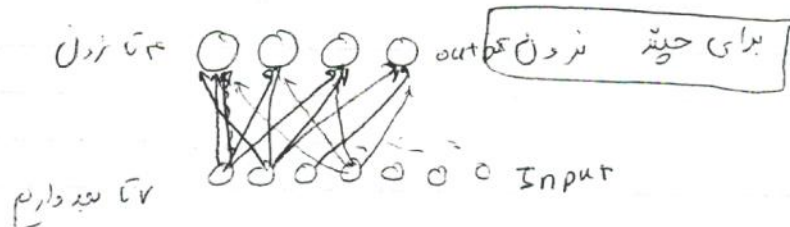
۱. بطور کلی برای تقریب خطی استفاده می شود. از یک سیم چند سطحی غیر خطی که به وسیله سیم های ورودی و خروجی به فضای P در فضای P بعدی تر اندازی شود.



در یک سیم غیر خطی نزدیک ما
خطی است که می توانیم
مدل کنیم.

۲- روش بپی و ولتر کردن خطی

معمولاً error در شبکه های آدراسی از یک نرون error $d - y$ ←
 $\boxed{\text{برای نرون} \xrightarrow{\text{error}} d - y = \epsilon}$ $\epsilon = d - y$



$$\epsilon_i = d_i - y_i \quad i = 1:m$$

$$\epsilon_{m \times 1} = d_{m \times 1} - y_{m \times 1}$$

جدول آموزش

$x_1, x_2, x_3, x_4, x_5, x_6, x_7$	D_1, D_2, D_3, D_4	y_1, y_2, y_3, y_4	$\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$
داده اول	0.5, 0.3, 0.3, 0.2	0.9, 0.4, 0.2, 0	0.15, 0.5, 0.2, 0.15
0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1		0.8, 0.7, 0.4, 1	

$\epsilon_1 = 0.9 - 0.5 = 0.4$
 مربوط به D_1
 $\epsilon_2 = 0.4 - 0.3 = 0.1$
 داده اول

منبر
 همه ورودی ها در وزن بچ می شوند با هم جمع می شوند و خروجی نود دوم بدست می آید مثل وزن x_1

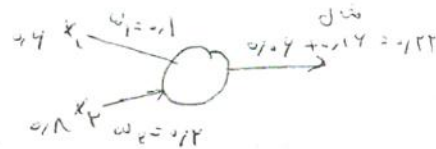
هدف: کمترین مقدار بسن $error_2$ ها $\leftarrow$ سوال چگونه کل این $error_2$ ها کمترین مقدار بسن (اول $error_2$ ها را بتوان ۲ی سویم چه رابطه جمع

می کنیم هدف این است $error_2$ ها کمترین مقدار بسنند

$$J(\omega) = \frac{1}{2mN} \sum_{i=1}^m \sum_{j=1}^N \epsilon_j^2(i)$$

مثال برای ۲ ورودی

x_1	x_2	D	y	$\epsilon = D - y$
0.2	0.4	0.4	0.58	0.12
0.4	0.8	0.8	0.72	0.08
0.1	0.8	0.1		



$$0.4 \times 0.1 = 0.04$$

$$0.8 \times 0.2 = 0.16$$

$$0.04 + 0.16 = 0.2$$

حسن توجه:

در توان ۲ خط های توخالی، توخالی می شوند و خط های پر را برابری شوند

$$E = \frac{1}{2mn} \sum_{i=1}^n \sum_{j=1}^m \epsilon_j^2(i)$$

$\sum \epsilon_j^2(i)$ یعنی ϵ^2 را به هم جمع کن

m تعداد نرون ها
 n تعداد آموزش ها
 انجام دادیم $n=100$
 به تعداد آگنوما این کار را کرداری شود

هر چه عبارت ریاضی کمتر باشد بهتر است

تعداد آموزش ها

w وزن

y
 $N \times m$

p ابعاد ورودی
 N : تعداد آموزش ها
 m : تعداد نرون ها

$x_N p$ به ابعاد ورودی

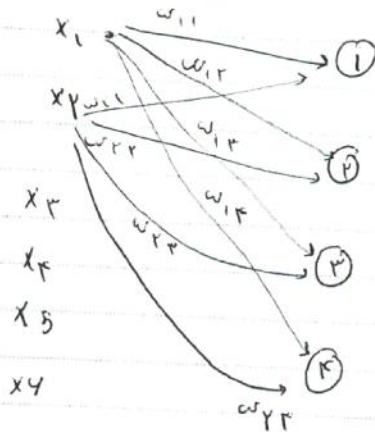
$D_{N \times m}$ هدف
 به N و m نظر
 به تعداد نرون ها

$$X_{N \times p} = \begin{bmatrix} x_{11} & \dots & x_{1p} \\ x_{21} & \dots & x_{2p} \\ \vdots & & \vdots \\ x_{N1} & \dots & x_{Np} \end{bmatrix}$$

ماتریس ورودی

$$D_{N \times m} = \begin{bmatrix} d_{11} & \dots & d_{1m} \\ d_{21} & \dots & d_{2m} \\ \vdots & & \vdots \\ d_{N1} & \dots & d_{Nm} \end{bmatrix}$$

ماتریس هدف



$$\begin{matrix} w_{ij} \\ \text{نورون } j \\ \text{ورودی } i \end{matrix}$$

$$\begin{matrix} w_{ij} \\ \text{نورون } j \\ \text{خروجی } i \end{matrix}$$

ورودی ۱	نورون ۱	نورون ۲	نورون ۳	نورون ۴
	w_{11}	w_{12}	w_{13}	w_{14}
ورودی ۲	w_{21}	w_{22}	w_{23}	w_{24}
⋮				

$$W_{p \times m} = \begin{bmatrix} w_{11} & \dots & w_{1m} \\ w_{21} & \dots & w_{2m} \\ \vdots & & \vdots \\ w_{p1} & \dots & w_{pm} \end{bmatrix}$$

ماتریس وزن ها

$$y = X W$$

معادله

ی بردار متناظر با d باشد در واقع است

$$D_{N \times m} = y_{N \times m}$$

چون $D_{N \times m}$ است

می شدن y و D مدتی رسیدن به هدف ما است برای همین با

باید w را بدست آوریم که D برابر y بشود.

$$X \cdot w = D \rightarrow x^{-1} X \cdot w = x^{-1} D \rightarrow I \cdot w = x^{-1} D$$

نتیجه

نکته!

اگر تعداد معادلات از تعداد مجهولات کمتر باشد معادله جواب دقیق ندارد.

مثلاً ۳ معادله ۴ تا مجهول

یا ۱۰ معادله ۲ تا مجهول

در هیچ دهمه

برای حل این معادلات باید نقطه را در نظر بگیریم و فاصله اش را با ۳ معادله



و اندازه بگیریم. دقیق ترین جوابی که می دهیم

را به MS

$$\omega_{p \times m} = (X_{p \times N}^T X_{N \times p})^{-1} X_{p \times N}^T D_{N \times m}$$

این روش در بین بهترین جوابی که می دهیم

(The error equitation =)

$$J(\omega) = \frac{1}{2N} \sum_m E_m^T E_m$$

این روش کم ترین خطا را دارد



performance index

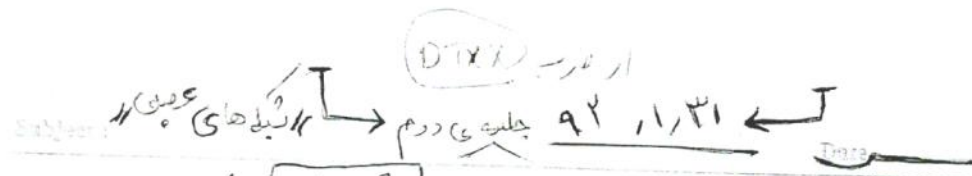
که ما داریم راهی برای محاسبه آن داریم در خودش ضرب کنیم حاصل می شود؟
تقریباً می شود مربع خطاها

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

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

$$= a_{11}^2 + a_{12}^2 + a_{21}^2 + a_{22}^2$$

$$\begin{bmatrix} a_{11}^2 + a_{12}^2 + a_{13}^2 \\ a_{11} \times a_{21} + a_{12} \times a_{22} + a_{13} \times a_{23} \\ a_{21}^2 + a_{22}^2 + a_{23}^2 \end{bmatrix}$$



MSE

برای یک نرون

For a single neuron, $m=1$:

error ترانسپوز

ارکس نرون از سبب آردین

دانشه با سیم

اوپر می تدر

$$\Rightarrow J(n) = \frac{1}{2N} E^T E_{N \times 1}$$

Replacing error in equation:

با سیمینی error در محادله:

$$J(w) = \frac{1}{2N} [(D - Xw)^T (D - Xw)]$$

جایی که همه دوم را باید ترانسپوز کنیم باید جای سیم در معادله

$$= \frac{1}{2N} [D^T - w^T X^T] [D - Xw]$$

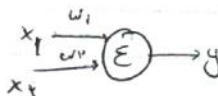
$$= \frac{1}{2N} [D^T D - D^T Xw - w^T X^T D + w^T X^T Xw]$$

تغییراتی که می بینیم

$$= \frac{1}{2N} [D^T D - 2D^T Xw + w^T X^T Xw]$$

آنها را می توانیم به صورت زیر بنویسیم

$$w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$



در نتیجه یک ماتریس $N \times 2$ داریم N تعداد داده ها آموزش است

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \\ 6 & 1 \\ 7 & 1 \\ 8 & 1 \\ 9 & 1 \end{bmatrix}$$

$N \times 2$

دانشه یعنی $N=10$ است

مثال

$$DD^T - D = D^T$$



۲۸؟



ماتریس هدف D هم ده تا سطر و ۲ ستون است

$$D = \begin{bmatrix} 1.1 \\ 1.8 \\ 3.2 \\ 4.1 \\ 4.8 \\ 5.7 \end{bmatrix}$$

D^T آن مندرجی نیست
 $N=20 \leftarrow$ ده تا سطر x و ده تا ستون دارند

$$DD^T = 1.1^2 + 1.8^2 + 3.2^2 + \dots = 38.38$$

$$\vec{D} (w_1, w_2) = \frac{1}{20} [38.38 - 2 \mid \frac{D^T \times x}{20} \mid w_1, w_2] +$$

$$[w_1, w_2] \begin{bmatrix} 38.38 & 4.5 \\ 4.5 & 10 \end{bmatrix} [w_1, w_2]^T = \frac{1}{20} [38.38w_1 + 24.5w_2 - 11.5w_1 - 11.5w_2 +$$

۲۸.۵ $w_1^2 + 4.95w_1w_2 + 1.5w_2^2$ x مجموع سطر استون x

در محور w_1 و w_2

خطی کمترین مقدار را در خودی دارد ، در حقیقت w_1 و w_2 در این نقطه دارا وزن ها مطلوب هستند به اینست می شود شبکه های عصبی من با این ورودی ها

به خروجی را ترکیب می کنند به این اهداف و وزن ها

وقتی شکل را رسم کنیم شکل ما ده را اینجا شده چرا در یک جهت

بسته اند؟ ندارد؟ چون داده در یک جهت ~~انجا~~ نه داشته در یک جهت ~~انجا~~ نه دارد. ولی اگر داده ها در یک جهت ~~انجا~~ داشته باشند نه تمام در یک جهت ~~انجا~~ دارند.

Subject:

Date:

هدف ما از آموزش نقطه‌ای که وزن مطلوب است (که گفت است) ما را می‌دهد.

که از رابطه $w_{p \times m}$ برای دهنده که w در تلف و دارد رابطه می‌دهد

$$w_{p \times N} = (X_{p \times N}^T X_{p \times N})^{-1} X_{p \times N}^T D_{N \times m} \leftarrow MSE$$

در آداین روش MSE وزن‌های دهنده ولی یک سری محدودیت دارد؟

برای شبکه‌های عصبی قابل اطمینان، مسئله ترکیبی عصبی با آداین باشد که

این روش MSE جواب می‌دهد؟ نمی‌خورد به دلیل زیر

$$d - y = e \rightarrow d - xw = e$$

$$d - y = e \rightarrow d - f(xw) = e$$

اگر N بزرگ باشد محاسبات زیاد می‌شود.

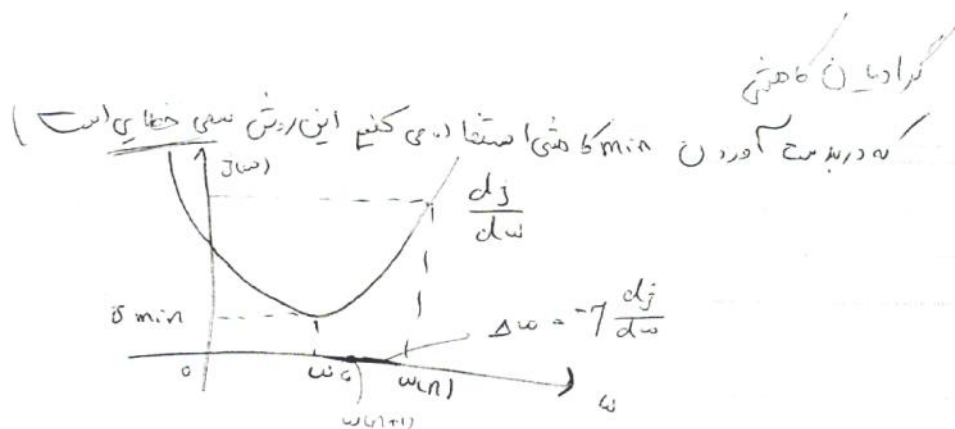
در روش **گرا دین** که منی $\square$ هر دو به هر دو تلاش می‌کنیم خطا را کاهش

به هم می‌زنیم. ما باید وزنی را بدست می‌آوریم که کمترین خطا را داشته باشد. فرض کنیم

ما خطای داشته باشیم حداقل ۲ باشد $\square$ از (w) مشتق می‌گیریم اگر

مشتق مثبت بود به جهت مثبت حرکت کند $\square$ مشتق منفی بود در جهت (منفی)

مثبت حرکت کند این مقدار را می‌دهیم که مشتق منفی شود (تلف باشد)



تغییرات وزن مشتق می گیریم از وزن $\leftarrow \frac{dJ}{dw}$ در η

$$\Rightarrow w(n+1) = w(n) + \Delta w(n) \quad J(w) = \frac{1}{2N} [D^T D - 2D^T Xw - w^T X^T X w]$$

گرایان هشی (مشتق به جزیء بعد دارد)

$$\Delta w(n) = -\eta \nabla J(w(n)) \quad \frac{\partial J}{\partial w} = \frac{1}{N} [-D^T X + X^T X w]$$

این را در گام آموزش نرخ آمدن می بینیم

$w(n)$

برای سه ترمون که خطای باشد در این شرایط $\rightarrow$

LMS

$$y = \sum w_i x_i$$

$$\epsilon = (d - y)$$

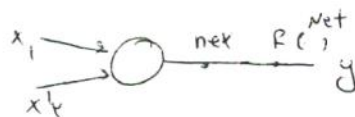
$$J = \frac{1}{2} \epsilon^2 = \frac{1}{2} (d - \sum w_i x_i)^2 \quad \text{مشتق بگیریم} \quad \frac{\partial J}{\partial w} = \frac{\partial J}{\partial \epsilon} \times \frac{\partial \epsilon}{\partial w}$$

$$\frac{\partial J}{\partial w_i} = -\epsilon \frac{\partial \epsilon}{\partial w_i} = - (d - \sum w_i x_i) x_i$$

خروجی $\rightarrow$ وزن $\rightarrow$ هدف

$$\frac{\partial J}{\partial w_i} = -\epsilon \frac{\partial \epsilon}{\partial w_i} = - (d - \sum w_i x_i) x_i$$

۲۰



$$net = \sum w_i x_i$$

$$y = f(net)$$

$$E = \frac{1}{2} (d - y)^2$$

تابع خطای هر سیستم

$$J = \frac{1}{2} e^2 = \frac{1}{2} (d - f(net))^2$$

خطا

$$\frac{\partial J}{\partial w_i} = \frac{\partial J}{\partial net} \frac{\partial net}{\partial w_i} = -(d - f(net)) f'(net) \frac{\partial net}{\partial w_i}$$

$$\frac{\partial J}{\partial w_i} = -(d - f(net)) f'(net) x_i$$

انواع آموزش شبکه های عصبی موجود است:

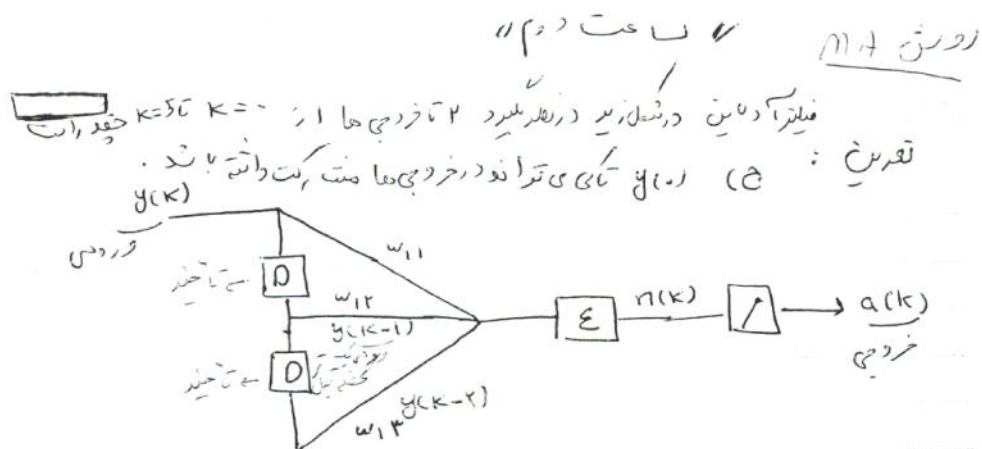
۱- ترتیبی (sequential mod)

۲- انزائی (online)، روی هر آنکو (Batch mod)

- ① یک ورودی را نشان می دهیم در همان لحظه هم می خواهیم اصلاحش کنیم (online / per pattern)
- ② یک سری ورودی را نشان می دهیم و خطاها را یادداشت می کنیم و تلاش می کنیم یک

جامه خطاها / رفع کنیم (epoch) (off line)

(per epoch)



$w_{11} = 2$
 $w_{12} = -1$
 $w_{13} = 3$

$\{y(k)\} = \{ \dots, 0, 0, 0, 5, -4, 0, 0, 0 \}$

$y(0) = 5$
 $y(1) = -4$
 $k=0$

این کلمه صفر است
 کلمه منفی
 $y(0)$
 $y(1)$

$y(k) = ?$ و متغیرها چیستند

$a(1) = ?$
 $a(0) = ?$
 $a(-1) = ?$

حل به صورت زیر می باشد

برای این صورت

$$a(k) = y(k)w_{11} + y(k-1)w_{12} + y(k-2)w_{13}$$

در این نوع ورودی ها هم برابر اند

$$a(0) = y(0)w_{11} + y(-1)w_{12} + y(-2)w_{13}$$

$$a(0) = 5 \cdot 2 = 10$$

$$w_{11} = 2$$

Step 1

Date

$$a(-1) = y(-1) \times w_{11} + y(-2) \times w_{12} + y(-3) \times w_{13} = 0$$

$$a(1) = y(1) \times w_{11} + y(0) \times w_{12} + y(1) \times w_{13} = -13$$

⋮

حل به صورت ماتریسی حل می کنیم

$$y = X \cdot w$$

X ← y

$$\begin{bmatrix} y_{1-1} & y_{1-2} & y_{1-3} \\ y_{2-1} & y_{2-2} & y_{2-3} \\ y_{3-1} & y_{3-2} & y_{3-3} \end{bmatrix} \begin{pmatrix} w_{11} \\ w_{12} \\ w_{13} \end{pmatrix}$$

ی برداری
ی لیست

$$\begin{pmatrix} 0 & 0 & 0 & 5 & -4 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

از اینجا شروع می کنیم

$$N(x) = \begin{pmatrix} 0 & 0 & 0 \\ 5 & 0 & 0 \\ -4 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -4 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 10 \\ -13 \\ 16 \\ 12 \\ 0 \end{pmatrix}$$

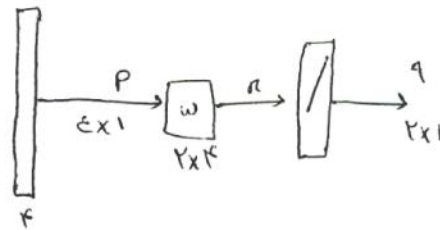
وزن ها

۴×۳ ۳×۱ ۴×۱

MICRO

۲۲

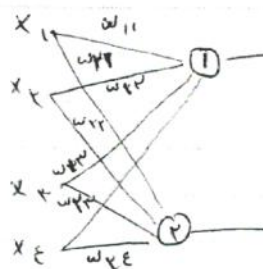
نقشه:



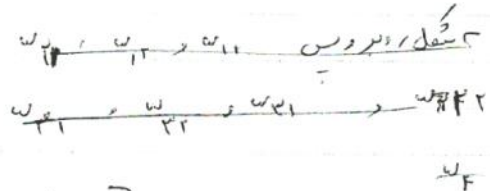
$$P_1 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ -1 \end{bmatrix}, \quad T_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$P_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \quad T_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

با استفاده از MSE یا قاعده شبکه مغزونی استفاده کنید و وزن ها را حساب کنید.



از دو شبکه معلوم می شود که ورودی دارد و خروجی می باشد به ساختارش می شود



$$\begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \end{bmatrix}$$

$$X = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \begin{matrix} P_1 \\ P_2 \end{matrix} \quad (2 \times 4)$$

$$T = \begin{bmatrix} +1 & -1 \\ 1 & 1 \end{bmatrix} \begin{matrix} T_1 \\ T_2 \end{matrix} \quad (2 \times 2)$$

Subject

Date

هدف $w = (X^T X)^{-1} X^T D \rightarrow$

$X_{(2 \times 4)} \rightarrow [(1 \times 2) (2 \times 4)]^{-1} (1 \times 2) (2 \times 2)$

ماتریس نه ماتریس 2×2 داریم آن است که برای آنکه ماتریس 2×4 که

از این خواهد جور در بیاید باید یکم توضیح تا ماتریس آن بد

$$w^T = \left(X^T \begin{bmatrix} 1 & 1 \\ -1 & +1 \\ 1 & -1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 \\ -1 & -1 \\ 1 & -1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

4×2 2×2

در اینجا به ما میگوید که X چگونه است درستی است

$$= \begin{bmatrix} \frac{1}{2} & 0 & 0 & -\frac{1}{4} \\ 0 & \frac{1}{2} & -\frac{1}{4} & 0 \end{bmatrix}$$

نت جواب

$$-1 \times \frac{1}{2} - \frac{1}{4} \Rightarrow -1$$

در ۱ دوی

حاصل کنیم t_1 را بد
که ها را باید

۲۵

← ترین →

تقریباً تپ را با روش گرادیان کاهشی با وزن ها و بایاس ها تغییر می دهیم

کدام وزن ها را در یک مرحله یا دو مرحله به دست می آوریم؟

وزن های چرخش $w = \begin{pmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \end{pmatrix}$

ماتریس	Y	T	Y	E
	1	1	0	1
	-1	-1	0	1
	1	1	0	1
	-1	-1	0	1
مرحله ۱	1	1	0.4	0.4
مرحله ۲	1	1	0.4	0.4

۱ را در مرحله ۱ ضرب می کنیم به ۰.۴

$$E = T - Y$$

$$w(n+1) = w(n) - \eta \Delta z(w(n))$$

در مرحله ۱

$$w(n+1) = w(n) - \eta E X$$

$$\eta = 0.1$$

$$w(1) = w(0) - \eta E(0) X$$

وزن در مرحله ۱

$$w(1) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} - 0.1 \times \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{pmatrix}$$

$$-0.1 \times \begin{pmatrix} 2 & -2 & 2 & 0 \\ 2 & 0 & 0 & -2 \end{pmatrix}$$

$$w(1) = \begin{pmatrix} 0 & -0.2 & -0.2 & 0 \\ -0.2 & 0 & 0 & 0.2 \end{pmatrix}$$

۲۶

Subject:

Date:

$$w(1) = w(0) - \eta x(1) \cdot x$$

می‌تواند به‌تر صورتی درج شود

$$w(2) = \begin{bmatrix} 0 & 0.2 & -0.2 & 0 \\ -0.2 & 0 & 0 & 0.2 \end{bmatrix} - 0.1 \times \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix}$$

$$x \cdot w = y = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 0 & -0.2 \\ 0.2 & 0 \\ -0.2 & 0 \\ 0 & 0.2 \end{bmatrix}$$

$$= \begin{bmatrix} 0.2 & -0.2 \\ 0.2 & -0.2 \end{bmatrix}$$

$$= \begin{bmatrix} 0.2 & -0.2 \\ 0.2 & -0.2 \end{bmatrix}$$

حل مجدد: درست شد

$$w = (x^T x)^{-1} x^T O$$

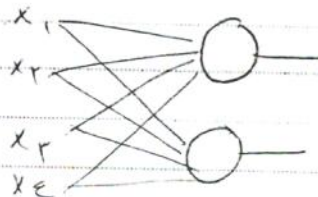
$$((4 \times 2) \times (2 \times 4))^{-1} \times (4 \times 2)$$

$$(8 \times 8) \times (2 \times 2) = 4 \times 2$$

$$x = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix}$$

$$I = 0 \text{ ماتریس } \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$w = \text{ماتریس وزن} = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \\ w_{31} & w_{32} \\ w_{41} & w_{42} \end{bmatrix}$$



© MICRO

۲۷

$$LMS: \Delta \omega_{p \times m}(n) = \gamma X_{p \times 1}^T(n) \varepsilon_{1 \times m}^T(n)$$

$$\begin{cases} \varepsilon_i = d_i - y_i & i=1:m \\ \varepsilon_{mx1} = d_{mx1} - y_{mx1} \end{cases}$$

ادش
// برای ن و ن (دو)

$$\omega_{(n+1)} = \omega_{(n)} + \gamma X_{(n)}^T \varepsilon_{(n)}$$

$$\omega_{Ex2} = \omega_{(Ex2)} - \gamma X_{Ex2}^T \varepsilon^T$$

$$\varepsilon_{Ex2} \varepsilon_{Ex2} \varepsilon_{Ex2} = \varepsilon_{Ex2}$$

$$\omega(1) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} - \gamma \begin{bmatrix} 1 & +1 \\ -1 & +1 \\ 1 & -1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 0 & 2 \\ -2 & 0 \\ 2 & 0 \\ 0 & -2 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & -0.2 \\ 0.2 & 0 \\ -0.2 & 0 \\ 0 & 0.2 \end{bmatrix}$$

$$\omega^T = \begin{bmatrix} 1/4 & 0 \\ 0 & 1/4 \\ 0 & -1/2 \\ 1/4 & 0 \end{bmatrix}$$

مرحله ۲ وزن در مرحله ۲ی شود (این در مرحله ۱)

$$\omega(2) = \begin{bmatrix} 0 & -0.2 \\ 0.2 & 0 \\ -0.2 & 0 \\ 0 & 0.2 \end{bmatrix} - \gamma \begin{bmatrix} 1 & +1 \\ -1 & +1 \\ 1 & -1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 1.4 & 0.6 \\ -0.2 & 1.4 \end{bmatrix} \varepsilon^T$$

$$\Rightarrow \begin{bmatrix} 0.18 & 2 \\ -2 & 0.18 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & -0.2 \\ 0.2 & 0 \\ -0.2 & 0 \\ 0 & 0.2 \end{bmatrix}$$

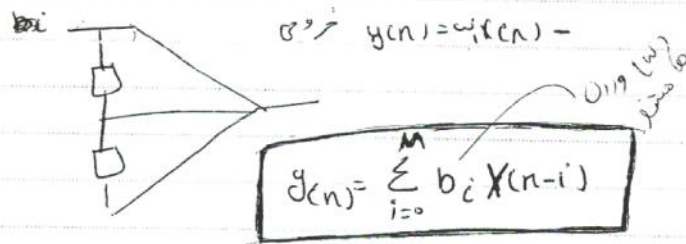
$$\begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix}_{2 \times 4} \begin{bmatrix} 0 & -0.2 \\ 0.8 & 0 \\ -0.2 & 0 \\ 0 & 0.2 \end{bmatrix}_{4 \times 2} = \begin{bmatrix} -0.2 & -0.2 \\ 0.8 & -0.2 \end{bmatrix}_y$$

$$\varepsilon = d - y = \begin{bmatrix} 1.2 & -0.2 \\ 0.2 & 1.2 \end{bmatrix}$$

تبدیل های عصبی به سمت بالای دارند.

در روش MA یک سیستم را می توانیم مدل کنیم بر مبنای ورودی ها و خروجی

خروجی ها را از ورودی ها بدست می آوریم



در حقیقت داریم خروجی را از ورودی تخمین می زنیم. y در مرحله دوم

می شود x_0, x_1, x_2 در مرحله دوم

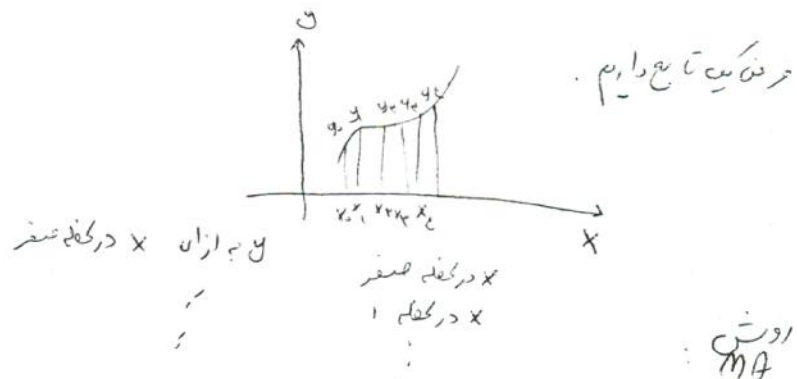
ضرایب b_0, b_1, b_2 می شوند w_0, w_1, w_2

همان وزن ها هستند.

نکته! ما می توانیم ماتریس هدف و ورودی ایجاد کنیم از روش وزن ضرایب معادلات را بدست می آوریم که همان وزن های هستند.

این فرم داریم
دو خط هم پهنه داشته تو نودها متشابه با این تو نود

clear
close all



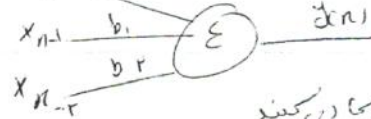
$$y(n) = b_0 x_n + b_1 x_{n-1} + b_2 x_{n-2} + \boxed{y(n-3)}$$

در هر مرحله باید بنویسیم
در بعضی توابع هم است
این محاسبه قبل تر
این محاسبه قبل تر

فرم شبکه
بیشتر شبکه
مای خوانسیم باید شبکه های ضرایب b_0, b_1, b_2 بدست آوریم؟
برای اینکه شبکه ی خطی مول زکای کنیم

$n=2$

هدف y هست که می داریم



$$\begin{pmatrix} x_1 & x_2 & x_3 \\ x_2 & x_3 & x_4 \\ x_3 & x_4 & x_5 \end{pmatrix}$$

ی اینجا می کنید
این
این

۳۰

$$d = \begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} \quad X = \begin{bmatrix} x_2 & x_1 & x_0 \\ x_3 & x_2 & x_1 \\ x_4 & x_3 & x_2 \\ \vdots & \vdots & \vdots \\ x_n & x_{n-1} & x_{n-2} \end{bmatrix} \quad (n-1) \times 3$$

$$d = \begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} \quad (n-1) \times 1$$

$$\begin{bmatrix} x_2 & x_1 & x_0 \\ x_3 & x_2 & x_1 \\ \vdots & \vdots & \vdots \end{bmatrix} \quad \begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ \vdots \end{bmatrix}$$

ناترین ورودی ناترین هدف

آنها را وزن های میان فرایب می کنند

مثلاً اگر نام ~~شماره~~ ^{ای}ی خواهیم که تحت وابستگی اول باید کار را معلوم

داریم فرایب $x_1 - x_2$ را ایجاد می کنند و هدف را هم مشخص می کنند

سیستم شبکه ای عصبی را آموزش می دهند و سپس فرایب را درست

می کنند. مثال

$$y_n = 3x_{n-2} + 2x_{n-1} + 1x_n$$

$$y_2 = 3x_2 + 2x_1 + x_0$$

که برای عدد Random و دهد در زیر حاشیه می کنند و سراسر هدف

p ایجاد می کنند و با این داده ها آموزش می دهند. در نهایت وزن ها را یاد

بشوند s و 2 و 1 به کار جواب شود شبکه ای درست است

clear
close all
num = [1 2 3] ^{بسیطره}
den = 1;
x = Random(1, 1000); ^{ستون}
y = Filter(num, den, x); ^{یک ستون}

کین رندوم ایجاد می کند

For i = 1 : 998
X(i, 1) = ... ^{در سطر اول ۱ تا ۹۹۸ برابر}
در سطر دوم ۱ تا ۹۹۸ در هر بار برابر
D(i, 1) = ... ^{ماتریس هدف را ایجاد کن}
end

$w = (X^T X)^{-1}$ ^{ماتریس وزن را اینجا حساب کرده}
که یک ماتریس ۱x3 است

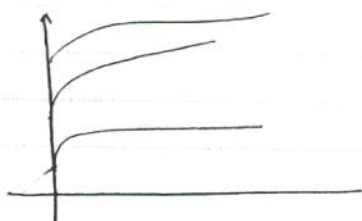
MSE ^{تایید روش}
N = 20
For i = 1 : N
mod(i, N) + 1
تمام شده رابطه را تقسیم بر ۲
و به عنوان اندیس استفاده می شود

از این ۹۹۸ تا ۱ ایجاد شده فقط از ۲ تا استفاده می کنیم بقیه استفاده نمی شوند

Date:

بعد جواب را plot کرده که گفته اگر وزن ها اویخته باشند
 $\eta = 0.01$ باشد اعداد error را رسم کرده که این خطا بعد از ۵۰۰ بار تکرار
 شبکه صفری شود.

مضایب μ_0, μ_1, μ_2 رسیده به ۲ و ۳ رسید است



تغییر $\eta = 0.01$ زودتر به جوابی رسم این error زودتر صفری شود باز است
 شدن، اما سرعت آموزش با آن رفته //

$N = 20$ وزن ها هم میند و سرعت یادگیری هم فرق نکرده
 $\eta = 0.01$

وزن Random باشد

$\eta = 0.01$ وزن ها Random
 زمان یادگیری فرقی نکرده

$\mu = 10$
 $\sigma = 1.1$
 ← آهسته یا در فته از سمت شیب دریا در فته

دادها به صورتش کم شوند اگرش نفقت می شود

LR (Linear Regression) = newlin(PR, LD)
 خطی است
 min و max ورودی را دارد
 به سطرین
 انتخاب می کند

$$LR = 6/01$$

$$LD = [E]$$

تابع $\sin(\pi T)$ را در ω_0 [2] در نقطه ω_0 داده ω_0 یاد کنید

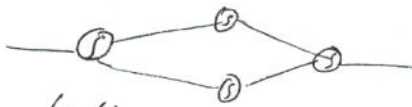
عوازمون را باروش ذیل تفهید کنید .

$$y = \sin \pi T$$

$$t = 0, 1, 2$$

از مقدار یک عدد یک عدد برود تا ۵۰ برسد

(1 2 1) این معبر شده ی عوی ایما (ی) کعد



۱) دولت برای جبران از حفظ کردن در بند های است (Validation)

①

۱۳/۲/۵۴

← شبیه سازی
نیروی خونی

" ضرورتی ز درسی ها قبل "

از درک و این تابع اعمال شده اش که تابع خطی است

که اگر در درج اول در ترتیب خطی استفاده می شود
۲- بیشترین هم انعام می دهد

دلیل مهم بودن مفهوم error

چون خطی نیست از آموزش به بین شبکه آموزش دیده است و ضعیف

$$\epsilon = d - y$$

که باید مقدارش کم باشد

میانگین مربعات MSE

که روی توابع غیر خطی قابل استفاده نبود

برای حل از روش گرادیان کاهشی استفاده می کنیم

$$w(n+1) = w(n) + \Delta w(n)$$

$$\Delta w(n) = -\eta \nabla J(w(n))$$

نزدیک به صفر

ص

گام آخر آموزش باید کم باشد.

در فرآیند آموزش باید به این نکات توجه کرد

→ MATLAB TOOLBOX ←

در مطلب دستور newline یک شبکه عصبی از نوع آدائین ایجاد می کند

ورودی ها $\leftarrow$ $\min$ و $\max$ ورودی ها $\leftarrow PR$

S تعداد خروجی ها

LR به نرخ یادگیری

$LR \leftarrow$ (نرخ یادگیری)

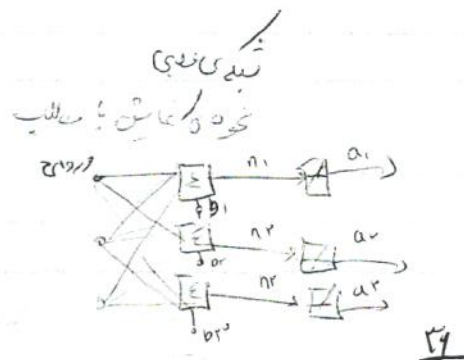
|| در جواب یک شبکه عصبی (لایه) ^{Linear} دهه ||

newlin (PR, S, ID, LR)
 اسم شبکه عصبی

1

$S = 2$ خروجی

$PR \leftarrow$ تعداد ورودی ها



(۲)

تپه عصبی (۱) Perceptron

این درختی برای دسته بندی استفاده می شود که کاربرد

نحوه کارش به شکل است ؟

این خط در فضا ترسیم می کند: $\frac{w_1x_1 + w_2x_2 + \dots + w_nx_n + bias}{b}$ خط است

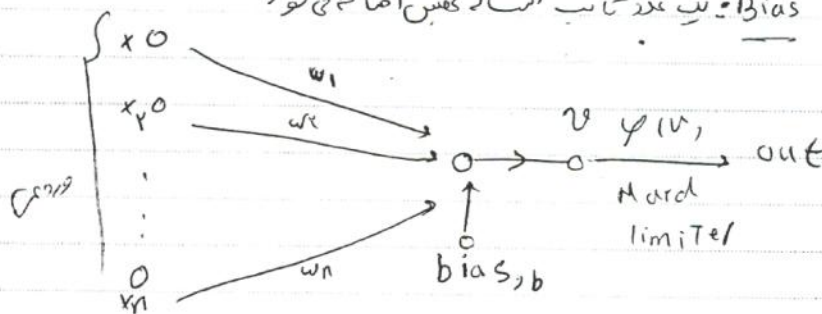
در فضا دوبندی می کند صفحه ای را

یعنی یک نرون کارش به شکل محاسبه فضا است

یک بهمت می شود دسته ای

و دیگر و دوم

Bias یک عدد ثابت است که بهش اضافه می شود

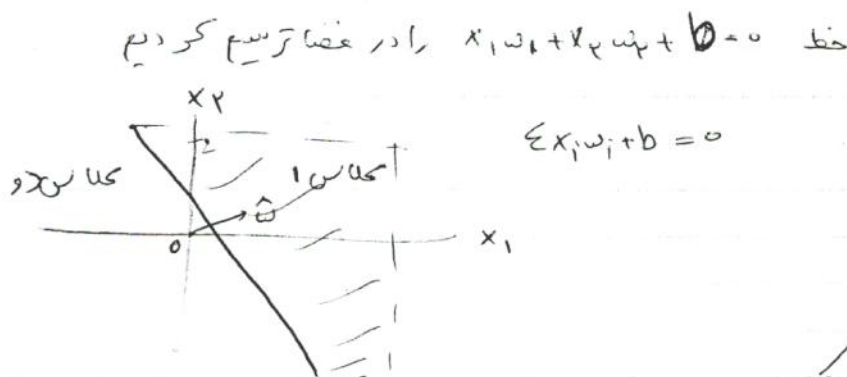
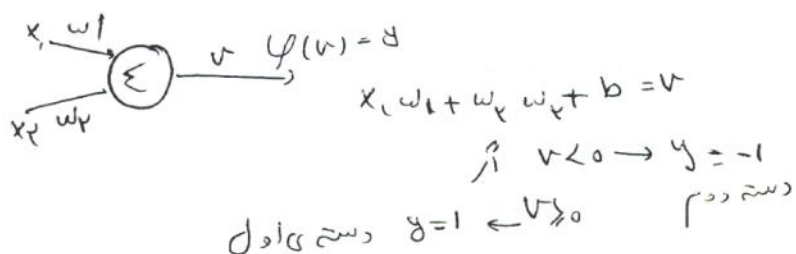
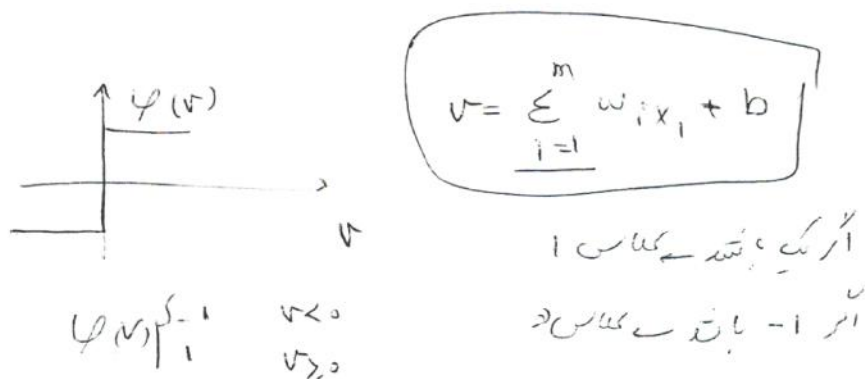


Hard limiter به خروجی یک است اگر ورودی از معنر ستر یا ستر ۱ - خروجی ۰ - در ورودی

در کار دین شود

© MICRO

ل



نکته!
 "خروجی یک است یا ۰ است نه فقط به عدد دسته بند می خورد"
 به سبب این

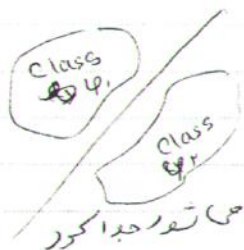
(۳)

Subject

Date

انتخاب وزن های پرسپترون

اگر ما وزن داشته باشیم و داده های ما به شکل یه دسته دسته باشند



خطی ترسیم کنیم و داده ها را تفکیک جدا کنیم



جدا نمی شود

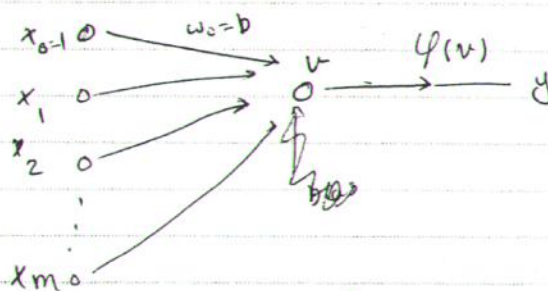
اگر ما به شکل یه ورودی بینیم و ورودی ها که x است همیشه یک باشد

$$v(n) = \sum_{i=0}^m w_i(n) x_i(n)$$

$m \leftarrow$ تعداد ورودی

ما می توانیم به شکل ماتریس بنویسیم

$$= w^T(n) x(n)$$



$$w_1 x_1 + w_2 x_2 + b x_0 = 0$$

$$\begin{pmatrix} w_1 & w_2 & w_3 \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = w_1 x_1 + w_2 x_2 + \dots$$

Theorems

چهار حالت برای شبکه عصبی پرسپترون وجود دارد

الف) سیدین نمی آید و زمانی که مثبت باشد به شطرنج به کلاس c_1 است و در ابتدا هم شطرنج است و وزن هم اینجا تغییر نمی کند.

$$① \quad w(n+1) = w(n) \text{ if } w^T x(n) > 0 \text{ and } x(n) \text{ belongs to class } c_1$$

مثال

$$② \quad w(n+1) = w(n) \text{ if } w^T x(n) \leq 0 \text{ and } x(n) \text{ belongs to class } c_2$$

↑ در این حالت اگر درست تشخیص داده باشد وزن ها تغییر نمی کند

ولی اگر ...
↓

$$③ \quad w(n+1) = w(n) - \eta(n) x(n) \text{ if } w^T x(n) > 0 \text{ and } x(n) \text{ belongs to class } c_2$$

$$④ \quad w(n+1) = w(n) + \eta(n) x(n) \text{ if } w^T x(n) \leq 0 \text{ and } x(n) \text{ belongs to class } c_1$$

نحوه عمل روش

۱- برای اصلاح خطاها در سطح داریم
- وزن اولیه را صفر در نظر بگیریم

$$\text{set } w(0) = 0$$

۲- ورودی به شبکه ی عصبی اعمال و خروجی را اندازه گیری کنیم بعد

(۳)

۳- معادله های یادگیری

Subject

Date

(۴)

$$y(n) = \text{sgn}(w(n)x(n))$$

$$w(n+1) = w(n) + \eta [d(n) - y(n)] x(n)$$

۴- به وزن ها را تنظیم می کنیم $d(n)$ $\rightarrow +1$ $\rightarrow$ class 1 $\rightarrow -1$ $\rightarrow$ class 2

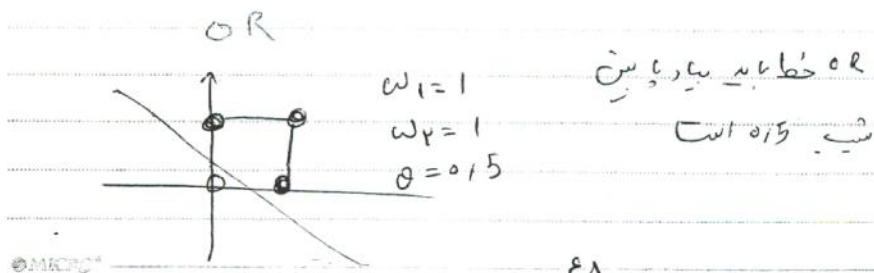
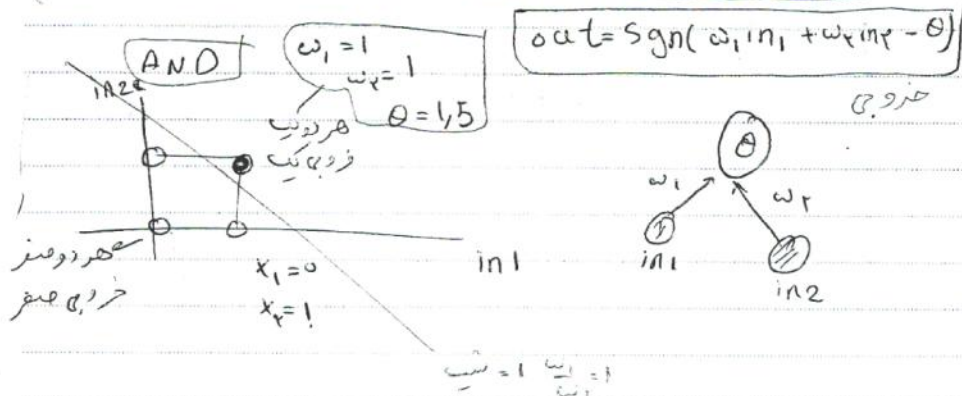
$$E(n) = d(n) - y(n) \text{ error}$$

$$w(n+1) = w(n) + \eta [d(n) - y(n)] x(n)$$

توضیح

۵- مقدار شرط ۴ برای شرط اول (همیشه)

معبر خط and or not در تدریس



برای یادگیری XOR برای شبکه های عصبی

برای بدست آوردن خط باید ورودی ها را در یک دایره قرار دهیم و

$$out = \text{sgn}(\sum_{i=1}^n w_i x_i - \theta)$$

in1 in2 out

0	0	0
0	1	0
1	0	0
1	1	1

$w_1 x_1 + w_2 x_2 - \theta < 0$ $w_1 x_1 + w_2 x_2 - \theta < 0$ $w_1 x_1 + w_2 x_2 - \theta < 0$ $w_1 x_1 + w_2 x_2 - \theta > 0$	$0 > 0$ $w_2 < \theta$ $w_1 < \theta$ $w_1 + w_2 > \theta$
--	---

این روش ها است در یادگیری

$$\sum_{i=1}^m w_i x_i + b = 0$$

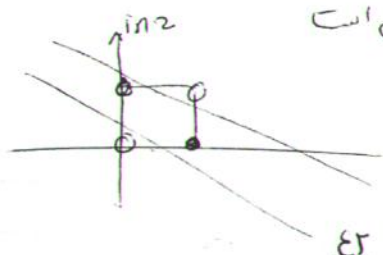
برای کدن دزن MSE فراموش
ریشه دوم ← MSE

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} * \begin{bmatrix} w_1 \\ w_2 \\ b \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} w_1 \\ w_2 \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

محدودیت شبکه پرسپترون

XOR ← محدودیت آن است

در R دو دسته است دو دسته 0 یک دسته اند دو دسته 1 و 0 یک دسته اند
البته برای هر فردی مشخص نیست ممکن است
دسته متفاوت بگیرند



راه حل شبکه ها چند لایه است

(5)

اگر پرسپترون تابع sgn باشد بصورت پله ای می شود

فرم یا 0 است یا 1 است ← فقط برای دسته بندی است

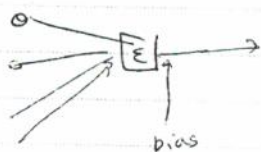


مدل Sigmoid می تواند مقادیری بین صفر و یک داشته باشد

می توان در سولساز از تن استفا ده کرد

نکته: پرسپترون در مطلب

ورودی ها R عدد دارند با هم جمع می شوند bias کسبه



افزای می شود

a ← مقدار لرون

در مطلب دستور new p این کار را انجام می دهد

$$\text{net} = \text{new p}(PR, S, TF, LP)$$

TF ← از نوع hard limiter ← پله ای است

LP ← طریق یادگیری ← Learn p ← از این کاشی است

۲۹

مثال

برنامه نویسی به تابع $\sin$ را مدل کنیم، ورودی x ، خروجی باشد

$\sin(x)$ (از روی داده آموزش اینجا داریم)

آنها یک خصوصیت دارند x

به یک خروجی دارند $\sin(x)$

→ یک سری داده آموزش از طریق

ماتریس را از ۰ تا ۲ π داده و ایجاد می کنند $x = 0:2:10$

mRile → تابعی که این را دارد که می تواند را بسوزم و نتایج او را

close all
clear all

$x = 0:2:10$ ؛ $y = \sin(x)$

داده ها در دسترس

$t = \sin(x)$

ماتریس های در دسترس

که می تواند ماتریس در دسترس

تعداد سرنها

$\text{plot}(t, 'ro')$ ؛ $\text{xlabel}(x)$ ؛ $\text{ylabel}(y)$

$\text{net} = \text{newff}(\text{minmax}(x), [5 1])$

↓ تعداد نرون در لایه اول

$\text{net.trainparam.epochs} = 1000$ ؛ $\text{net.trainparam.lr} = 0.1$

$\text{net} = \text{train}(\text{net}, x, t)$ ؛

۴۴

(۴)

Subject :

Date :

$$y_b = \text{sim}(\text{net}, x_i)$$

تورینگ

تبدیل عصبی Net روی آموزش ها و در فرجه های فاینال ده

hold on

$$\text{plot}(x, y_b, 'b^*')$$

$$\text{net} = \text{Train}(\text{net}, x_{st});$$

$$x_1 = 2 * \pi * \text{rand}(1, 20)$$

$$y = \text{sin}(\text{net}, x)$$

$$\text{plot}(x, y, 'r^*')$$

امکان مادرستیت - چون تبدیل عصبی در ردی را از قبل دید، خودی را می دهد.

برای مشخص قبل از * x_1 افغانه کنیم

(با تغییر کردن های تدریس خطاها را کاهش داد و آموزش بدیم تبدیل عصبی

مثلاً بیایم ببینیم 20 - 40 - 200

از کمی بقیه تبدیل عصبی ما حفظ کرده!

// اگر در performance خطا از میزان دهده ولی نمودار خطا نشان دهد

به این معنی که تبدیل عصبی ما حفظ کردن و یادگیری و overfitting رخ داده

چون ما تعداد ترون ها را زیادتر کنیم 200

MICRO

~~نویسندگان~~ ~~محققان~~

راه حل جلوگیری از over fitting چیست؟

باید براساس متوقف کنیم تا زمانی که نتواند

Neuff

برای تولید داده ها و نیز برای امتحان کردن باید سیستم

Close all

داده ها validation به صورت

Clear all

وقتی این داده ها را تولید کنیم می توانیم اندازه بگیریم و دارد

$x = 0:0.2:2 \times \pi$;

یا می تواند حفظ کند

$t = \sin(x)$;

$x_1 = 2 \times \pi \times \text{rand}(1, 20)$;

داده ها به صورت تصادفی کمتر از داده ها

$y_1 = \sin(x_1)$

validation است

$pt.p = x_1$;

بینی این می تواند به ندرت به صورت تصادفی داده ها به صورت تصادفی

$pt.T = y_1$;

بینی این می تواند به ندرت به صورت تصادفی داده ها به صورت تصادفی

$\text{plot}(x, t, 'ro')$;

$\text{net} = \text{newff}(\text{minmax}(x), [1 \ 4 \ 1])$;

$\text{net} = \text{trainparameter}(\text{epochs} = 100)$;

$y_b = \text{sim}(net, x)$

hold on

$\text{net} = \text{train}(net, x, t, [], [], pt)$;

$x_2 = 2 \times \pi \times \text{rand}(1, 20)$

۴۶

(1)

Subject :

Date

$y = \sin(\text{net}, x_2)$

$\text{plot}(x_2, y, 'r*')$

$y_r = \sin(\text{net}, x_1);$

$\text{plot}(x_1, y_r, 'b*');$

۲ دی

$y = \sin x_1 + \cos x_2$

close all

clear all

$x_1 = 0:0.2:2 \times \pi;$

$x_2 = 0:0.2:2 \times \pi;$

For $ii = 1: \text{length}(x_1)$

For $jj = 1: \text{length}(x_2)$

~~$x(ii, jj) = \sin(x_1(ii)) + \cos(x_2(jj))$~~

$x(ii, (jj-1) * \text{length}(x_1) + jj) = x_1(ii);$

$x(2, (ii-1) * \text{length}(x_1) + jj) = x_2(jj);$

~~$t = \sin(x_1(ii)) + \cos(x_2(jj))$~~

© MICRO

۴۷

```

t((ii-1)*length(x1)+jj)=sin(x1(ii))+cos(x2(jj))/2;
% برای t1(ii,jj)=(sin(x1(ii))+cos(x2(jj)))/2;
% این اشیا
% کردم
end
end

```

mesh(t1)

حالت آموزش دوم

```
net=newff(minmax(x),[4 1], 'tansig')
```

```
net=train(net,x,t)
```

```
y=sim(net,x)
```

```
for ii=1:length(x1)
```

```
for jj=1:length(x2)
```

```
ym(ii,jj)=y((ii-1)*length(x1)+jj);
```

```
end
```

```
end
```

```
figure(2);
```

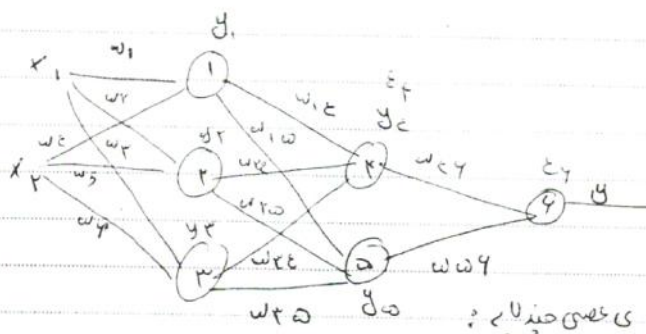
```
mesh(ym)
```

تا

Subject: شبکه های عصبی → تاریخ Date:

نیمه دوم خرداد ۱۳۹۱

مثال ۳



نمودار صورت شبکه ی عصبی خرد

در شبکه عصبی که برای $d = y - \hat{y}$ داریم



در این حالت برای آموزش به روش های اعم است

اما در خرد لایه های میانی هدف را نداریم که بر حساب باید از روش

بین اینها خطا استفا ده کنیم

ما به ϵ در خردی داریم که می توانیم ضریبش کنیم $\epsilon_4 = \epsilon \times w_{44}$ در تلفات قبل

$$\epsilon_5 = \epsilon_4 \times w_{54}$$

MICRO

$$E_1 = E_1 \times w_{11} + E_2 \times w_{12}$$

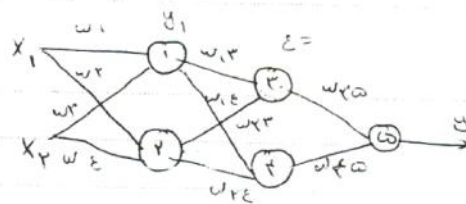
$$E_2 = E_1 \times w_{21} + E_2 \times w_{22}$$

$$E_3 = E_1 \times w_{31} + E_2 \times w_{32}$$

ا) حل کنید: فرض کنید این شبکه عصبی را با داده های جدول زیر آموزش دهید.

ع ج ا

مثال:



x_1	x_2	d
1	1	1
0.1	0	1
0.5	0.3	0.4

وزن اولیه های نرون ها را یک بگیرید. تابع غیر خطی نرون های یک

$$F(v) = \frac{1}{1 + e^{-v}} \quad \text{تجربا، ۱-۴}$$

تکلیف عصبی را با داده های جدول فوق و $\eta = 0.1$ آموزش دهید.



رون آ در نقطه صحت است. تابع غیر خطی است.

$$x_1 = 1, x_2 = 1$$

$$y_1 = \frac{1}{1 + e^{-2}}$$

$$v = x_1 w_{11} + x_2 w_{21} = 2$$

Subject :

Date

$$y_p = \frac{1}{1 - e^{-r}}$$

$$y_r = \left(\frac{1}{1 + e^{-r}} \right) + \left(\frac{1}{1 + e^{-r}} \right) = \frac{2}{1 + e^{-r}}$$

$$y_p = \frac{1}{1 + e^{-\left(\frac{2}{1 + e^{-r}} \right)}}$$

$$y_c = y_p$$

$$y = y_c = y_c \times 1 + y_r \times 1 = 2y_p = 1.7 \quad y_p = 0.17$$

$$d - y = \varepsilon \rightarrow 1.7 - 1 = 0.7 \rightarrow \varepsilon$$

(نکته: گاهی اوقات به دلیل این که تغییرات عدد میسر نیست، این را به صورت زیر می نویسند)



پس این وزن را به 0.1 تغییر می دهیم

$$w(n+1) = w(n) + \eta x^T \varepsilon(n)$$

آموختن لایه ها:
در لایه آخر وزن در آموزش 0.1 می دهیم

$$\begin{bmatrix} w_{10} \\ w_{20} \end{bmatrix} = \begin{bmatrix} w_{10} \\ w_{20} \end{bmatrix} - \eta x^T \varepsilon(n)$$

(1) (2) 0.1 × [1] × 0.7

میانگین میانگین میانگین

[0.07] [0.07]

$$\begin{bmatrix} w_{12} \\ w_{22} \end{bmatrix}_{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0.127 \\ 0.107 \end{bmatrix} = \begin{bmatrix} 0.873 \\ 0.893 \end{bmatrix}$$

$$y_2 = 0.18$$

$$\begin{bmatrix} w_{12} \\ w_{22} \end{bmatrix}_{(1)} = \begin{bmatrix} w_{12} \\ w_{22} \end{bmatrix}_{(0)} - \eta x^T \varepsilon^T F'(v) X$$

$$F'(v) = \frac{+e^{-v}}{(1+e^{-v})^2} \quad \begin{bmatrix} w_{12} \\ w_{22} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - 0.1 \times 0.7 \times 0.127 \times \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.9913 \\ 0.9913 \end{bmatrix}$$

$$F'(v) = \frac{+e^{-1.17}}{(1+e^{-1.17})^2} = 0.127$$

$$\varepsilon = 0.18 - 0.107 = 0.073$$

در مرحله بعد $x_1 = 0.18$ و $x_2 = 0$

در مرحله بعد $x_1 = 0$ و $x_2 = 0$

کتابخانه
شبکه های عصبی

ساعت دوم شبکه های عصبی

شبکه های عصبی در بدن

در علم دارو یک سری دروس و یک سری هدف داریم و مشخص است که

کدام دروس مربوط به کدام هدف است. دسته بندی را خوبان انجام بدهیم

مثال

دروس	بیم
دروس	سره

بدون معلم: مشخص کنید کدام بیماری و کدام سره است. بخش یکتا را در ده

را در دهت کن. را اختیار با شنبه که عصبی است که کدام بیمار را برکتیم. بدی

در این دسته بندی استخوان کف

در یک جدول داریم در یک طرف خصوصیات داده ها قرار دارد و طرف دیگر

تعداد در دهت را می بینیم.

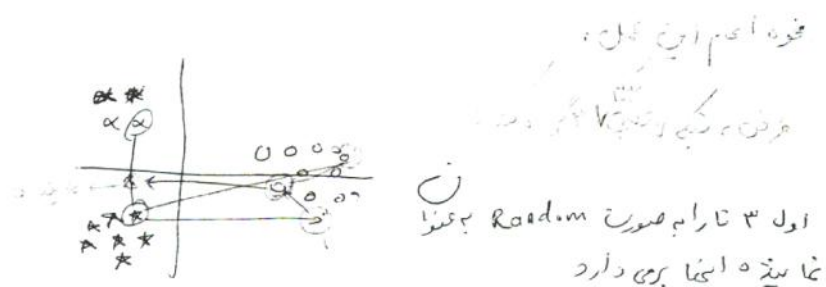
مثال بار هیوانات

دوربین	خزنده	آبی	پرند	متر	آمل	پرند	فولاد	فولاد
--------	-------	-----	------	-----	-----	------	-------	-------

رکت

MICRO

مدرسه مثل این است که اختیار را به تپه های عصبی داده ایم پس می آید دسته بندی می کنند
که ما انجام نداده ایم.



می باید بررسی می کنند مثلاً یک داده به فاصله اش با داده های دیگر حقیر است بعد می آید
این داده را با داده ای که کمترین فاصله دارد می گیرند. بعد می آید می بینیم را انتخاب می کنند
و وسط را می بینیم نماینده برده می گیرند باز برای گروه بندی می کنند و از نماینده می گیرند
که از نماینده جدید را می گیرند و از آنجا می بینیم این کار را ادامه می دهیم
تا نماینده داده را گرفته اند. این روش بدون معلم می باشد.

گیر کردن در
مسئله گزافه سازی. کلاسیم محلی بود

over Fitting چیست؟

Subject:

Date _____

فرین، ۳ نقطه آموزش برا Sin در نظر رفتم صلا ۳ عددی

داریم ۳ تا خودی داریم اگر با ۱۰ تا زن بجواییم آموزش دهیم

جاء انهم غوايا لشد ٣ تا نقطه ما را حفظي كند

حقوق نقدی کردن ها، ~~نقدی کردن~~ زیاده است. می باید نقاط را

که دادیم فقط خوب حفظ می‌کند ولی نقاط دیگر را یاد نمی‌گیرد و این روش
استیادمی که خود وسطا استیادی شود.

مثال برعکس به اندازه زدن ما بخلاف داده ها آموزش

Sine

الفبای دست نویس! $a \neq b$ را نشانی کن با طراحی یک

○	○	○	
○		○	
○		○	○

(a)

شکل عقی که وارد رشت - خروم رشت می شوند و یا

مکمل شدہ ہے

	2	0	0
0			0
0			0
		0	

33

