۵ـ برنامه‌ریزی برای ایجاد چندین مرکز تحقیقاتی بین‌المللی در شهرهای مختلف دانشگاهی که زمینهٔ بررسیهای علمی را دارند و پذیرش دانشجویان خارجی و دعوت از استادان و محققان کشورهای خارج برای تدریس و همکاری با مؤسسات علمی داخلی (دوره‌های کوتاه‌مدت) و در صورت پذیرش دانشجو، ایجاد ارتباط علمی با بسیاری از کشورها».

در استمرار بخشیدن به حرکت همگانی کردن علم در جهت توسعهٔ فرهنگ علمی و نقش تعیین کنندهٔ دانشگاهها در روند علمی کشور، پیشنهادهای لازم در جهت رفع موانع موجود بر سر راه ارتقاء سطح علمی دانشگاهها، به ترتیب موارد زیر مطرح شد:

۱ـ «در حال حاضر بار آموزشی فعالیتها و درگیریهای غیردانشگاهی اعضای هیأت علمی بسیار زیاد، وقت‌گیر و بعضاً کشنده است و این عوامل باعث شده است که روحیهٔ تحقیق در اکثر همکاران تا حد زیادی از بین برود. برای رفع این معضل موارد زیر پیشنهاد می‌شود:

ـ تأمین مایحتاج مادی کادر هیأت علمی به طور مستقیم و غیرمستقیم

ـ حذف پرداختهایی مانند حق‌التدریس و حق‌التحقیق

ـ در ترفیع پایه تجدید نظر شود (حدود ده درصد حقوق به ازای هر پایه ترفیع مناسب به نظر می‌رسد) و ترفیع بسته به میزان فعالیتهای دانشگاهی سالانه در محدودهٔ بین ۰ تا ۲ باشد.

۲ـ تجدید نظر در جایگاه وزارت فرهنگ و آموزش عالی و دانشگاهها

ـ جا دارد که وزارت فرهنگ (به جای درگیر شدن در مسائل جزیی و آیین‌نامه نویسی و ابلاغ آن) نقش ارشادی، ستادی، ارزیابی و هدایت دانشگاهها را به عهده بگیرد (در حال حاضر هیچ گونه سیستم ارزیابی مؤثری در امور آموزشی، پژوهشی و اجرایی در دانشگاهها و وزارتخانه فعال نیست).

ـ لازم است با تقویت هیأت امنای دانشگاهها، به آنها و مدیریتی که دارند در تمام امور استقلال بیشتری داده شود تا بتوانند از امکانات موجود منطقه‌ای خود به نحو مطلوب‌تری استفاده کنند. رهایی دانشگاهها از ضوابط و بورکراسی دولتی مسألهٔ مهم دیگری است

۳ـ ضرورت استفادهٔ بهینه از امکانات موجود: در این رابطه لازم است یک شبکهٔ اطلاع‌رسانی بین دانشگاهها و مراکز آموزشی ـ پژوهشی ایجاد شود که مجهز به بانک اطلاعاتی جامعی باشد. با استفاده از چنین شبکه و امکاناتی مانند Faxe-mail می‌توان مقاله یا سایر مدارک علمی را سریعاً مبادله کرد.

۴ـ ضرورت سرمایه‌گذاری ملی برای شکل‌گیری و تقویت بعضی از زمینه‌های تحقیقاتی راهبردی: به عنوان مثال در دانشگاهها، در زمینهٔ علوم زیردریا و بخش رادار، پتانسیل تحقیقاتی خوبی وجود دارد که نیاز به سرمایه‌گذاری ملی دارد».

شبکه‌های عصبی مصنوعی یك ابزار قوی برای تحلیل سازه‌های پرالمان

(قسمت دوم)
علی کاوه
استاد
حسین یوسف‌پور
کارشناس ارشد
دانشکدهٔ عمران
دانشگاه علم و صنعت

مقدمه

امروزه، شبکه‌های عصبی از سه دیدگاه مختلف بیولوژیکی، کامپیوتری و کاربردی مورد بررسی قرار می‌گیرد. بیولوژیستها با طرح نظریهٔ سلولی و بررسی نحوهٔ عملکرد سلولهای مغز اولین گام را در مورد شبکه‌های عصبی برداشته‌اند. پس از آن، مهندسان کامپیوتر این سلولها را برای کامپیوترهای خود مدلسازی کردند و در پی آنها، مهندسان فنی از جمله مهندسان سازه و مکانیک، شبکه‌های مزبور را برای کاربرد مقاصد خود مورد استفاده قرار دادند.

شبکه‌های عصبی بیولوژیکی و مصنوعی زمانی بیشتر جلب توجه می‌کند که بین لایه‌های سلولی تشکیل دهندهٔ شبکه، تقسیم کار صورت گرفته و هر لایه از سلولها وظیفهٔ مشخصی را به عهده می‌گیرند. سلولهای مربوط به لایهٔ ورودی اطلاعات را از محیط بیرون دریافت‌می‌کنند، سلولهای میانی روی اطلاعات ورودی پردازش انجام می‌دهند و لایهٔ خروجی، واکنش شبکه را اعلام می‌کند. ساختار ساختمانی سلولهای تشکیل دهندهٔ شبکه‌های عصبی موجب به وجود آمدن چهار سیستم کلی برای شبکه‌های عصبی می‌شود که عبارتند از:

۱- سیستم هب - ویدرو
۲- سیستم کوهونن
۳- سیستم انتسابی هاپفیلد
۴- سیستم مستقیم

هر کدام از سیستم‌های فوق، روشهای بخصوصی برای تشکیل ماتریس وزن و یادگیری شبکه بکار می‌گیرند که در بخش اول این مقاله، دو سیستم یادگیری مورد بررسی قرار گرفته است. چنانچه در هرکدام از سیستم‌های فوق شرایط ایده‌آل، برقرار باشد شبکه با یک بار تکرار آموزش خواهد دید. این شرایط ایده‌آل متعامد بودن فضای زوج‌الگوهای آموزشی است. ولی از آنجا که در اغلب مسائل این شرایط تأمین نمی‌شود لذا، یادگیری شبکه از طریق روشهای تقریبی انجام می‌گیرد و شبکه برای آموزش خود احتیاج به تکرار زوج‌الگوها دارد. همچنین دو روش یادگیری E.P و S.P نیز مطرح است که توضیحات کافی در این مورد در مثال ارائه شده است. تشخیص الگو و کلاسیفایر موضوع بحثی است که در آن شبکه‌های عصبی از دیدگاه منطق فازی مورد بررسی قرار می‌گیرد. برای آشنایی بیشتر با این مسأله، در این مقاله چند مثال ساده مورد بررسی قرار گرفته است.

روشهای یادگیری در شبکه‌های عصبی

بر اساس نوع یادگیری، تقسیم‌بندی زیر را می‌توان مورد توجه قرار داد:

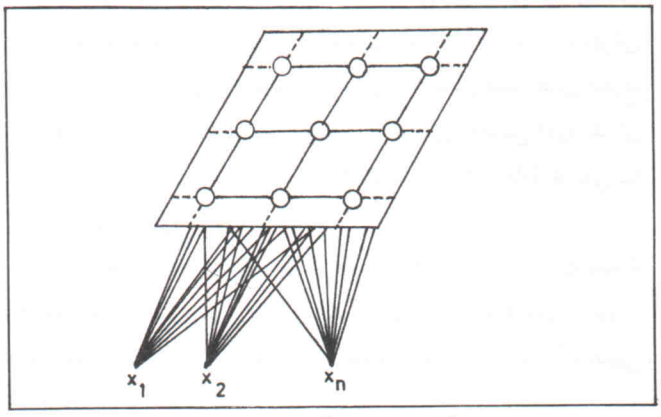
۱- یادگیری هم واقعه
۲- یادگیری بر اساس عملکرد
۳- یادگیری رقابتی
۴- یادگیری در شبکه‌های برگشتی

نرم‌افزارهای ارائه شده برای شبکه‌های عصبی، چیزی حدود ۱۶ سیستم را برای یادگیری ارائه کرده‌اند، که اساس همهٔ آنها بر پایهٔ ۴ سیستم اصلی فوق بنیاد نهاده شده‌اند که در قسمت اول این مقاله دو سیستم را مورد بررسی قرار داده‌ایم. حال، به طور اجمال سیستم سوم را مورد بررسی قرار می‌دهیم.

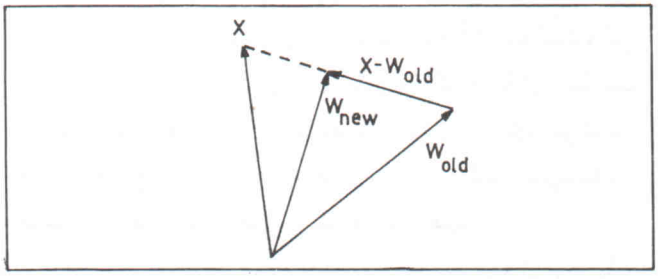
یادگیری رقابتی:

در این نوع یادگیری بر خلاف روشهای قبلی، فقط وزن اتصال چند پردازنده تغییر می‌کند. در این شیوهٔ از یادگیری، یک رقابت جمعی بین پردازنده‌های عصبی به وجود آمده و پردازنده‌ای که در این رقابت برنده می‌شود به اصلاح وزنهایش می‌پردازد. در شکل ۱، شبکهٔ عصبی که به لایهٔ کوهونن موسوم است نشان داده شده‌است.

شکل ۱، یک لایهٔ کوهونن



شکل ۲، تعبیر هندسی کوهونن



این لایه از n پردازنده تشکیل شده است و هریک از این پردازنده‌ها از واحد ورودی انشعاب دهندهٔ، $(x_1,x_2,...,x_n)$ تغذیه می‌شود. پس‌از اعمال ورودی هر پردازنده، تابع شدت ورودی خود را طبق روابط زیر محاسبه می‌کند.

$$I_i=D(W_i,X)$$
$$W_i=(w_{i1},w_{i2},...,w_{in})^T$$
$$X=(x_1,x_2,...,x_n)^T$$

در این رابطه، $D(W_i,X)$ تابعی است که فاصلهٔ بین دو بردار W_i,X را اندازه‌گیری می‌کند. برای فاصلهٔ دو بردار تعاریف زیر در نظر گرفته می‌شود:

فاصلهٔ اقلیدوسی با رابطهٔ $D(W_i,X)=|W_i-X|$

فاصلهٔ قوس کروی بارابطهٔ $D(W_i,X)=1-W_iX$

در تعاریف زیر،برای فاصلهٔ دو بردار از فاصلهٔ اقلیدوسی استفاده می‌شود. پس از محاسبهٔ تمام I_i هاتوسط پردازنده‌ها،رقابت بین آنها شروع و پردازنده‌ای که کوچکترین I_i را داشته باشد به عنوان برنده انتخاب می‌شود، زیرا بردار وزن این پردازنده (W_i) نزدیکترین بردار وزن به بردار ورودی (X) است. 1 پس از تعیین پردازندهٔ برنده، خروجی بردار

اصلاح کننده برابر و برای بقیهٔ پردازنده‌ها برابر 0 خواهد بود. در این مرحله، قانون یادگیری برای اصلاح وزنها به صورت زیر اعمال می‌شود:

$$W_{i,new}=W_{i,old}+j(t)\times[X-W_{i,old}]\times Z_i$$

Z_i: خروجی پردازندهٔ iام

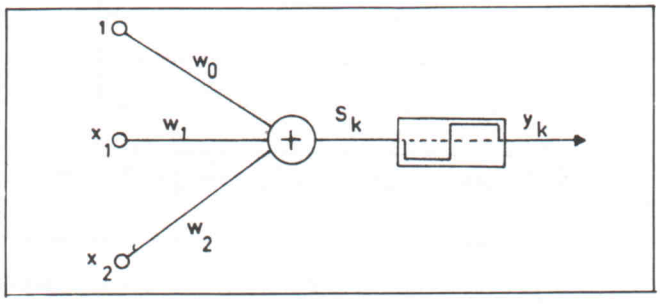
$j(t)$: ضریب زمانی برای تربیت شبکه

مقدار ضریب زمانی معمولاً در ابتدای تربیت شبکه برابر 0.8 و با رسیدن وزنهای شبکه به حالت بهینه به مقدار 0.1 و یا کمتر کاهش می‌یابد.

رده‌بند خطی:

شکل زیر، ساختار یک شبکه موسوم به پرسپترون را نشان می‌دهد:

شکل ۳، ساختار یک پرسپترون



$$S_k=w_0+w_1x_1+w_2x_2 \quad \begin{cases} \text{if } S_k>0 \rightarrow Y_k=1 \\ \text{if } S_k<0 \rightarrow Y_k=0 \end{cases}$$

موارد زیر را می‌توان در نظر گرفت:

۱- پرسپترون فضای دو بعدی (X_1,X_2) را به دو ناحیه با معادلهٔ زیر تبدیل می‌کند:

$$w_0+w_1x_1+w_2x_2=0 \rightarrow x_2=\frac{w_1}{w_2}x_1-\frac{w_0}{w_2}$$

از دیدگاه تشخیص الگو این دو ناحیه متناظر با دو کلاس مختلف در فضای تغییر است.

۲- استفاده از یک نرون با خروجی "1" که موسوم به تحریک مستقل یا بایاس (bais) ضروری است، چرا که در صورت عدم استفاده از آن، مرز جدا کنندهٔ الگوها از مبداء مختصات می‌گذرد و حل مسأله را غیر ممکن می‌سازد.

۳- استفاده از تحریک مستقل، امکان قرار دادن مرز جداگانه را در هر ناحیه از فضا (با هر ضریب زاویه و هر عرض‌از مبدأ دلخواه) فراهم می‌کند.

۴- ضریب پرسپترون یا پارامترهای مرزهای جدا کننده در طی فرآیند تربیت شبکه حاصل می‌شود.

۵- در صورتی که بُعد فضای نمایش افزایش یابد، مرز جداکننده یک خط نخواهد بود. در فضای سه بعدی، این مرز یک صفحه و در فضاهایی با بعد بیشتر، یک فوق صفحه است. در این فضاها، مسائل تفکیک‌پذیر خطی به آن دسته از مسائل گفته می‌شود که الگوهای مربوط به دو کلاس مختلف توسط یک فوق صفحه قابل تفکیک باشد.

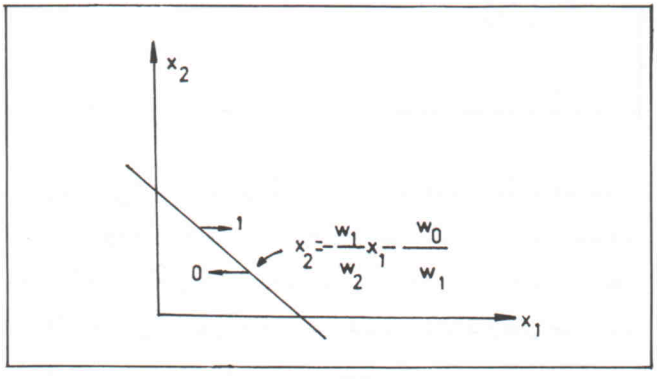
برای نمونه، معادلهٔ فوق صفحهٔ تولیدی توسط پرسپترون در فضای n بعدی به صورت زیر خواهد بود:

$$w_0+w_1x_1+w_2x_2+\ldots+w_nx_n=0$$

۶- عمده‌ترین ضعف پرسپترون، عدم توانایی این ساختارها در حل مسائلی است که الگوهای آن به صورت خطی نباشد. برای حل اینگونه مسائل باید از رده‌بندهای غیر خطی استفاده کرد.

مثال ۱: یک نمونه از مسائل رده‌بندی، تابع "AND" منطقی است. در جدول زیر، مؤلفهٔ اول هر زوج یک الگوی دو بعدی و مؤلفهٔ دوم کلاس مربوط به الگوهاست.

شکل ۴، جداسازی الگوهای تابع "AND" به کمک توابع خطی



x_1	x_2	x_1 & x_2
0	0	0
0	1	0
1	0	0
1	1	1

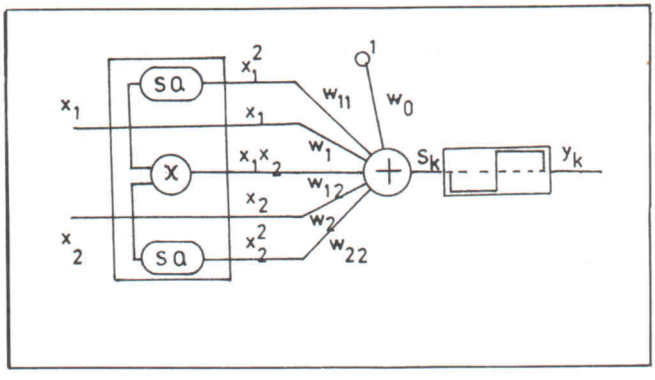
جدول ۱، زوج الگوهای آموزشی برای تابع منطقی "AND"

رده بند غیر خطی:

در یک نوع از این رده‌بندها، الگوی ورودی قبل از اعمال به رده‌بند خطی تبدیل می‌شود. این تبدیل توسط یک شبکهٔ پیش‌پردازنده، تحت تبدیلات غیر خطی قرار می‌گیرد. یکی از این پیش‌پردازش کننده‌ها توابع مبین چندجمله‌ای (Polynomial discriminant function) است. در نوع دیگری از این رده‌بندها، از شبکهٔ چندلایه استفاده می‌شود.

مثال ۲: در شکل ۵ یک شبکهٔ پیش‌پردازنده که به صورت سری با یک پرسپترون قرار دارد، نشان داده شده‌است. ورودی شبکه پیش‌پردازندهٔ مؤلفه‌های الگو و خروجی آن خود مؤلفه‌های الگو و ترکیبات غیرخطی آنها از قبیل مربع مؤلفه‌ها و حاصل‌ضرب مؤلفه‌ها هستند.

شکل ۵، یک شبکهٔ پیش پردازنده سری با پرسپترون

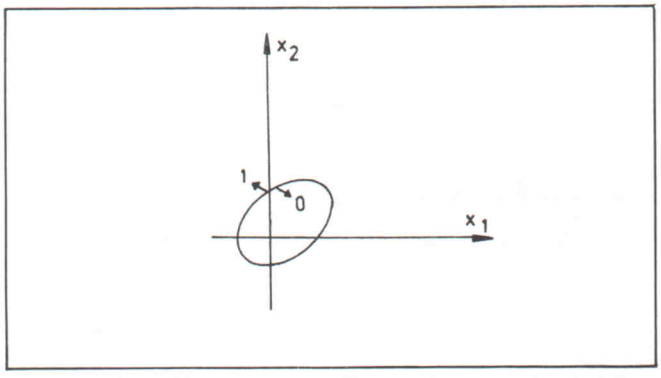


در صورتی‌که بردار ستونی X به شکل $X=[x_1,x_2,1]^T$ تعریف شود، که در آن مؤلفهٔ اول همان مؤلفه‌های الگوست، آنگاه نقش شبکهٔ پیش‌پردازنده، حاصل ضرب ماتریسی XX^T و خروجیهای شبکه مؤلفه‌های متمایز ماتریس حاصل خواهدبود. معادلهٔ مرزهای جداکننده برای این قبیل رده‌بندی به صورت زیر است:

$$w_{11}x_1+w_{12}x_2x_2+w_{22}x_2+w_1x_1+w_2x_2+w_0=0$$

از رابطهٔ فوق چنین نتیجه می‌شود که مرزهای جداکننده در فضای دو بعدی دیگر خط نیست بلکه یک مقطع مخروطی است. حال، با انتخاب مناسب وزنهای پرسپترون می‌توان مرزهای جداکننده را به شکل مطلوب برای جداسازی الگوهای مختلف درآورد. با تربیت شبکه روی زوجهای موجود با روش (کمترین مقدار مربعات) می‌توان به این وزنها دست یافت.

شکل ۶، جداسازی الگوریتم‌های "XOR" به کمک توابع چندجمله‌ای



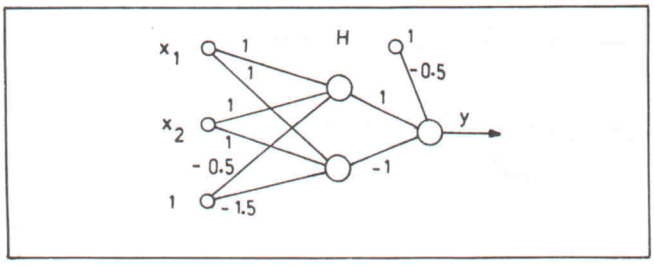
x_1	x_2	x_1"XOR"x_2
0	0	0
0	1	1
1	0	1
1	1	0

جدول ۲، زوج الگوهای آموزشی برای تابع منطقی "XOR"

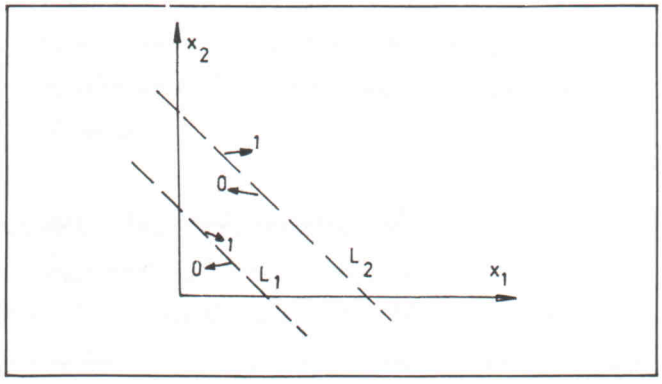
به عمل توابع چند جمله‌ای روی مؤلفه‌های الگوی ورودی، تقویت نمایش الگوها نیز می‌گویند. در مثال فوق، این عمل معادل با انتقال الگوها از فضای نمایش اولیه دو بعدی به الگوهای متناظر در فضای نمایش پنج بعدی است. به گونه‌ای که، الگوهای متناظر در این فضای جدید به کمک یک فوق صفحه قابل تفکیک است.

مثال ۳: با استفاده از ۳ پرسپترون که مطابق شکل در دو لایه قرار می‌گیرد، می‌توان تابع را اثبات کرد. در این شکل، تابع جمع‌کننده و محدودکنندهٔ سخت درون دایره‌ها فرض می‌شود.

شکل ۷، تحقق تابع "XOR" با استفاده از دو لایه پرسپترون



شکل ۸، خطوط مرزی تولید شده توسط پرسپترونهای لایهٔ اول



هرکدام از پرسپترونهای لایهٔ اول، معادلهٔ یک مرز جداکننده را در فضای الگو تعیین می‌کند که در این حالت عبارتند از:

$$L_1: x_1+x_2-0.5=0$$
$$L_2: x_1+x_2-1.5=0$$

این خطوط که در شکل نیز نشان داده شده‌اند، خروجی دو پرسپترون لایهٔ اول را مشخص می‌کند و هر الگوی واقع در فضای نمایش در یک طرف این دو خط قرار می‌گیرند. در مرحلهٔ بعدی این اطلاعات در اختیار پرسپترون لایهٔ دوم قرار می‌گیرند، تا با توجه به موقعیت هر الگو نسبت به دو خط، کلاس مربوط را مشخص کند. مقدار خروجی هر کدام از پرسپترونهای لایهٔ اول در ستون دوم جدول زیر به ازاء الگوهای مختلف ورودی در جدول زیر ارائه شده همچنین، خروجی لایهٔ دوم در ستون سوم آمده‌است. وزنهای شبکهٔ بالا تنها وزنهای ممکن "XOR" نیستند،بلکه وزنهای دیگری را نیز می‌توان پیدا کرد که مسأله "XOR" را حل کند. این وزنها در طی فرآیند تربیت شبکه روی ۴ زوج موجود برای تابع "XOR" حاصل می‌شود.

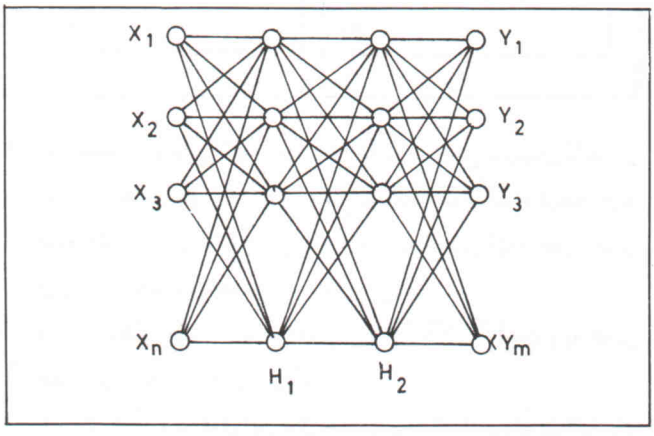
X_1	X_2	L_1	L_2	C
0	0	0	0	0
0	1	1	0	1
1	0	1	0	1
1	1	1	1	0

جدول ۳، مراحل آموزش شبکه با در نظر گرفتن دو لایه پرسپترون

شبکه‌های عصبی چند لایه به عنوان رده‌بند:

استفاده از شبکه‌های عصبی یک لایه در تشخیص الگو به مسائل تفکیک‌پذیر خطی محدود می‌شوند. در مواردی که الگوهای مربوط به دو کلاس خروجی تفکیک‌پذیر خطی نباشند یا تعداد کلاسهای فضای تفسیر بیشتر از ۲ باشد، استفاده از شبکه‌های عصبی چندین لایه امری ضروری خواهد بود. فرض کنیم که تابع انتقال نرونها محدود کنندهٔ سخت باشد و در تمام لایه‌ها خروجی یک نرون تحریک مستقل با خروجی ثابت وجود داشته باشد، حال اگر n نرون ورودی مقادیر حقیقی به خود بگیرند، آنگاه الگوهای ورودی متناظر با بردارهای واقع در فضای R_n خواهد بود.

شکل ۹، یک شبکهٔ عصبی مستقیم با دو لایهٔ پنهان



هریک از نرونهای لایهٔ پنهان اول یک فوق صفحه را در فضای R_n مشخص می‌کند. هر فوق صفحه فضا را به دو نیم فضا تقسیم می‌کند و خروجی هر نرون در این لایه تعلق الگوی ورودی را به یکی از دو نیم فضای تولید شده توسط فوق صفحهٔ مربوط به آن نرون مشخص می‌کند.

خروجی نرونهای لایهٔ پنهان اول به نرونهای لایهٔ پنهان دوم اعمال می‌شود. از لحاظ هندسی، وظیفهٔ سلولهای لایهٔ پنهان دوم، ترکیب فوق صفحات تولیدی توسط نرونهای لایهٔ پنهان اول و در نتیجه تولید فوق حجمهایی در فضای نمایش الگوی R_n است.

۱- فوق حجمهایی که توسط نرونهای این لایه تولید می‌شوند از نوع محدب هستند. این به این معنی که هر دو نقطه متعلق به فوق حجم را می‌توان توسط یک خط مستقیم به هم وصل کرد بدون‌آنکه مرزهای فوق حجم قطع شوند.

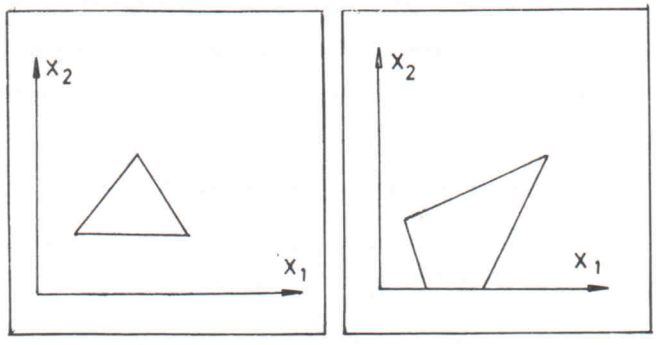
۲- فوق حجمهای تولید شده می‌توانند از نوع باز یا بسته باشند. فوق حجم بسته از هر طرف محصور به یک فوق صفحه خواهد بود. برای تشکیل یک فوق حجم بسته در فضای R_n به $(n+1)$ فوق صفحه نیاز خواهد بود که هیچ دوتای آنها موازی نباشند. بنابراین، برای تشکیل فوق حجمهای بسته،$(n+1)$ نرون در لایهٔ پنهان اول لازم

خواهد بود.

۳- فوق حجمهای فضای دو بعدی مساحت یا ناحیهٔ و فوق حجمهای فضای سه بعدی حجم هستند. در شکل زیر، نواحی باز و بسته توسط نرونهای لایهٔ پنهان دوم به وجود آمده‌اند.

شکل ۱۰، فوق حجم یا ناحیهٔ تولیدی باز در فضای دو بعدی

شکل ۱۱، فوق حجم یا ناحیهٔ تولیدی بسته در فضای دو بعدی



۴- در حقیقت،نرونهای لایهٔ دوم از اطلاعات خروجی نرونهای لایهٔ اول در مورد تعلق الگوی ورودی به نیم فضاهای تولید شده توسط فوق صفحات لایهٔ اول استفاده کرده و با ترکیب منطقی آنها، تعلق الگوی ورودی به یک فوق حجم را مشخص می‌کند.

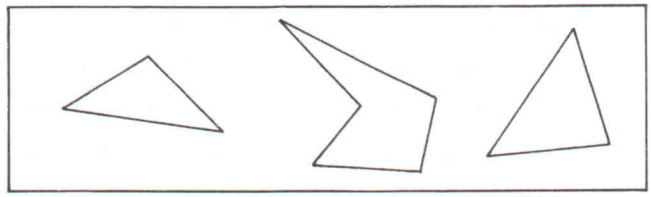
۵- ترکیب منطقی مورد استفاده اغلب از نوع "AND" است و به همین جهت این لایه را لایهٔ "و" می‌گویند.

۶- حداکثر تعداد وجه‌های این فوق حجمها برابر تعداد نرونهای لایهٔ اول پنهان است.

۷- خروجی نرونهای لایهٔ دوم پنهان به نرونهای لایهٔ خروجی اعمال می‌شود و بنابر این، نرونهای لایهٔ دوم پنهان با توجه به فوق حجمهای تولیدی وارد نرونهای خروجی شده و با آن ترکیب می‌شوند که از ترکیب فوق حجمها، می‌توان هر فوق حجم دلخواهی را تشکیل داد.

۸- فوق حجم حاصل لزوماً محدب نبوده و می‌توان از چند فوق حجم مجزا، متداخل یا متقاطع تشکیل شود. به عبارت دیگر، لایهٔ قبل در مورد تعلق الگوی ورودی به فوق حجمهای حاصل ویا ترکیب منطقی آنها تعلق الگوی ورودی را به ترکیبی از فوق حجمها مشخص می‌کند.

شکل ۱۲، نمونه‌هایی از ترکیب فوق حجمها یا نواحی توسط نرونهای لایهٔ خروجی



این ترکیب منطقی از نوع "OR" بوده و به لایه، لایهٔ "یا" گفته می‌شود.

۹- با توجه به اینکه با یک شبکهٔ چهار لایه که دارای تعداد مناسبی نرون در هر لایه است می‌توان هر نوع فوق حجم دلخواهی را به وجود آورد در مسألهٔ تشخیص الگو، در هیچ موردی بیشتر از دو لایهٔ پنهان مورد نیاز نیست.

روشهای کاهش احتمال کمینه‌های محلی:

یکی از مشکلاتی که ممکن است در تربیت شبکه‌های عصبی چند لایه پیدا شود کمینه‌های محلی (Local Minima) تابع متوسط مربعات خطاست. بردار وزن متناظر با این کمینه‌های محلی وزن بهینهٔ سیستم نخواهد بود زیرا وزنهای دیگری وجود دارند که به ازای آنها مقدار خطای سیستم کوچکتر خواهد شد. حرکت از کمینه‌های محلی در هر جهت باعث افزایش مقدار خطا خواهد شد. برای کاهش احتمال وقوع این کمینه‌ها، از روشهای زیر می‌توان استفاده کرد:

۱- **کاهش تدریجی میزان یادگیری:** چنانچه میزان یادگیری در ابتدا بزرگ انتخاب شود، تغییرات وزن بزرگی را برای دستیابی به کمینهٔ قطعی تابع خواهد داشت. پس از اینکه خطای سیستم به اندازهٔ لازم کوچک شد و در حوالی جواب بهینه قرار گرفتیم، میزان یادگیری را کاهش داده و به سمت جواب بهینه نزدیک می‌شویم. کمینه‌های محلی معمولاً خیلی به کمینهٔ قطعی نزدیک نیستند و بنابراین، روش فوق‌الذکر برای فرار از کمینه‌های محلی مناسب است.

۲- **افزایش نرونهای لایهٔ پنهان:** هنگامی که شبکه توانایی کافی برای تفکیک کلاسهای مختلف را نداشته باشد به گونه‌ای که کلاسهای مختلف را به عنوان یک کلاس در نظر بگیرد، در این صورت به یک کمینهٔ محلی خواهد رسید. در این حالت، برای بالا بردن توانایی شبکه در ایجاد مرزهای جداکنندهٔ مطلوب، مقدار نرونهای لایهٔ پنهان را باید افزایش داد.

۳- **استفاده از جملهٔ ممنتم:** در روش، انتشار خطا به عقب (BP) با اضافه کردن جملهٔ ممنتوم خواهیم داشت:

$$\Delta W_{ij}(n) = \alpha \delta_i y_j + \eta \Delta W_{ij}(n-1)$$

که در این رابطه η : ضریب ممنتوم، و $\Delta W_{ij}(n-1)$: میزان اصلاح وزن و W_{ij} مقدار وزن در یک مرحله است. معمولاً η را در حدود ۰/۹ انتخاب می‌کنند.

تربیت اضافی در شبکه‌ها:

تربیت اضافی تکرار بیش‌از زوجهای آموزشی بوده و عقیده بر آن است که تربیت اضافی شبکه‌ها که اصطلاحاً overtraining نامیده می‌شود، قابلیت تعمیم شبکه راکاهش خواهد داد و لذا، بایداز تربیت بیش از اندازهٔ شبکه پرهیز کرد. به عبارت دیگر، وقتی یک شبکه نسبتاً تربیت و به مقدار نسبی V دست یافت، با ادامه دادن آموزش شبکه مقدار V کاهش ولی در عوض مقدار V_G افزایش می‌یابد.

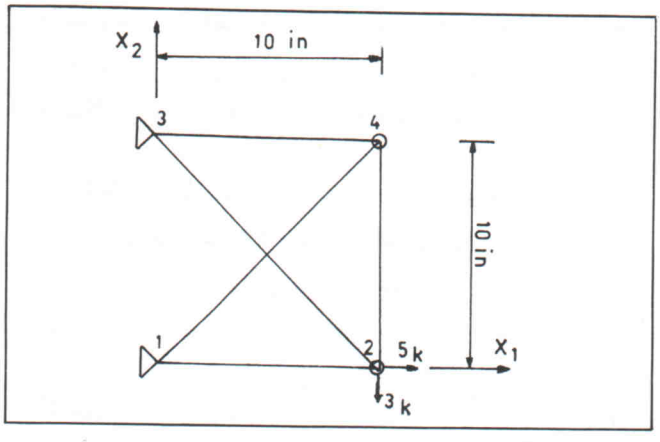
مثال ۴: یک خرپا متشکل از پنج عضو مطابق شکل ۱۳ در نظر می‌گیریم، اگر در این خرپا:

VOL:بیان کنندهٔ کمترین مقدار عددی برای ارضاع حالت پایداری باشد با در نظر گرفتن مختصات مربوط به X_1,X_2می‌توان تابعی به صورت زیر برای گره ۲ تعریف کرد:

$$VOL = fun(X_1,X_2)$$

برای دسترسی به اطلاعات مورد نیاز یک شبکهٔ ۳۱×۳۱ در نظر گرفته شده واز این اطلاعات برای تقریب چندجمله‌ایها و آموزش شبکه استفاده و نتایج حاصل از این عملیات در شکل ۱۴ نشان داده شده‌است. حالتهایی که برای تقریب مد نظر بوده عبارتند از:

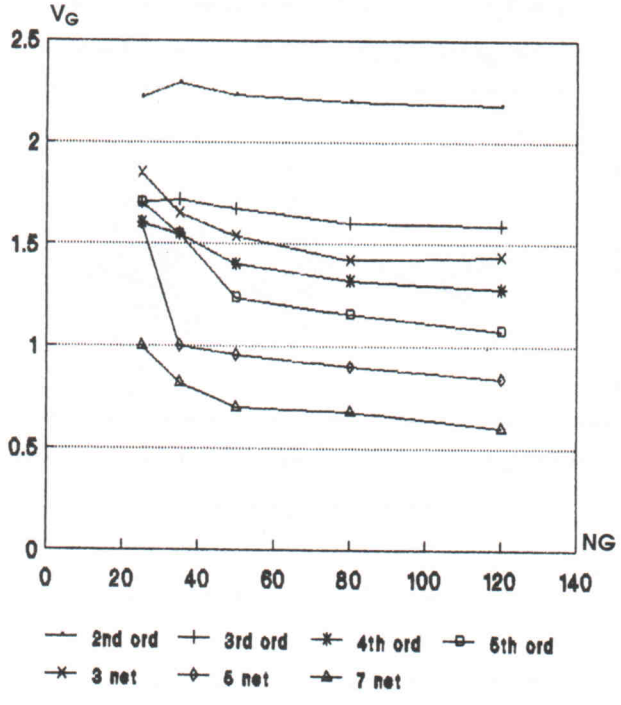
شکل ۱۳، خرپای متشکل از ۵ عضو



۱- استفاده از چندجمله‌ایها: چهار چند جمله‌ای با مراتب ۲، ۳، ۴ و ۵ مد نظر بوده و بدیهی است که برای هر کدام از چند جمله‌ایها تعداد مشخصی پارامتر نامعینی وجود دارد. به عنوان نمونه، برای چندجمله‌ای مرتبهٔ دوم مطابق زیر با 6 پارامتر سروکار خواهیم داشت:

$$VOL=AX^2_1+BX^2_2+CX_1X_2+DX_1+EX_2+F$$

شکل ۱۴، نتایج حاصل از تقریب معادلهٔ پایداری با استفاده از چند جمله‌ایها و شبکه‌های عصبی



نتایج حاصل از بررسی‌های خرپای پنج عضوی توسط شبکه‌های عصبی و چندجمله‌ایها

روشن است که با افزایش مراتب چندجمله‌ای تعداد پارامترهای نامعین و تقریبهای حاصل افزایش خواهد یافت. لذا مقدار عددی کاهش می‌یابد. V_G یک پارامتر بی‌بعد است که مطابق فرمول زیر مقدار خطا را از روش کمترین مربعات به ما خواهد داد:

$$V_G=\frac{\sqrt{\sum_{1}^{NG}(y_i-\hat{y}_i)^2/NG}}{\bar{y}_G}*100$$

$$V=\frac{\sqrt{\sum_{1}^{N}(y_i-\hat{y}_i)^2}}{\bar{y}}*100$$

y_i: مقدار به دست آمده از تقریب

y_i: مقدار منتظره

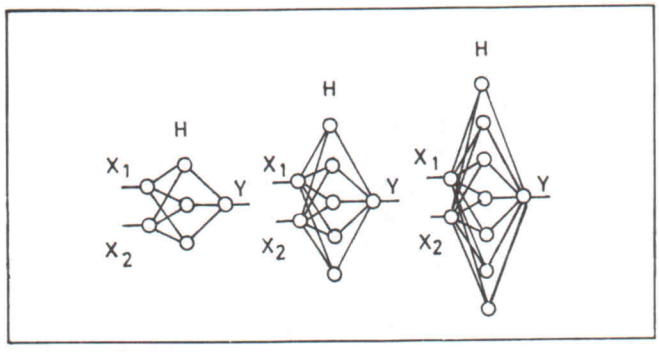
NG: تعداد الگوهای در نظر گرفته شد

y_G: مقدار متوسط برای الگوی مورد نظر

۲- استفاده از روش شبکه‌های عصبی: با در نظر گرفتن سه نوع شبکهٔ عصبی مسأله بررسی و نتایج حاصل در شکل ۱۴ نشان داده شده‌است. این شبکه‌های عصبی دارای گره‌های متفاوتی در لایهٔ

مخفی خود می‌باشند که جزئیات مربوطه در شکل ۱۵ نشان داده شده‌است.

شکل ۱۵، شبکهٔ عصبی با تعداد ۳، ۵، و ۷ گره در لایهٔ مخفی که بترتیب مترادف با ۱۳، ۲۱ و ۲۹ پارامتر نامعینی است.



۳- **نتیجه گیری:** دقت هرکدام از روشهای ارائه شده بستگی به مقدار پارامتر نامعینی آن دارد و مستقل از روش انتخابی است.

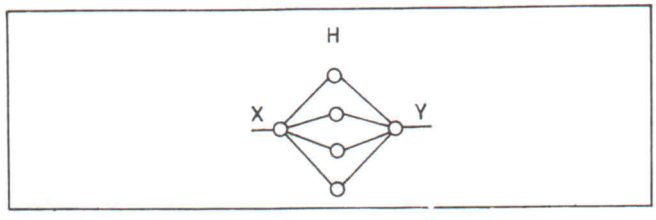
مثال ۵: تابع یک متغیره با ضابطهٔ زیر را در نظر می‌گیریم:

$$y = 2x + \sin(\pi x) + \sin(2\pi x) \quad , x\in[0,1]$$

شبکهٔ عصبی همانند آن برای مدلسازی رفتار تابع در دو حالت بررسی شده‌است:

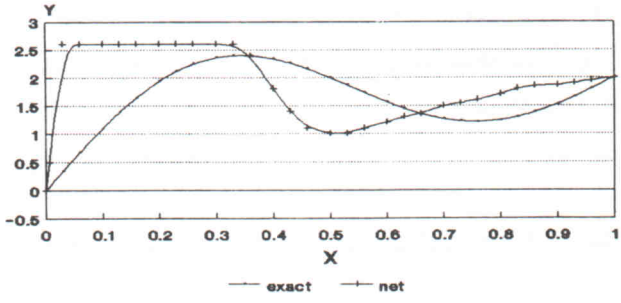
۱- **چهار گره در لایهٔ مخفی:** برای این حالت، پارامتر نامعینی شبکه برابر ۱۳ بوده که جزئیات آن در شکل ۱۶ نشان داده شده است.

شکل ۱۶، یک شبکهٔ عصبی با ۴ گره در لایهٔ مخفی



تعداد ۴ زوج‌الگوی آموزشی برای تربیت شبکه انتخاب شده است. ولی به همان ترتیب که در شکل ۱۷ نیز مشخص است، واکنش شبکه دستخوش تغییرات زیادی بوده و به غیر از نقاط تربیتی این تغییرات قابل توجه است. برای توجیه علت آن می‌توان گفت که تعداد پارامتر نامعینی ۱۳ است در حالی که تعداد زوج‌الگوهای آموزشی ۴ عدد انتخاب شده‌است و بالطبع این اختلاف یک حالت ناپایداری را به شبکه تحمیل خواهد کرد.

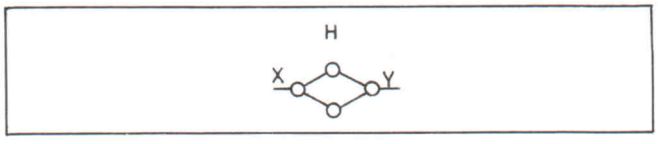
شکل ۱۷، مقایسهٔ نتایج حاصل از واکنش شبکهٔ عصبی با در نظر گرفتن ۴ گره در لایهٔ مخفی و ۴ زوج الگوی آموزشی با مقادیر حقیقی



نتایج حاصل با در نظر گرفتن تابع یک بعدی که در آن تعداد گره‌های لایهٔ مخفی برابر ۴ و تعداد زوج الگوی آموزشی آموزش شبکه برابر ۴ می‌باشد

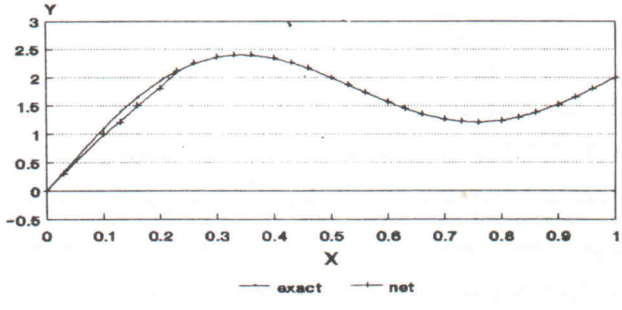
۲- **دو گره در لایهٔ مخفی:** برای این حالت، پارامتر نامعینی شبکه برابر ۷ بوده که جزئیات آن در شکل ۱۸ نشان داده شده است.

شکل ۱۸، یک شبکهٔ عصبی با ۲ گره در لایهٔ مخفی



در این حالت ۷ زوج الگوی آموزشی تربیت شبکه را بر عهده دارند. برخلاف حالتهای قبلی، واکنش شبکه با مقادیر حقیقی همخوانی دارد و این به دلیل انتخاب صحیح تعداد زوج‌الگوها است. در شکل ۱۹ واکنش شبکه بامقادیر حقیقی مقایسه شده‌است.

شکل ۱۹، مقایسهٔ نتایج حاصل از واکنش شبکهٔ عصبی با در نظر گرفتن ۲ گره در لایهٔ مخفی و ۷ زوج‌الگوی آموزشی با مقادیر حقیقی



نتایج حاصل با در نظر گرفتن تابع یک بعدی که در آن تعداد گره‌های لایهٔ مخفی برابر ۲ و تعداد زوج الگوی آموزشی آموزش شبکه برابر ۷ می‌باشد

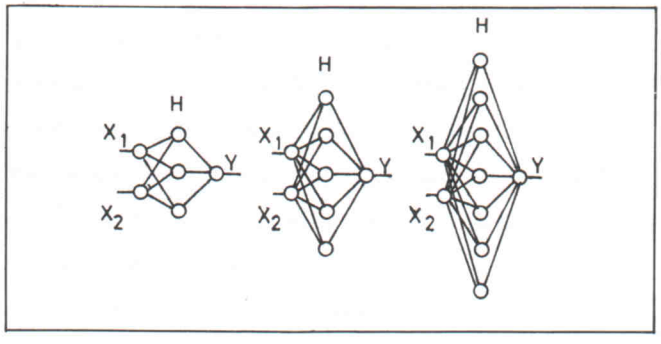
۳- **نتیجه گیری:** در یک شبکه باید تعداد زوج‌الگوهای آموزشی حداقل به تعداد پارامترهای نامعینی انتخاب شده باشد.

مثال ۶: تابعی دو متغیره با ضابطهٔ زیر را مد نظر قرار دهید:

$$y = 10x_1^4 - 20x_2x_1^2 + 10x_2^2 + x_1^2 - 2x_1 + 5$$

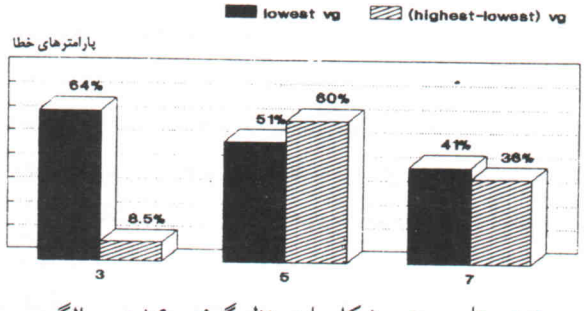
سه ساختار مختلف از شبکه‌های عصبی برای تقریب این تابع مورد استفاده قرار گرفته‌است. در شکل ۲۰ این ساختارها را نشان داده شده‌اند:

شکل ۲۰، شبکهٔ عصبی با تعداد۳،۵و۷گره در لایهٔ مخفی



هر شبکه با استفاده از ۱۶ زوج‌الگوی آموزشی تربیت شده و نتایج حاصل در شکل ۲۱ نشان داده شده‌است. در این شکل، مقادیر مربوط به ماکزیمم مقدار V_G برای هرکدام‌از شبکه‌ها که در سه مرحله به دست آمده نشان داده شده است. به عبارت دیگر، مقدار V_G در هر مرتبه از آموزش شبکه محاسبه و ماکزیمم مقدار آن در دیاگرام نشان داده شده‌است. همچنین در کنار آن، کمترین مقدار V_G نیز نشان داده شده‌است.

شکل ۲۱، تغییرات پارامتر خطا در شبکه‌های حاوی ۵،۳ و۷گره در لایهٔ مخفی



منحنی تابع موزی شکل با در نظر گرفتن ۱۶ زوج الگوی آموزشی برای شبکه

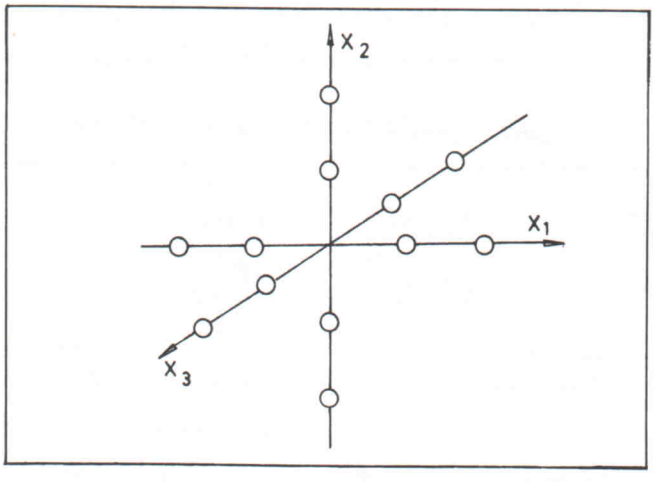
نتیجه‌گیری: برای دسترسی به حالت پایدار، باید حداقل تعداد زوج‌الگوهای آموزشی انتخابی برابر تعداد پارامترهای نامعینی‌باشد.

مثال ۷: تابع زیر را در نظر می‌گیریم:

$$y=1+X_1+X_2+X_3+X_1^2+X_1X_2+X_1X_3+X_2^2+X_2X_3+X_3^2$$

۱- **روش چندجمله‌ای ها:** مطابق شکل ۲۲، از ۱۲ زوج الگوی آموزشی برای تقریب چندجمله‌ای استفاده شده است.

شکل ۲۲



ولی اطلاعات برای دستیابی به مقادیر با توجه به مقادیر X_1,X_2,X_3 قابل دسترس نیست. بنابراین،یک چندجمله‌ای مرتبهٔ دوم کامل برای تقریب تابع غیرممکن است. در این زمینه، بهترین حالت استفاده از قانون هب و به کارگیری، شبه معکوس ماتریس برای حصول ماتریس وزنهاست. همچنین برای تقریب می‌توان از یک چندجمله‌ای مطابق زیر استفاده کرد:

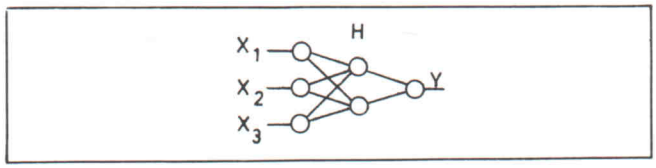
$$y=b_0+b_1X_1+b_2X_2+b_3X_3+b_4X_1^2+b_5X_2^2+b_6X_3^2$$

در این حالت، به زوج‌الگوی زیادی نیاز داریم و مقدار V_G برای مقادیر انتخابی به سمت 36.4% میل می‌کند.

۲- **استفاده از شبکه‌های عصبی:** یک شبکهٔ عصبی با دو گره در لایهٔ مخفی در نظر گرفته شده که در این شبکه، ۱۲ زوج‌الگوی آموزشی برای شبکه انتخاب شده است.

اگرچه تعداد زوج‌الگوهای آموزشی بزرگتر از تعداد پارامترهای نامعینی در شبکه انتخاب شده ولی حالت ناپایداری در شبکه مشاهده می‌شود، به گونه‌ای که مقدار V_G از 32.9% به 93.5% افزایش پیدا کرده

شکل ۲۳، یک شبکهٔ عصبی با دو گره در لایهٔ مخفی



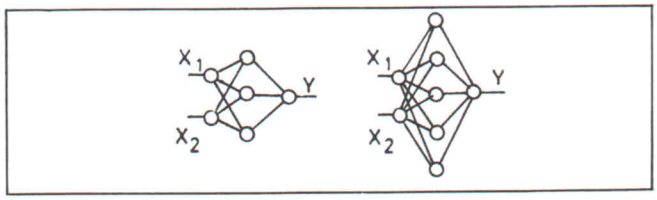
است. این امر نشان می‌دهد که داشتن تعداد زوج الگوی بیشتر از پارامتر نامعینی یک شرط لازم بوده ولی کافی نیست.

مثال ۸ : تابع مثال ۶ را دوباره در نظر گیرید:

$$y = 10x_1^4 - 20x_2x_1^2 + 10x_2^2 + x_1^2 - 2x_1 + 5$$

شبکه‌های در نظر گرفته شده برای بررسی نتایج حاصل از تربیت اضافی مطابق شکل ۲۴ شامل ۳ و ۵ گره در لایهٔ مخفی است.

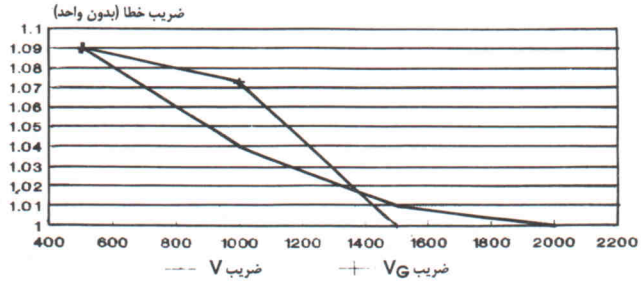
شکل ۲۴، شبکهٔ عصبی با تعداد ۳و۵ گره در لایهٔ مخفی



شکل ۲۵ نشان دهندهٔ تغییرات V و V_G با افزایش تربیت شبکه با تعداد ۵ گره در لایهٔ مخفی است.

در شکل ۲۶ نیز مراتب بالا برای شبکه‌ای با ۳ گره در لایهٔ مخفی بررسی شده‌است. در این شکل، مقادیر V و V_G با افزایش تربیت به سمت ۱ میل می‌کند.

شکل ۲۵، تغییرات V و V_G بر حسب تعداد مراحل یادگیری برای شبکهٔ حاوی ۵ گره در لایهٔ مخفی

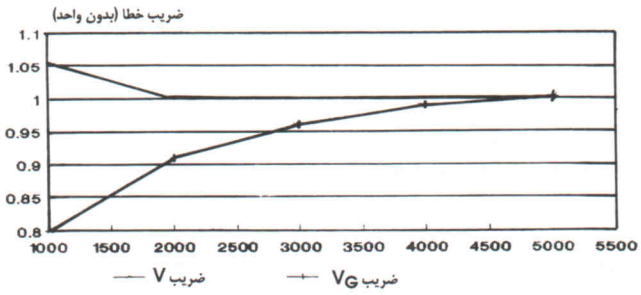


نتایج حاصل برای تابع موزی شکل با ۱۶ زوج الگوی آموزشی

مثال ۹: انجام یک حلقه‌بندی وفقی (adaotive) در المانهای محدود:

در حلقه‌بندی و ایجاد پاره‌سازه‌ها، تعداد المانهایی که از یک المان درشت حاصل می‌شود، از اهمیت زیادی برخوردار است.در این موارد، اگر روش حلقه‌بندی وفقی مورد استفاده قرار گیرد، حلقه‌هایی که درشت‌تر هستند دوباره حلقه‌بندی می‌شوند. در روش تحلیل دینامیکی،

شکل ۲۶، تغییرات V و V_G بر حسب تعداد مراحل یادگیری با تعداد ۳ گره در لایهٔ مخفی



نتایج حاصل از تربیت اضافی شبکه با ۱۶ زوج الگوی آموزشی

مراحل زمانی به صورت گامهای زمانی مختلف تقسیم‌بندی می‌شود و برای تحلیل آن از روش المانهای محدود به صورت موازی (parallel) استفاده می‌شود. و به طور خلاصه، در یک حلقه‌بندی اصلاح شده ما با پاره‌سازه‌هایی مواجه خواهیم شد که در آنها یک بالانس نیرویی بین المانها برقرار است. الگوریتم (BP) در شبکه‌های عصبی از قابلیت تشخیص وفقی برخوردار است و به گونه‌ای موثر می‌تواند در پیش‌بینی تعداد المانهای تولید شده از یک المان مورد استفاده قرار گیرد. اطلاعات مربوط به داده‌ها که توسط این اطلاعات آموزش شبکه انجام می‌گیرد، از سه قسمت تشکیل شده‌است

۱- **مشخصات هندسی المان:** این مشخصات برای هر المان بخصوص، منحصر به فرد است.

۲- **پارامترهای گرهی یا ابعاد المانهایی که باید تولید شوند.**

۳- **تعداد المانهای ایجاد شده**

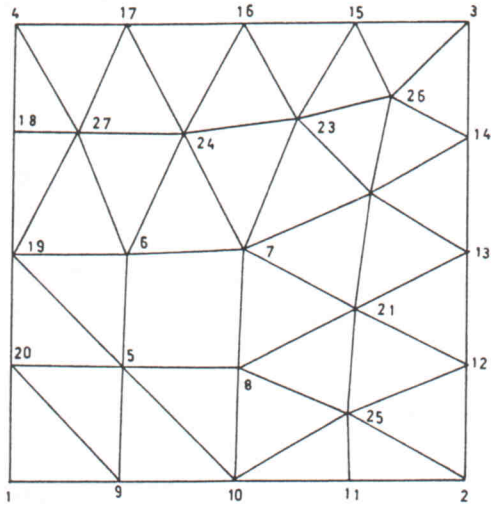
برای یک المان مثلثی، مشخصات هندسی شامل طول اضلاع و l_1 , l_2 , l_3 زوایای داخلی θ_3 , θ_2 , θ_1 است.

پارامترهای گرهی در حقیقت بیان کنندهٔ ابعاد مثلثی است که باید ایجاد شود. این پارامترها نیز با نشان داده می‌شود.پارامترهای گرهی و طول اضلاع ممکن است بر حسب یکی از پارامترها سنجیده شود. بنابراین، سنجیدن واندازه‌گیری پارامترها توسط δ_1 به این معنی است که δ_1 را برابر واحد در نظر گیریم. باید دقت کرد که این مقدار ثابت در مرحلهٔ آموزش شبکه و انجام دادن عملیات متوالی به حد کافی کوچک شده و عملاً" در مراحل آموزش بعدی شبکه تأثیری نداشته‌باشد. همچنین، برای خصوصیات مربوط به مشخصات هندسی تنها طول اضلاع کافی است و به این ترتیب، می‌توان ورودیهای مربوط به زوایای داخلی را حذف کرد. بنابراین، اطلاعات مربوط به داده‌ها عبارت خواهند بود از اندازهٔ طول اضلاع سه ضلع و دو مقدار مربوط به پارامترهای گرهی هر المان جزء. پاسخ شبکه نیز شامل تعداد مثلثهای تولید شده در حلقه‌بندی اصلاح شده است.

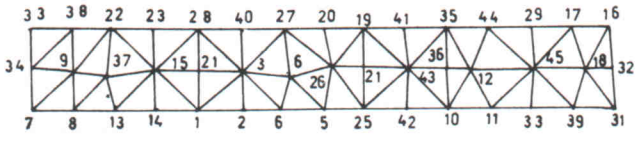
چگونگی مراحل آموزش شبکه:

برای انجام دادن عملیات آموزش شبکه دو المان،شکلی راکه به طور نامناسبی حلقه‌بندی شده‌اند مورد بررسی قرار می‌دهیم. این المانها در اشکال ۲۷ و ۲۸ نشان داده شده‌اند. المانهای مزبور در نقاط مختلفی بارگذاری شده، که جداول زیر مقدار و جهت بارها را نشان می‌دهد:

شکل ۲۷، یک حلقه‌بندی شامل ۳۴ المان



شکل ۲۸، یک حلقه‌بندی شامل ۵۶ المان

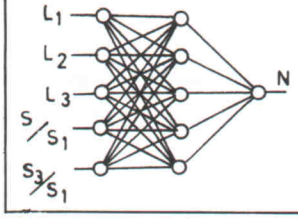


گره	مقدار	حالت بار
16	-100	Py

جدول ۴، حالت بارگذاری مربوط به حلقهٔ ۵۶ المانی

گره	مقدار	حالت بار
4,5,6,7,12,13,18,19,20	100	Px
3,15,16,17	100	Py

جدول.۵، حالت بارگذاری مربوط به حلقهٔ ۳۴ المانی



شکل ۲۹، شبکهٔ عصبی انتخاب شده برای آموزش حلقه‌بندی

فایل اطلاعاتی مربوط به دو حلقه‌بندی زمینه را باهم ترکیب و یک فایل اطلاعاتی برای شبکه تشکیل داده‌ایم. این فایل اطلاعاتی، شامل ۴۰۷ زوج الگوی آموزشی و هر الگوی آموزشی حاوی اندازهٔ سه ضلع،مقدار دو پارامتر گرهی و تعداد المانهای تولید شده است. شبکهٔ عصبی انتخاب شده در لایهٔ ورودی دارای ۵ گره، در لایهٔ مخفی دارای ۵ گره ودر لایهٔ خروجی ۱ گره است.

برای تربیت شبکه از دو روش متداول استفاده شده‌است:

۱- روش Single Pattern: در این روش، یک الگوی آموزشی را در نظر گرفته و این الگو تا به شبکه تعلیم داده‌نشود به سراغ الگوی آموزشی دیگر نمی‌رویم. به عبارت دیگر، الگوها را به صورت تکی آموزش می‌دهیم.

۲- روش Epoch Pattern: در این روش، الگوها به صورت گروهی آموزش داده می‌شود. به عبارت دیگر، ۴۰۷ الگوی آموزشی را از الگوی اول شروع و تا الگوی آموزشی ۴۰۷ آموزش داده می‌شود و این سیکل را تا موقع تعلیم یافتن شبکه ادامه می‌دهیم. با توجه به ساختار شبکهٔ در نظر گرفته شده مطابق شکل، عملیات مربوط به آموزش شبکه با استفاده از هر دو روش صورت گرفته و جدولی به دست آمده‌است که حاوی اطلاعاتی از قبیل، ماکزیمم ومینیمم وزن شبکه، میزان یادگیری، ضریب ممنتوم، خطای متوسط مربعات، تعداد سیکلهای تکراری و مدت زمان تعلیم شبکه است.

الگوهای‌گروهی	الگوهای فردی	
۰/۳	۴	حداکثر وزن
-۰/۳	-۴	حداقل وزن
۰/۰۵	۱/۲۵	میزان یادگیری
۰/۰۵	۰/۳	ضریب ممنتوم
۰/۰۲	۰/۰۲	خطای متوسط مربعات
۴۶۸۸	۳۶۲	تعداد سیکلهای تکراری
۲۹۶	۱۰۸	مدت زمان تعلیم شبکه

جدول ۶، نتایج حاصل از آموزش شبکه به دو روش (S.P) و (E.P)با مقدار R.M.S برابر ۲ درصد

الگوهای گروهی	الگوهای فردی	
۴	۴	حداکثر وزن
۰/۰۰۰۰۱	۰/۰۰۰۰۱	حداقل وزن
۰/۰۵	۲/۵	میزان یادگیری
۰/۰۱۵	۰/۴	ضریب ممنتوم
۰/۰۱۹۲۷۰	۰/۰۱۲۱۷۶	خطای متوسط مربعات
۲۱۰۰۰	۷۰۰۰	تعداد سیکلهای تکراری
۱۲۶۳	۲۰۲۵	مدت زمان تعلیم شبکه

جدول ۷، نتایج حاصل از آمرزش شبکه به دو روش (S.P) و (E.P) باهمگرایی تا حد ممکن

اولین کنفرانس سالانهٔ انجمن کامپیوتر ایران

۴ـ۷ دیماه ۱۳۷۴

دانشکدهٔ مهندسی کامپیوتر ـ دانشگاه صنعتی شریف

فراخوان دوم مقاله

انجمن کامپیوتر ایران با هدف رشد و اعتلای دانش و فن کامپیوتر، اولین کنفرانس سالانهٔ کامپیوتر خود را با مشارکت دانشکدهٔ مهندسی کامپیوتر دانشگاه صنعتی شریف برگزار می‌کند. موضوعهای قابل ارائه به کنفرانس شامل مباحث زیر (به ترتیب حروف الفبا) می‌شود:

- تکنولوژی اطلاعات و آثار در جامعه
- اطلاع‌رسانی
- استانداردهای کامپیوتری در ایران
- امنیت اطلاعات در سیستم‌های کامپیوتری
- الگوریتم‌ها و نظریهٔ محاسبات
- بانک‌های اطلاعاتی
- پردازش موازی و سیستم‌های توزیع شده
- پردازش صوت، تصویر و علائم رقمی
- پردازش زبان فارسی
- خودکارسازی (اتوماسیون) اداری
- سیستم‌های اطلاعات مدیریت
- سیستم‌های اطلاعات جغرافیایی
- شبکه‌های کامپیوتری و انتقال داده‌ها
- شبکه‌های عصبی
- طراحی و ساخت معماریهای پیشرفتهٔ کامپیوتری
- کاربرد کامپیوتر در صنعت و خودکارسازی
- نرم‌افزارهای مهندسی
- نظام آموزشی رشتهٔ علوم و مهندسی کامپیوتر در ایران
- هوش مصنوعی و سیستم‌های خبره
- سایر موضوعهای مرتبط

*** شرایط ارائه مقاله**

الف) ارسال چهار نسخه اصل مقاله یا خلاصهٔ مبسوط آن (Extended Summary) شامل: یک صفحهٔ جداگانه برای اسامی و آدرس نویسندگان و عنوان مقاله، چکیدهٔ مقالهٔ نهایی، کلمات کلیدی (Keywords)، خلاصهٔ مقاله (حداقل ۱۰۰۰ کلمه) و مراجع (چنانچه نام نویسنده در مرجع باشد با علامت* جایگزین شود). پذیرش نهایی مقاله منوط به دریافت و ارزیابی مثبت اصل مقاله خواهد بود.

ب) مقاله باید نشان دهندهٔ کار جدید بوده و قبلاً در کنفرانس یا مجله‌ای منتشر نشده و از کیفیت بالایی برخوردار باشد.

پ) زبان رسمی کنفرانس فارسی است هر چند مقالات انگلیسی نیز پذیرفته می‌شود.

ت) ارتباط یا نزدیکی مقاله با عناوین فوق باید مشخص شود.

ث) مهلت دریافت اصل یا خلاصهٔ مبسوط مقالات ۱۳۷۴/۴/۱۷ (۸ جولای ۱۹۹۵) است.

ـ تاریخ اعلام نتایج: ۱۳۷۴/۶/۱۸ (۹ سپتامبر ۱۹۹۵)

ـ تاریخ دریافت اصل مقالات: ۱۳۷۴/۷/۱۷ (۱۹ اکتبر ۱۹۹۵)

*** آدرس دبیرخانهٔ کنفرانس**

تهران ـ خیابان آزادی ـ دانشگاه صنعتی شریف ـ دانشکدهٔ مهندسی کامپیوتر ـ ص. پ ۹۵۱۷ـ۱۱۳۶۵

فاکس: ۶۰۱۲۹۸۳ (۲۱) (۹۸)، تلفن: ۶۰۰۵۳۱۰ (۲۱) (۹۸)، پست الکترونیکی csicc95 @ ce.Sharif. ac. ir

---

* - Computer Society of Iran