Archive of SID

کاربرد شبکه‌های عصبی مصنوعی به عنوان ابزاری آماری در تحلیل داده‌های پزشکی

دکتر انوشیروان کاظم نژاد[1]، سیدمهدی سادات هاشمی، دکتر سقراط فقیه زاده، بهروز کاوه‌ئی

چکیده مقاله

مقدمه. شبکه‌های عصبی مصنوعی رفتار مغز را شبیه سازی می‌کنند و توانایی این را دارند که با استفاده از اطلاعات، عمل پیش بینی و تشخیص الگوها و طبقه بندی را انجام دهند. لذا بطور بالقوه رقیبی برای مدلسازیهای مرسوم آماری نظیر رگرسیون و آنالیز تشخیصی هستند.

روشها. در این قسمت عملکرد نرونهای طبیعی و تعمیم آن را برای نرونهای مصنوعی تشریح شده است. همچنین روش پس انتشار خطا که یکی از مهمترین الگوریتمهای یادگیری برای شبکه‌های عصبی می‌باشد نیز توضیح داده شده است.

نتایج. طی دو مثال شبیه سازی شده و یک مثال واقعی شبکه‌های عصبی مربوطه با استفاده از الگوریتم پس انتشار آموزش دیدند و مدلهای رگرسیونی نیز بر آنها برازش شد و نتایج بدست آمده مورد مقایسه قرار گرفت.

بحث. شبکه‌های عصبی را می‌توان به عنوان روشی غیر پارامتری برای مدلبندی داده‌ها در نظر گرفت و از مدلسازی‌های انجام شده در این مقاله چنین بر می‌آید که شبکه‌های عصبی بطور بالقوه تواناتر از مدلهای رگرسیونی هستند اما بطور نمادین خیلی پیچیده‌تر از مدلهای رگرسیونی می‌باشند.

● *واژه‌های کلیدی: شبکه‌های عصبی، داده‌های پزشکی، رگرسیون خطی، رگرسیون غیر خطی.*

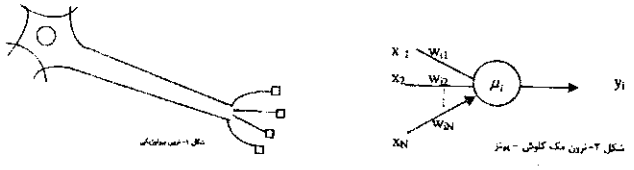
مقدمه

در سالهای اخیر شبکه‌های عصبی مصنوعی توجه بسیاری از محققین رشته‌های مختلف را بخود جلب نموده‌اند و این بدان دلیل است که این شبکه‌ها رفتار مغز را بصورت مصنوعی شبیه سازی نموده و در تصمیم گیریها از آن استفاده می‌کنند. مغز این توانایی را دارد که از اطلاعات جزئی پیشینه، نتایج آینده را استنباط کند و الگوهای مختلف را حتی در صورتی که با اغتشاش توأم باشند، تشخیص دهد و خاطرات گذشته را بسرعت بیاد آورد. از دیدگاه آماری شبکه‌های مصنوعی به این دلیل جالب هستند که بطور بالقوه توانایی پیش بینی و گروه بندی اطلاعات را دارا می‌باشند (۱).

شبکه‌های عصبی در طیف وسیعی از مسائلی کاربرد دارد که علم آمار در تجزیه و تحلیل آنها نقش ایفا می‌کند و مقالات و گزارشهای علمی بسیار زیادی در این زمینه نیز وجود دارد. شبکه‌های عصبی رقیبی برای روشهای مرسوم آماری نظیر رگرسیون چندگانه و چند متغیره، آنالیز تشخیصی، رگرسیون لجستیک و مدلهای سریهای زمانی هستند (۲).

در این مقاله سعی می‌شود تا زمینه‌ای کلی از شبکه‌های عصبی ارائه و برای عینی شدن نحوه بکارگیری این مدلها در علم پزشکی مثالی عملی ارائه شود تا چشم اندازی مختصر از نحو بکارگیری شبکه‌ها را در ذهن پژوهشگران علوم پزشکی ایجاد کند.

نرون کوچکترین واحد پردازشگر در مغز است (شکل ۱). مغز انسان در حدود 10^{11} نرون را شامل می‌باشد و در آن هر نرون به چندین نرون دیگر متصل است و شبکه‌ای پیچیده از پردازشگرهای موازی را بوجود می‌آورد. (۱) مسئله مدلسازی مغز از روی مدلسازی تک نرون عصبی شروع شد و مدل نرون وارن کلوش و والتر پیتز (۱۹۴۳) اولین تلاش برای این کار بود (شکل ۲).



نرونها با استفاده از دوایر و اتصالات بین آنها بوسیله پیکان نشان داده شده است. هر اتصال دارای وزنی است که وزن سیناپس نامیده می‌شود و بوسیله W_{ij} نشان داده شده و بعنوان میزانی از قدرت اتصال از واحد jام به واحد iام در نظر گرفته شده است. ورودی هر واحد مجموع وزنی خروجی دیگر واحدهای متصل به آن است. بنابراین ورودی واحد iام عبارت است از:

$$(۱) \qquad netinput_i = \sum_j w_{ij} \times output_j + w_i$$

و w_i مقدار آستانه واحد iام نام دارد.

پس از این یک تابع محرک بروی این مجموع عمل می‌کند و مقدار این تابع مقدار محرک آن واحد نامیده می‌شود. تابع محرک معمولا غیر خطی است. توابع سیگموئیدی (S شکل) نظیر:

$$(۲) \qquad f(x) = tanh(x) \ , \ f(x) = \frac{e^x}{1+e^x}$$

از این قبیل هستند.

مغز یادگیری را بر اساس قدرت اتصالات سیناپس انجام می‌دهد که در شبکه‌های عصبی مصنوعی ضرایب وزنی این نقش را ایفا می‌کنند. (چیزی

---

۱ - دانشیار گروه آمار زیستی، دانشکده پزشکی، دانشگاه تربیت مدرس

www.SID.ir

Archive of SID

شبیه ضرایب در مدل رگرسیونی) بنابراین یادگیری در یک شبکه عصبی با ساده‌ترین تعریف ممکن است عبارت از پیدا کردن مقادیر این وزنها. سه دسته عمده از انواع یادگیری عبارتند از یادگیری با ناظر، بدون ناظر و تشدیدی، یادگیری با ناظر هنگامی انجام می‌شود که پاسخ مربوطه به هر ورودی مشخص باشد. پس انتشار خطا یکی از عمده‌ترین روشهای یادگیری با ناظر است که در اینجا شرح داده خواهد شد.

پس انتشار خطا

الگوریتم پس انتشار خطا روشی برای پیدا کردن وزنها در یک شبکه پیش خور چند لایه است. فرض کنیم مجموعه‌ای از الگوهای ورودی (مشاهدات) در دسترس باشند که خروجیهای هر یک (پاسخها) مشخص و هدف آموزش دادن شبکه بصورت ناظر باشد. پس از آموزش، چنین شبکه‌ای می‌تواند برای پیش بینی پاسخی متناظر با یک الگوی ورودی جدید مورد استفاده قرار بگیرد و این نظیر یکی از استفاده‌هایی است که از مدلهای رگرسیون در آمار می‌شود.

برای اجرای آموزش بایستی ابتدا تابع هدفی را تعریف کنیم و سپس آنرا بر اساس وزنها بصورت بهینه برای اینکار درآوریم. برای این کار مجموع مربعات خطا که بصورت زیر تعریف می‌شود مناسب بنظر می‌رسد.

$$E=\frac{1}{2}\sum_{p=1}^{n}\sum_{k=1}^{O}\left(y_{pk}-\hat{y}_{pk}\right)^2 \qquad (3)$$

که در آن اندیس n تعداد الگوهای (مشاهدات) است و اندیس k، نمایانگر واحد خروجی، y پاسخ مشاهده شده و $\hat{y}$ پاسخ پیش بینی شده توسط مدل هستند در صورتیکه تنها یک خروجی داشته باشیم O=1 و معادله ۳ بصورت زیر در می‌آید.

$$E=\frac{1}{2}\sum_{p=1}^{n}\left(y_p-\hat{y}_p\right)^2 \qquad (4)$$

برای درک یادگیری به روش پس انتشار خطا با چگونگی گذر رو به جلوی اطلاعات در شبکه آغاز می‌کنیم. فرایند با مقادیر ورودی در لایه ورودی شروع می‌شود. نرونهای این لایه هیچ کار خاصی روی اطلاعات ورودی انجام نمی‌دهند و آنها را به نرونهای لایه مخفی تحویل می‌دهند. ورودی به j امین نرون لایه مخفی عبارت است از

$$h_{pj}=\sum_{i=1}^{N} w_{ji}x_{pi} \qquad (5)$$

که در آن N تعداد کل ورودیها، w_{ji} وزن از iامین نرون ورودی به j امین نرون مخفی است. و x_{pi} مقدار i امین ورودی برای p امین الگوست با عمل تابع فعالیت نرون داریم.

$$v_{pj}=f(h_{pj})=\frac{e^{h_{pj}}}{1+e^{h_{pj}}} \qquad (6)$$

بطور مشابه kامین واحد خروجی زیر را دریافت می‌کند.

$$g_{pk}=\sum_{j=1}^{M} W_{kj}v_{pj} \qquad (7)$$

که در آن M تعداد نرونهای مخفی و W_{kj} وزن از نرون مخفی j به نرون خروجی k را نشان می‌دهد. بنابراین کمیت زیر مقدار خروجی خواهد بود.

$$\hat{y}_{pk}=f(g_{pk})=\frac{e^{g_{pk}}}{1+e^{g_{pk}}} \qquad (8)$$

هدف پیدا کردن W_{kj} ها است و تابع هدف ما همان E در معادله (۴) است. بنابراین مشتقات جزئی تابع هدف نسبت به وزنهای مذکور میزان تغییرات آنرا نسبت به هر وزن نشان می‌دهد لذا می‌توان وزنها را در راستایی به حرکت در آورد که باعث کاهش پیدا کردن تابع هدف شود. به لحاظ ریاضی این مطلب به صورت زیر در می‌آید.

$$\Delta W_{kj}=-\alpha\frac{\partial E}{\partial W_{kj}} \qquad (9)$$

که در آن α میزان یادگیری گفته می‌شود و در واقع اندازه گام برای رسیدن به حالت بهینه می‌باشد. کار را با بدست آوردن عبارتی برای محاسبه وزنها از نرونهای مخفی به نرونهای خروجی آغاز می‌کنیم؛ یعنی همان W_{kj}ها با قرار دادن معادلات (۵) و (۸) در معادله (۳) بدست می‌آوریم.

(۱۰)

$$E=\frac{1}{2}\sum_{p=1}^{n}\sum_{k=1}^{O}\left(y_{pk}-f\left(\sum_{j=1}^{M}W_{kj}f\left(\sum_{i=1}^{N}W_{ji}x_{pi}\right)\right)\right)^2$$

بر اساس قاعده زنجیر داریم:

$$-\alpha\frac{\partial E}{\partial W_{kj}}=-\alpha\frac{\partial E}{\partial \hat{y}_{pk}}\frac{\partial \hat{y}_{pk}}{\partial g_{pk}}\frac{\partial g_{pk}}{\partial W_{kj}} \qquad (11)$$

اما داریم $\frac{\partial E}{\partial \hat{y}_{pk}}=-\left(y_{pk}-\hat{y}_{pk}\right)$

و $\frac{\partial \hat{y}}{\partial g_{pk}}=\hat{y}_{pk}\left(1-y_{pk}\right)$ و $\frac{\partial g_{pk}}{\partial W_{kj}}=v_{pj}$

پس با قرار دادن این مقادیر در معادله (۹) بدست می‌آوریم.

(۱۲)

$$\Delta W_{kj}=-\alpha\left[(-1)\left(y_{pk}-\hat{y}_{pk}\right)\right]\hat{y}_{pk}\left(1-\hat{y}_{pk}\right)v_{pj}$$

و از این مقدار برای به هنگام کردن وزنها از نرونهای لایه مخفی به نرونهای خروجی بصورت زیر استفاده می‌شود.

www.SID.ir

Archive of SID

$$W_{kj}^{t+1} = W_{kj}^{t} + \Delta W_{kj} \quad (۱۳)$$

برای به هنگام در آوردن وزنهای W_{ji} که ورودیها را به نرونهای مخفی متصل می‌کنند با منطق مشابهی داریم

$$\Delta W_{ji} = -\alpha \frac{\partial E}{\partial W_{ji}} \quad (۱۴)$$

و با بسط این عبارت بر اساس قاعده زنجیره خواهیم داشت

$$-\alpha \frac{\partial E}{\partial w_{ji}} = -\alpha \sum_{k=1}^{O} \frac{\partial E}{\partial \hat{y}_{pk}} \frac{\partial \hat{y}_{pk}}{\partial g_{pk}} \frac{\partial g_{pk}}{\partial v_{pj}} \frac{\partial v_{pj}}{\partial h_{pj}} \frac{\partial h_{pj}}{\partial w_{ji}} \quad (۱۵)$$

و مشابه با حالت قبلی نتیجه نهایی عبارت خواهد بود از:

$$\Delta W_{ji} = \alpha \sum_{k=1}^{O} \left(y_{pk} - \hat{y}_{pk}\right) \hat{y}_{pk} \left(1 - \hat{y}_{pk}\right) W_{kj} v_{pj} \left(1 - v_{pj}\right) x_{pi} \quad (۱۶)$$

شبکه‌های عصبی و رگرسیون

در آمار از رگرسیون برای مدلبندی روابط بین متغیرها استفاده می‌شود. در آن متغیرهای مستقل با x_iها نشان داده می‌شوند و می‌توانند تحت کنترل پژوهشگر باشند و یا اینکه توسط وی مشاهده شوند. پاسخ یا متغیر وابسته با y نشان داده می‌شود. هدف رگرسیون پیش بینی یا طبقه بندی پاسخ y از روی مجموعه x_i هاست.

فرم عمومی مدل رگرسیون به این صورت است (مک کولاف و نلدر ۱۹۸۹)

$$\eta = \sum_{i=0}^{N} \beta_i x_i \quad (۱۷)$$

که در آن β_iها ضرایب رگرسیونی و N تعداد متغیرهای مستقل است. این مدل دارای سه مولفه است.

۱. مؤلفه تصادفی y با میانگین μ و واریانس σ^2

۲. مؤلفه سیستماتیک η که ترکیبی خطی از x_iهاست.

۳. تابع پیوند $\eta = h(\mu)$ که میانگین را به مؤلفه سیستماتیک پیوند می‌دهد.

در صورتیکه مؤلفه تصادفی دارای توزیع نرمال صفر و واریانس σ^2 باشد آنگاه بنا به قضایای استنباط آماری تابع h(0) تابعی همانی خواهد بود و مدل بصورت زیر در می‌آید.

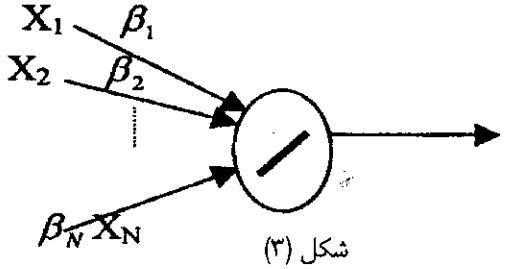
$$y_p = \beta_0 + \sum_{i=1}^{N} \beta_i x_{pi} + \varepsilon_p \quad (۱۸)$$

که در آن $\varepsilon_p \sim N(0,\sigma^2)$. برای پیدا کردن ضرایب β_i در این معادله مجموع مربعات خطای زیر را مینیمم کرد.

$$E = \sum_{p=1}^{n} \left(y_p - \sum_{i=0}^{N} \beta_i x_{pi} \right)^2 \quad (۱۹)$$

این مسئله هم ارز است با یک شبکه پیش خور تک لایه با یک نرون خروجی (شکل ۳).

یک تفاوت رگرسیون خطی چندگانه با این شبکه عصبی در این است که در حالت رگرسیون چندگانه یک فرم بسته برای برآورد وجود دارد اما در شبکه عصبی از روشهای تکراری استفاده می‌شود. یعنی مشابه آنچه در رگرسیونهای تعمیم یافته صورت می‌گیرد.



شکل (۳)

در مدلهای رگرسیونی یک فرم تابعی به داده‌ها تحمیل می‌شود. مثلا در ساده‌ترین حالت فرض می‌شود که پاسخ ترکیبی خطی از متغیرهای مستقل است و هرگاه این فرض درست نباشد منجر به خطای سیستماتیک در پیش بینی‌ها خواهد شد.

اما راه دیگر این است که هیچ شکل تابعی خاصی را فرض نکنیم و بگذاریم که شکل تابعی از درون داده‌ها بیرون بیاید و این یکی از عوامل اساسی توانمند بودن شبکه‌های عصبی است. از یک شبکه عصبی پیش خور چند لایه به عنوان تقریب زن جهانی یاد می‌شود زیرا چنانچه توابع محرک آن سیگموئید باشند قادر است که هر تابعی را با هر درجه دقت مطالب تقریب بزند. [5] بنابراین هرگاه هیچ ایده‌ای بروی رابطه بین متغیرهای وابسته و مستقل وجود نداشته باشد استفاده از یک شبکه عصبی بسیار مناسبتر خواهد بود.

اما در شبکه‌های عصبی مشکلاتی نیز وجود دارند مثلا چه تعداد نرون برای لایه‌های مخفی لازم است؟ یا میزان یادگیری α چقدر باید انتخاب شود، مقادیر اولیه وزنها چه باشند؟ چه چیزی می‌تواند تابع اجرایی باشد و بالاخره اینکه آموزش در چه زمانی باید متوقف شود. [1-4]

مثالهای شبیه سازی شده و کاربردی در زمینه پزشکی

برای نشان دادن نحوه استفاده از شبکه‌های عصبی دو مثال را با داده‌های شبیه سازی شده و یک مثال را با استفاده از داده‌های واقعی در علم پزشکی ارائه می‌دهیم.

مثال یک: یک مسئله خطی ساده که رابطه‌ی واقعی آن بصورت زیر است را در نظر می‌گیریم.

$$y = 3 + 5x$$

www.SID.ir

Archive of SID

تعداد ۱۰۰ مقدار برای X از حوزه مقادیر ۷۰-۲۰ بصورت تصادفی انتخاب شدند سپس با استفاده از یک توزیع نرمال با میانگین ۰ و انحراف استاندارد ۱۰، تعداد صد خطای تصادفی e و به مقادیر y هایی که از روی xها محاسبه شده بودند اضافه گردید مدل رگرسیون خطی برای این مسئله به شکل زیر می‌باشد.

$$E(y \mid x)=\alpha+\beta x+e \quad (۲۰)$$

پس از برازش مدل رگرسیونی بدست آوردیم:

$$MSE=91.575 \quad , \quad R^2=0.986 \; , \; \beta=5.003 \quad , \quad \alpha=3.425$$

برای حل این مسئله با استفاده از شبکه‌های عصبی یک شبکه عصبی پیش خور دو لایه که دارای چهار نرون مخفی با توابع محرک سیگموئیدی بود در نظر گرفته شد و با استفاده از الگوریتم پس انتشار خطا مورد آموزش قرار گرفت. لازم به ذکر است که شبکه‌ای با این پیکربندی برای چنین مسئله‌ای خیلی عام است با این وجود یاد آور می‌شود که در کاربردهای عملی در واقع هیچ اطلاعی درباره رابطه تابعی مورد نظرمان را نداریم. برای شبکه ما پس از آموزش $MSE=89.623$ بدست آمد که نشان می‌دهد در این حالت تفاوت زیادی بین رگرسیون خطی و شبکه عصبی وجود ندارد. در این حالت هم رگرسیون و هم شبکه هر دو به حالت واقعی نزدیک هستند. اما مدل رگرسیون خطی بسیار سریعتر برازش می‌گردد و سهل التفسیرتر است.

مثال دو: این مثال را کمی پیچیده‌تر انتخاب کردیم تا برتری شبکه‌های عصبی را در مدلبندی روابط غیر خطی نشان دهیم رابطه تابعی به صورت زیر انتخاب شد:

$$y=20\exp^{-0.5}[\,\mathrm{Ln}\,(0.9x+0.2)\,+1.5]$$

باز هم نظیر مثال قبل ۱۰۰ تا x بصورت تصادفی این بار از مقادیر بین ۰ تا ۱ انتخاب شدند و مقادیر خطاهای تصادفی از توزیع نرمال با میانگین ۰ و انحراف معیار استاندارد ۰/۵ بدست آمدند که به مقدار yهای محاسبه شده اضافه گردیدند. همان مدل رگرسیون خطی (۲۰) و همان شبکه با پیکر بندی مثال قبل برای این مسئله نیز بکار رفت. اما این بار بجای ۴ نرون مخفی ۸ نرون برای آن بکار گرفته شد.

در این مسئله رگرسیون خطی معمولی هیچ ارتباطی بین x و y را نشان نداد. (البته در چنین مواردی از تبدیلهای مناسب برای خطی کردن رابطه بین متغیرها استفاده می‌کنیم اما باید توجه داشت که این مثال از نوع شبیه سازی است و در حالت کلی معمولا از رابطه واقعی بین x و y بی اطلاع هستیم) برای آن MSE = ۱۸۲۵.۳۲۱ در صورتیکه برای شبکه عصبی MSE = ۶۵.۲۷۱ بدست آمد که از مقایسه این دو پی به این نکته می‌بریم که در صورت وجود روابط غیر خطی بین متغیرها که به سادگی از روی ساختار داده‌ها قابل حدس زدن نباشد مدل شبکه عصبی پیش بینی بهتری را ارائه می‌دهد.

مثال سه: این مثال نسبت به دو مثال قبلی کمی پیچیده‌تر است زیرا متغیر پاسخ در آن از نوع دو حالته می‌باشد. داده‌ها از طرح تحقیقاتی مشاوره پس از زایمان در بیمارستانهای دانشگاهی شهر تهران ۷۶-۷۵ بدست آمده‌اند و شامل ۴۱۷۷ خانم هستند که برای زایمان به آن بیمارستانها مراجعه کرده بدند در این متغیرهای ورودی عبارتند از سن مادر، تعداد فرزند دختر و تعداد فرزند پسر و متغیر خروجی عبارت است از نتیجه حاملگی (خواسته یا ناخواسته) پس از حذف موارد گمشده و عدم پاسخ تعداد ۴۰۱۲ داده باقی مانده است. معماری شبکه در این حالت بصورت ۳:۳:۱ در نظر گرفته شد یعنی شبکه‌ای پیش خور با سه نرون ورودی، سه نرون در لایه مخفی و یک نرون در لایه خروجی و توابع محرک همه آنها نیز سیگموئید می‌باشد بدیهی است که با این فرض خروجی شبکه مقداری بین ۰ و ۱ می‌تواند باشد که ۱ نشانه ناخواسته بودن و ۰ نشاندهنده خواسته بودن حاملگی است. در این حالت روش مرسوم آماری برای تحلیل این داده‌ها استفاده از رگرسیون لجستیک است. برای مقایسه بین این دو، داده‌ها به دو قسمت مساوی تقسیم شد و یکی از آنها به عنوان داده‌های آموزشی مورد استفاده قرار گرفت که هر دو مدل رگرسیون لجستیک و شبکه عصبی بر اساس آنها برازش گردیدند سپس برای بررسی اعتبار هر یک از این دو مدل در پیش‌بینی صحیح نتیجه از داده‌های دیگر که داده‌های آزمون نام دارند استفاده شد.

در این حالت معیار مقایسه دو مدل بر مبنای مساحت زیر منحنی ROC (Reciever Operating Characteristic) قرار داده می‌شود که مقدار آن از ۰ تا ۱ متغیر است و هر چه این مقدار به یک نزدیکتر باشد نمایانگر توان بیشتر مدل در پیش بینی نتیجه است در مثال ما این مقدار برای رگرسیون لجستیک ۰/۷۵۸ و برای مدل شبکه عصبی ۰/۸۲۳ بدست آمد. البته لازم به ذکر است که با افزایش نرونهای لایه مخفی می‌توان این مقدار را باز هم افزایش داد ولی در اینصورت خطر این وجود خواهد داشت که شبکه عصبی خطاهای تصادفی مستتر در داده‌ها را نیز فرا بگیرد.

نتیجه‌گیری

شبکه‌های عصبی بسرعت در حال گشودن جای خود در میان روشهای تحلیل داده‌ها هستند، در این مقاله سعی کردیم تا توانایی آنها را در پیش‌بینی نشان دهیم و تشابهات آنها را با مدلهای آشناتر رگرسیونی تشریح کنیم. اما شبکه‌های عصبی بطور نمادین خیلی پیچیده‌تر از مدلهای رگرسیونی اعم از خطی و غیر خطی هستند.

همچنین در این مقاله جزئیات اجرای الگوریتم پس انتشار خطا بر اساس کارهای انجام شده در این زمینه ارائه شده و بر اساس آن می‌توان شبکه‌های پیش‌خور را بگونه‌ای آموزش داد که قادر به تقریب زدن هرگونه تابعی باشند.

www.SID.ir

Archive of SID

بحث

شبکه‌های عصبی را می‌توان به عنوان روشی غیر پارامتری برای مدلبندی داده‌ها در نظر گرفت. در اینجا سعی شد که کاربرد عمل شبکه‌ها در تحلیل داده‌های پزشکی نیز با ارائه مثالی عملی نشان داده شود وطی مثالهای ارائه شده هر دو حالت که در آنها پاسخها پیوسته یا گسسته دو حالته باشند مورد بررسی قرار گرفت و در هر حالت توانایی شبکه‌های عصبی نسبت به روشهای رگرسیون خطی (معمولی و لجستیک) مورد بحث قرار گرفت. بدیهی است که با توجه به وسعت علم شبکه‌های عصبی ارائه همه جزئیات مربوط به آنها در قالب یک مقاله امریست نشدنی. لذا ما در این مقاله تنها به معرفی مهمترین و پر استفاده ترین رده از آنها یعنی شبکه‌های پیشخور به همراه الگوریتم پس انتشار خطا برای آموزش دادن آنها پرداختیم، خواننده علاقه‌مند می‌تواند برای دستیابی به اطلاعات بیشتر به منابع درج شده در این مقاله مراجعه کند و با توجه به روند رو به رشد استفاده از شبکه‌های عصبی در علوم مختلف به ویژه علوم پزشکی و بهداشتی امید می‌رود تا در آینده با تعمیم روشهای استنباط آماری به این علم معیارهایی برای قضاوت در مورد نحوه برازش مدلهای عصبی و آزمون کردن آنها فراهم آید.

مراجع

1- *Dayhoff J. Neural Networks Architectures: An Introduction. USA. VNR. 1994.*

2- *Werner V. Neural Networks and Logstic Regression. Computation Statistic & Data Analysis. 1996;21.683-701.*

3- *Smith M.Neural Networks for Statistical Modeling. USA. VNR. 1993.*

4- *Hagan M T, Demuth H B . Beal, M. Neural Network Design. USA. PWS Publishing Co. 1995.*

5- *Morgan P, Curry B, Beynon, M. Comparing Neural Approximation for different Functional Forms. Expert Systems. 1999;16(2),60-71.*

www.SID.ir