Archive of SID

کاربرد شبکه های عصبی مصنوعی در پیش‌بینی حاملگی ناخواسته و مقایسه آنها با روشهای مرسوم آماری

سید مهدی سادات هاشمی[1]، انوشیروان کاظم نژاد[2]، بهروز کاوه‌ئی[3]

چکیده

در همه علوم از مدل سازی آماری برای پیش‌بینی و تعیین ارتباطات بین متغیرها استفاده می‌شود و انتخاب مدل مناسب و اجرای آن بستگی به ماهیت متغیرها و پیش فرضهای اولیه دارد. شبکه‌های عصبی مصنوعی روش نوینی برای مدلسازی هستند که با توجه به ساختار هوشمند و انعطاف پذیرشان رقیب مدل سازی‌های مرسوم آماری شده‌اند. و چه در زمینه نظری و چه در زمینه کاربردی رو به پیشرفت هستند . ابتدا به معرفی شبکه‌های عصبی مصنوعی پرداخته و سپس از آنها برای پیش‌بینی حاملگی ناخواسته استفاده کرده‌ایم و در این ضمن توانایی آنها را نسبت به سایر روشهای مدل سازی با استفاده از تحلیل ROC نشان داده‌ایم. با استفاده از سن زن ، تعداد فرزند زنده دختر و تعداد فرزند زنده پسر به عنوان پیش‌بین گر و بر اساس تحلیل ROC،مساحت زیر منحنی ROCرا برای رگرسیون لجستیک،پرابیت،آنالیز تشخیصی و پرسپترون ۳:۲:۱ و۳:۳:۱ به ترتیب ۰/۷۵۸ و۰/۷۵۷ و ۰/۶۳۰و۰/۷۶۸و۰/۸۲۳ محاسبه کردیم. پیش‌بینی حاملگی ناخواسته با استفاده از شبکه‌های عصبی مصنوعی دقیقتر از مدلهای مرسوم آماری است.

واژه های کلیدی:شبکه‌های عصبی مصنوعی،پیش‌بینی،مدلسازی آماری،حاملگی ناخواسته.

مقدمه

امروزه در همه علوم، مدل سازی های آماری نظیر مدلهای رگرسیونی (خطی و غیر خطی)، تحلیل خوشه ای و تحلیل تشخیصی بطور گسترده ای مورد استفاده قرار می گیرد. هدف اصلی در این مدلسازیها بر دو پایه قرار دارد: تعیین روابط بین متغیرها و پیش بینی.انتخاب روش مدل سازی و تحلیل آن به ماهیت متغیرها و شرایط حاکم بر مسئله بستگی دارد. مثلاً رگرسیون خطی معمولی و یا مدلهای تحلیل واریانس هنگامی که با پاسخهای پیوسته مواجه باشیم می تواند روش مناسبی برای پیشبرد اهداف مطالعه باشد و در حالتهای پیچیده تری که با پاسخهایی گسسته در کنار مجموعه‌ای از پیش بین گرها روبرو

۱ – دانشجوی دکترای آمار زیستی ـ دانشگاه تربیت مدرس

۲ – دانشیارگروه آمار زیستی ـ دانشگاه تربیت مدرس

۳ – عضو هیئت علمی سازمان سنجش آموزش کشور

هستیم احتمالاً مدلهای خطی تعمیم‌یافته (GLM) نظیر رگرسیون لجستیک و پرابیت و یا مدلهای لگاریتم خطی چاره ساز خواهند بود. هر چه متغیر پاسخ حساسترباشد دقت پیش بینی صحیح نیز دارای اهمیت بیشتری خواهد بود . بخصوص در حیطه علوم پزشکی و بهداشتی که در آنها سلامت افراد جامعه مد نظر است، بر اهمیت این مسئله افزوده می گردد. بنابراین اگر روشهایی دقیقتر از روشهای موجود در دسترس باشند، بطور طبیعی با استقبال محققین این علم مواجه خواهند شد. مدلهای شبکه های عصبی ، الگوریتمهای ژنتیک و فازی از جمله این روشها هستند.

توجه ما در این مقاله بر دو حیطه قرار داردکه عبارتنداز :

www.SID.ir

Archive of SID

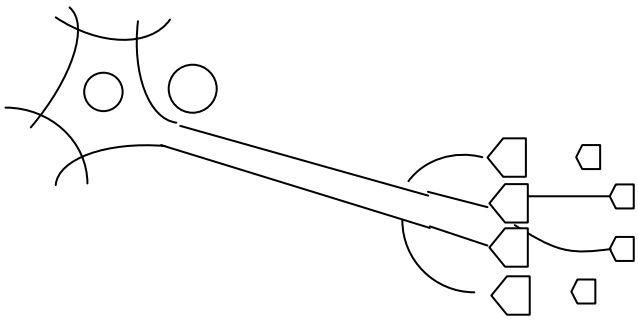
۱)معرفی شبکه‌های عصبی مصنوعی و مقایسه آنها با سایر روشهای مدلسازی آماری ۲) پیش بینی حاملگی ناخواسته . روشهای مرسومی که نتایج آنها در این مقاله با شبکه های عصبی مقایسه شده اند عبارتند از رگرسیون لجستیک، رگرسیون پرابیت و تحلیل تشخیصی. از آنجا که همه جزئیات مربوط به شبکه های عصبی در یک مقاله نمی گنجد، برای جزئیات بیشتر می توان به منابع رجوع کرد.

ساختار شبکه های عصبی مصنوعی

شبکه های عصبی مصنوعی با الهام از سیستم بیولوژیکی مغز ساخته شده‌اند[۱،٤] زیرا دستیابی به توانایی های شگفت انگیز مغز و بازسازی مصنوعی آنها از آرزوهای دیرینه بسیاری از پژوهشگران بوده است. اگرچه به دلیل ساختار پیچیده و مبهم مغز توانایی انجام کامل این کار دور از ذهن بنظر می رسد. اما تلاشها برای رسیدن به این مقصود همچنان ادامه دارد. ماحصل این تلاشها تا امروز شبکه های عصبی مصنوعی نام گرفته است.سلولهای بنیادین مغز که نرون نام دارند هر یک پیچیدگی یک ریز پردازنده را دارا بوده اما سرعت محاسباتی بسیار کندتری نسبت به ریز پردازنده ها دارند.

نرون عصبی یک ماشین تابعی غیر خطی (که می تواند خطی نیز باشد) است. لذا شبکه ای که از اجتماع این نرونها تشکیلمی شود غیر خطی و کاملاً پیچیده خواهد بود[۱]. حال این پرسش بوجود می آید که آیا می توان نرونهایی مصنوعی مشابه با نرونهای طبیعی را بوجود آورد که بتوان از اجتماع آنها برای حل مسائل پیچیده استفاده کرد؟ برای پاسخ این سوال به نحوه عملکرد یک نرون بیولوژیکی می پردازیم.

شکل (۱) چنین نرونی را نشان می دهد.



شکل ۱ – نرون بیولوژیکی

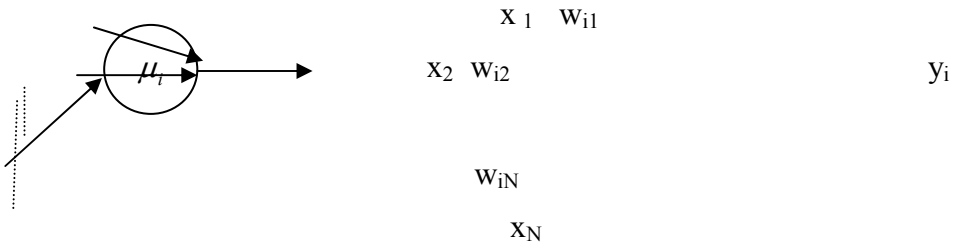
نرون بیولوژیکی از بدنه، هسته ، دندریتها ، سیناپسها و آکسون تشکیل شده است.

هر نرون توسط سیناپسها به سایر نرونها متصل می شود و با دریافت سیگنالهای سایر نرونها از طریق سیناپسها و دندریتهایش تحریک می شود. از ترکیب این سیگنالها و فراتر رفتن آنها از مقدار آستانه مشخصی که سطح فعالیت نامیده می شود نرون شلیک می کند و با اینکار سیگنال خروجیش را توسط آکسون ودندریت به سایر نرونهایی که به آن متصل هستند ارسال می کند. در این بین برخی از سیگنالهای ورودی سلول به عنوان بازدارنده و برخی دیگر به عنوان محرک عمل می کند و بدنه سلول که این سیگنالها را دریافت می کند آنها را ترکیب نموده و سپس هسته سلول تعیین می کند که نرون سیگنالش را شلیک کند یا خیر؟

با این توضیح مختصر از نرون بیولوژیکی ، به نرون مصنوعی می‌پردازیم و تشابهات بین این دو را شرح می دهیم. نرون مصنوعی آنطور که اولین بار توسط «مک کلوش» و «پیتز» ارایه شد بصورت شکل (۲) است.

www.SID.ir

Archive of SID

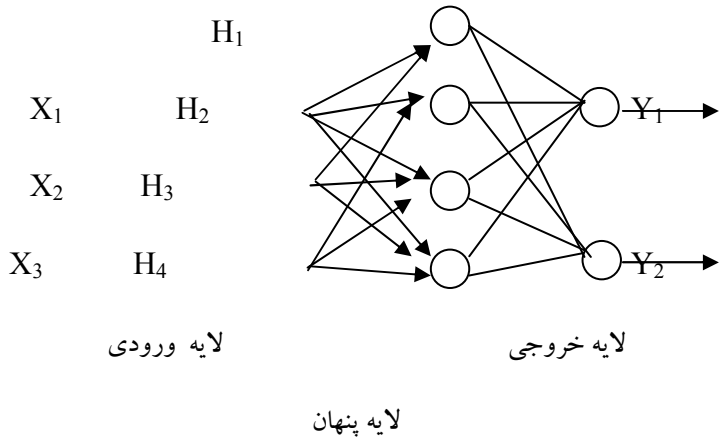


شکل ۲ – نرون مصنوعی

در اینجا ورودی های x_i متناظر با سیگنالهای ورودی به نرون بیولوژیکی، w_i ها متناظر با شدت این سیگنلها ، $\sum$ جمع کننده سیگنالهای وزنی نظیر بدنه سلول و تابع محرک f که بر روی این مجموع اثر می کند نقش هسته سلول را ایفا می کند و نیز خروجی y نظیر سیگنال خارج شده از نرون بیولوژیکی می‌باشد. همانطور که مشاهده می شود نرون مصنوعی ساختاری شبیه معادله رگرسیونی دارد زیرا اگر تابع فعالیت همانی باشد خواهیم داشت $y = \sum w_i x_i$ که همان ارتباط آشنای خطی رگرسیون بین xها و y است. اما تابع محرک f لزوماً همیشه خطی و یا همانی نیست و بنا به شرایط مسئله مورد بررسی فرق می کند[۱،۷].

یک شبکه عصبی بطور کلی دارای سه نوع لایه می باشد[٤]:

۱-لایه ورودی،۲ – لایه میانی یا پنهان،۳ – لایه خروجی(شکل۳)



شکل ۳- شبکه عصبی

تعداد لایه های پنهان می تواند بیش از یکی باشد. همچنین توجه به این نکته اهمیت دارد که توابع محرک f_i لزوما در لایه ها و نرونهای مختلف یکسان نمی باشند. لایه های پنهان در حالت بیولوژیکی نرونهای واسطه ای از آغاز تا رسیدن به هدف هستند و در این مسیر سیستم پردازشگر آنها خروجی نرونهای اولیه را تا رسیدن به مقصد در جهت مناسبی با مسئله دچار تحول و تغییر می‌سازند.

اما یادگیری که از ارکان شبکه های عصبی می باشد عبارت است از فرایند تنظیم وزنهای شبکه که در سه حیطه کلی قرار می گیرد[۳]:

۱- یادگیری با ناظر

www.SID.ir

Archive of SID

۲- یادگیری خود سازمانده یا بدون ناظر

۳- یادگیری تشدیدی یاتقویتی

بکارگیری هر کدام از اینها پیرو شرایط حاکم بر موضوع مورد بررسی قرار می گیرد.برای خواننده علاقمند مجموعه کاملی از روشهای یادگیری در [۲،۳] معرفی شده اند.

شبکه عصبی پرسپترون

پرسپترون پر استفاده ترین و در عین حال یکی از ساده ترین انواع شبکه های عصبی است. در ساده ترین حالت این شبکه از سه لایه تشکیل شده است(مانند شکل ۳)[۷].

حس گرهای ورودی در اولین لایه بصورت تصادفی و یا اختصاصی به سلولهای لایه میانی متصل هستند. نرونهای لایه میانی نیز بصورت تصادفی یا اختصاصی به هم و یا برخی از نرونهای لایه خروجی متصل هستند.

نماد $I : H : O$ پرسپترونی با I ورودی ، H نرون میانی و O نرون خروجی را نشان می دهد.

هر گاه پرسپترون با الگویی از ورودیها مهیا باشد، سلولهای میانی آن فعال شده و برخی از سلولهای لایه خروجی را تحریک می‌کنند. هر کدام از سلولهای خروجی که قویترین ورودی را داشته باشد شبکه را تولید خواهد کرد و خروجی شبکه، پاسخ آن سلول خواهد بود. این پاسخ می تواند عملی مشخص و یا طبقه‌بندی الگویی ورودی بر حسب تعلق داشتن به یک رده باشد.

مقادیر نرونهای لایه ورودی در وزن متناظر آنها ضرب شده و با اعمال تابع محرک همانی روی آنها ، حاصل را به اولین لایه میانی ارسال می کنند. سپس هر یک از نرونها ی این لایه، پس از جمع بندی خروجیهای موزون شده توسط نرونهای ورودی ، تابع محرک لجستیک $f(x) = \frac{e^x}{1+e^x}$ روی حاصل جمع فوق عمل می کند و خروجیها را به لایه بعدی که می تواند لایه پنهان دیگر و یا لایه خروجی باشد ارسال می کند . در لایه خروجی هر نرون پس از دریافت خروجی لایه قبلی باز هم با استفاده ازتابع لجستیک،ورودی از لایه قبلی را مورد پردازش قرار می‌دهد. بنابراین امکان تمایز الگوهایی را بوجود می‌آورد که با مدلهای خطی معمول قابل بیان نیستند[۳].

یادگیری شبکه پرسپترون از نوع با ناظر می باشد. و معمولا با روش پس انتشار خطا مورد آموزش قرار می گیرد. اگر O_{ij} مقدار پیش بینی شده توسط شبکه برای iامین ورودی درi امین نرون لایه خروجی و D_{ij} مقدار واقعی آن باشد، آنگاه خطای اجرایی شبکه در نگاشت ورودی به خروجی توسط $E = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{k}(D_{ij} - O_{ij})^2$ اندازه گیری می شود که در آن n تعداد بردارهای ورودی یا مشاهدات و k تعداد نرونهای لایه خروجی است. همانگونه که گفته شد در روش پس انتشار خطا وزنهای پرسپترون بگونه ای تنظیم می شوند که E حداقل گردد. این عمل را می توان به کمک روشهای تکراری نظیر نیوتن – رافسون انجام داد. در هر مرحله تکرار، تغییرات در وزنها از رابطه زیر بدست می آید:

$$\Delta w = -\lambda \frac{\partial E}{\partial w}$$

به λ در این عبارت میزان یادگیری اطلاق می شود[۷].

تحقیقات مختلف نشان داده اند که یک شبکه پرسپترون می تواند باتعداد مناسب لایه های میانی که دارای نرونهای کافی باشند، هر تابع ریاضی را تقریب بزند. از اینرو نام تقریب گر جهانی را برای آن انتخاب نموده اند[۵].

www.SID.ir

Archive of SID

پیش‌بینی حاملگی ناخواسته

برای نشان دادن توانایی پرسپترون در پیش‌بینی حاملگی ناخواسته و ارجحیت آن نسبت به مدلهای مرسوم رگرسیون لجستیک، رگرسیون پرابیت و تحلیل تشخیصی از داده های مربوط به طرح مشاوره پس از زایمان در کلینکهای دانشگاهی شهر تهران استفاده می کنیم که توسط مرکز ملی تحقیقات بهداشت باروری طی سالهای جمع آوری شده است . در این مطالعه ٤١٥٥ خانم که دارای حداقل یک فرزند بوده و طی سالهای ٧٦-١٣٧٥ برای زایمان به ١٢ بیمارستان دانشگاهی شهر تهران مراجعه کرده بودند به تصادف انتخاب شده و مورد مصاحبه قرار می‌گرفتند، هر بیمارستان به عنوان یک خوشه و خانمهای مراجعه کننده در طی این سالها به عنوان واحدهای نمونه‌ای در هر خوشه محسوب شده‌اند. برای پیشبرد اهداف این مقاله پس از بررسیهای اولیه از بین سئوالات پرسشنامه چهارمورد از مهمترین آنها برای بررسی حاملگی ناخواسته انتخاب شدند که کدگذاری و شرح آنها مطابق جدول زیر است:

کد	شرح	نام متغیر
۰=خواسته ۱=ناخواسته	وضعیت حاملگی فعلی(وابسته)	Y
	سن مادر در هنگام بارداری بر حسب سال(مستقل)	X1
	تعداد فرزند زنده دختر(مستقل)	X2
	تعداد فرزند زنده پسر(مستقل)	X3

٤-١) مدل آنالیز تشخیصی

برای داده های این مدل به صورت $z = a_1x_1 + a_2x_2 + a_3x_3$ می باشد که در آن z مقدار تابع تشخیص برای یک فرد خاص و a_i ها ضرایب تابع تشخیص هستند.

٤-٢) مدل رگرسیون لجستیک

بصورت $\ln(\frac{p}{1-p}) = b_0 + b_1x_1 + b_2x_2 + b_3x_3$ است که در آن b_i ضریب رگرسیون لجستیک و p احتمال ناخواسته بودن حاملگی می باشد.

٤-٣) مدل رگرسیون پرابیت

بصورت $\varphi^{-1}(p) = c_0 + c_1x_1 + c_2x_2 + c_3x_3$ است که در آن φ^{-1} تابع وارون توزیع تجمعی نرمال استاندارد و p احتمال ناخواسته بودن حاملگی است.

٤-٤) شبکه پرسپترون سه لایه ٣:٢:١

شبکه پرسپترون کاملا متصل (یعنی هر نرون لایه قبلی به همه نرونهای لایه بعدی متصل است) سه لایه را با سه نرون در لایه ورودی ، ٢ نرون در لایه میانی و یک نرون در لایه خروجی در نظر گرفته ایم و با روش پس از انتشار خطا آن را آموزش می‌دهیم.

٤-٥) شبکه پرسپترون سه لایه ٣:٣:١

این شبکه را نیز کاملاً مستقل در نظر گرفته‌ایم ، دارای سه نرون

www.SID.ir

Archive of SID

ورودی ، سه نرون دولایه میانی و یک نرون در لایه خروجی. این شبکه نیز با روش پس انتشار خطا مورد آموزش قرار می‌گیرد.

روش بررسی

پس از حذف داده های گمشده(۱۴۳ مورد) ،موارد باقیمانده را به تصادف به دو قسمت آموزشی (شامل ۲۰۰۶ داده) و آزمونی (شامل ۲۰۰۶ داده) تقسیم می کنیم. داده های آموزشی را برای برازش مدلها بکار می بریم و از داده های آزمونی که در برازش مدلها شرکت نداشته اند به منظور سنجش کیفیت مدلهای برازش‌شده استفاده می کنیم. برای برازش مدلهای رگرسیون لجستیک، پرابیت و مدل آنالیز تشخیصی از نرم افزار SAS نسخه ۸ و برای برازش شبکه های عصبی از نرم افزارStatistica 2000 استفاده کرده ایم. همچنین برای مقایسه نتایج پیش بینی مدلها از تحلیل ROC (Reciver Operating Characteristic) استفاده شده است و از آنجا که نرم افزار مناسبی برای این کار وجود نداشت لذا با استفاده از زبان برنامه نویسی $Delphi5.00$ نرم افزار مناسب تهیه و استفاده شد.

نتایج

از آنجا که در این تحقیق هدف پیش‌بینی هرچه دقیقتر حاملگی ناخواسته می‌باشد،لذا بایستی بتوان با استفاده از معیارهای آماری مناسب مقایسه‌ای را بین مدلهای پیش‌بین‌گر مختلف انجام داد تا در صورت لزوم بتوان یکی از آنها را به عنوان مدل بهینه انتخاب نموده و مورد استفاده قرار داد.بنابراین با توجه به دو سطحی بودن متغیر پاسخ یک ملاک مناسب برای سنجش کیفیت مدلهای برازش شده و تعیین توان پیش بینی صحیح آنها استفاده ازسطح زیر منحنی مشخصه عملکرد دریافتی (ROC) است.این منحنی عبارت است از حساسیت بر حسب یک منهای ویژگی(1-Specifity، Sensitivity)، و مساحت زیر آن بعنوان ملاکی برای سنجش توانایی پیش‌بینی صحیح مدل مورد استفاده قرار می‌گیرد و مقدار این مساحت عددی بین صفر و یک است که هرچه به یک نزدیکتر باشد نشانه تواناتر بودن مدل در پیش‌بینی می‌باشد . با استفاده از نرم‌افزار تهیه شده، مساحت ناحیه زیر منحنی ROC برای رگرسیون لجستیک، رگرسیون پرابیت ، آنالیز تشخیصی ، پرسپترون ۳:۲:۱ و پرسپترون ۳:۳:۱ بترتیب: ۰/۷۵۸ ، ۰/۷۵۷ ، ۰/۶۳۰ ، ۰/۷۶۸ و ۰/۸۲۳ محاسبه شده است.

بحث و نتیجه گیری

در مطالعه انجام شده توسط فراهانی و سادات هاشمی[۸] برروی حاملگی ناخواسته در شهر تهران،با استفاده از الگوی رگرسیون لجستیک چندجمله‌ای (Multinomial Logistic Regression) متغیرهای سن زن در هنگام حاملگی،سواد زن،سواد همسر، نظر همسر در مورد استفاده از روش پیشگیری،تعداد فرزند زنده دختر و تعداد فرزند زنده پسر به عنوان عوامل موثر بر وضعیت حاملگی مشخص شدند اما چنانچه هدف پیش‌بینی وضعیت حاملگی باشد همانطور که از این مقاله مشهود است تنها سه مورد از متغیرها می‌توانند پیش‌بینی قابل قبولی را ارائه کنند و نیازی به پیچیده‌تر کردن مدل نمی‌باشد همچنین از مساحت ناحیه‌های زیر منحنی ROC برازش شده از مدلهای مختلف مشخص است که بین گروه رگرسیون لجستیک پرابیت و شبکه عصبی ۳:۲:۱ تفاوت زیادی وجود ندارد اما این مقدار برای مدل آنالیز تشخیصی نسبتاً کمتر است و این احتمالاً به دلیل انحراف از پیش‌فرضهای اولیه است که این مدلها مستلزم به در نظر گرفتن آنها هستند (مثلاً Xها بایستی نرمال چند متغیره باشند). همانطور که ملاحظه می‌شود این مقدار برای شبکه عصبی ۳:۳:۱ بیشتر

www.SID.ir

Archive of SID

از بقیه مدلهاست و نشان می‌دهد با اضافه شدن یک نرون به لایه میانی وضعیت تا چه حد می‌تواند تغییر کند. با اضافه کردن تعداد سلولهای لایه میانی تا آنجا می‌توان پیش رفت که شبکه صد در صد حساس و ویژه شود. اما در اینجا خطری که مدل شبکه را تهدید می‌کند این است که شبکه، خطاهای تصادفی را که در الگوهای ورودی وجود دارند نیز یاد بگیرد، در این صورت شبکه بیش آموزش دیده (Over Training)شده است و اعتبار لازم را برای پیش بینی نخواهد داشت[۲].

تاکنون استنباط هایی برای شبکه‌های عصبی نظیر آنچه که در علم آمار برای مدل سازی و ارزیابی آنها وجود دارد ، ارایه نشده است و مدل سازی بیشتر بر اساس الگوریتم رسیدن به جوابهای مطلوب و حداقل خطا صورت می‌گیرد. بنابر این تا حدودی تابع سلیقه و انتخابهای اولیه محقق قرار دارد. در این زمینه برخی سوالات نیز وجود دارد که همچنان ذهن محققان را بخود مشغول کرده است و مهمترین آنها عبارتند از :

۱ - معماری مناسب و توپولوژی بهینه برای یک شبکه عصبی به چه صورت است ؟

۲ - تعداد لایه‌های پنهان چند مورد می‌تواند باشد؟

۳ - تعداد بهینه نرونهای هر لایه پنهان را به چه صورت می‌توان برآورد کرد؟

البته امید می‌رود که با سعی و تلاش اندیشمندان، این علم تقریباً پویا ، از آنچه که امروز هست فراتر رفته و به مرحله با ثبات‌تر و مستحکم‌تری پا بگذارد.

سپاسگزاری

بدینوسیله از همکاری صمیمانه خانمها: دکتر شایسته جهانفر عضو هیئت علمی دانشگاه علوم پزشکی ایران و فریده خلج آبادی فراهانی کارشناس محترم مرکز ملی تحقیقات علوم پزشکی کشور ، که اطلاعات لازم برای تحلیلهای ارایه شده در این مقاله را در اختیار ما گذاشتند، تشکر و قدردانی می‌گردد.

منابع:

۱ – منهاج ، محمد باقر.:***مبانی شبکه‌های عصبی*** ، مرکز نشر پروفسور حسابی ، تهران،۱۳۷۷

2- Smith، M.،***Neural Networks for statistical modeling***،VNR،USA،1993

3- Hagan.M،T.،Demuth،H،B.،***Neural Networks Design***،PSW،USA،1996

4- Dayhoff،J، E.،***Neural Network Architectures***، VNR، USA،1990

5-Morgan، P.، Carry، B.،Baynon، M.،Compering ***Neural Networks Application for different functional 6 - form expert system***،1999،16(2):60-72

6 - West، D.،West، V.،***Model Selection for Medical Diagnostic Design Support System***:a breast cancer detection case،Artificial Intelligence in Medicine، 2000، 20،183-204

7-Bard، W.، Man an endia، M.، ***Underestanding Neural Networks as Statistical. Tools,*** American Statistician، 1996، 50(4)،284-293

۸ – فراهانی ـ ف. سادات هاشمی ـ م. ***عوامل مؤثر بر حاملگی ناخواسته در تهران***.مجله حکیم،دوره۵،شماره ۳،پاییز ۱۳۸۱

www.SID.ir