نشریه تحقیقاتی فنی ـ مهندسی استقلال، دانشگاه‌صنعتی‌اصفهان شماره ۱۲ اسفندماه ۱۳۷۱

تشخیص اعداد و علامات دستنویس فارسی به کمک شبکه‌های عصبی

فریبرز سبحان منش* ـ حسین علوی**

چکیده:

در این مقاله، پس از مروری بر مسئله تشخیص الگو و نقش شبکه‌های عصبی در این مورد، از این شبکه‌ها به عنوان یک ابزار محاسباتی قوی در تشخیص اعداد و علامات دستنویس فارسی استفاده‌می‌کنیم. بدین منظور شبکه‌های چند لایه تغذیه مستقیم را مورد بررسی قرارمی‌دهیم و سپس از این ساختار برای رسیدن به مقاصد موردنظر استفاده‌می‌کنیم. ابتدا پارامترهای مختلف ساختار فوق از قبیل تعداد نرونهای لایه‌های ورودی، پنهان و خروجی را بهینه‌سازی نموده و سپس به ساده‌کردن و هرس‌کردن[1] شبکهٔ بهینه می‌پردازیم.

مقدمه:

تشخیص حروف و کلمات تایپ‌شده و دستنویس به‌کمک کامپیوتر یا OCR[2] مدتهاست که ذهن محققین هوش مصنوعی را به‌خود مشغول ساخته‌است. در زمینه تشخیص حروف و کلمات لاتین، پیشرفت زیادی حاصل‌شده و نرم‌افزارهای گوناگون معرفی و به بازار عرضه‌شده‌است (برای نمونه به مراجع [۱] تا [۳] رجوع شود). تشخیص حروف و کلمات عربی و فارسی نیز فعالیتهایی را در سالهای اخیر به‌خود اختصاص داده‌است (مثلاً مراجع [۴] تا [۶]).

---

* مربی بخش مهندسی و علوم کامپیوتر ـ دانشکده مهندسی دانشگاه شیراز

** استادیار دانشکده برق و کامپیوتر ـ دانشگاه‌صنعتی‌اصفهان

1. Pruning

2. Optical Character Recognition

[Downloaded from iutjournals.iut.ac.ir on 2024-07-23]

[DOR: 20.1001.1.22287698.1371.12.1.4.9]

روشهای معمول در تشخیص الگوها به‌طور اعم و نوشته‌ها به‌طور اخص را می‌توان به دو دستهٔ روشهای ساختاری[1] و روشهای آماری[2] تقسیم‌بندی کرد [۷]. در روشهای ساختاری برای تشخیص الگو از خواص هندسی و توپولوژیکی آن نظیر قوسها، حلقه‌ها، زوایا و امتدادهای موجود در شکل الگو استفاده می‌شود. در روشهای آماری به هر الگو یک بردار در یک فضای برداری نسبت داده می‌شود. این فضای برداری را فضای ویژگی[3] الگو می‌نامند و مؤلفه‌های بردار ویژگی الگو با استفاده از اندازه‌گیری‌های متعددی که روی الگو انجام‌می‌شود به دست می‌آیند. فضای ویژگی الگو با استفاده ازاطلاعاتی که قبلاً در اختیار سیستم تشخیص الگو قرارداده شده، به طبقات ونواحی متعددی تقسیم شده‌است. هدف از تشخیص یک الگوی بخصوص تعیین طبقه یا ناحیه‌ای است که بردار الگو در آن قراردارد. به این عمل، طبقه‌بندی[4] و به سیستم تشخیص الگو، طبقه‌بندی‌کننده[5] می‌گویند. اطلاعات لازم جهت تقسیم‌بندی فضای ویژگی الگو شامل توابع چگالی احتمال شرطی و غیرشرطی و مشترک طبقات و الگوهاست. اما این توابع همواره موجود نیستند و تنها اطلاعاتی که از طبقات و الگوها در دسترس است تعدادی زوج متناظر به صورت طبقه یا الگو است. در روشهای آماری سعی می‌شود با استفاده از این اطلاعات توابع چگالی احتمال لازم تقریب‌زده شود، اما تقریب توابع چگالی شرطی کار مشکلی است؛ خصوصاً اگر ابعاد فضای ویژگی بزرگ باشد.

روش دیگری که به‌لحاظ سادگی و دقت، توجه زیادی را در سالهای اخیر به خود جلب کرده، استفاده از شبکه‌های عصبی مصنوعی است. هر شبکه عصبی مصنوعی از تعداد زیادی پردازندهٔ ساده تشکیل‌شده‌است که به‌دلیل تشابه با شبکهٔ عصبی بیولوژیکی، نرون نامیده‌می‌شوند. برحسب نوع ساختاری که برای این شبکه‌ها انتخاب می‌شود، هر پردازنده می‌تواند به تعدادی یا به تمام پردازنده‌های شبکه متصل شود. ارتباط بین هر دو پردازنده با یک وزن خاص صورت می‌گیرد، به طوری که w_{ij} وزن اتصال بین خروجی نرون jام و ورودی نرون iام است. بسته به اینکه نرون jام در تحریک نرون iام اثر محرک یا بازدارنده داشته‌باشد، w_{ij} مثبت یا منفی خواهدبود و اگر نرون iام از نرون jام هیچگونه سیگنالی دریافت نکند، w_{ij} برابر صفر خواهدشد. با توجه به توضیحات فوق مدل مصنوعی یک نرون به‌صورت نشان داده‌شده در شکل ۱ ارائه شده‌است ([۸] و[۹]).

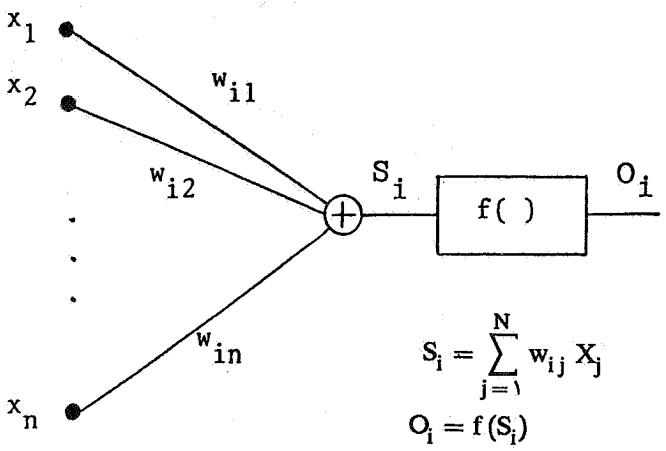
---

1. Structural
2. Statistical
3. Feature space
4. Classification
5. Classifier

[Downloaded from iutjournals.iut.ac.ir on 2024-07-23]

[DOR: 20.1001.1.22287698.1371.12.1.4.9]

برای تابع f فرمهای مختلفی را می‌توان فرض‌کرد. اگر f را تابع محدودکنندهٔ سخت[1] انتخاب‌کنیم،



شکل ۱ـ مدل یک نرون مصنوعی

مدل مصنوعی نرون با مدل بیولوژیکی آن کاملاً منطبق‌می‌شود، اما از لحاظ محاسباتی استفاده از تابع محدودکنندهٔ سخت باعث از دست‌رفتن اطلاعات مفیدی دربارهٔ مجموع وزن‌دار ورودیها می‌شود و بنابراین در شبکه‌های عصبی مصنوعی فرمهای دیگر تابع f نیز به کار گرفته‌می‌شود که معمولترین آنها تابع غیر خطی سیگموید با معادله $f(s_i) = \tanh(s_i)$ است.

شبکه‌های چند لایه تغذیه مستقیم و نقش آنها در تشخیص الگو [۱۰]

یکی از ساختارهای معروف در شبکه‌های عصبی، شبکه‌های چند لایه تغذیه مستقیم است (شکل۲). در این ساختار پردازنده‌های عصبی در چند لایهٔ مجزا و پشت سرهم قرارداده‌می‌شوند و پردازنده‌های هر لایه فقط مجاز به دریافت سیگنال از پردازنده‌های لایهٔ ما قبل و ارسال سیگنال به پردازنده‌های لایهٔ بعد هستند. نرونهای لایهٔ سمت چپ راکه با x_i نشان داده‌شده‌اند، نرونهای ورودی

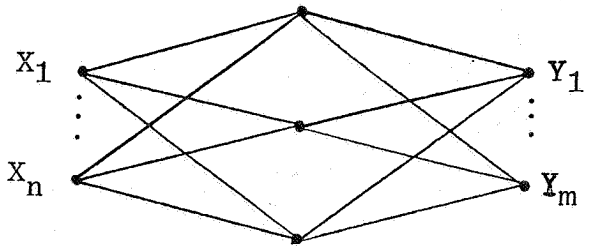
---

1. Hard Limiter

[Downloaded from iutjournals.iut.ac.ir on 2024-07-23]

[DOR: 20.1001.1.22287698.1371.12.1.4.9]

و نرونهای لایهٔ سمت راست راکه با y_j نشان داده شده‌اند، نرونهای خروجی می‌گویند. این نرونها



شکل ۲ـ یک شبکه چند لایه تغذیه مستقیم

ارتباط شبکه را با دنیای خارج برقرار می‌کنند. نرونهای لایهٔ میانی راکه ورودی و خروجی آنها منحصر به نرونهای شبکه می‌شود، نرونهای پنهان می‌نامند. برای تربیت یک شبکهٔ عصبی از تعدادی زوج تربیتی (طبقه مطلوب ، الگو) استفاده می‌شود. هدف از تربیت شبکه، اصلاح وفقی وزنهای شبکه به قسمی است که تابع متوسط مربعات خطای شبکه روی مجموعهٔ زوجهای تربیتی می‌نیمم شود. الگوریتم تربیتی شبکه‌های چند لایهٔ تغذیه مستقیم به الگوریتم انتشار خطا به عقب[1] موسوم است که به اختصار به آن اشاره‌می‌شود. بر طبق این قانون الگوهای تربیتی بکرّات با ترتیب تصادفی به ورودی اعمال‌شده و پس از هربار اعمال، وزنها تغییر داده‌می‌شوند. تغییر وزن w_{ij} به‌صورت زیر انجام می‌پذیرد:

$$w_{ij}^{new} = w_{ij}^{old} + a\delta_i y_j$$

$$\delta_i = \begin{cases} f'_i(s_i)(d_i - y_i) & \text{نرون } i \text{ام در لایهٔ خروجی} \\ f'_i(s_i)\sum_n \delta_n W_{ni} & \text{نرون } i \text{ام در لایهٔ پنهان} \end{cases}$$

در این رابطه، y_j خروجی واقعی نرون jام، y_i و d_i به ترتیب خروجیهای واقعی و مطلوب نرون iام و s_i مجموع وزن‌دار ورودیهای نرون iام هستند. a نرخ یادگیری است که سرعت همگرایی وزنها

---

1. Back Propagation (BP)

[Downloaded from iutjournals.iut.ac.ir on 2024-07-23]

[DOR: 20.1001.1.22287698.1371.12.1.4.9]

و پایداری الگوریتم را کنترل‌می‌کند. هرچه a بزرگتر باشد سرعت همگرایی و احتمال‌ناپایداری و واگراشدن وزنها بیشتر می‌شود. معمولاً $0.01 < a < 1$ انتخاب می‌شود و از $a = 0.1$ به عنوان حدس اولیه استفاده می‌شود. یکی از مشکلات الگوریتمهای تربیتی شبکه‌های چند لایه، به تله افتادن وزنها در یک می‌نیمم محلی[1] و در نتیجه جلوگیری از رسیدن به می‌نیمم قطعی تابع متوسط مربعات خطاست. برای جلوگیری از این امر در الگوریتم BP از یک جملهٔ مومنتوم با ثابت مومنتوم η، که معمولاً حدود 0.9 اختیار می‌شود، به یکی از دو صورت زیر استفاده می‌شود:

$$\Delta w_{ij}(n) = a\,\delta_i\,y_j + \eta\,\Delta w_{ij}(n-1)$$

$$\Delta w_{ij}(n) = a\,(1-\eta)\,\delta_i\,y_j + \eta\,\Delta w_{ij}(n-1)$$

دراین دو رابطه $\Delta w_{ij}(n)$، تغییر وزن w_{ij} به ازای الگوی تربیتی فعلی و $\Delta w_{ij}(n-1)$ تغییر وزن w_{ij} به ازای الگوی تربیتی قبلی است.

استفاده از شبکه‌های عصبی جهت تشخیص اعداد دست‌نویس فارسی[۱۱]

بعد از معرّفی شبکه‌های عصبی و نقش آنها در تشخیص الگو، شبکه‌های چند لایهٔ مستقیم را جهت تشخیص اعداد دستنویس فارسی انتخاب‌می‌کنیم و به بهینه‌سازی پارامترهای مختلف این شبکه از قبیل تعداد نرونهای ورودی، پنهان و خروجی می‌پردازیم. طبق تعریف، شبکهٔ بهینه شبکه‌ای‌است که با کمترین پیچیدگی، بیشترین دقت را در تشخیص اعداد داشته‌باشد. اگر تعداد نرونهای لایهٔ پنهان یک شبکه بسیار کم انتخاب شوند، با توجه به تعدادکم درجات آزادی در یادگیری، شبکه قادر به یادگیری نخواهد بود. از طرف دیگر اگر تعداد نرونهای پنهان شبکه خیلی زیاد باشد، شبکه سعی در حفظ و ذخیره‌کردن الگوهای تربیتی خود خواهد داشت و از قدرت تعمیم[2] و درون یابی[3] آن کاسته می‌گردد (منظور از تعمیم، توانایی شبکه در تشخیص الگوهای جدیدی است که در مجموعهٔ تربیتی‌اش نبوده‌اند امّا شباهتی با آنها دارند، این پدیده، به تنظیم بر نویز[4] و یا تربیت

---

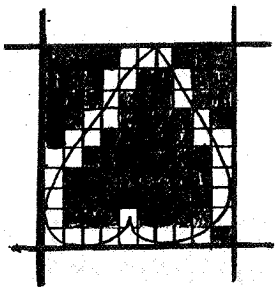
1. Local Minimum
2. Generalization
3. Interpolation
4. Tuning to the Noise

[Downloaded from iutjournals.iut.ac.ir on 2024-07-23]

[DOR: 20.1001.1.22287698.1371.12.1.4.9]

بیش از حد[1] معروف است). بنابراین برای هر منظور و کاربرد خاص، تعداد بهینه‌ای از نرونهای پنهان وجود دارد که باید به کمک شبیه‌سازی شبکه‌های مختلف و اندازه‌گیری میزان تعمیم آنها تعیین شود.

تعداد نرونهای ورودی با نوع کدگذاری[2] مورد استفاده برای الگوهای ورودی و استخراج ویژگی الگو ارتباط دارد. دراین طرح، جهت انتقال الگوهای دستنویس به کامپیوتر از دیجیتایزر به عنوان سنسور ونرم‌افزار Auto Cad به عنوان محرک[3] دیجیتایزر استفاده‌شد. عمل استخراج ویژگی باید به‌گونه‌ای انجام‌شود که بردار ویژگی الگو از اندازه و موقعیت الگو روی دیجیتایزر مستقل باشد.برای این منظور دو خط افقی بالا و پایین الگو و دو خط عمودی از چپ و راست الگو به سمت الگو حرکت داده‌می‌شوند تا بر الگو مماس‌شوند و بدین‌ترتیب یک پنجره بر الگو محیط‌گردد. سپس گوشهٔ چپ و پایین پنجره را به مبدأ مختصات منتقل‌کرده ومساحت پنجره را به $M \times N$ مستطیل کوچک تقسیم‌می‌کنیم و عبور یا عدم‌عبور الگو از هر مستطیل را به ترتیب با دو عدد باینری ۱ یا ۰ نشان‌می‌دهیم (شکل۳). در پایان یک بردار $M \times N$ بعدی باینری از هر الگو خواهیم‌داشت که به عنوان برداروِیژگی[4] آن الگو مورد استفاده‌قرارمی‌گیرد. این نوع کدگذاری راکدگذاری شبکه‌ای می‌نامند. این روش فقط برای دو رقم ۱ و ۰ و علامت منها (-)، که دارای شکل یک بعدی هستند، نتیجهٔ مطلوبی



شکل ۳ـ کدگذاری شبکه‌ای (عدد۵)

1. Over-Training
2. Coding
3. Driver
4. Feature Vector

[Downloaded from iutjournals.iut.ac.ir on 2024-07-23]

[DOR: 20.1001.1.22287698.1371.12.1.4.9]

به دست‌نمی‌دهد. برای رفع این اشکال طول یا عرض پنجرهٔ محیط بر این الگوها راافزایش می‌دهیم. نمونه‌ای ازکدهای تولیدی توسط این روش برای الگوهای دستنویس ۰ تا ۹ و پنج علامت +، −، × ، ÷ و / (ممیز) در شکل۴ نشان داده‌شده‌است.

با استفاده از روش فوق تعداد ۷۲۰ الگوی دستنویس (۴۸الگو از هر عدد و علامت) از ۱۵ نفر با ۷ فرمت مختلف ۵×۵، ۶×۵، ۶×۶، ۷×۶، ۷×۷، ۸×۷، ۸×۸ برداشته‌شد که از این تعداد ۵۲۰ الگو برای تربیت شبکه و ۲۰۰ الگو برای تست شبکه (اندازه‌گیری میزان تعمیم شبکه) درنظرگرفته‌شد. کدهای ورودی ۵×۵، ۶×۵ به علت فقدان قدرت تفکیک کافی برای دو الگوی دستنویس ۲و۴، کدهای یکسانی را نتیجه‌دادند و بنابراین فرمت‌های مزبور حذف‌شدند و انتخاب فرمت بهینهٔ ورودی از میان ۵ فرمت دیگر صورت‌گرفت.

تعیین تعداد بهینهٔ نرونهای لایهٔ خروجی نیز با نوع کد مورد استفاده برای طبقاتِ خروجی ارتباط دارد. باتوجه به تعداد اعداد و علامتهایی که باید مورد تشخیص قرارگیرند. تعداد طبقاتِ خروجی برابر ۱۵ است. برای کدگذاری این طبقات از ۱۵ بردار مقدماتی ۱۵ بعدی استفاده‌می‌کنیم. دراین صورت کد مربوط به طبقه i ام، یک بردار ۱۵ بعدی است که مولفهٔ i ام آن ۱ و بقیهٔ مؤلفه‌هایش صفراست (یعنی c_i =(0,0,...,0,1,0,...,0,0)). دلیل استفاده از این نوع کد، ماکزیمم‌کردن فاصلهٔ همینگ[1] بین کدهای طبقاتِ مختلف‌است. ثابت می‌شود که تنها در صورت استفاده از این نوع کد، طبقه‌بندی کننده عصبی به سمت طبقه‌بندی کننده بهینهٔ بیز[2] میل می‌کند [۸]. در این حالت هر نرون خروجی نمایندهٔ یک طبقه خواهدبود و خروجی هر نرون به عنوان تقریبی از امکان[3] قرارگرفتن الگوی ورودی درآن طبقه تفسیر خواهدشد. اگر از کدهای ۴ بیتی باینری یا ۸ بیتی ASCII استفاده‌شود، بدلیل فاصلهٔ همینگ کوچک بین کدهای طبقاتِ مختلف خطای تشخیص الگوها زیادمی‌شود. با توجه به توضیحات فوق برای پیاده‌سازی کد مربوط به طبقاتِ خروجی روی شبکهٔ عصبی، به ۱۵ نرون در لایهٔ خروجی احتیاج خواهیم‌داشت.

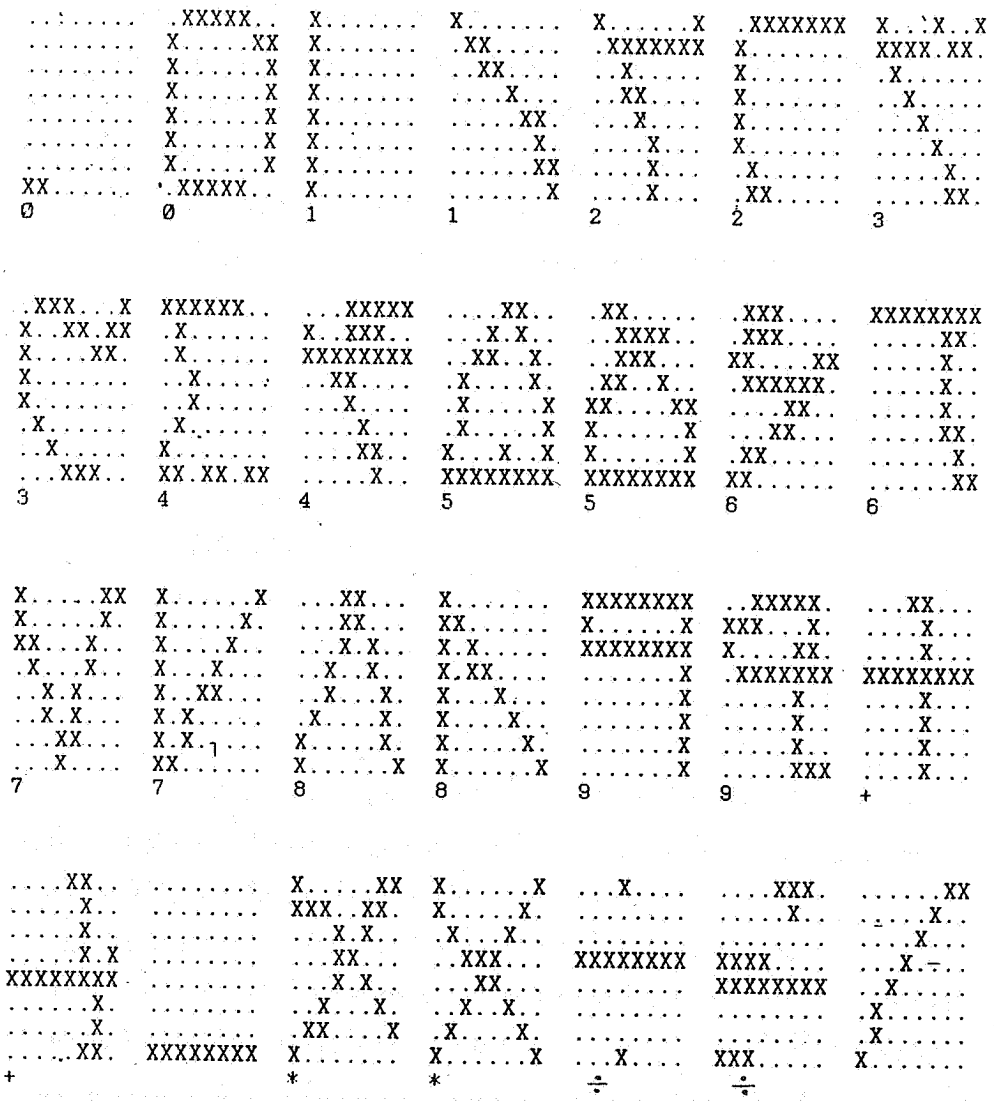
برای تعیین تعداد بهینهٔ نرونهای لایهٔ پنهان و فرمت بهینهٔ ورودی از میان ۵ فرمت ۶×۶، ۷×۶، ۷×۷، ۸×۷، ۸×۸، از شبیه‌سازی شبکه‌های مختلف و اندازه‌گیری میزان تعمیم و درون‌یابی شبکه‌های تربیت‌شدهٔ فوق روی یک مجموعهٔ واحد تست (شامل ۲۰۰الگو) استفاده‌می‌کنیم و به

---

1. Hamming Distance
2. Bayes
3. Likelihood

[Downloaded from iutjournals.iut.ac.ir on 2024-07-23]

[DOR: 20.1001.1.22287698.1371.12.1.4.9]



شکل ۴- نمونه‌هایی از کدهای حاصل از اعمال کدگذاری شبکه‌ای روی الگوهای دست‌نویس

[Downloaded from iutjournals.iut.ac.ir on 2024-07-23]

[DOR: 20.1001.1.22287698.1371.12.1.4.9]

ازای هر کدام از۵ فرمت ورودی، شبکه‌ای را که دارای کمترین خطا در تشخیص الگوهای مجموعهٔ تست باشد انتخاب‌می‌کنیم.

شبیه‌سازی شبکه‌های مختلف و نتایج بدست‌آمده

در شبیه‌سازی شبکه‌ها، تابع انتقال نرونها را تابع سیگموید با معادلهٔ $f(x) = (1+e^{-x})^{-1}$ درنظرمی‌گیریم و با استفاده از الگوریتم تربیتی BP با نرخ یادگیری ۰/۱ و ثابت مومنتوم ۰/۹ و تلورانس تربیتی ۰/۱، به‌تربیت شبکه‌های مختلف می‌پردازیم. لازم به تذکراست که با انتخاب تلورانس تربیتی ۰/۱، مقادیر قابل قبول خروجی به‌جای ۱، فاصله ۰/۹ تا ۱ و بجای ۰، فاصله ۰ تا ۰/۱ خواهدبود. باتوجه به میل‌کردن مجانبی تابع سیگموید به سمت ۱، انتخاب تلورانس تربیتی غیر صفر ضروری است.

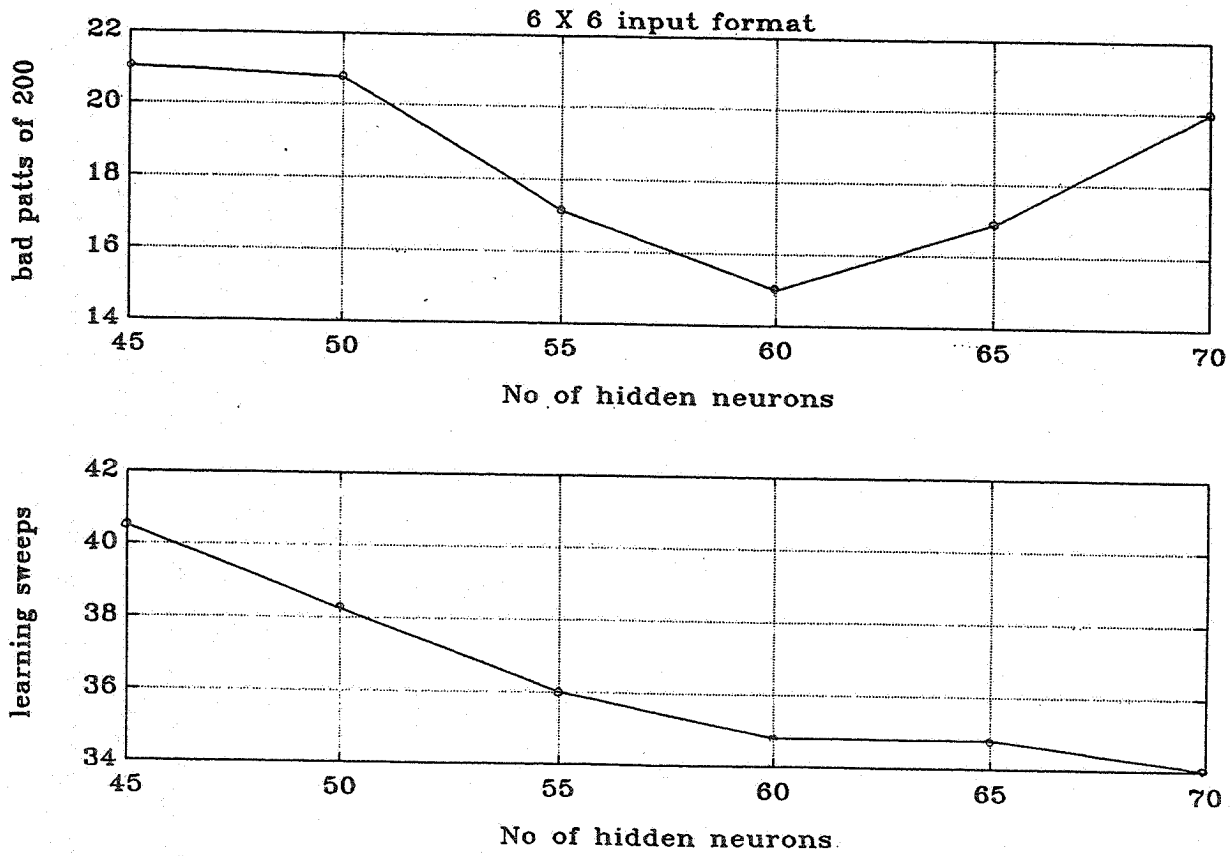
به ازای هریك از ۵ فرمت ورودی مختلف، تعداد نرونهای پنهان را تغییرداده و به‌تربیت شبکه‌های حاصل می‌پردازیم. منحنی تغییرات خطای شبکه‌های مختلف در تشخیص ۲۰۰ الگوی مجموعهٔ تست و منحنی تغییرات زمان همگرایی این شبکه‌ها بر حسب تغییرات تعداد نرونهای لایهٔ پنهان در اشکال ۵ تا ۹ آمده‌اند.

لازم به تذکراست که تربیت هر شبکه با ۷ وزن اولیهٔ مختلف انجام‌گرفته‌است و نقاط نشان داده‌شده روی منحنی متوسط آماری نتایج حاصل از ۷ بار تربیت هستند. نکته قابل‌توجه دیگر در تربیت شبکه‌ها، لزوم ارائه تصادفی الگوهای تربیتی به شبکهٔ در حال تربیت است. اگر در ارائه الگوهای تربیتی به شبکه، نظم وتربیت خاصی برقرارباشد، توانایی یادگیری شبکه پایین‌می‌آید وحتی ممکن است به صفر برسد. مثلاً اگر تمام الگوهای مربوط به یك طبقه خاص را به‌طور متوالی به شبکه بدهیم، آنگاه شبکه در هنگام یادگیری این الگوها هر آنچه را که قبلاً در مورد الگوهای طبقاتِ دیگر یادگرفته‌است، از یاد می‌برد. همان‌طور که از شکلهای ۵ تا ۹ پیداست به ازای هر فرمت ورودی، منحنی خطای شبکه‌ها، به ازای تعداد بهینه‌ای از نرونهای پنهان می‌نیمم می‌شود. هم‌چنین مشاهده‌می‌کنیم که با افزایش تعداد نرونهای لایهٔ پنهان، زمان همگرایی شبکه کاهش‌می‌یابد و این به علت افزایش درجات آزادی شبکه (وزنهای شبکه) است. مشخصات شبکه‌های انتخاب شده به ازای هر یک از ۵ فرمت مختلف ورودی در جدول ۱ آمده‌است. این شبکه‌ها متناظر با می‌نیمم منحنی تغییرات خطا هستند.

[Downloaded from iutjournals.iut.ac.ir on 2024-07-23]

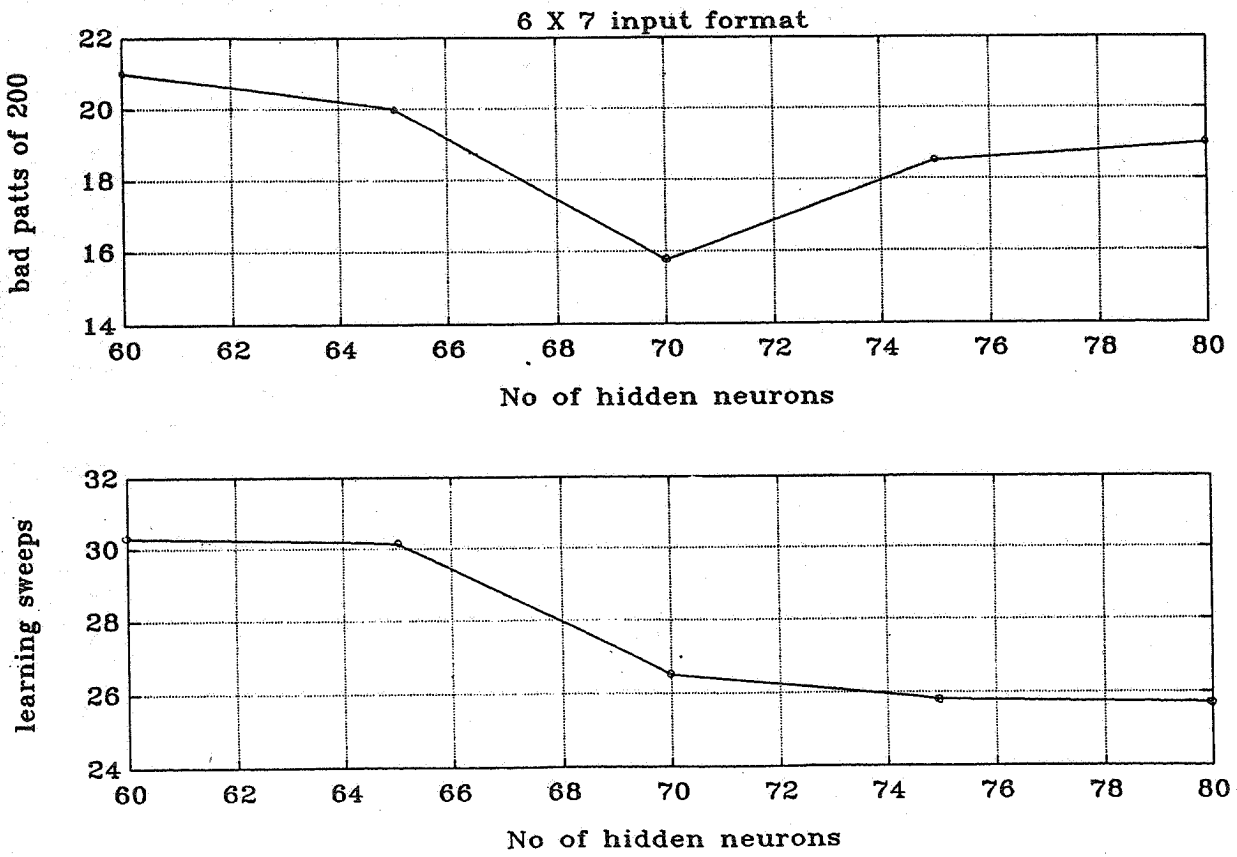
[DOR: 20.1001.1.22287698.1371.12.1.4.9]

[DOR: 20.1001.1.22287698.1371.12.1.4.9] [Downloaded from iutjournals.iut.ac.ir on 2024-07-23]



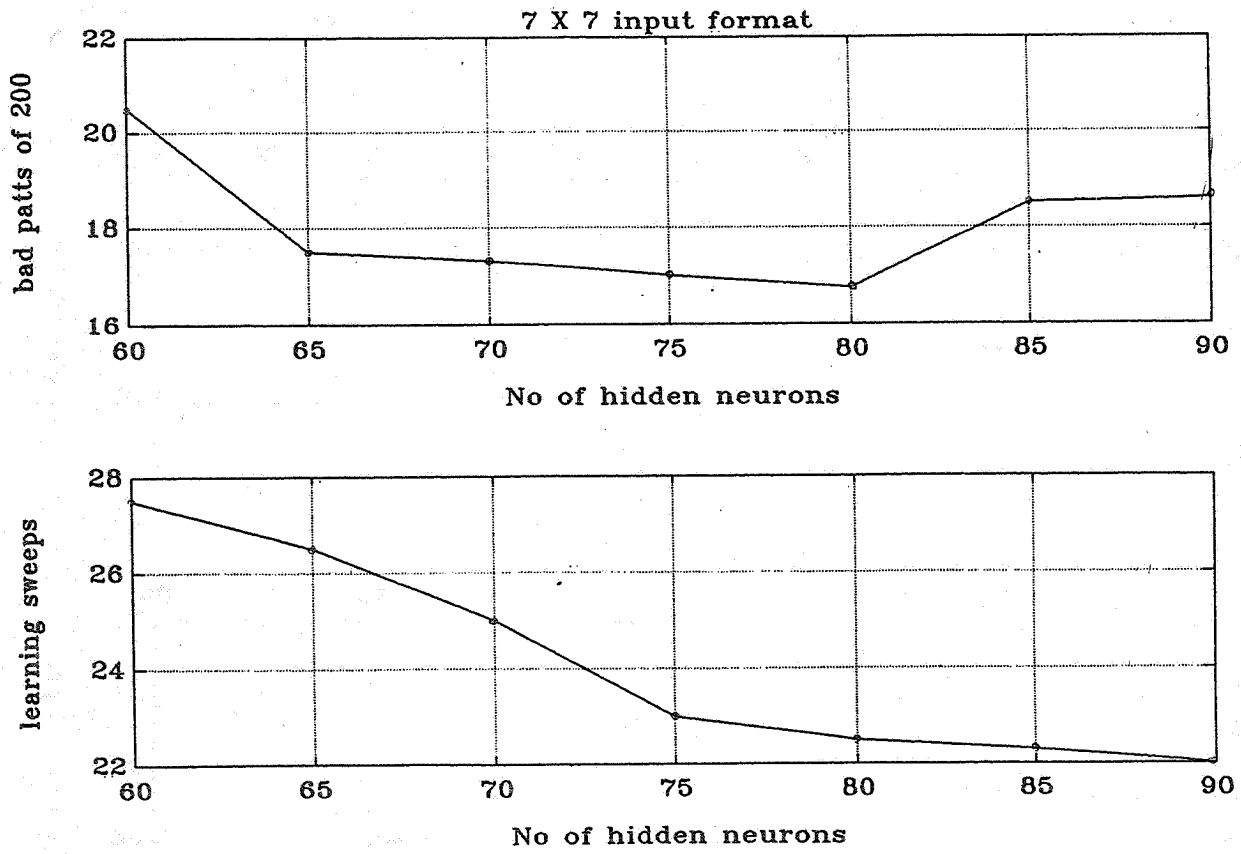
شکل ۵ـ منحنی تغییرات خطا و زمان همگرایی شبکه بافرمت ورودی ۶×۶، باتغییر تعدادنرونهای پنهان

[DOR: 20.1001.1.22287698.1371.12.1.4.9] [Downloaded from iutjournals.iut.ac.ir on 2024-07-23]



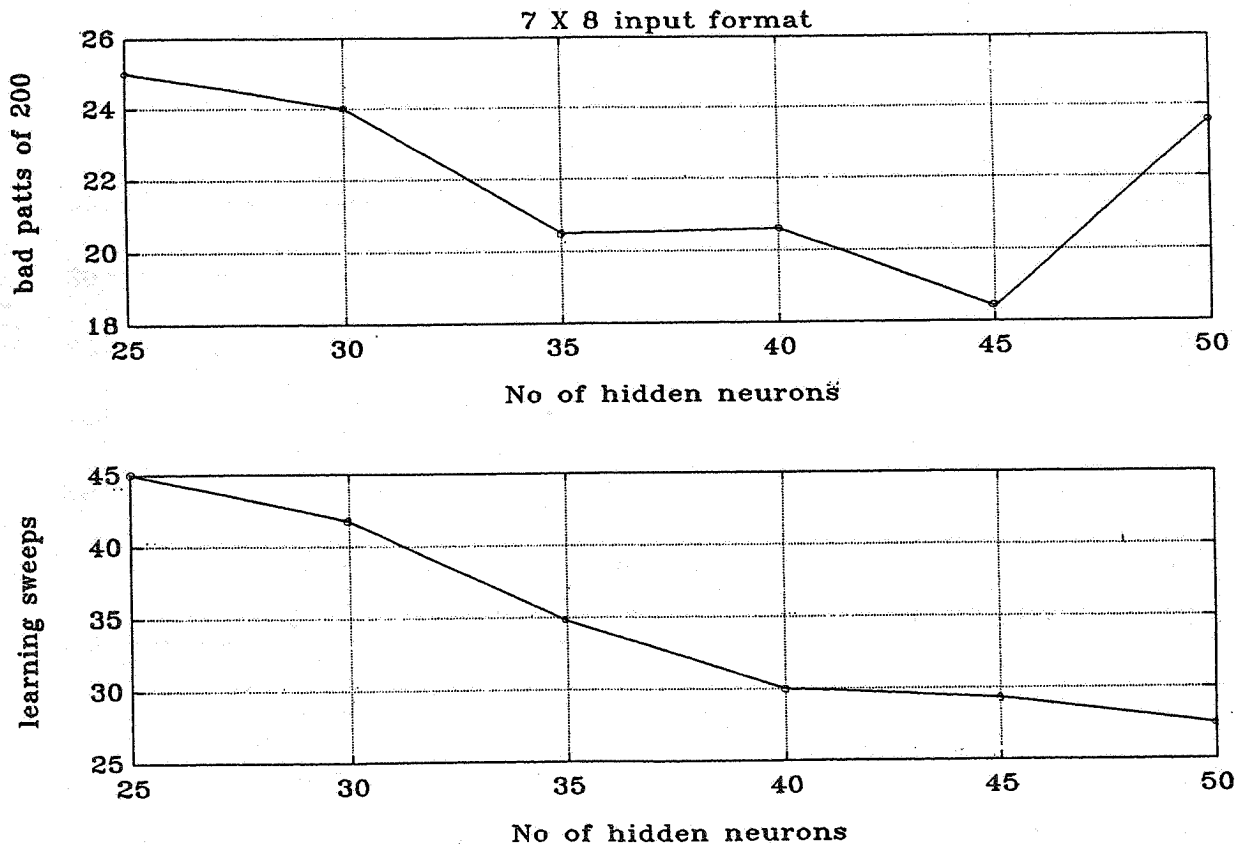
شکل ۶ـ منحنی تغییرات خطا و زمان همگرایی شبکه بافرمت ورودی ۷×۶، باتغییرتعدادنرونهای پنهان

[DOR: 20.1001.1.22287698.1371.12.1.4.9] [Downloaded from iutjournals.iut.ac.ir on 2024-07-23]



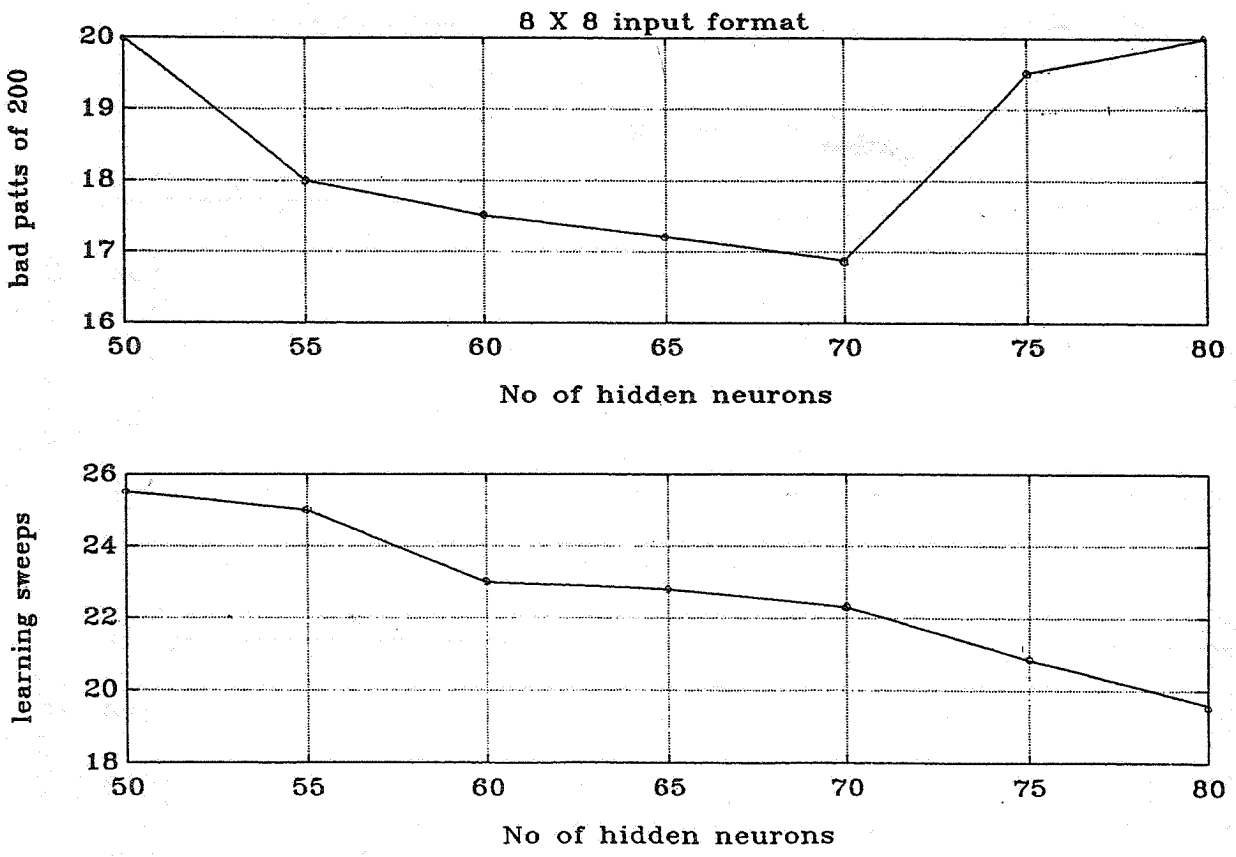
شکل ۷ـ منحنی‌تغییرات خطا و زمان همگرایی شبکه بافرمت ورودی ۷×۷، باتغییرتعدادنرونهای پنهان

[DOR: 20.1001.1.22287698.1371.12.1.4.9] [Downloaded from iutjournals.iut.ac.ir on 2024-07-23]



شکل ۸ـ منحنی تغییرات خطا و زمان همگرایی شبکه بافرمت ورودی ۸×۷، باتغییر تعدادنرونهای پنهان

[DOR: 20.1001.1.22287698.1371.12.1.4.9] [Downloaded from iutjournals.iut.ac.ir on 2024-07-23]



شکل ۹ـ منحنی تغییرات خطا و زمان همگرایی شبکه بافرمت ورودی ۸×۸، باتغییرتعدادنرونهای پنهان

برای انتخاب فرمت بهینهٔ ورودی، قدرت تعمیم و درون‌یابی شبکه‌های فوق را در تشخیص الگوهای طبقاتِ مختلف اندازه‌گیری کردیم. بدین‌منظور یک مجموعهٔ تست شامل ۷۵۰ الگو (از هر طبقه ۵۰ الگوی دستنویس) تهیه شد و درصد خطای شبکه‌ها در تشخیص الگوهای طبقاتِ مختلف به‌دست‌آمد. نتایج حاصل از این اندازه‌گیری در جدول ۲ آمده‌است. همان‌طور که از اعداد مندرج در این جدول پیداست، شبکهٔ با فرمت ورودی ۶×۶ به ازای الگوهای سه طبقه (۳، ۰، -)، شبکهٔ با فرمت ورودی ۷×۶ به ازای الگوی طبقه (-)، شبکهٔ با فرمت ورودی ۷×۷ به ازای الگوی طبقه (۳) و شبکهٔ با فرمت ورودی ۸×۷ به ازای الگوهای طبقاتِ (۰، ۳) خطای زیادی دارند. اما تغییرات خطا به ازای الگوهای طبقاتِ مختلف در شبکهٔ با فرمت ورودی ۸×۸ کمتراست، هم‌چنین درصد خطای کل شبکهٔ با فرمت ورودی ۸×۸ نیز از سایر شبکه‌ها کمتراست بنابراین فرمت ورودی ۸×۸ را به‌عنوان فرمت مناسب ورودی انتخاب می‌کنیم.

جدول ۱ـ شبکه‌های انتخاب شده به ازای فرمتهای ورودی مختلف

زمان همگرایی	تعداد خطا از ۲۰۰	لایه پنهان بهینه	فرمت ورودی
۳۴/۸	۱۵	۶۰ نرون	۶×۶
۲۶/۵	۱۵/۷۵	۷۰ نرون	۶×۷
۲۲/۵	۱۶/۷۵	۸۰ نرون	۷×۷
۲۹/۲۸	۱۸/۴۲	۴۵ نرون	۷×۸
۲۲/۳	۱۶/۸۷	۷۰ نرون	۸×۸

به‌منظور تعیین تعداد بهینهٔ نرونهای لایهٔ پنهان و هم‌چنین کاهش خطای شبکهٔ بهینه، تعداد الگوهای تربیتی را از ۵۲۰ الگو به ۸۰۰ الگو افزایش می‌دهیم و مجدداً با متغیرگرفتن تعداد نرونهای لایهٔ پنهان، منحنی تغییرات خطای شبکه‌های مختلف را به‌دست‌می‌آوریم (شکل ۱۰). همان‌طور که از این شکل پیداست می‌نیمم منحنی خطا به ازای ۴۰ نرون پنهان اتفاق می‌افتد و بنابراین تعداد بهینهٔ نرونهای لایهٔ پنهان برابر ۴۰ انتخاب‌می‌شود. بعد از تعیین تعداد بهینهٔ نرونهای لایهٔ پنهان، دقت شبکه را در تشخیص الگوهای مجموعهٔ تست شامل ۷۵۰ الگو (۱۵ الگو از هر طبقه) اندازه‌گیری کردیم که نتایج حاصل از این اندازه‌گیری برحسب درصد در جدول ۳ آمده‌است.

همانطور که از اعداد مندرج در جدول ۳ پیداست، متوسط میزان دقت شبکهٔ بهینه در تشخیص

[Downloaded from iutjournals.iut.ac.ir on 2024-07-23]

[DOR: 20.1001.1.22287698.1371.12.1.4.9]

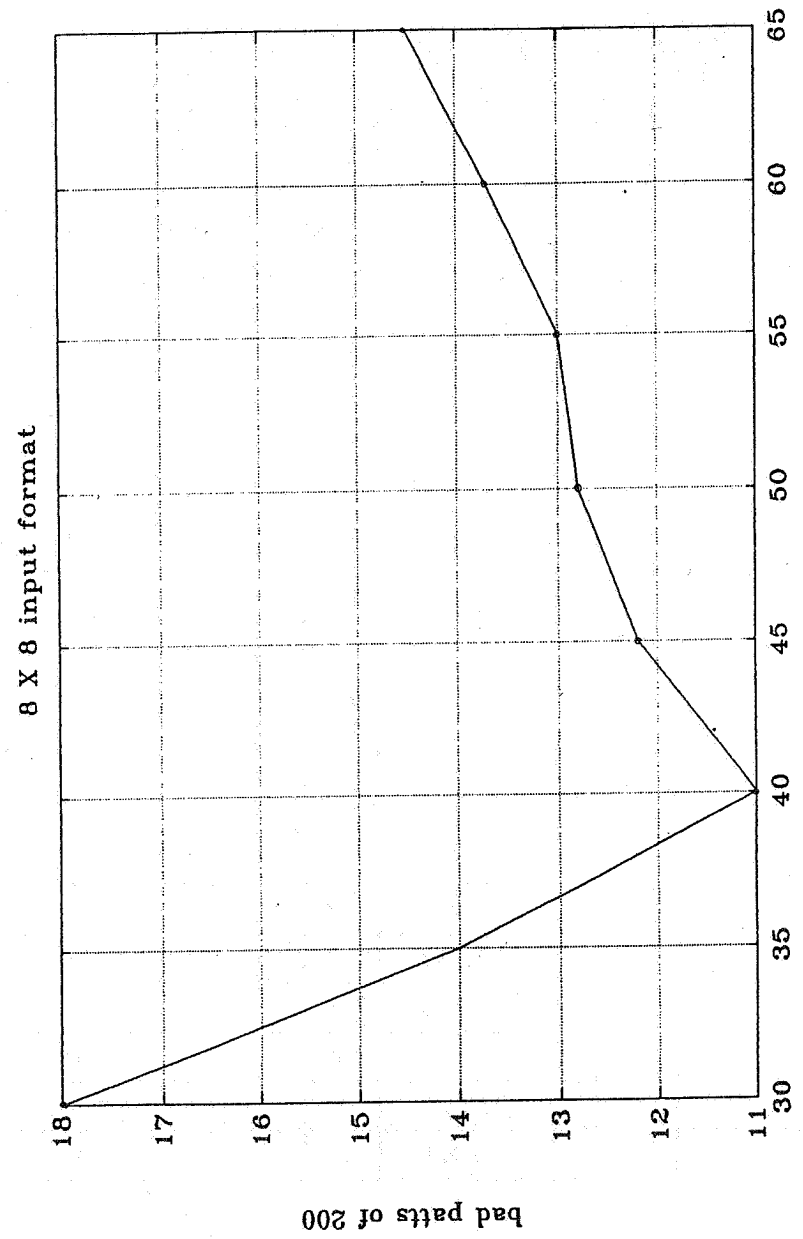
جدول ۲ـ درصد خطای شبکه‌های انتخاب‌شده روی الگوهای طبقاتِ مختلف

فرمت ۸×۸	فرمت ۷×۸	فرمت ۷×۷	فرمت ۶×۷	فرمت ۶×۶	کلاس
۱۶	۴۸	۲۶	۳۰	۵۲	۰
۱۶	۸	۱۰	۱۲	۱۰	۱
۲۰	۳۲	۲۸	۳۰	۲۰	۲
۳۰	۳۸	۴۲	۳۰	۴۸	۳
۱۶	۱۸	۱۶	۱۰	۸	۴
۱۶	۱۶	۲۶	۱۰	۸	۵
۲۴	۱۴	۱۴	۱۴	۱۶	۶
۰	۴	۱۰	۲	۰	۷
۱۸	۱۸	۱۴	۱۸	۰	۸
۶	۴	۴	۲	۲	۹
۲۴	۱۶	۱۴	۱۸	۱۴	+
۰	۰	۰	۵۸	۵۴	−
۲۶	۱۴	۲۲	۱۲	۲۲	×
۱۰	۲۸	۱۲	۲۲	۱۶	÷
۰	۰	۰	۲	۰	/
٪۱۴/۵	٪۱۷/۲	٪۱۵/۹	٪۱۸	٪۱۸	کل خطا

اعداد دستنویس فارسی در حدود ۹۵٪ است که نشان‌دهندهٔ توانایی بالای شبکه در تعمیم و تشخیص الگوهای دستنویس فارسی است.

[Downloaded from iutjournals.iut.ac.ir on 2024-07-23]

[DOR: 20.1001.1.22287698.1371.12.1.4.9]



شکل ۱۰-منحنی تغییرات خطای شبکه با فرمت ورودی ۸×۸

[DOR: 20.1001.1.22287698.1371.12.1.4.9]
[Downloaded from iutjournals.iut.ac.ir on 2024-07-23]

جدول ۳ـ دقت شبکه با فرمت ورودی ۸×۸ و ۴۰ نرون پنهان

متوسط دقت	/	÷	×	−	+	۹	۸	۷	۶	۵	۴	۳	۲	۱	۰	الگو
۹۴/۸	۱۰۰	۹۸	۹۰	۱۰۰	۹۸	۹۸	۹۰	۹۶	۹۶	۹۲	۹۰	۹۲	۹۲	۹۰	۱۰۰	٪دقت

ساده‌کردن و هرس‌کردن شبکهٔ بهینه

باتوجه به‌اینکه وزن اتصال بین دو نرون در یک شبکهٔعصبی معیاری از میزان همکاری و همبستگی بین دو نرون است، به‌نظرمی‌رسد که حذف وزنهای نزدیک صفر تأثیر چندانی بر عملکرد شبکه نگذارد. برای تحقیق در این امر، ابتدا وزنهای با قدرمطلق کوچکتر از ۰/۱، سپس ۰/۲ ، ۰/۳، ...،۱ از شبکهٔ بهینه حذف‌شدند و در هر مرحله خطای شبکه‌های حاصل روی یک مجموعهٔ تست شامل ۶۰۰ الگوی دستنویس اندازه‌گیری‌شد. نتایج حاصل از این اندازه‌گیریها در جدول ۴ آمده است.

جدول ۴ـ خطای حاصل از حذف وزنها روی ۶۰۰ الگوی تست

خطا با تلورانس ۰/۵	خطا	تعداد وزنهای حذف‌شده	سطح آستانهٔ حذف وزنها
۶۵	۴۰	۰	۰
۶۵	۴۰	۱۷۷	۰/۱
۶۹	۴۲	۳۶۵	۰/۲
۶۶	۴۲	۵۲۷	۰/۳
۷۱	۴۱	۷۱۵	۰/۴
۷۸	۴۱	۸۹۸	۰/۵
۷۸	۴۵	۱۰۷۲	۰/۶
۹۳	۴۷	۱۲۴۷	۰/۷
۱۱۰	۵۲	۱۳۹۲	۰/۸
۱۲۱	۵۱	۱۵۶۲	۰/۹
۱۳۶	۵۳	۱۶۸۶	۱

ستون دوم این جدول تعداد وزنهای حذف‌شده از شبکهٔ بهینه را نشان‌می‌دهد. در ستون چهارم نیز

[Downloaded from iutjournals.iut.ac.ir on 2024-07-23]

[DOR: 20.1001.1.22287698.1371.12.1.4.9]

تعداد تشخیص‌های غلط شبکه با درنظرگرفتن تلورانس ۰/۵ آمده‌است. در این صورت هنگامی یک الگوی دستنویس به‌طور صحیح تشخیص‌داده‌می‌شود که خروجی نرون مربوط به طبقه آن الگو از ۰/۵ بیشتر و خروجی نرونهای مربوط به طبقاتِ دیگر از ۰/۵ کمتر باشد. درغیر این صورت شبکه در تشخیص الگو دچار خطا خواهدشد. لازم به تذکر است که معیار تشخیص صحیح شبکه در تعلق الگوی x به طبقه c_i، که تا این مرحله مورد استفاده قرارمی‌گرفت ماکزیمم بودن خروجی نرون iام در لایهٔ خروجی نسبت به دیگر نرونهای لایهٔ خروجی بود و اعداد مندرج در ستون سوم جدول ۴ نیز بر این اساس به‌دست‌آمده‌اند. از اعداد مندرج در این جدول درمی‌یابیم که حذف وزنهای با قدرمطلق کوچکتر از ۰/۵ تأثیری بر عملکرد شبکه نمی‌گذارد (اعداد ستون سوم) اما میزان اعتماد به عملکرد شبکه کاهش می‌یابد (باتوجه به اعداد ستون چهارم)، بنابر این نتیجه می‌گیریم که تعداد ۸۹۸ وزن از کل ۳۲۱۵ وزن شبکهٔ بهینه قابل حذف است.

باتوجه به اینکه حذف تعدادی از وزنهای شبکه، بردار وزنها را از می‌نیمم تابع متوسط مربعات خطای شبکه روی مجموعهٔ تربیتی منحرف می‌کند، شبکه‌های حاصل از حذف وزنها مجدداً روی مجموعهٔ تربیتی شامل ۸۰۰ الگو تحت تربیت قرارگرفتند و پس از تربیت مجدد، خطای این شبکه‌ها در تشخیص ۶۰۰ الگوی مجموعهٔ تست اندازه‌گیری شد. نتایج حاصل از این اندازه‌گیری در جدول ۵ آمده‌است. همان‌طور که از اعداد مندرج در جدول پیداست، برای شبکه‌های با سطح آستانهٔ حذف وزن بیشتر از ۰/۵ تربیت مجدد باعث کاهش خطا گردیده اما برای شبکه‌های با سطح آستانهٔ کمتر از ۰/۵ تربیت مجدد تأثیری نداشته‌است.

اقدام دیگری که در جهت ساده‌سازی شبکهٔ بهینه انجام‌گرفت، کوانتیزه‌کردن وزنهای شبکه بود. بدین‌صورت که وزنهای شبکه به ترتیب با سه رقم اعشار، ۲ رقم اعشار، ۱ رقم اعشار و بدون رقم اعشار درنظر گرفته‌شد و خطای شبکه‌های حاصل از کوانتیزه‌کردن وزنها روی مجموعهٔ تست ۶۰۰ الگویی اندازه‌گیری شد که نتایج حاصل در جدول ۶ آمده‌است. همان‌طور که از اعداد مندرج در این جدول پیداست، با ۱ رقم اعشار نیز شبکه دارای خطای قابل قبول است که از این نکته می‌توان در پیاده‌سازی سخت‌افزاری شبکه استفاده‌کرد.

اندازه‌گیری قابلیت اعتماد تشخیص

برای جداسازی تشخیص‌های غیرقابل اعتماد شبکه، از آنتروپی خروجی می‌توان بهره جست [۹]. بدین‌ترتیب که ابتدا اندازهٔ خروجی نرونهای خروجی نسبت به مجموع آنها تراز می‌شود به

[Downloaded from iutjournals.iut.ac.ir on 2024-07-23]

[DOR: 20.1001.1.22287698.1371.12.1.4.9]

قسمی که بتوان خروجی نرونها را به عنوان احتمال وقوع طبقات مختلف (P_i) تفسیرکرد. سپس آنتروپی نرونهای لایهٔ خروجی طبق فرمول $H = -\sum_{i=1}^{15} P_i \ln(P_i)$ برای هر الگو محاسبه می‌شود. اگر مقدار آنتروپی از ۰/۷ بیشتر شد تصمیم طبقه‌بندی کننده عصبی قابل اعتماد نیست. اما اگر آنتروپی کوچکتر از ۰/۷ بود، تصمیم شبکه عصبی قابل قبول است. سطح آستانهٔ فوق به روش سعی و خطا و با مطالعهٔ تعداد زیادی الگوی دستنویس و تصمیم شبکه در برابر آنها تعیین‌شد که با

جدول ۵ـ خطای شبکه‌های دوباره تربیت‌یافته روی ۶۰۰ الگوی تست

خطاباتلورانس۰/۵	خطا	سطح‌آستانه‌حذف‌وزنها
۶۵	۴۰	۰/۱
۶۹	۴۰	۰/۲
۶۶	۴۰	۰/۳
۶۸	۴۱	۰/۴
۷۱	۳۹	۰/۵
۷۶	۳۷	۰/۶
۸۰	۳۶	۰/۷
۸۰	۳۹	۰/۸
۸۳	۴۰	۰/۹
۹۰	۵۷	۱

جدول ۶ـ خطای شبکه‌های حاصل از کوانتیزه‌کردن وزنها روی ۶۰۰ الگوی تست

تلورانس ۰/۵	خطا	ارقام اعشار
۶۵	۴۰	۳
۶۶	۴۱	۲
۶۵	۴۳	۱
۱۱۱	۵۳	۰

مقدار به‌دست‌آمده از تئوری نیز مطابقت دارد؛ زیرا حداکثر ابهام در تصمیم‌گیری طبقه‌بندی‌کننده

[Downloaded from iutjournals.iut.ac.ir on 2024-07-23]

[DOR: 20.1001.1.22287698.1371.12.1.4.9]

هنگامی اتفاق می‌افتد که مقادیر دو نرون خروجی نزدیک هم باشند. اگر در این حالت خروجی دیگر نرونها در حدود صفر فرض‌شود، خروجی ترازشدهٔ آن دو نرون در حدود ۰/۵ خواهدبود. در این حالت مقدار آنتروپی برابر خواهدشد با:

$$H = -2 \times 0.5 \ln 0.5 = 0.693$$

نتیجه‌گیری

در این مقاله پس از مروری بر مسئله تشخیص الگو و شبکه‌های عصبی مصنوعی، ساختار شبکه‌های چند لایهٔ تغذیه مستقیم انتخابی برای تشخیص اعداد دستنویس فارسی معرفی‌شد. آنگاه بهینه‌سازی ساختار فوق برای کاربرد موردنظر انجام‌گرفت که در نهایت یک شبکهٔ عصبی بهینه با دقت تشخیص ۹۵٪ به دست‌آمد. سپس ساده‌کردن شبکه و حذف تعدادی از وزنهای آن مورد توجه قرارگرفت. نتایج به دست‌آمده نشان‌می‌دهد که می‌توان تعداد قابل توجهی از وزنهای شبکهٔ بهینه را حذف‌کرد بدون اینکه در عملکرد شبکهٔ بهینه تأثیر چندانی داشته‌باشد. عملکرد شبکه در مقابله با نویز کوانتیزاسیون وزنها نیز مورد بررسی قرارگرفت. نتایج حاصله حاکی از اثر ناچیز این نویز بر عملکرد سیستم است.

[Downloaded from iutjournals.iut.ac.ir on 2024-07-23]

[DOR: 20.1001.1.22287698.1371.12.1.4.9]

مراجع:

1. Marzoco, F. N., "Computer Recognition of Handwritten First Names", *IEEE Trans. Electron. Computer*, Vol. 14,pp.210-217, Apr. 1965.

2. Suen, C. Y., Berthold, M. & Mori, S., "Automatic Recognition of Handprinted Characters, the State of the Art", *Proc. of IEEE*, Vol. 68, No. 4, pp. 469-487, 1980.

3. Tappert,C. C., Suen, C. Y. & Wakahara, T., "The State of the Art in on-line Handwriting Recognition", *IEEE Trans. on Pattern Analysis and Machine Intelligence,* Vol. 12, no. 8, pp. 787-804, Aug. 1990.

4. Mahmoud, M. Y., El-Hamalaway, M. A.& Flamy, A. A.,"A Statistical Approach For Arabic Character Recognition", *Proc. 9th National Con. &Exhib.*, pp. 1. 301-1. 307, Riyadh, Saudi Arabia, Oct. 1987.

5.Parhami, B. & Taraghi, M., "Automatic Recognition Of Printed FarsiText",*Pattern Recognition*, Vol. 14, No. 14, pp. 395-403, 1981.

6. Tolba, M. F., Wahab, S. M.& Salem, A., "A Recognition Algorithm For Arabic Printed Characters", *Proc. ISTED Int. Symp on Appliad Information*, Grindeweld, Switzerland, pp. 128-131, Feb. 1987.

7. Duda, R. O. & Hart, P. E., *Pattern Classification and Scene Analysis*, John Wiley, New York, 1973.

8. Nielsen, H. R., *Neuro Computing*, Addison Wseley, New York, 1990.

[Downloaded from iutjournals.iut.ac.ir on 2024-07-23]

[DOR: 20.1001.1.22287698.1371.12.1.4.9]

9. Beal, R.& Jackson, T., *Neural Computing, an Introduction,* Adam Hilger,Bristel,1990.

10. Pao, Y. H., *Adaptive Pattern Recognition and Neural Networks,* Addison Wesley, New York , 1989.

۱۱. سبحان منش، فریبرز "تشخیص اعداد دستنویس فارسی به کمک شبکه‌های عصبی"، پایان‌نامهٔ کارشناسی ارشد، دانشکده‌برق و کامپیوتر دانشگاه‌صنعتی‌اصفهان، بهمن‌ماه ۱۳۷۰.

[Downloaded from iutjournals.iut.ac.ir on 2024-07-23]

[DOR: 20.1001.1.22287698.1371.12.1.4.9]