محمد راوری

دانشجوی کارشناسی ارشد مهندسی پزشکی دانشگاه آزاد اسلامی واحد مشهد

فارغ التحصیل مهندسی برق(مخابرات) دانشگاه فردوسی مشهد

Email: mo_ravari@yahoo.com

آدرس منزل: مشهد، خیابان سناباد، خیابان مولوی جنوبی پلاک۱۱۲

تلفن: ۰۹۱۵۵۰۶۰۴۴۶

طبقه بندی خودکار اعداد دست نوشته با استفاده از شبکه های عصبی MLP

محمد راوری
دانشجوی کارشناسی ارشد مهندسی پزشکی
دانشکدهٔ مهندسی دانشگاه آزاد اسلامی مشهد
mo_ravari@yahoo.com

سعید راحتی قوچانی
استادیار گروه مهندسی برق
دانشکدهٔ مهندسی دانشگاه آزاد اسلامی مشهد
quchani@yahoo.com

کلمات کلیدی: بازشناسی الگو Pattern Recognition،شبکه های عصبی مصنوعی Artificial Neural Network(ANN)، پرسپترون چند لایه MultiLayer Perceptron (MLP) ، اعداد دست نوشته Handwritten Digits

چکیده:

شیوه ها و تکنیکهای زیادی جهت تشخیص ارقام digit recognition چه از نوع ماشینی و چه از نوع دست نوشته Handwritten Digits تا به حال بر روی آن کار شده است . در این مقاله ضمن معرفی بعضی از آنها ، به بررسی میزان عملکرد شبکه های عصبی MLP=Multilayer Perceptron با استفاده از الگوریتم تصحیح ضرایب Error Back propagation با سه لایه، که لایه ورودی ۶۴ گره ، لایه خروجی ۱۰ گره، و لایه وسطی(hidden layer) از ۱ تا 7 امتحان شده است که در حالت 7 لایه hidden و با بیش از 100 epochs برای دادهای آموزش Training به ٪۹۹,۵۸۶ ،برای دادهای اعتبارسنجی Validation به ٪۹۲,۲۸۳ ، برای داده های Writer-dependent به ٪۹۱,۰۹۲ و برای داده های Writer-independent به ٪۸۸,۱۴۷ دست پیداکردیم.

طبقه بندی خودکار اعداد دست نوشته با استفاده از شبکه های عصبی MLP

کلمات کلیدی: بازشناسی الگو Pattern Recognition،شبکه های عصبی مصنوعی Artificial Neural Network(ANN)، پرسپترون چند لایه MultiLayer Perceptron (MLP) ، اعداد دست نوشته Handwritten Digits

۱- مقدمه

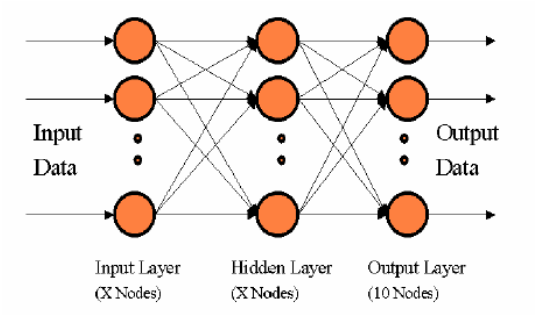
در این مقاله ، به بررسی طبقه بندی خودکار (automatic classification) اعداد دست نوشته (که این تشخیص دهنده می تواند جهت تشخیص اعداد تایپی نیز بکار رود که در آن حالت خیلی ساده تراست.) با استفاده از شبکه های عصبی MLP خواهیم پرداخت.

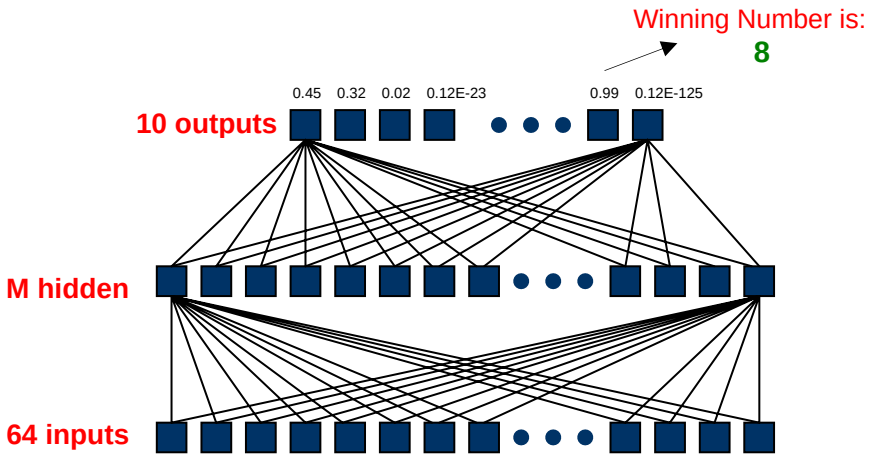
کاربردهای سیستمهای مبتنی بر بازشناسی اعداد (digit recognition) بسیار وسیع است. برای مثال:

۱. تشخیص خودکار کدپستی automated zip code recognition جهت مرتب کردن نامه ها بر اساس کدپستی مقصد
۲. چکهای بانکی bank checks
۳. فرمهای سرشماری census forms
۴. فرمهای مالیات tax forms
۵. دست نوشته های صفحه گسترده handwritten spreadsheets
۶. و به طور کلی در کلیه سیستمهای مبتنی بر OCR (Optical Character Recognition) کاربرد دارد.

که در این موارد می توانیم نمونه های(patterns) شیفت پیدا کرده (shifted)، تغییر مقیاس یافته(scaled)، کج و تغییرشکل یافته(distorted) با کمی اوریبی(skew) و حتی بر روی هم نوشته شده (overwritten) را پیدا کرد.

شیوه به کار گرفته شده بر پایه شبکه های عصبی چند لایه (multilayer neural network) [1] می باشد که با الگوریتم پس انتشار خطا (Back propagation) آموزش(train) می بیند.





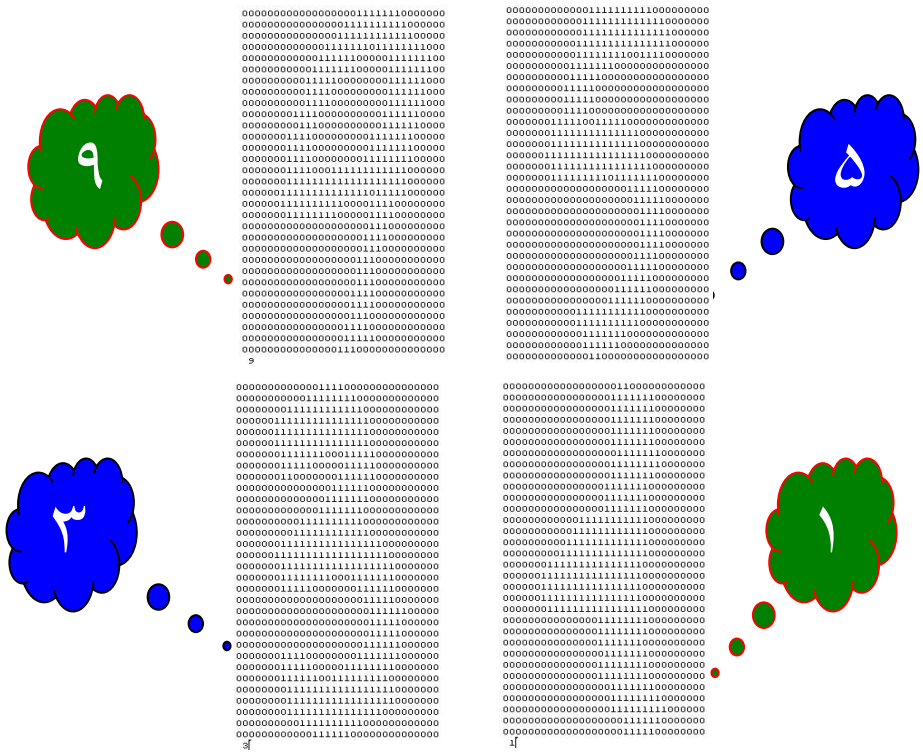
ساختار کلی شبکه

۲- مشخصات data set و پیش پردازشهای صورت گرفته

اصل این data set در سال ۱۹۹۴ توسط انستیتوی بین المللی استاندارد و تکنولوژی تهیه گردیده است.

National Institute of Standards and Technology (NIST)

که شیوه کار بدین ترتیب بوده است که از اشخاص خواسته شده بود بر روی فرم های از پیش چاپ شده (preprinted form) اقدام به نوشتن اعداد بکنند که این اعداد نوشته شده پس از اسکن و انتقال به کامپیوتر توسط برنامه پیش پردازشگر NIST به صورت تصاویر bitmap نرمالیزه در می آمده است. که نمونه ای از آن در زیر مشاهده می گردد.



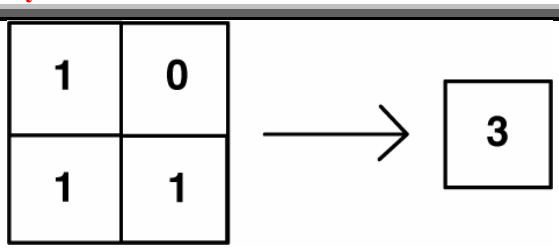
شامل اعداد دست نوشته توسط ۴۳ نفر می باشد. که از این ۴۳ نفر ۳۰ نفر برای training set و ۱۳ نفر دیگر برای test set در نظر گرفته شده است.

همانطور که در بالا مشاهده می شود اعداد به صورت تصاویر باینری ۳۲*۳۲ می باشد. که اگر خواسته باشیم این تصاویر را بدون پیش پردازش به ورودی شبکه عصبی داده شود تعداد گره های ورودی ۱۰۲۴=۳۲*۳۲ خواهد بود که این تعداد گره علاوه بر اینکه زمان آموزش شبکه را بشدن بالا می برد ، باعث حساسیت شدید شبکه نسبت به نویز نیز خواهد بود.

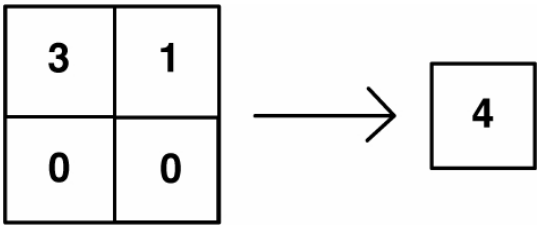
۱-۲ پیش پردازشهای صورت گرفته

جهت برطرف کردن مشکلات ذکر شده در بالا پیش پردازشی به صورت زیر بر روی تصاویر باینری انجام می شود.

بدین صورت که ابتدا هر چهار خانه مجاور را در نظر گرفته و آنها را به شیوه زیر به یک خانه تبدیل می کنیم که مینیمم آن ۰ و ماکزیمم ۴ می تواند باشد.



سپس پس از این مرحله در مرحله بعد مشابه حالت قبل هر چهارخانه مجاور را در نظر گرفته و آنها را به شیوه زیر به یک خانه تبدیل می کنیم که مینیمم آن صفر و ماکزیمم آن ۱۶ می باشد.

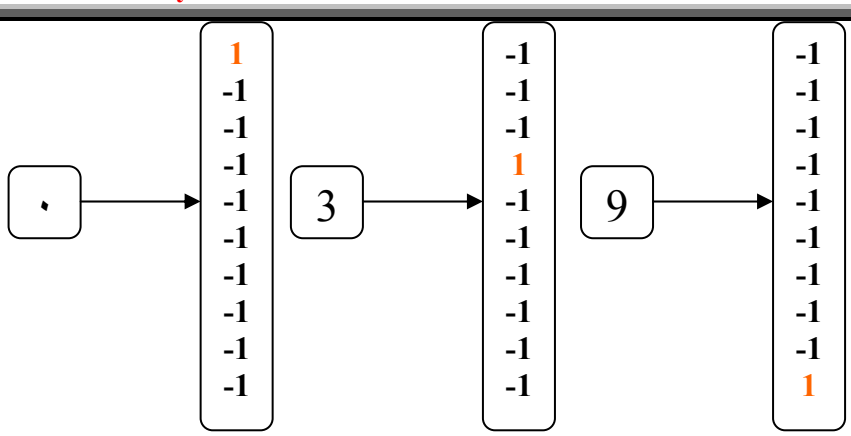


با اعمال پیش پردازش فوق تعداد ویژگیها (attributes) از ۱۰۲۴ به ۶۴ تقلیل می یابد. که این عمل باعث نصف شدن رزولوشن(Resolution is halved) و دوبرابر شدن عمق تصویر(depth is double)می شود.

که با این عمل تعداد گره های ورودی برای آموزش(train) کاهش یافته و پاسخ شبکه را نسبت به تغییرات جزئی(subtle differences) تصویر بدون تغییر می کند.

پیش پردازش دیگری که صورت گرفت ، این بود که با استفاده از دستور prestd در MATLB مقادیر ویژگی ها که بین ۱و۱۶ بود با توجه میانگین و انحراف معیار آنها نرمالیزه کردم تا هم شبکه سریعتر آموزش ببیند و هم اینکه مقادیر بسیار بزرگ یا کوچک سبب به هم خوردن روند آموزش شبکه نگردد.

و همچنین با توجه به اینکه خروجی شبکه دارای ۱۰ گره می باشد کلیه کلاسها را با توجه به اعداد ۰ تا ۹ که داشتند به ماتریسی ۱۰ تایی با ارقام ۱و۱-، که ۱ در موقعیت متناظر با رقم کلاس قرار می گرفت. که در زیر چند نمونه مثال آمده است:



۲-۲ تعداد نمونه ها و مشخصات آماری آنها

در کل این dataset دارای ۵۶۲۰ نمونه می باشد .

که از این تعداد ۳۸۲۳ نمونه در فایل optdigits.tra می باشد. که حدود نیمی از این داده ها را برای actual training ، یک چهارم آن برای اعتبارسنجی(validation) و یک چهارم باقیمانده برای writer-dependent testing مورد استفاده قرار گرفته است .

و مجموعه تست شامل ۱۷۹۷ نمونه در فایل optdigits.tes می باشد که برای writer-independent testing استفاده می شود که در واقع این تست چون نویسنده های آنها مستقل از نویسنده های داده های آموزش می باشد معرف کیفیت واقعی شبکه می باشد.

Name	*Number of Instances:*
optdigits-orig.traTraining	*1934*
optdigits-orig.cvValidation	*946*
optdigits-orig.wdepWriter-dependent	*943*
optdigits-orig.windepWriter-independent	*1797*

توزیع کلاسها یا همان اعداد در دادهای تست و آموزش

class	tra	cv	Wdep	Windep
0	189	87	100	178
1	198	97	94	182
2	195	92	93	177

3	199	85	105	183
4	186	114	87	181
5	187	108	81	182
6	195	87	95	181
7	201	96	90	179
8	180	91	109	174
9	204	89	89	180
sum	1934	946	943	1797

۳- روش آزمایش

با توجه به اینکه تابحال روشهای زیادی بر روی این DATASET اعمال گردیده است و نتایج آن اعلام گردیده است، در اینجا از شبکه های عصبی MLP با سه لایه استفاده گردیده است. و همانطور که می دانیم در عملکرد شبکه های عصبی عوامل متعددی تاثیرگذار است که در یک دید کلی به سه دسته می توان تقسیم کرد. که اولین عامل توپولوژی شبکه می باشد و بعد از آن نوع انتخاب اجزای شبکه و در نهایت الگوریتم یادگیری شبکه می باشد.

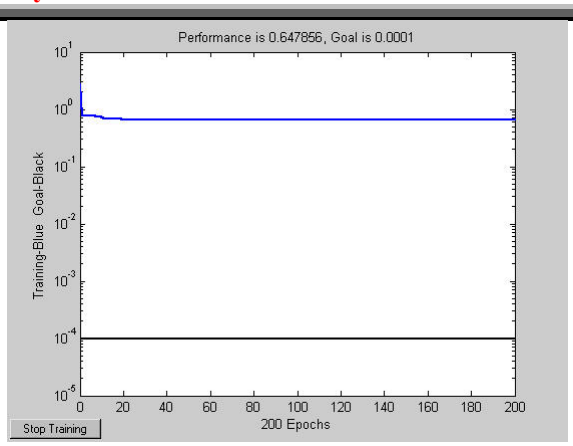
لذا با استفاده از نرم افزار MATLAB 7.1 و استفاده از toolbox ، شبکه های عصبی Neural Network بر روی سیستمی با پردازشگر 2.6MHz و با 512Mbyte RAM این dataset بررسی گردیده است.

٤- نتایج آزمایش

باتوجه به اینکه پس از پیش پردازش بر روی داده ها ما دارای ۶۴ ویژگی شدیم ، شبکه عصبی ما دارای ۶۴ گره ورودی و ۱۰ گره خروجی می باشد ، به گونه ای که در گره های خروجی گره ای برنده است که بیشترین مقدار را، بین ۱ و ۱- بتواند کسب کند.

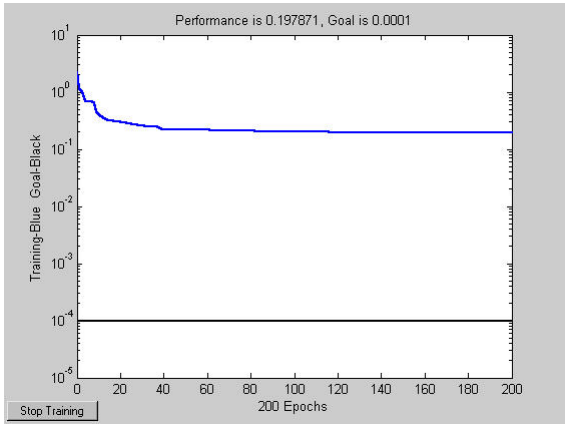
در ابتدا ما جهت آموزش شبکه کل داده های ۳۰ نفر یعنی ۳۸۲۳ نمونه را برای train شبکه استفاده می کنیم و برای test آن از داده های ۱۳ نفر دیگر یعنی ۱۷۹۷ نمونه استفاده می کنیم که این ۱۳ نفر متمایز از آن ۳۰ نفر می باشند یعنی این آزمایش را برای داده های جدید از افراد جدید امتحان می کنیم. یعنی همان writer-independent testing.

که در اینجا با تغییر تعداد لایه های پنهان (hidden layer) به میزان بهبود شبکه توجه می کنیم.



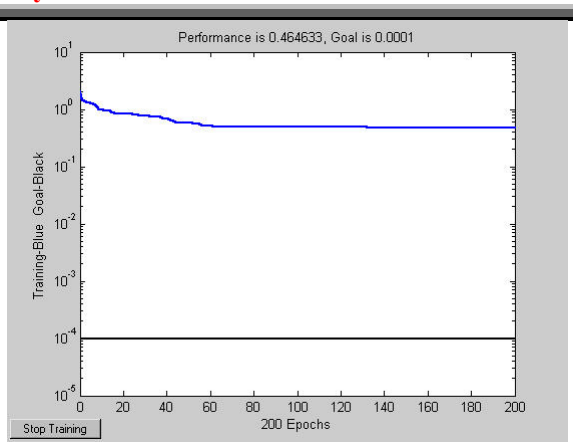
با یک لایه میانی که پس از 200 Epochs به نتایج زیر رسید.

	Train data	Test data
درصد موفقیت	۱۸,۷۵۵	۱۸,۰۸۶



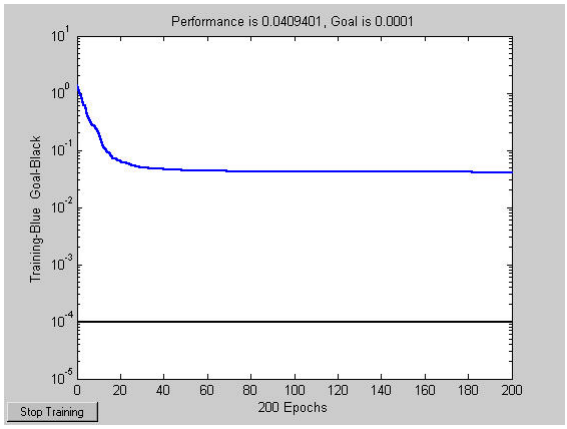
با دو لایه میانی که پس از 200 Epochs به نتایج زیر رسید.

	Train data	Test data
درصد موفقیت	۶۷,۴۳۴	60.991



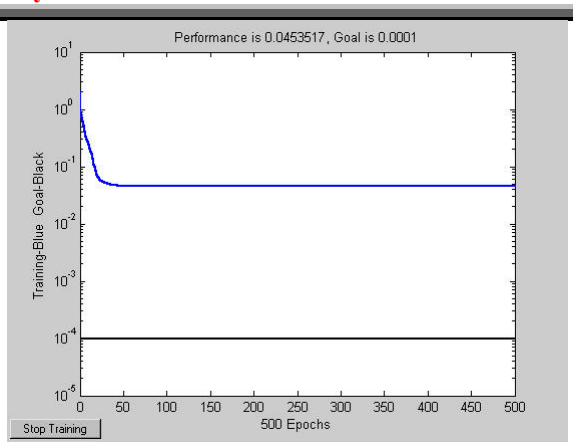
با سه لایه میانی که پس از 200 Epochs به نتایج زیر رسید.

	Train data	Test data
درصد موفقیت	۱۹,۵۱۳	6.956



با چهار لایه میانی که پس از 200 Epochs به نتایج زیر رسید.

	Train data	Test data
درصد موفقیت	۹۲,۲۰۵	۸٤,۵۳



با چهار لایه میانی که پس از 500 Epochs به نتایج زیر رسید.

	Train data	Test data
درصد موفقیت	93.042	***83.973***

که نتیجه آخری که با ۴ لایه میانی و 500 Epochs بود با توجه به PC مورد استفاده، Learn شبکه در حدود ۲ ساعت به طول انجامید.

که مراحل فوق به صورت خلاصه در زیر آمده است:

1	2	3	4	5
18.755	67.434	19.513	92.205	93.042

Training Test

1	2	3	4	5
18.086	60.991	6.956	84.53	83.973

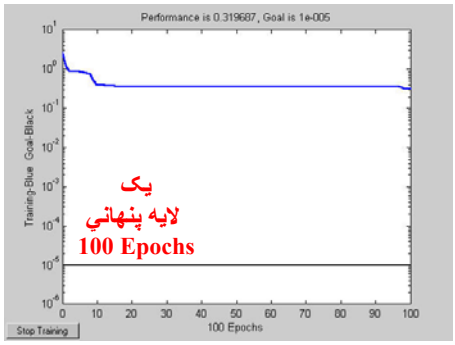
Writer-independent testing

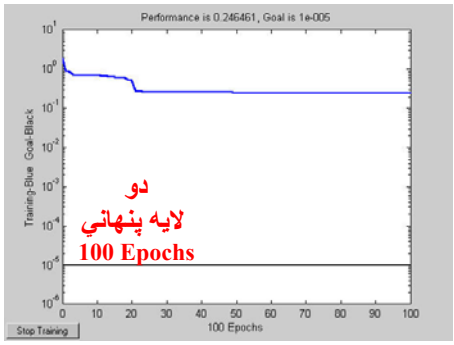
با توجه به زیاد بودن تعداد دادهای آموزشی ۳۸۲۳ نمونه، هنگامی که تعداد لایه های میانی را ۵ در نظرگرفتم با پیغام out of memory از جانب نرم افزار MATLAB مواجهه گردیدم.

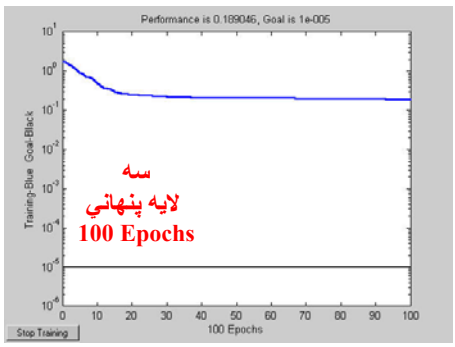
لذا در مرحله بعد ۳۸۲۳ نمونه را که مربوط به ۳۰ نفر می باشد به سه دسته مطابق جدول زیر تقسیم کردم و داده های ۱۳ نفر دیگر یعنی ۱۹۷۹ را برای writer-independent testing مورد استفاده قرار گرفت.

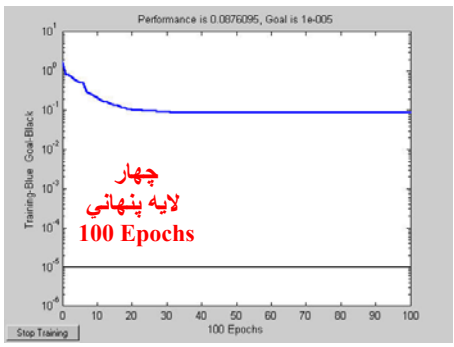
Name	Number of Instances:
optdigits-orig.traTraining	1934
optdigits-orig.cvValidation	946
optdigits-orig.wdepWriter-dependent	943
optdigits-orig.windepWriter-independent	1797

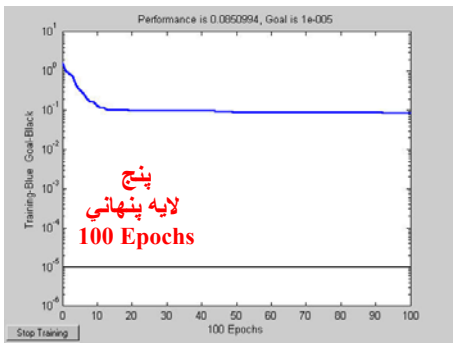
که نتایج آن به شرح زیر است:

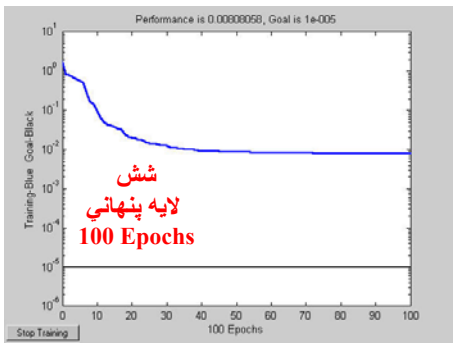


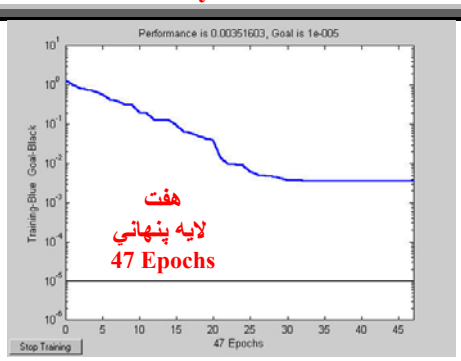


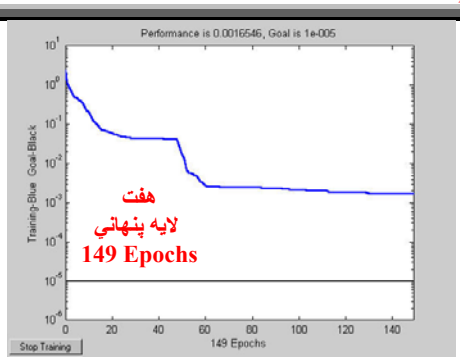












که در جدول زیر خلاصه نتایج آورده شده است. لازم است در اینجا نیز بیان شود که در این حالت باز هم با افزایش تعداد لایه های میانی به ۸ لایه با پیغام out of memory نرم افزار MATLAB مواجهه گردیدم.

نوع آزمایش	۱ لایه با 100 Epochs	2 لایه با 100 Epochs	3 لایه با 100 Epochs	4 لایه با 100 Epochs	5 لایه با 100 Epochs	6 لایه با 100 Epochs	7 لایه با 47 Epochs	7 لایه با 149 Epochs
Training	35.0052	48.9659	69.2347	79.9897	79.3692	98.1386	99.1210	99.586
Validation	33.4038	48.3087	62.5793	75.5814	76.0042	91.7548	91.2262	92.283
Writer-dependent	34.8887	44.1145	63.5207	74.1251	72.6405	89.7137	92.1527	91.092
Writer-independent	34.5019	44.8525	58.6533	75.2365	73.1219	88.2582	81.4135	88.147

۵- بحث و نتیجه گیری

از آنجایی که شیوه ها و تکنیکهای زیادی جهت تشخیص ارقام digit recognition چه از نوع ماشینی و چه از نوع دست نوشته Handwritten Digits تا به حال بر روی آن کار شده است . در این مقاله سعی بر آن بود که با استفاده از dataset موجود (Optical Recognition of Handwritten Digits) UCI Machine Learning Database/ (NIST)) ضمن تسلط پیداکردن بر شبکه های عصبی به بررسی میزان عملکرد شبکه های عصبی MLP=Multilayer Perceptron با استفاده از الگوریتم تصحیح ضرایب Error Back propagation با سه لایه، بپردازیم و به توانمندی آن به خاطر عملکرد غیرپارامتری و غیرخطی آن پی ببریم .

که با توجه به محدودیت های برنامه MATLAB و همچنین زیاد بودن داده های آموزشی training ، عملا نتوانستیم تعداد لایه ها را از ۷ لایه میانی افزایش دهم و با اینحال با ۷ لایه میانی و صرف حدود ۲ ساعت زمان با استفاه از کامپیوتری با مشخصات 2.6MHz و 512MByte RAM به نتایج قابل قبولی رسیدیم.

به طوریکه که در حالت 7 لایه hidden و با بیش از 100 epochs ، به ٪۹۹,۵۸۶ برای دادهای آموزش Training ، ٪۹۲,۲۸۳ برای دادهای اعتبارسنجی Validation ، ٪۹۱,۰۹۲ برای داده های Writer-dependent و ٪۸۸,۱۴۷ برای داده های Writer-independent دست پیداکردیم. که این میزان دقت با توجه به کم بودن تعداد لایه های میانی قابل توجه است.

٦- تشکر و قدردانی

در پایان از استاد گرامی، جناب آقای دکتر راحتی به خاطر راهنمایی های ایشان، و همچنین دکتر آذرنوش کمال تشکر را دارم.

۷- مراجع و منابع

- Handwritten Numeral Recognition via Neural Networks with Novel Preprocessing Schemes-Matthew Ziegler ziegler@virginia.edu Department of Electrical and Computer Engineering
- http://www.ece.virginia.edu/~mmz4s/papers/ECE682project_neuralnets.pdf
- Pattern Recognition and Data Clustering-J.-S. Roger Jang (張智星) CS Dept., Tsing Hua Univ., Taiwan - http://www.cs.nthu.edu.tw/~jang- jang@cs.nthu.edu.tw http://www.google.com/url?sa=T&start=16&url=http%3A//neural.cs.nthu.edu.tw/jang/slide/pr/pr.ppt
- D. C. Alvarez, F.M. Rodriguez, X.F. Hermida. "Printed and Handwritten Digits Recognition Using Neural Networks." Original publication source unkown; paper available at http://wgpi.tsc.uvigo.es/pub/papers/icsp98_1.pdf
- Help نرم افزار MATLAB
- dataset را از لینک زیر می توانید download کنید:

ftp://ftp.ics.uci.edu/pub/machine-learning-databases/optdigits/