
فصل اول:

کلیات

۱-۱- مقدمه

سنگ مخزن^۱ فضاهای خالی ظرف زیرزمینی نفت و گاز را تأمین می‌کند. این فضاها را خلل و فرج بین ذرات کانی‌ها ایجاد می‌کنند. هر سنگی که به اندازه کافی فضای خالی برای تجمع هیدروکربور داشته باشد و هنگامی که چاهی در آن حفر می‌شود سیال ذخیره شده را در چاه تخلیه کند می‌تواند سنگ مخزن باشد. سنگ‌ها کم و بیش دارای فضای خالی یا تخلخل می‌باشند ولی تنها سنگ‌هایی که دارای خلل و فرج بیشتر و مرتبط به یکدیگرند سنگ مخزن‌های مفید را تشکیل می‌دهند. بنابراین مهم‌ترین خاصیت فیزیکی سنگ مخزن دارا بودن تخلخل و تراوایی است. هر سنگ مدفون چه رسوبی^۲، چه آذرین^۳ و چه دگرگونی^۴ در صورت داشتن این ویژگی‌ها می‌تواند نقش سنگ مخزن را داشته باشد. در طبیعت بیشترین نفت و گاز در سنگ‌های ماسه‌ای و کربناته قرار دارند چون این سنگ‌ها دارای تخلخل و تراوایی کافی بوده و در حوضه‌های رسوبی، یعنی در بخشی از پوسته زمین قرار گرفته‌اند که نفت و گاز در آن تولید می‌گردد. علاوه بر این، سنگ‌های آذرین و دگرگونی اغلب به علت تبلور، فاقد تخلخل و تراوایی کافی می‌باشند [۱].

۱-۲- تخلخل و عوامل مؤثر بر آن

تخلخل هر سنگ نسبت حجم فضاهای خالی موجود در آن به حجم کلی سنگ است و با علامت ϕ نشان داده می‌شود. این نسبت بدون واحد است و معمولاً به صورت عدد کسری (کمتر از یک) و یا درصد بیان می‌شود.

تخلخل سازندهای زیرسطحی می‌تواند در بازه وسیعی تغییر کند. کربنات‌های متراکم (سنگ آهک و دولومیت) و تبخیری‌ها (نمک، انیدریت، ژیپس و ...) ممکن است تخلخل صفر را نشان دهند.

1- Reservoir rock
2- Sedimentary
3- Igneous
4- Metamorphic

ماسه‌سنگ‌های خوب تحکیم‌یافته، تخلخلی بین ۱۰ تا ۱۵٪ دارند. ماسه‌سنگ‌های تحکیم‌نیافته تخلخلی حدود ۳۰٪ یا بیشتر دارند. شیل‌ها و رس‌ها ممکن است بیش از ۴۰٪ تخلخل داشته باشند [۳].

لورسن^۱ تخلخل ۵-۰٪ را ناچیز، سنگ مخزن با تخلخل ۱۰-۵٪ را ضعیف، ۱۵-۱۰٪ را متوسط، ۲۰-۱۵٪ را خوب و بیش از ۲۰٪ را عالی محسوب می‌کند [۱].

از جمله عوامل مؤثر در تخلخل عبارتند از:

۱- نحوه کنار هم قرار گرفتن دانه‌ها: با فرض کروی بودن دانه‌ها و یکسان بودن اندازه آن‌ها، سیستم تبلور بر مقدار تخلخل تأثیر می‌گذارد، به طوری که در سیستم مکعبی تخلخل در حدود ۴۸٪، در انباشتگی رومبوئدریک در حدود ۴۰٪ و در انباشتگی هگزاگونال در حدود ۲۶٪ محاسبه می‌گردد [۱۶].

۲- اثر ابعاد و جورشدگی^۲ دانه‌ها: در صورتی که سنگ از ذرات با ابعاد مختلفی تشکیل شده باشد، دانه‌های ریز فضای بین ذرات درشت را پر نموده و در نتیجه تخلخل کاهش می‌یابد [۱۶].

۳- اثر سیمان‌شدگی: در برخی از سنگ‌ها، دانه‌ها توسط خمیره‌ای از جنس سیلیس یا آهک به یکدیگر متصل می‌شوند که در این صورت تخلخل می‌تواند حتی تا حدود صفر درصد نیز کاهش یابد [۱۶].

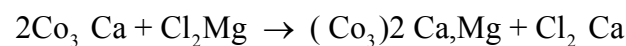
۴- اثر گردشگی دانه‌ها: درجه صیقلی بودن و زاویه‌دار بودن دانه‌ها در تخلخل اثر به‌سزایی دارد به طوری که هر چه ذرات صیقلی‌تر (گردتر) باشند حجم فضاهای خالی بین آن‌ها بیشتر از حالتی خواهد بود که دانه‌ها زاویه‌دار بوده و در هم فرو روند [۱۶].

۵- اثرات انحلال ذرات: حلالیت کانی‌ها توسط چرخه آب، تخلخل را افزایش می‌دهد به طوری که در سنگ‌های کربناته این امر عامل مهمی برای ایجاد و توسعه فضاهای خالی است. اثر مذکور غالباً در سنگ‌هایی که منشاء شیمیایی یا بیوشیمیایی دارند دیده می‌شود. تخلخل ناشی از فرایند انحلال را تخلخل حفره‌ای^۳ می‌نامند [۱۶].

1- A.I.Levorsen
2- Sorting
3- Vuggy porosity

۶- اثرات ناشی از تنش‌های تکتونیکی: وجود تنش‌های تکتونیکی در یک منطقه سبب ایجاد درزه، شکاف و شکستگی در سنگ می‌شود. تخلخل ایجاد شده در اثر شکاف، زیاد نیست و از ۱-۲٪ تجاوز نمی‌کند ولی وجود درزه و شکاف در سنگ مخزن سبب ازدیاد تراوایی می‌گردد. این امر در مورد سنگ مخزن‌های کربناته ایران نیز صادق است. در مخازنی که نفت از سنگ‌های آذرین و دگرگونی استخراج می‌شود، نفت تنها در درزه‌ها و شکاف‌ها جمع می‌شود مانند میدان نفتی اماریلو در تگزاس [۱ و ۱۶].

۷- دولومیتی شدن: سنگ مخزن دولومیتی، توان تولیدی بیشتری نسبت به سنگ مخزن آهکی دارد. این امر به علت تخلخل و تراوایی بیشتری است که اغلب، سنگ دولومیت نسبت به سنگ آهک داراست. به این جهت مکانیسم دولومیتی شدن سنگ آهک که سبب ازدیاد تخلخل می‌گردد مورد توجه پژوهشگران نفتی بوده است. فرمول این تبدیل به صورت زیر است:



در نتیجه این تبدیل، تخلخل به میزان ۱۱/۲٪ افزایش می‌یابد [۱].

۱-۳- انواع تخلخل

۱- تخلخل کل یا مطلق (ϕ_t): نسبت حجم کلیه فضاهای خالی موجود در سنگ به حجم کل آن است. تخلخل کل ممکن است اولیه^۱ (ϕ_1) یا ثانویه^۲ (ϕ_2) و یا ترکیبی از این دو باشد. تخلخل اولیه حجم آن دسته از فضاهای خالی است که در حین تشکیل سنگ بین دانه‌ها و یا بلورهای تشکیل دهنده سنگ به وجود می‌آید و به این جهت به تخلخل بین دانه‌ای^۳ و بین بلوری^۴ نیز معروف است. مقدار این تخلخل به شکل، اندازه و جورشدهگی ذرات تشکیل دهنده بستگی دارد. این نوع تخلخل غالباً در سنگ‌های آواری یا تخریبی مشاهده می‌شود. تخلخل ثانویه درصدی از فضاهای خالی موجود در سنگ است که بعد از تشکیل آن، بر اثر تنش‌های تکتونیکی و یا فرایندهای انحلال (شیمیایی و بیوشیمیایی) به وجود می‌آید [۱۶].

-
- 1- Primary porosity
 - 2- Secondary porosity
 - 3- Intergranular porosity
 - 4- Inter crystalline porosity

۲- تخلخل پیوسته^۱ (φ_{con}): درصدی از کل حفره‌های موجود در سنگ است که به همدیگر مرتبطند. در صورت عدم ارتباط حفره‌ها به یکدیگر، مقدار آن می‌تواند بسیار کمتر از تخلخل مطلق باشد، مثل پومیس (سنگ پا) که نزدیک به ۵۰٪ تخلخل مطلق دارد ولی تخلخل پیوسته آن نزدیک به صفر است [۱۶].

۳- تخلخل نهفته^۲ (φ_{pot}): قسمتی از تخلخل پیوسته است که به علت باریکی بیش از حد مجاری ارتباطی (موئینه بودن)، سیال نمی‌تواند از آن عبور کند. به عنوان مثال در مخازن نفتی و گازی حداقل قطر فضاهای موئینه به ترتیب باید بیش از ۲۰ و ۵ میکرومتر باشد تا امکان حرکت سیالات در آن‌ها فراهم گردد. در شیل‌ها که تخلخل پیوسته بسیار بالایی دارند (۴۰ تا ۵۰ درصد در شیل‌های متراکم و تا ۹۰ درصد در رسوبات لجنی جوان) اندازه کانال‌های ارتباطی به حدی موئینه است که سیالات موجود به علت خاصیت جذب مولکولی، نمی‌توانند حرکت نمایند [۱۶].

۴- تخلخل مؤثر^۳ (φ_e): تخلخل پیوسته‌ای است که در آن امکان حرکت سیال و در نتیجه دسترسی و استحصال آن وجود دارد و بنابراین تخلخل ناپیوسته، خلل و فرج اشغال شده توسط آب جذبی و آب محصور در شیل‌ها را شامل نمی‌شود. با استفاده از نمونه‌برداری از مغزه‌ها و یا اطلاعات نمودارهای چاه با اعمال اصلاحات لازم می‌توان مقدار تخلخل مؤثر را برآورد نمود. در واقع این نوع تخلخل، مورد نظر مفسر نمودارهای چاه‌نگاری است زیرا تراوایی سنگ و توان بهره‌دهی آن به میزان تخلخل مؤثر بستگی دارد [۱، ۲، ۱۶].

در جدول ۱-۱ تخلخل کل و تخلخل مؤثر بعضی از سنگ‌ها آمده است. این ارقام صرفاً مقادیر شاخص هستند و ممکن است تا حد قابل توجهی تغییر کنند.

جدول ۱-۱: تخلخل کل و تخلخل مؤثر بعضی از سنگ‌ها [۲].

سنگ	تخلخل کل	تخلخل مؤثر
شن	٪۴۵	٪۴۰
ماسه	٪۴۰	٪۳۰
ماسه سیلتی	٪۳۲	٪۵
سیلت	٪۳۶	٪۳
رس	٪۴۷	٪۰
چاک	٪۳۰	٪۲-۵
ماسه سنگ درز و شکافدار	٪۲۰	٪۲-۱۵
گرانیت درز و شکافدار	٪۲	٪۰/۱-۲
سنگ آهک درز و شکافدار	٪۰/۵-۱۷	٪۰/۱-۱۲

- 1- Connected porosity
2- Potential porosity
3- Effective porosity

۱-۴- اندازه‌گیری تخلخل

۱-۴-۱- اندازه‌گیری مستقیم از روی مغزه‌ها

برای یک نمونه متخلخل سه بخش قابل اندازه‌گیری وجود دارد که عبارتند از حجم کلی سنگ (V_b)، حجم خمیره یا ماتریکس سنگ (V_{ma}) و حجم منافذ و خلل و فرج سنگ (V_p). برای تعیین تخلخل مغزه، بایستی دو تای این کمیت‌ها محاسبه شود. چندین روش آزمایشگاهی برای محاسبه تخلخل مغزه وجود دارد که روش انتخابی به نوع سنگ و زمان لازم برای به دست آوردن اطلاعات بستگی دارد [۱۲]. در ادامه به توضیح اجمالی بعضی از آن‌ها می‌پردازیم.

۱- روش مجموع سیالات: در این روش از دو نمونه کاملاً یکسان سنگ استفاده می‌شود. حجم کلی یکی از نمونه‌ها با غوطه‌ور کردن آن در جیوه و به دست آوردن وزن آن در این حالت و استفاده از قانون ارشمیدس به دست می‌آید. حجم آب و نفت موجود در یکی از نمونه‌ها با استفاده از تقطیر تعیین می‌شود. در نمونه دوم، حجم منافذ اشغال شده از گاز با تزریق سیال به درون آن محاسبه می‌شود. مجموع این سه حجم V_p را به دست می‌دهد. این روش دقت بالایی ندارد [۱۲ و ۲۹].

۲- روش قانون بویل- ماریوت: در این روش نمونه‌ای خشک از سنگ، در ظرفی قرار می‌گیرد و ظرف با مقداری گاز پر می‌شود. سپس با انبساط گاز و استفاده از قانون انبساط گاز بویل- ماریوت ($PV = c$) حجم ذرات سنگ به دست می‌آید. اگر نمونه خوب خشک نباشد، تخلخل کمتر از مقدار واقعیش تعیین خواهد شد [۱۲، ۲۹، ۳۵ و ۳۷].

۳- روش تزریق جیوه: در این روش، جیوه تحت فشار زیاد به نمونه تزریق می‌شود. حجم جیوه تزریق شده به صورت تابعی از فشار اندازه‌گیری می‌شود. حجم تزریقی در بیشینه فشار، V_p است. اگر بیشینه فشار برای اشباع کامل نمونه از جیوه کافی نباشد، تخلخل کمتر از مقدار واقعیش تعیین خواهد شد [۲۹ و ۳۷].

۴- خرد کردن^۱ نمونه: در این روش نمونه به طور کامل خرد می‌شود و حجم ذرات سنگ به طور دقیق اندازه‌گیری می‌شود. اگر با خرد کردن، منافذ و خلل و فرج سنگ به طور کامل از بین نرود، تخلخل کمتر از مقدار واقعیش تعیین خواهد شد [۲۹].

1- Disaggregation

این که در یک اندازه‌گیری آزمایشگاهی، کدام یک از انواع تخلخل تعیین شده است به روش مورد استفاده و نحوه آماده‌سازی نمونه بستگی دارد.

در ارتباط با اندازه‌گیری‌های مغزه مسائل و مشکلاتی وجود دارد. مغزه‌گیری و به دست آوردن اطلاعات از مغزه مشکل، وقت‌گیر و بسیار پرهزینه است و به همین دلیل محاسبه تخلخل از روی مغزه در همه چاه‌ها کاری غیرمعمول است. تغییر شرایط مؤثر روی مغزه در آزمایشگاه نسبت به عمق برداشت و در نتیجه تغییر در خصوصیات فیزیکی آن باید مورد توجه قرار گیرد. همچنین ممکن است خواص مغزه کوچک برداشته شده در یک عمق معین، نشان‌دهنده خواص حجم کافی از سازند در این عمق نباشد و با هم متفاوت باشند [۱۹]. علاوه بر این، اندازه‌گیری‌های آزمایشگاهی بر روی مغزه در معرض خطا قرار دارند [۲۹]. بر حسب اهمیت و تأثیرگذاری، یک کارشناس ممکن است از برخی از این موارد و خطاها صرف نظر کند و برخی را مورد توجه قرار دهد و در صدد تصحیح آن‌ها برآید.

۱-۴-۲- اندازه‌گیری غیرمستقیم از روی داده‌های نگارها

ثبت پیوسته تغییرات یک مشخصه فیزیکی بر حسب عمق در سازندهایی که چاه از آن‌ها عبور می‌کند را نگار^۱ می‌نامند [۱۹]. در اکتشافات نفتی، از چاه‌نگاری برای ارزیابی دقیق خصوصیات و ویژگی‌های مخازن نفتی از قبیل تخلخل استفاده می‌شود. علاوه بر سهولت و مزایای اقتصادی روش‌هایی که تخلخل را بر اساس اندازه‌گیری‌های نگارها تعیین می‌کنند، این روش‌ها اطلاعات را به صورت پیوسته در اختیار کارشناس قرار می‌دهند و در یک عمق مشخص از حجم بزرگتری از سازند تأثیر می‌پذیرند [۱۹]. نگاری وجود ندارد که مستقیماً بتواند تخلخل را محاسبه کند اما متخصصین از روش‌های مختلفی استفاده می‌کنند تا تخلخل را از داده‌های نگارها محاسبه کنند. با استفاده از داده‌های یک نگار، یا با استفاده از ترکیبی از نگارها و یا با استفاده از تحلیل‌های آماری داده‌های چندین نگار می‌توان تخلخل را به دست آورد. با استفاده از یک نگار فقط می‌توان تخلخل را محاسبه کرد ولی با استفاده از نگارهای بیشتر علاوه بر تخلخل می‌توان درصد کانی‌ها را نیز به خوبی تعیین نمود [۲۹]. در این جا چندین نگار را که متأثر از تخلخل هستند و از آن‌ها برای محاسبه این کمیت استفاده می‌شود، شرح می‌دهیم.

1- Log

۱-۲-۴-۱- نگار صوتی

در این نوع نگار، ثبت پارامترهایی مورد نظر است که در ارتباط با فرستادن پالس‌های موج به داخل سازند به دست می‌آید. یکی از این پارامترها، مدت زمان عبور موج و رسیدن آن از فرستنده به گیرنده است. زمانی که به وسیلهٔ سوند ثبت می‌شود (T_t) تابع سه عامل اصلی زیر است:

۱- مدت زمان سیر موج در خمیره سنگ (T_{ma}).

۲- مدت زمان سیر موج در سیال پرکننده فضاهای خالی (T_f).

۳- حجم خلل و فرج سنگ و یا همان تخلخل (φ).

بنابراین می‌توان نوشت [۱۶]:

$$T_t = (1-\varphi)T_{ma} + \varphi T_f \quad (1-1)$$

با دانستن نوع خمیره و سیالی که فضاهای خالی را پر کرده، می‌توانیم تخلخل را پیدا کنیم. رابطهٔ فوق برای سازندهای فشرده و تمیز فاقد شیل صادق است، در غیر این صورت تصحیحات لازم باید انجام شود.

چون موج کوتاهترین مسیر را انتخاب می‌کند، در صورتی که در سازند حفره و شکستگی (تخلخل ثانویه) وجود داشته باشد، موج صوتی تمایل به نادیده گرفتن آن‌ها و عبور از نقاط دیگر را داشته و بنابراین نگار صوتی فقط تخلخل اولیه را ثبت می‌کند.

در صورتی که سازند فشرده و متراکم نباشد از ضریب فشردگی (C_p) که معمولاً بین ۰/۸ تا ۱/۲ تغییر می‌کند، استفاده می‌کنیم و φ را در $1/C_p$ ضرب می‌کنیم [۱۶].

اگر سازند با ناخالصی‌های شیلی که تخلخل ناتراوای زیادی دارند همراه باشد، این تخلخل غیرمؤثر باید حذف شود. ابتدا حجم شیل (V_{sh}) با استفاده از نگارهای رادیواکتیو تعیین می‌شود و سپس رابطهٔ تخلخل به صورت زیر تغییر می‌کند [۱۶]:

$$T_t = [(1-\varphi)-V_{sh}]T_{ma} + \varphi T_f + V_{sh}T_{sh} \quad (2-1)$$

با توجه به سرعت موج در سازند و فرکانس آن، شعاع بررسی نگار بین ۱۲ سانتی‌متر تا حدود یک متر تغییر می‌کند. در مناطق آبدار پدیده تهاجم اثری در نگار صوتی ندارد ولی در مناطق هیدروکربوری، وجود نفت یا گاز سبب می‌شود که زمان ثبت شده بیشتر شود و در نتیجه، تخلخل بیشتر از مقدار واقعی آن تعیین شود. در این موارد به منظور تصحیح تخلخل از روابط زیر استفاده می‌شود:

$$\varphi_{gas} = \varphi_{sonic} \times 0.7 \quad (3-1)$$

$$\varphi_{oil} = \varphi_{sonic} \times 0.9 \quad (4-1)$$

علاوه بر این موارد، عوامل محیطی شامل دما، فشار و اثرات چاه (قطر چاه، نوع سیال حفاری، کیک گل، کیفیت دیواره چاه و موقعیت دستگاه در چاه) نیز بر قرائت‌ها تأثیر می‌گذارد و باید مورد توجه قرار گیرد و در صورت نیاز تصحیح شود [۱۶ و ۱۹].

۱-۲-۲-۴- نگار چگالی یا نگار گاما-گاما^۱

وقتی که پرتوهای گامای مصنوعی به سازند تابیده می‌شوند، بر حسب انرژی‌شان می‌توانند سه نوع برهم‌کنش با سازند داشته باشند: اثر فتوالکتریک، پراکندگی کامپتون و تولید جفت.

در نگار چگالی از اثر کامپتون استفاده می‌شود. پرتو گامای با انرژی متوسط در برخورد با سازند دچار اثر کامپتون می‌شود. گامای پراکنده شده توسط گیرنده‌ها دریافت می‌شود و شمارش می‌گردد. دستگاه، چگالی الکترونی^۲ (ρ_e) سازند را بررسی خواهد کرد زیرا هر چه تعداد برخوردها بیشتر و یا به عبارت دیگر تعداد الکترون‌ها در واحد حجم سازند زیادتر باشد، شدت ثبت شده کمتر خواهد بود. از طرف دیگر، چگالی الکترونی نسبت مستقیم با چگالی سازند (ρ_b) دارد [۱۶]:

$$\rho_e = \left(\frac{2Z}{A} \right) \cdot \rho_b \quad (5-1)$$

در این رابطه Z عدد اتمی و A جرم اتمی عناصر و ترکیبات تشکیل دهنده سنگ است. در اغلب عناصر و ترکیبات، چگالی الکترونی به چگالی ظاهری حاصل از دستگاه خیلی نزدیک است بنابراین مقدار ρ_b به دست می‌آید.

چگالی کل به سه عامل وابسته است: چگالی خمیره سنگ (ρ_{ma})، چگالی سیالات پرکننده فضاهای خالی (ρ_f) و تخلخل سازند.

$$\rho_b = (1 - \varphi) \rho_{ma} + \varphi \rho_f \quad (6-1)$$

1- Gamma-gamma tools
2- Electron density
3- Bulk density

بنابراین با دانستن چگالی خمیره و چگالی سیال سازند با استفاده از ρ_b حاصل از نگار، می‌توانیم تخلخل را محاسبه کنیم. مانند نگار صوتی، در نگار چگالی نیز اثرات مربوط به ناخالصی‌های شیلی، اثر گاز و اثرات محیطی باید اصلاح شود [۱۶].

۱-۴-۲-۳- نگار نوترون^۱

چنان‌که از نام آن پیداست، برای چاه‌نگاری به این روش از پرتو نوترون استفاده می‌شود. نوترون پراثرژی در برخورد با سازند دو نوع برهم‌کنش خواهد داشت:

الف- پراکندگی: در این پدیده، نوترون‌های سریع چشمه مصنوعی سوند با انرژی اولیه‌ای بین ۴ تا ۶ Mev سازند را بمباران می‌کنند. نوترون‌های متحرک در برخورد با هیدروژن‌های سیالات سازند، بیشترین انرژی را به آن‌ها منتقل می‌کنند و خود بیشترین افت انرژی را خواهند داشت.

ب- جذب: با کاسته شدن از انرژی جنبشی نوترون‌ها در برخوردهای پیاپی، در نهایت وقتی که سطح انرژی آن‌ها به سطح انرژی نوترون گرمایی رسید، یعنی انرژی و سرعت آن‌ها به 0.025 eV و 2500 m/s کاهش یافت، توسط هسته‌های بعضی از عناصر جذب می‌شوند و یک یا چندین پرتو گاما منتشر می‌کنند.

آن‌چه که در سوندهای نگار نوترون اندازه‌گیری می‌شود، شاخص هیدروژنی سیال است که با مقدار هیدروژن موجود در واحد حجم متناسب است. شاخص هیدروژنی آب شیرین ۱ است و با افزایش نمک، از آن کاسته می‌شود. شاخص هیدروژنی نفت خیلی نزدیک به آب است. گازها به دلیل چگالی هیدروژنی کم، شاخص هیدروژنی کوچکی دارند، به طوری که اثر آن بر اندازه‌گیری بیش از حد انتظار است و باعث می‌شود که در سازندهای گازدار، تخلخل بسیار کمتر از مقدار واقعی آن برآورد شود که به آن اثر تهی‌شدگی^۲ گویند و باید اصلاح شود [۱۶].

به دلیل وجود آب روزه‌ای در لایه‌های شیلی، تخلخل غیرمؤثر آن به وسیله نگار نوترونی به حساب می‌آید که باید با استفاده از رابطه زیر اصلاح شود [۱۹]:

$$\varphi_{corr} = \varphi_{log} - (V_{sh} \varphi_{sh}) \quad (7-1)$$

در ضمن، تصحیحات محیطی لازم نیز باید اعمال شود.

1- Neutron porosity hydrogen index (NPHI)

2- Excavation effect

۱-۴-۲-۴- نگار پرتو گامای طبیعی^۱

موادی که به طور طبیعی از خود پرتوهای α ، β و γ ساطع می‌کنند رادیواکتیو طبیعی نامیده می‌شوند. از جمله این مواد، عناصر پتاسیم، توریم و اورانیوم هستند. پرتوهای α و β قدرت نفوذ چندانی ندارند و به این دلیل در عملیات چاه‌نگاری کاربردی ندارند اما پرتو گاما به دلیل قدرت نفوذ بالاتری که نسبت به آن دو دارد، نقش اساسی در چاه‌نگاری و از جمله نگار پرتو گامای طبیعی دارد. در سنگ‌های رسوبی عملاً فقط عناصر پتاسیم، توریم و اورانیوم در پرتوزایی گاما مشارکت دارند. با توجه به این که این سه عنصر به مقدار زیادی در کانی‌های رسی تمرکز می‌یابند، جهت برآورد حجم شیل سازندهای غیر تمیز رسوبی، می‌توان از نگار پرتو گاما بهره برد [۱۹].

واحد ثبت پرتو گاما A.P.I^۲ است که به این صورت تعریف می‌شود: اگر ۱ میکروگرم رادیم در یک تن سنگ داشته باشیم مقدار تابش آن A.P.I ۱۶/۵ است. به علت بزرگ بودن A.P.I از GAPI که ۰/۰۵ A.P.I است استفاده می‌شود [۳].

۱-۵- کاربرد روش شبکه‌های عصبی مصنوعی^۳ در تعیین تخلخل

روش‌هایی که به طور معمول از آن‌ها برای تعیین تخلخل و یا هر ویژگی مخزنی دیگر از روی نگارهای چاه استفاده می‌شود، شامل فرض‌ها و شرایطی هستند که در صورت اقتناع آن‌ها، نتایج حاصل صحیح است و در غیر این صورت می‌بایست تصحیحات لازم به روشی مناسب اعمال شود. همچنین روابط و معادلاتی که از آن‌ها در این روش‌ها استفاده می‌شود، شامل ضرایب و عباراتی هستند که به نقطه قرائت و سنگ‌شناسی آن بستگی دارد و به طور کلی ناشناخته‌اند و بایستی با تحلیل نمونه‌های سنگی تعیین شوند، مثلاً ناخالصی رسی، نوع سیال منافذ، چگالی دانه‌ای در روابط مربوط به نگار چگالی و زمان گذر موج در دانه‌ها در روابط مربوط به نگار صوتی [۲۰ و ۲۹].

چنانچه از روش شبکه عصبی برای تخمین تخلخل استفاده شود، حداقل می‌توان گفت بعد از آن که شبکه مناسب طراحی شد، در مرحله اجرا^۴ و کاربرد، ما را از فرمول‌ها و روابطی که نیاز به تعیین

1- Gamma ray (GR)
2- American petroleum institute
3- Artificial neural networks (ANNs)
4- Operation

ضرایب و عبارات لازم دارند و همچنین تصحیحات ناشی از عدم تحقق فرض‌ها و شرایط حاکم بر روش مورد استفاده بی‌نیاز می‌کند، زیرا شبکه این موارد را در مرحله آموزش^۱ آموخته و در مرحله تعمیم^۲، توانایی شبکه در اعمال آن‌ها مورد آزمون قرار گرفته است.

واضح است در مواردی که راه‌حل‌های متعارف می‌توانند مسأله را به طور مؤثر حل کنند، لزومی ندارد که از روش شبکه عصبی یا هر روش دیگر هوش مصنوعی استفاده شود، اما مسائل متعددی وجود دارد که برای آن‌ها، این روش راه‌حلی منحصر به فرد است [۲۰ و ۲۱]. روش شبکه عصبی دارای پتانسیل بالایی برای تولید نتایج تقریباً دقیق از پایگاه داده‌های بزرگ می‌باشد و در شرایط زیر از قابلیت و توانایی بیشتری نسبت به سایر روش‌ها برخوردار است و می‌تواند به طور مؤثر به کار رود: هنگامی که این احتمال وجود دارد که داده‌ها خطای زیادی داشته باشد، وقتی که داده‌ها نویزی باشد، الگوهای تصمیم‌گیری بسیار پیچیده و مبهم باشد به طوری که مغز انسان و روش‌های آماری استاندارد قادر به استنباط آن‌ها نباشد، داده‌ها کاملاً غیرخطی، آشفته و نابسامان باشد به گونه‌ای که همه پارامترهای دخیل در یک فرآیند بخصوص شناخته شده نباشد و یا این که رابطه پارامترها برای مدل کردن سیستم بسیار پیچیده باشد و در نتیجه امکان استفاده از مدل‌سازی ریاضی وجود نداشته باشد. داده‌های مربوط به علوم زمین، در اغلب موارد، چنین شرایط و پیچیدگی‌هایی را دارند و به همین دلیل در این علوم از روش شبکه عصبی مصنوعی به طور روزافزون استفاده می‌شود [۲۰ و ۲۱]. ویژگی‌ها و مزایایی از روش شبکه عصبی که آن را در مواجهه با داده‌هایی با شرایط مذکور توانمند و ممتاز می‌سازد عبارتند از:

۱- قابلیت یادگیری: قابلیت یادگیری یعنی توانایی تنظیم پارامترهای شبکه وقتی که محیط شبکه تغییر می‌کند و شبکه شرایط جدید را تجربه می‌کند، با این هدف که اگر شبکه برای یک وضعیت خاص آموزش دید و تغییر کوچکی در شرایط محیطی آن (وضعیت خاص) رخ داد، شبکه بتواند با آموزش مختصر برای شرایط جدید نیز کارآمد باشد. دیگر این که اطلاعات در شبکه‌های عصبی در پارامترها ذخیره می‌گردد و هر نرون^۳ در شبکه، به صورت بالقوه از کل فعالیت سایر نرون‌ها متأثر می‌شود. در نتیجه، اطلاعات از نوع مجزا از هم نبوده بلکه متأثر از کل شبکه می‌باشد [۱۷].

1- Training
2- Generalization
3- Neuron

۲- پراکندگی اطلاعات در تمام شبکه: اطلاعات و دانشی که شبکه فرا می‌گیرد در پارامترهای شبکه مستتر می‌باشد. رابطه یک به یک بین ورودی‌ها و پارامترها وجود ندارد. می‌توان گفت که هر پارامتر مربوط به همه ورودی‌هاست ولی به هیچ یک از آن‌ها به طور منفرد و مجزا مربوط نیست. به عبارت دیگر هر نرون در شبکه، از کل فعالیت سایر نرون‌ها متأثر می‌باشد. بر این اساس چنان چه بخشی از نرون‌های شبکه حذف شوند و یا عملکرد غلط داشته باشند باز هم احتمال رسیدن به پاسخ صحیح وجود دارد. اگر چه این احتمال برای تمام ورودی‌ها کاهش یافته ولی برای هیچ یک از بین نرفته است [۱۷].

۳- قابلیت تعمیم: پس از آن که الگوهای اولیه به شبکه آموزش داده شد، شبکه می‌تواند در مقابل یک ورودی آموزش داده نشده^۱ قرار گیرد و یک خروجی مناسب ارائه نماید. این خروجی بر اساس مکانیسم تعمیم که همانا چیزی جز فرایند درونیابی^۲ نیست به دست می‌آید. به عبارت روشن‌تر، شبکه تابع را یاد می‌گیرد، الگوریتم را می‌آموزد و یا رابطه تحلیلی مناسبی را برای تعدادی نقاط در فضا به دست می‌آورد [۱۷].

۴- پردازش موازی: هنگامی که شبکه عصبی در قالب سخت‌افزار پیاده می‌شود، نرون‌هایی که در یک تراز قرار می‌گیرند می‌توانند به طور هم‌زمان به ورودی‌های آن تراز پاسخ دهند. این ویژگی باعث افزایش سرعت پردازش می‌شود. در واقع در چنین سیستمی، وظیفه کلی پردازش بین پردازنده‌های کوچک‌تر توزیع می‌گردد [۱۷].

۵- مقاوم بودن^۳: در یک شبکه عصبی، رفتار کلی شبکه، برآیند رفتارهای محلی نرون‌های متعدد است. این ویژگی باعث می‌شود تا خطاهای محلی از چشم خروجی نهایی دور بماند. به عبارت دیگر، نرون‌ها در یک روند همکاری، خطاهای محلی یکدیگر را تصحیح می‌کنند. این خصوصیت باعث افزایش قابلیت مقاوم بودن و تحمل‌پذیری خطا در سیستم می‌گردد [۱۷].

اشکال اصلی در استفاده از روش شبکه عصبی مصنوعی در مسائل گوناگون و از جمله تخمین تخلخل، نیاز به مجموعه‌ای با تعداد کافی از الگوهای تحقق‌یافته است به طوری که نشان‌دهنده تمام حالات ممکن مسأله باشد تا از آن‌ها در آموزش و آزمون شبکه استفاده شود. تولید این مجموعه مستلزم کار و

1- Unseen data
2- Interpolation
3- Robustness

تلاش فراوان است. البته این اشکال در بین همه روش‌های مدل‌سازی داده‌های واقعی مشترک است [۲۰]. علاوه بر این، روش مشخصی برای تعیین ساختار بهینه شبکه معرفی نشده و به همین دلیل باید از روش سعی و خطا^۱ برای یافتن آن استفاده کرد.

۱-۶- مروری بر مطالعات انجام شده

مطالعات زیادی در زمینه کاربرد شبکه‌های عصبی در تحلیل کمی خصوصیات و ویژگی‌های مخزن با استفاده از داده‌های نگارها انجام شده است که در این جا به چند مورد آن اشاره می‌کنیم. هوانگ^۲ و همکارانش از روش شبکه عصبی مصنوعی به منظور تخمین تخلخل و تراوایی در میادین هیدروکربوری فلات قاره شرق کانادا بهره گرفته‌اند. مطالعات آن‌ها نشان می‌دهد که روش شبکه عصبی نسبت به روش‌های رگرسیون خطی چندمتغیره^۳، رگرسیون غیرخطی چندمتغیره^۴ و تخمین با استفاده از روابط تجربی و نیمه‌تجربی، نتایج بسیار بهتری دارد [۳۱ و ۳۲]. هلی^۵ و همکارانش نیز با استفاده از شبکه عصبی تخلخل و تراوایی را در یکی از مخازن دریای شمال با دقت کافی که بیشتر نیازهای عملی را جوابگو می‌باشد، ارزیابی کرده‌اند [۳۰]. در کشورمان نیز افراد متعددی در قالب پایان‌نامه‌های دانشجویی و مقالات، در این زمینه تحقیق نموده‌اند. به عنوان مثال، جلالی لیجایی، نبی‌بیدهدندی و میرزایی با استفاده از روش شبکه عصبی مصنوعی و رگرسیون خطی چندمتغیره، تخلخل سنگ مخزن آسماری را در میدان نفتی گچساران ارزیابی کرده‌اند. نتایج آن‌ها حاکی از برتری قابل توجه روش شبکه عصبی نسبت به روش رگرسیون خطی چندمتغیره است به طوری که با استفاده از روش شبکه عصبی در مرحله تعمیم، شیب بهترین خط برازش داده شده بین تخلخل مغزه و تخمین ۰/۹۱ و ضریب همبستگی نقاط نسبت به این خط ۰/۹۴ به دست آمده، در حالی که با استفاده از روش رگرسیون خطی در مرحله تعمیم، این کمیت‌ها به ترتیب ۰/۳۸ و ۰/۶۵ به دست آمده است [۵]. در تحقیقی دیگر، با استفاده از روش شبکه عصبی مصنوعی نگارهای چگالی و شاخص فتوالکتریک در بازه عمقی سازند آسماری در یک چاه قدیمی برای مخزن ماسه‌سنگی میدان نفتی

1- Trial and error approach

2- Huang

3- Multiple linear regression (MLR)

4- Multiple nonlinear regression (MNLN)

5- Helle

اهواز توسط خاکبازان، فتحیان‌پور و باقری تخمین زده شده است. در مرحله اعتبارسنجی، در مورد نگار چگالی، خطای تخمین $0/068$ گرم بر سانتی‌متر مکعب و ضریب همبستگی حول بهترین خط برازش داده شده $0/9$ و در مورد نگار شاخص فتوالکتریک خطای تخمین $0/36$ بور بر الکترون و ضریب همبستگی $0/88$ به دست آمده است [۸].

علی‌رغم تمام توانایی‌ها و امتیازات روش شبکه عصبی مصنوعی، با توجه به این که برای تمام حالاتی که قبلاً شبکه عصبی آن‌ها را تجربه نکرده است و به شبکه عرضه نشده‌اند، تخمین 100% درست غیرممکن است، هم‌چنان تلاش‌ها برای ارتقاء این روش ادامه دارد. استفاده از سیستم‌های چندشبکه‌ای^۱ که از آن‌ها به شبکه عصبی کمیته‌ای یا ماشین کمیته‌ای^۲ نیز تعبیر می‌شود و می‌توان گفت ترکیبی از چندین شبکه‌اند که به طور منفرد و مجزا آموزش دیده‌اند، در راستای این تلاش‌هاست.

پارمانتو^۳ و همکارانش (۱۹۹۶) نشان دادند که استفاده از میانگین‌گیری ساده چندین شبکه با ساختار و مجموعه آموزشی یکسان و پارامترهای اولیه تصادفی متفاوت، می‌تواند به مقدار قابل توجهی واریانس^۴ خطا را کاهش دهد و نسبت به روش انتخاب بهترین شبکه، کارایی بیشتری دارد [۲۱]. ناftالی^۵ و همکارانش (۱۹۹۷) به بررسی عملکرد ماشین کمیته‌ای پرداختند. بنابر نتایج آن‌ها، ماشین کمیته‌ای واریانس خطای تخمین را کاهش می‌دهد ولی کاهش در بایاس^۶ خطا ایجاد نمی‌کند، در حالی که هاشم^۷ روش ترکیب خطی بهینه نامقید را به منظور حذف بایاس پیشنهاد کرد [۲۰، ۲۱] و [۲۷]. شارکی^۸ و همکارانش نیز رهیافت آزمون و انتخاب^۹ را به منظور انتخاب ترکیبات آنسامبلی^{۱۰} کارآمد شرح داده‌اند [۳۶].

-
- 1- Multiple network systems
 - 2- Committee machine (CM)
 - 3- Parmanto
 - 4- Variance
 - 5- Naftaly
 - 6- Bias
 - 7- Hashem
 - 8- Sharkey
 - 9- Test and select approach
 - 10- Ensemble combinations

میزان به کارگیری شبکه‌های عصبی کمپته‌ای در علوم مربوط به زمین و از جمله چاه‌نگاری، نسبت به شبکه‌های عصبی منفرد بسیار محدودتر است. چند مورد از آن‌ها عبارتند از: فانگ^۱ و همکارانش از ترکیب مدوله‌ای^۲ شبکه‌های عصبی که نوعی از ماشین کمپته‌ای است برای تشخیص جنس و درصد کانی‌های سنگ‌های مجاور یک چاه از روی نگارهای چاه‌پیمایی استفاده کرده‌اند و نتایج را با شبکه پس‌انتشار خطای^۳ منفرد مقایسه نموده‌اند. مقایسه نشان می‌دهد که ترکیب مدوله‌ای شبکه‌های عصبی مدت زمان آموزش و میانگین مربعات خطا^۴ را نسبت به شبکه پس‌انتشار خطای منفرد به میزان زیادی کاهش داده است [۲۴]. آلفانا بات^۵ در رساله^۵ دکتری خود (۲۰۰۲)، با استفاده از شبکه‌های عصبی کمپته‌ای به تخمین تخلخل، تراوایی و اشباع سیالات مخزنی در میداین نفتی دریای شمال پرداخته است. وی برای تخمین تخلخل از ترکیب آنسامبلی شبکه‌ها که نوعی دیگر از ماشین کمپته‌ای است، بهره گرفته. در ترکیب آنسامبلی مورد استفاده، نتایج ۹ شبکه منفرد با هم ترکیب شده است. نتیجه آنسامبل حاصل، از نتایج هر یک از ۹ شبکه جزء بهتر است و به خوبی توانسته است تخلخل را تخمین بزند [۲۰]. در تخمین تراوایی از روش مدوله‌ای استفاده گردیده. دامنه تراوایی به سه زیردامنه تقسیم و هر یک از مدول‌های سیستم با استفاده از یکی از زیردامنه‌ها آموزش داده شده است. با هدف افزایش دقت، به جای استفاده از یک شبکه عصبی در هر مدول، از یک ترکیب آنسامبلی متشکل از ۵ شبکه عصبی منفرد در هر مدول استفاده شده است. مقایسه نتایج روش شبکه عصبی منفرد با نتایج این ماشین کمپته‌ای نشان می‌دهد که در تخمین تراوایی، ماشین کمپته‌ای دقت بیشتر و خطای کمتری داشته است [۲۰]. برای محاسبه اشباع هر یک از سه سیال مخزنی، از یک ماشین کمپته‌ای استفاده شده که ترکیب آنسامبلی ۹ شبکه پرسپترون^۶ با ساختار ۴-۴-۱ می‌باشد. این ترکیب آنسامبلی یکی از سه بلوک ساختاری یک ترکیب مدوله‌ای را تشکیل می‌دهد. ترکیب آنسامبلی نتایج ۹ شبکه به خوبی توانسته است خطا را نسبت به شبکه‌های منفرد جزء کاهش دهد [۲۰].

1- Fung

2- Modular neural networks (MNN)

3- Error back-propagation (BP)

4- Mean of squares of errors (MSE)

5- Alpna Bhatt

6- Perceptrons

چن^۱ و لین^۲ نیز از یک ماشین کمپته‌ای متشکل از سه فرمول تجربی برای تخمین تراوایی استفاده کرده‌اند. آن‌ها از دو روش میانگین‌گیری ساده و میانگین‌گیری وزنی به منظور ترکیب فرمول‌ها بهره گرفته‌اند. در روش میانگین‌گیری وزنی، ضرایب بهینه فرمول‌ها با استفاده از روش الگوریتم ژنتیک^۳ تعیین شده است. نتایج نشان می‌دهد که میانگین مربعات خطای تخمین الگوهای آموزش با استفاده از میانگین‌گیری وزنی ۶٪ نسبت به بهترین فرمول تجربی کاهش پیدا کرده است. همچنین میانگین مربعات خطای تخمین الگوهای آزمون توان تعمیم، با استفاده از میانگین‌گیری وزنی و میانگین‌گیری ساده به ترتیب ۶٪ و ۱۵٪ نسبت به بهترین فرمول تجربی کاهش یافته است [۲۲].

۱-۷- ضرورت و هدف از مطالعه حاضر

با توجه به مطالعات انجام شده و مزایای غیر قابل انکار روش شبکه‌های عصبی مصنوعی در حل مسائل پیچیده علوم و مهندسی، به ویژه قابلیت یادگیری و تعمیم، دسته‌بندی و انجام پردازش بسیار زیاد در زمان کم، از این روش برای تخمین تخلخل از روی نگارهای چاه استفاده شد. همان طور که قبلاً نیز گفته شد، تخمین تخلخل از روش‌های دیگر مانند مغزه‌ها و یا حتی نگارهای چاه، مشکل و نیازمند انجام تصحیحات لازم و در اختیار داشتن اطلاعات کافی از عمق مورد نظر است، اما استفاده از روش شبکه عصبی در تخمین تخلخل، کارا و با نتایج قابل قبول همراه است و تا حدود زیادی ما را از اطلاعات زمین‌شناسی عمقی و تصحیحات ضروری بی‌نیاز می‌کند.

اگر چه بهترین شبکه عصبی منفرد حاصل، بهترین الگوی بین ورودی‌ها و خروجی‌ها را تولید می‌کند، اما این که بتواند خروجی همه الگوهای آزمون ممکن را درست تعیین کند، نامحتمل است و بنابراین جستجوی روشی برای افزایش دقت روش شبکه عصبی توجیه منطقی پیدا می‌کند. همان طور که قبلاً نیز اشاره کردیم یکی از روش‌هایی که به این منظور پیشنهاد شده است استفاده از ماشین کمپته‌ای می‌باشد.

در رهیافت ماشین کمپته‌ای تأکید بر آن است که از همه اطلاعات مفید، حداکثر استفاده شود و نه این که با انتخاب یک شبکه به عنوان بهترین شبکه، از اطلاعات مفید سایر شبکه‌ها صرف نظر شود.

1- Chen

2- Lin

3- Genetic algorithm (GA)

عدم کفایت و محدود بودن مطالعات انجام شده در زمینه ماشین کمپته‌ای، به خصوص درباره کاربرد آن در مطالعات مخزن، باعث شده است تا نقاط قوت و ضعف روش‌های مختلف ساخت آن‌ها مبهم باشد، به همین دلیل تحقیقات نظری و کاربردی بیشتر با هدف رفع این ابهامات ضروری است. لازم است به طور صریح و روشن بیان کنیم که هدف از این مطالعه تخمین تخلخل سنگ مخزن با استفاده از دو رهیافت شبکه‌های عصبی منفرد و شبکه‌های عصبی کمپته‌ای است تا در نهایت با مقایسه نتایج، ببینیم که آیا در این مورد مطالعاتی، ماشین کمپته‌ای می‌تواند نتایج روش شبکه‌های عصبی منفرد را بهبود دهد یا خیر و در صورت مثبت بودن پاسخ، چه عواملی بر میزان بهبود و ارتقاء نتایج تأثیر می‌گذارند؟

متأسفانه در کشورمان تحقیقی در مورد کاربرد ماشین کمپته‌ای در تخمین و تشخیص ویژگی‌ها و خصوصیات مخزنی مشاهده نشد و به نظر می‌رسد که این تحقیق گام اولی است که در این زمینه برداشته می‌شود. امیدواریم در قالب این پایان‌نامه، با مطالعه در این مورد، مقداری از این خلأ را، اگر چه بسیار اندک، جبران کنیم و این تحقیق بتواند نقطه آغازی مفید و مناسب برای پژوهش‌های آینده باشد.

۱-۸- ساختار پایان‌نامه

در فصل دوم، تحت عنوان مبانی شبکه‌های عصبی مصنوعی، نکات و موارد اصلی پیرامون آن‌ها را بیان کرده‌ایم. اگرچه سعی ما در این فصل ایجاز و اختصار بوده است، اما در آوردن بعضی از مطالب که مقدمه‌ای بر مباحث اصلی بوده‌اند و حذف آن‌ها به وضوح و پیوستگی مطلب زیان می‌رساند، ناگزیر بوده‌ایم.

در فصل سوم با عنوان مبانی شبکه‌های عصبی مصنوعی کمپته‌ای، نتیجه مطالعات خود را در زمینه ماهیت ماشین کمپته‌ای، ترکیب آنسامبلی و مدوله‌ای شبکه‌ها، روش‌های تولید این نوع ترکیبات و روش تحلیلی ترکیب خطی بهینه MSE-OLC شرح داده‌ایم.

فصل چهارم، شرح کارها و نتایجی است که با استفاده از شبکه‌های عصبی مصنوعی و ترکیب آنسامبلی آن‌ها به منظور تخمین تخلخل سنگ مخزن گازی کنگان میدان عظیم هیدروکربوری پارس جنوبی انجام شده است.

در فصل پنجم نتایجی را که از تحقیقمان در مورد ماشین کمیت‌های و نقاط قوت و ضعف آن به دست آورده‌ایم بیان نموده‌ایم. همچنین مطالعات و کارهایی را که می‌تواند به منظور تبیین و وضوح بیشتر موضوع در آینده انجام شود، پیشنهاد کرده‌ایم.

فصل دوم:

مبانی شبکه‌های عصبی مصنوعی

۲-۱- مقدمه

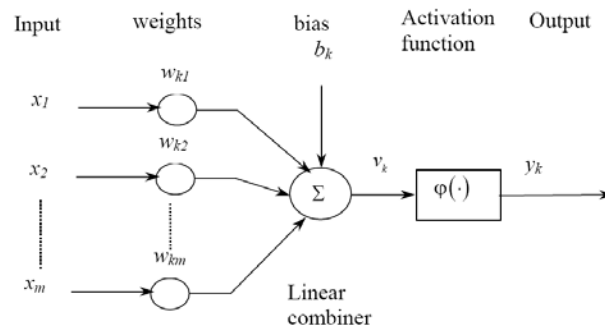
الگو و الهام‌بخش شبکه‌های عصبی مصنوعی، سیستم عصبی موجودات زنده است که به گونه‌ای کاملاً متفاوت از کامپیوترهای دیجیتالی مرسوم، محاسبات را انجام می‌دهد. مغز یک سیستم پردازش داده بسیار پیچیده است که از واحدهای ساختاری به نام سلول عصبی یا نرون تشکیل شده است.

به طور مشابه، شبکه عصبی مصنوعی، سیستمی پویا و غیرخطی است که از تعداد زیادی واحد پردازش به نام نرون مصنوعی و اتصالات بین این واحدهای پردازش تشکیل می‌شود. این شبکه‌ها برخلاف کامپیوتر که نیازمند دستورهای کاملاً صریح و مشخص است، به مدل‌های ریاضی محض نیاز ندارد، بلکه مانند انسان تجربه کسب کرده و سپس نتیجه این تجربیات را تعمیم می‌دهد. امروزه این شبکه‌ها تبدیل به ابزاری قدرتمند و عمومی شده‌اند که برای حل مسائل بسیار پیچیده از قبیل تخمین، تشخیص الگو، طبقه‌بندی و خوشه‌بندی به کار می‌روند [۷].

برای حل هر مسأله‌ای، شبکه‌های عصبی سه مرحله را طی می‌کنند: آموزش، تعمیم و اجرا. آموزش فرایندی است که طی آن شبکه، الگوی موجود در ورودی‌ها و ارتباط بین ورودی‌ها و خروجی‌های مجموعه الگوهای آموزشی را فرا می‌گیرد. در اثر آموزش شبکه، وزن‌های داخلی که بر روی ورودی‌های هر نرون اعمال می‌شود، تغییر می‌کند و به وضعیت مناسب می‌رسد. به این منظور هر شبکه عصبی از مجموعه‌ای از قوانین یادگیری که نحوه یادگیری را تعریف می‌کنند، استفاده می‌کند. تعمیم توانایی شبکه برای ارائه جواب قابل قبول در قبال ورودی‌هایی که در مجموعه آموزشی نبوده‌اند، است. استفاده از شبکه برای انجام عملکردی که به آن منظور طراحی شده است را اجرا گویند [۷].

۲-۲- ساختار شبکه‌های عصبی مصنوعی

همان‌طور که ذکر شد شبکه‌های عصبی مصنوعی از یک سری واحدهای ساختمانی اولیه به نام نرون تشکیل می‌شوند. شکل ۱-۲ ساختمان یک نرون را نشان می‌دهد.



شکل ۱-۲: نمایشی از یک نرون مصنوعی [۲۰].

قسمت‌های مختلف این نرون به شرح زیر است:

۱) سیناپس‌ها که هر یک با وزنی مختص به خودش تعیین می‌شود. سیگنال x_j در ورودی سیناپس j ام، با ضرب شدن در وزن w_{kj} با نرون k ام، ارتباط پیدا می‌کند. (۲) یک جمع‌کننده که سیگنال‌های وزن‌دار را با هم جمع می‌کند. (۳) یک تابع انتقال یا فعال‌ساز که مقدار خروجی جمع‌کننده را به خروجی نرون تبدیل می‌کند. (۴) همچنین هر نرون بایاسی دارد (b_k) که وظیفه آن کاهش و یا افزایش ورودی تابع انتقال است. به مجموعه وزن‌های سیناپسی و بایاس نرون‌ها، پارامترهای شبکه گفته می‌شود.

خروجی نرون k ام به صورت زیر بیان می‌شود:

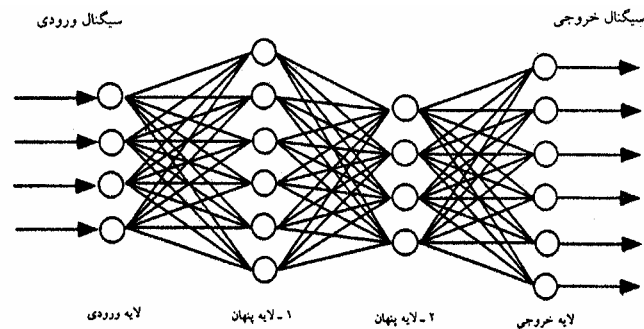
$$y_k = \varphi(v_k) = \varphi(u_k + b_k) \quad (1-2)$$

که در آن:

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (2-2)$$

شبکه‌های عصبی ساختار لایه‌ای دارند و نرون‌ها در این لایه‌ها، تعبیه شده‌اند. اولین لایه، لایه ورودی است که صرفاً ورودی‌ها را در قسمت خروجی خود کپی می‌کند و کار خاص دیگری انجام نمی‌دهد.

لایه آخر لایه خروجی است و لایه‌های بین لایه ورودی و لایه خروجی را لایه‌های میانی و یا پنهان می‌نامند. به جز لایه ورودی تمام لایه‌های دیگر شبکه در کار پردازش شرکت می‌کند.



شکل ۲-۲: نمایشی از ساختار لایه‌ای شبکه‌های عصبی [۷].

بر اساس نحوه اتصال نرون‌ها به یکدیگر، شبکه‌های عصبی به دو گروه بزرگ تقسیم می‌شوند: ۱- شبکه‌های با تغذیه پیش‌رو^۱ که در آن‌ها تنها یک جریان یک‌طرفه از لایه ورودی به سمت لایه خروجی وجود دارد و هیچ مسیر برگشتی ندارند. این نوع شبکه ساده‌ترین و پرکاربردترین نوع شبکه عصبی است و برای مقاصد پیش‌بینی، بسیار مناسب است. ۲- شبکه‌های با تغذیه برگشتی^۲ که حداقل یک حلقه برگشتی فعال دارند. شبکه‌ای که خروجی هر نرون آن به تمام نرون‌های دیگر آن لایه وارد می‌شود مثالی از این دسته است. وجود حلقه برگشتی اثر زیادی بر قابلیت یادگیری شبکه می‌گذارد [۷].

۲-۳- آموزش شبکه‌های عصبی مصنوعی

آموزش فرایندی است که منجر به دستیابی به پارامترهای بهینه برای هر نرون شبکه می‌شود. هدف آموزش آن است که پارامترهایی تولید کند که خروجی شبکه را در نزدیک‌ترین حد ممکن به خروجی مطلوب قرار دهد.

1- Feed forward networks
2- Feed backward networks

در سال ۱۹۵۸ زُرنبلت^۱ و تعدادی از محققین با هدف حل مسائل تشخیص الگو، نوعی از شبکه‌های عصبی با نام پرسپترون را ایجاد کردند که در ساختار نرون‌های آن از تابع انتقال پله‌ای استفاده شد و تفاوت مهم آن با شبکه‌هایی که پیش از آن طراحی شده بود، استفاده از یک قانون یادگیری به منظور آموزش شبکه بود. زرنبلت ثابت کرد که این قانون آموزش همیشه به پارامترهای صحیحی برای شبکه همگرا می‌شود، البته به این شرط که چنین پارامترهایی وجود داشته باشند. این قانون به صورت ساده و اتوماتیک انجام می‌شد و مزیت قابل توجهی داشت، زیرا تا پیش از آن پارامترهای شبکه می‌بایست به صورت دستی تعیین می‌شد [۲۰].

الگوریتم‌های آموزشی شبکه‌های عصبی به دو گروه عمده تقسیم می‌شوند:

۱- آموزش باناظر^۲: در این نوع آموزش، به شبکه آموخته می‌شود که بین ورودی و خروجی الگوهای مجموعه آموزشی ارتباط برقرار کند. این کار از طریق حداقل کردن اختلاف بین نتایج حاصل از شبکه و خروجی‌های مطلوب انجام می‌شود، بنابراین وجود مجموعه آموزشی ضروری است.

۲- آموزش بدون ناظر^۳: در این نوع آموزش، خروجی‌های الگوهای آموزشی به شبکه عرضه نمی‌شود و پارامترها، فقط بر اساس ورودی‌ها تعیین می‌شود. در این حالت، هدف یافتن الگویی بین ورودی الگوهای آموزشی است که بتواند بر اساس خصوصیات مشابه، آن‌ها را گروه‌بندی کند. روش اصلی انجام آموزش بدون ناظر، خوشه‌بندی است. این شبکه‌ها را خودسازمان‌ده^۴ نیز می‌گویند.

انتخاب یک مجموعه آموزشی مناسب که دامنه وسیعی از الگوهای تحقق‌یافته را در بر گیرد از اهمیت زیادی برخوردار است، زیرا کیفیت عملکرد نهایی شبکه عصبی مستقیماً تابع الگوهای آموزشی به کار رفته در آموزش آن است. مهم‌ترین شرط لازم برای مجموعه آموزشی آن است که تمام دامنه مقادیر ورودی‌ها و خروجی‌ها را پوشش دهد. در بسیاری از موارد، پوشش مناسب ورودی‌ها به دامنه تغییراتشان، موجب پوشش مناسب به خروجی‌ها می‌شود. همچنین بهتر است تعداد الگوهای آموزشی برای هر محدوده از خروجی‌ها یکسان باشد. اگر شبکه‌ای به قصد پیش‌بینی مواردی با فراوانی کم طراحی می‌شود، باید مطمئن شد که مجموعه آموزشی به تعداد کافی، الگوهایی از موارد نادر دارد. به

1- Frank Rosenblatt
2- Supervised learning
3- Unsupervised learning
4- Self organization

این دلیل، نمونه‌گیری تصادفی از الگوهای موجود نامناسب است زیرا الگوهایی با فراوانی بیشتر، از حضور الگوهایی با فراوانی کمتر جلوگیری می‌کنند و نقش آن‌ها را کم‌رنگ می‌سازند. برای رفع این مشکل باید از الگوهای نادر به طور نظام‌مند، نمونه‌های اضافی برداشت کرد. به طور کلی، مجموعه آموزشی باید به گونه‌ای انتخاب شود که الگوی صحیح را در اختیار شبکه بگذارد [۷].

پیش‌پردازش ورودی‌ها و خروجی‌ها نیز به آموزش شبکه و شناخت الگوها کمک شایانی می‌کند. روش‌های مختلفی برای این کار وجود دارد. یکی از این روش‌ها، تغییر مقیاس داده‌های آموزشی به گونه‌ای است که تمام ورودی‌ها و خروجی‌ها بین صفر و یک قرار گیرند. روش دیگر تبدیل ورودی‌ها و خروجی‌ها به توزیع گوسی استاندارد است که در آن میانگین توزیع، صفر و انحراف معیار ۱ است. هر چه تعداد متغیرهای ورودی شبکه بیشتر شود نرون‌های میانی بیشتری لازم است تا شبکه الگوهای بین متغیرها را در حد کافی بشناسد. از طرفی با افزایش تعداد نرون‌های میانی، به الگوهای آموزشی بیشتری نیاز است تا شبکه به جواب‌های نامناسب هم‌گرا نشود. همچنین زمان لازم برای هم‌گرایی شبکه نیز بیشتر خواهد بود. حذف متغیرهایی که اثر مثبت و معنی‌داری در افزایش قدرت پیش‌بینی و تخمین ندارند موجب افزایش قدرت شبکه خواهد شد. در چنین مواردی استفاده از روش پیش‌پردازش آنالیز مؤلفه‌های اصلی^۱ مزیت بسیاری دارد [۷ و ۲۳].

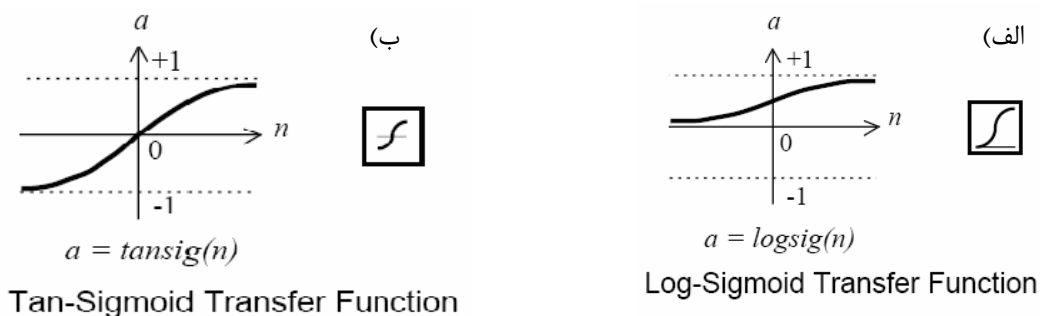
۲-۴- الگوریتم آموزش پس‌انتشار خطا

رایج‌ترین روش آموزش باناظر، الگوریتم پس‌انتشار خطاست. یادگیری این الگوریتم بر پایه قانون تصحیح خطا بنا شده است. یادگیری از طریق این روش دو مرحله دارد: مرحله پیش‌روی و مرحله بازگشت. در مرحله پیش‌روی، ورودی‌ها به صورت لایه به لایه در شبکه پیش می‌رود و در پایان یک سری خروجی به عنوان جواب حقیقی شبکه به دست می‌آید. در این مرحله پارامترهای شبکه ثابت است. در مرحله بازگشت، تفاضل پاسخ حقیقی شبکه و پاسخ مطلوب که خطا نامیده می‌شود، در جهت مخالف اتصالات در شبکه منتشر می‌شود و پارامترهای شبکه بر اساس قانون تصحیح خطا، به گونه‌ای تغییر می‌کنند که پاسخ شبکه به پاسخ مطلوب نزدیک‌تر شود. این فرایند تا وقتی که معیار

1- Principal component analysis

پایان آموزش تحقق پیدا کند، ادامه می‌یابد. چندین معیار برای پایان آموزش به کار می‌رود که هر کدام مزایای خاص خود را دارد. یکی از این معیارها آن است که تغییرات پارامترهای شبکه بسیار کوچک شود و تقریباً، دیگر تغییری از خود نشان ندهند. معیار دیگری که می‌تواند مورد استفاده قرار گیرد آن است که آموزش به تعداد دوره‌های خاصی مثلاً ۱۰۰۰ بار تکرار شود. هر دوره آموزش^۱ شامل یک بار ارائه کامل همه الگوهای مجموعه آموزشی به شبکه است. کاهش خطای شبکه تا مقداری مشخص برای مجموعه الگوهای آموزشی نیز می‌تواند معیاری برای توقف باشد [۷].

ساختار شبکه‌ای که به طور معمول با الگوریتم پس‌انتشار خطا به کار می‌رود، شبکه چندلایه پیش‌رو است. نکته مهم آن است که به هنگام استفاده از الگوریتم پس‌انتشار خطا توابع انتقال نرون‌های شبکه بایستی مشتق پذیر باشند. سه مورد از این توابع که معمولاً از آنها استفاده می‌شود، عبارتند از توابع سیگموئیدی لجستیک^۲، تانژانت هیپربولیک و خطی. دو نوع اول این توابع در شکل ۲-۳ نشان داده شده‌اند.



$$a(n) = \frac{e^n - e^{-n}}{e^n + e^{-n}}$$

$$a(n) = \frac{1}{1 + e^{-n}}$$

شکل ۲-۳: الف) تابع انتقال سیگموئیدی لجستیک (ب) تابع انتقال تانژانت هیپربولیک [۲۳].

هورنیک^۳ و همکارانش (۱۹۸۹) اعلام کردند که اگر یک شبکه پرسپترون سه‌لایه با الگوریتم آموزشی پس‌انتشار خطا که تابع انتقال لایه میانی آن، تابع سیگموئیدی و تابع انتقال لایه خروجی آن، تابع خطی است، در لایه میانی تعداد کافی نرون داشته باشد، می‌تواند هر تابعی با نقاط ناپیوستگی محدود را با هر دقت دلخواه تخمین بزند [۲۰].

1- Epoch
2- Logistic-sigmoid function
3- Hornik

در محیط نرم‌افزار MATLAB برای آموزش شبکه‌ها با استفاده از الگوریتم پس‌انتشار خطا، ۱۳ تابع مختلف طراحی شده است. در همه این توابع از گرادیان تابع عملکرد که در شبکه‌های پیش‌رو، معمولاً میانگین مربعات خطا در نظر گرفته می‌شود، استفاده شده تا با تصحیح پارامترها، تابع عملکرد کمینه شود. این توابع عبارتند از تابع پس‌انتشار خطای تفریق گرادیان^۱، تابع پس‌انتشار خطای تفریق گرادیان با تکانه^۲، تابع پس‌انتشار خطا با سرعت یادگیری متغیر^۳، تابع پس‌انتشار خطا با سرعت یادگیری متغیر و با تکانه^۴، تابع پس‌انتشار خطای بهبودپذیر^۵، تابع گرادیان مزدوج فلچر-ریویس^۶، تابع گرادیان مزدوج پولک-ریبیر^۷، تابع گرادیان مزدوج پاول-بیل^۸، تابع گرادیان مزدوج مقیاسی^۹، تابع شبه‌نیوتنی بی‌اف‌جی‌اس^{۱۰}، تابع شبه‌نیوتنی متقاطع یک مرحله‌ای^{۱۱}، تابع لونبرگ-مارکوارت^{۱۲} و تابع مرتب‌سازی بیزین^{۱۳}. این توابع و پارامترهای مربوط به آن‌ها در پیوست الف معرفی شده‌اند.

تابع پس‌انتشار خطای تفریق گرادیان و تابع پس‌انتشار خطای تفریق گرادیان با تکانه کندترین توابع هستند. تابع پس‌انتشار خطا با سرعت یادگیری متغیر و تابع پس‌انتشار خطا با سرعت یادگیری متغیر و با تکانه نیز که از سرعت آموزش متغیر استفاده می‌کنند، کند هستند [۲۳].

بعضی از این توابع، در شرایط و مسائل معینی بهتر عمل می‌کنند. به طور کلی، وقتی که مسأله، تخمین مقدار یک تابع است و شبکه زیاد بزرگ نیست (مثلاً تعداد پارامترهای شبکه ۱۰۰ یا ۲۰۰ باشد)، تابع آموزش لونبرگ-مارکوارت از همه توابع سریع‌تر همگرا می‌شود. همچنین در بسیاری مواقع این تابع از هر تابع دیگری، میانگین مربعات خطای کمتری به دست می‌دهد. این ویژگی در حالتی که آموزش بسیار دقیق مطلوب باشد، قابل توجه است [۲۳]. به هر حال، وقتی که تعداد پارامترهای شبکه افزایش می‌یابد، قابلیت این تابع کاهش پیدا می‌کند. به طور کلی در مسائل تشخیص الگو، کارایی تابع لونبرگ-مارکوارت کم است. حافظه‌ای که این تابع در حین عملیات به کار

-
- 1- Traingd
 - 2- Traingdm
 - 3- Traingda
 - 4- Traingdx
 - 5- Trainrp
 - 6- Traincgf
 - 7- Traincgp
 - 8- Traincgb
 - 9- Trainseg
 - 10- Trainbfg
 - 11- Trainoss
 - 12- Trainlm
 - 13- Trainbr

می‌گیرد، نسبت به سایر روش‌ها بیشتر است. تدابیری برای کاهش این حجم حافظه اندیشیده شده است اما استفاده از آن‌ها زمان اجرا را افزایش می‌دهد [۲۳].

تابع پسانتشار خطای بهبودپذیر، اگر چه در تخمین مقدار یک تابع کارکرد خوبی ندارد، در مسائل تشخیص الگو سریع‌ترین تابع است. در مقایسه با دیگر توابع، حافظه نسبتاً کمتری برای انجام عملیات به کار می‌برد. اگر دقت مطلوب در مسأله بیشتر شود، مزیت آن نسبت به دیگر توابع کاهش می‌یابد [۲۳].

توابع گرادیان مزدوج، به ویژه تابع گرادیان مزدوج مقیاسی، در مسائل گوناگون اعم از تخمین تابع و تشخیص الگو به خصوص در مواردی که شبکه بزرگ است و پارامترهای زیادی دارد، کارایی مناسبی دارند. تابع گرادیان مزدوج مقیاسی در تخمین توابع، سرعتی در حدود تابع لونبرگ-مارکوارت دارد و حتی برای شبکه‌های بزرگ، سریع‌تر همگرا می‌شود و در مسائل تشخیص الگو، قابلیتی مانند تابع پسانتشار خطای بهبودپذیر دارد با این برتری که با افزایش دقت مطلوب، کارایی آن به اندازه این تابع افت نمی‌کند. الگوریتم‌های گرادیان مزدوج در حل مسائل، حافظه متوسطی به کار می‌گیرند [۲۳].

قابلیت تابع شبه‌نیوتنی بی‌اف‌جی‌اس مشابه تابع لونبرگ-مارکوارت است ولی نسبت به آن حافظه کمتری نیاز دارد. با افزایش اندازه شبکه، محاسبات آن به طور مضاعفی افزایش می‌یابد [۲۳].

تابع آموزش مرتب‌سازی بیزین، در آموزش شبکه‌ها به روش مرتب‌سازی استفاده می‌شود و پارامتر مرتب‌سازی بهینه را به طور اتوماتیک تعیین می‌کند [۲۳].

۲-۵- ساختار بهینه شبکه‌های پرسپترون چندلایه

در طراحی شبکه عصبی، دو سؤال عمده وجود دارد: ۱- چه تعداد الگو برای آموزش نیاز است؟ ۲- بهترین تعداد نرون و لایه میانی شبکه چقدر است؟ این دو سؤال پاسخ دقیقی ندارد و به مسأله‌ای که شبکه برای حل آن طراحی شده یا به کار رفته است بستگی دارد^۱. به طور کلی می‌توان گفت که هر چه الگوهای آموزشی بیشتر و نرون‌های میانی کمتر باشد، بهتر است [۲۱ و ۳۰]. تعداد الگوی آموزشی خیلی کم و یا تعداد نرون میانی خیلی زیاد باعث می‌شود که شبکه، الگوهای آموزشی را به

1- Application dependent

خاطر بسپارد، در نتیجه، در مرحله آموزش عملکرد خوبی دارد و خطا به مقدار بسیار کوچکی می‌رسد اما وقتی که با الگوهایی به غیر از الگوهای آموزشی آزموده می‌شود، عملکرد ضعیفی دارد و خطای پاسخ زیاد است. به بیان دیگر، شبکه دچار بیش‌برازش^۱ شده و توانایی تعمیم کمی دارد. تعداد نرون‌های میانی باید کافی باشد تا شبکه مسأله را به خوبی بیاموزد ولی نرون و لایه میانی بیشتر از حد لزوم، احتمال بیش‌برازش و کاهش توان تعمیم را افزایش می‌دهد.

ساختار بهینه با استفاده از رهیافت سعی و خطا از میان تعدادی شبکه آموزش دیده انتخاب می‌شود. این روش، روشی وقت‌گیر است و تضمین نمی‌کند که حاصل کار بهترین نتیجه باشد، به همین دلیل تحقیق در مورد عملکرد ماشین‌های کمیت‌های به عنوان یکی از روش‌هایی که به منظور رفع این مشکل پیشنهاد شده است ضرورت می‌یابد [۲۱].

۲-۶- بایاس و واریانس خطا

اگر برای تخمین یک تابع شبکه‌ای طراحی شده باشد و برای بردار ورودی \vec{x} ، تابع هدف، $r(\vec{x})$ و تخمین شبکه، $F(\vec{x})$ باشد، میانگین مربعات خطای تخمین F به صورت $E[(F(\vec{x}) - r(\vec{x}))^2]$ نوشته می‌شود که در آن E عملگر امید ریاضی است. تفکیک میانگین مربعات خطای تخمین به صورت رابطه زیر بایاس و واریانس را می‌دهد [۲۰ با تغییر]:

$$E[(F(\vec{x}) - r(\vec{x}))^2] = (E[F(\vec{x}) - r(\vec{x})])^2 + E[(F(\vec{x}) - r(\vec{x}) - E[F(\vec{x}) - r(\vec{x})])^2] \quad (۳-۲)$$

MSE Bias term Variance term

بایاس، میانگین خطای تخمین و واریانس، پراکندگی خطای تخمین حول بایاس است. بایاس می‌تواند مقادیر مثبت و منفی داشته باشد ولی واریانس همواره مثبت است زیرا امید ریاضی یک عبارت همواره مثبت می‌باشد.

معادله ۲-۳ را می‌توان به صورت زیر نوشت:

$$\text{میانگین مربعات خطا} = \text{مجدور بایاس خطا} + \text{واریانس خطا}$$

انحراف معیار همان جذر واریانس است. بنابراین:

$$\text{میانگین مربعات خطا} = \text{مجدور بایاس خطا} + \text{مجدور انحراف معیار}$$

۲-۷- توان تعمیم

توان تعمیم مناسب عبارت است از آن که شبکه آموزش دیده بتواند برای الگوهایی که در سری آموزش نبوده‌اند، پاسخ را با دقت مناسب برآورد کند. یکی از مشکلاتی که در طی آموزش شبکه ممکن است پیش بیاید، بیش‌برازش و یا همان توانایی تعمیم کم است. اگر شرایط به گونه‌ای باشد که شبکه تا کاهش خطا به مقدار مطلوب برای الگوهای آموزش و یا به تعداد معینی دوره، آموزش ببیند ولی دچار بیش‌برازش نشود، این روش آموزش را فراآموزش^۱ نامند [۲۳]. در غیر این صورت باید به دنبال روشی باشیم که احتمال بیش‌برازش را کاهش دهد. به این منظور دو راه‌حل پیشنهاد شده است:

۲-۷-۱ آموزش به روش مرتب‌سازی^۲: در این روش تابع عملکرد که معمولاً میانگین مربعات خطا (MSE) در نظر گرفته می‌شود، تغییر می‌کند و به صورت زیر انتخاب می‌شود [۲۳]:

$$F = \gamma MSE + (1 - \gamma) MSW \quad (۴-۲)$$

که در آن:

$$MSW = \left(\frac{1}{n} \right) \sum_{j=1}^n w_j^2 \quad (۵-۲)$$

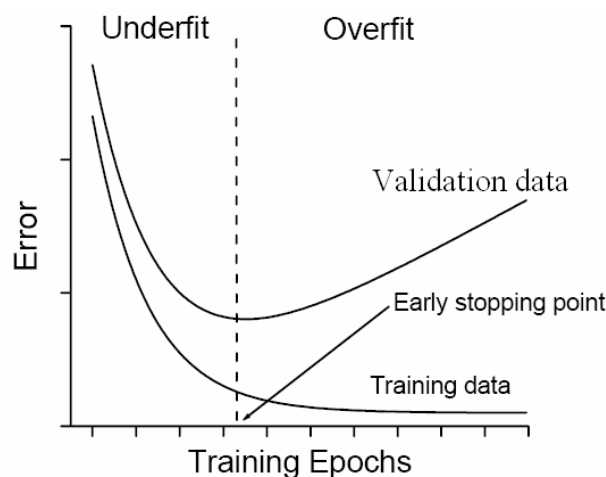
در این دو رابطه MSW میانگین مربعات پارامترهای شبکه^۳ و γ پارامتر مرتب‌سازی است که مقدار آن بین صفر و یک انتخاب می‌شود. انتخاب تابع عملکرد به این شکل باعث می‌شود که پارامترهای شبکه مقادیر کوچک‌تر و در نتیجه، پاسخ‌های آن هموارتر شوند و احتمال بیش‌برازش کاهش یابد [۲۳].

باید دقت شود که مقدار مناسب و بهینه‌ای برای γ در نظر گرفته شود، زیرا اگر مقدار آن خیلی بزرگ فرض شود، ممکن است بیش‌برازش به وجود آید و اگر خیلی کوچک فرض شود، شبکه به اندازه کافی با الگوهای آموزشی برازش نخواهد داشت [۲۳]. روش‌هایی وجود دارد که γ بهینه به طور اتوماتیک تعیین می‌شود. در نرم‌افزار MATLAB تابع آموزشی Trainbr پیش‌بینی شده است که γ بهینه را به طور اتوماتیک تعیین می‌کند. در استفاده از این تابع، زمانی که تعداد پارامترهای مؤثر^۴ شبکه، مجموع

-
- 1- Overtraining
 - 2- Regularization
 - 3- Mean of squares of the network weights and biases (MSW)
 - 4- Effective parameters

مربعات خطا^۱ و مجموع مربعات پارامترهای شبکه^۲ در چندین تکرار به مقدار نسبتاً ثابتی برسند، شبکه همگرا شده است و آموزش پایان می‌یابد [۲۳].

۲-۷-۲- آموزش به روش توقف سریع^۳: در این روش الگوهای آموزش به دو دسته تقسیم می‌شوند. از دسته اول برای آموزش شبکه استفاده می‌شود. گروه دوم، سری اعتبارسنجی است که در فرایند تعیین پارامترهای شبکه نقشی ندارند ولی در حین آموزش، خطای شبکه برای آن‌ها مرتباً ثبت می‌شود. معمولاً این خطا در مراحل اولیه آموزش کاهش می‌یابد. در مرحله‌ای که شبکه شروع به بیش‌برازش کند خطای سری اعتبارسنجی شروع به افزایش می‌کند. وقتی که این خطا به تعداد خاصی تکرار افزایش یابد آموزش متوقف می‌شود و پارامترهای شبکه مقادیری را که در کمینه خطای سری اعتبارسنجی داشتند می‌گیرند [۲۳]. در شکل ۲-۴ آموزش به روش توقف سریع به صورت شماتیک نشان داده شده است.



شکل ۲-۴: تغییرات خطا در سری آموزش و اعتبارسنجی در آموزش با روش توقف سریع [۲۰ با اندکی تغییر].

بر اساس تجربه وقتی که از شبکه برای تخمین مقدار یک تابع استفاده می‌شود، روش مرتب‌سازی کارایی بهتری دارد زیرا این روش نیازی به سری اعتبارسنجی ندارد و از همه الگوهای آموزش، در فرایند آموزش و تعیین پارامترهای شبکه استفاده می‌کند. این برتری در مواقعی که تعداد الگوهای آموزش کم است از اهمیت ویژه‌ای برخوردار است [۲۳].

-
- 1- Sum of squares of errors (SSE)
 - 2- Sum of squares of the network weights and biases (SSW)
 - 3- Early stopping

نکته مهم این است که اگر تعداد پارامترهای شبکه بسیار کم‌تر از تعداد الگوهای سری آموزش باشد، شانس برای بیش‌برازش وجود نخواهد داشت و یا احتمال آن بسیار کم خواهد بود. بنابراین اگر بتوانیم به راحتی الگوهای بیشتری را برای آموزش شبکه جمع‌آوری کنیم، آن‌گاه نیازی به استفاده از دو روش فوق نیست [۲۳].

فصل سوم:

مبانی شبکه‌های عصبی مصنوعی کمیته‌ای

۳-۱- مقدمه

در حال حاضر، شبکه پرسپترون سه لایه با یک لایه ورودی، یک لایه میانی و یک لایه خروجی که با الگوریتم پس انتشار خطا آموزش می بیند، شبکه معمول و متعارفی است که در اغلب مسائل از آن استفاده می شود. توانایی تعمیم به تعداد نرون های میانی و تعداد الگوهای آموزشی وابستگی زیادی دارد. به هر حال این ایراد وجود دارد که توانایی تعمیم شبکه ها، برای ورودی هایی که در سری آموزشی نبوده اند، محدود است و امکان خطا در تخمین شبکه وجود دارد. اگر بپذیریم که برای همه الگوهای آزمون ممکن، تعمیم ۱۰۰٪ غیرممکن است، دلیل خوب و قانع کننده ای برای جستجوی روش هایی که با استفاده از آنها عملکرد شبکه های عصبی مصنوعی را بهبود دهیم داریم. سطح میانگین مربعات خطای یک شبکه چند لایه می تواند یک کمینه مطلق و تعداد زیادی کمینه محلی داشته باشد. در فرایند همگرایی، احتمال زیادی وجود دارد که شبکه به جای این که به کمینه مطلق همگرا شود در یکی از کمینه های محلی به دام بیفتد. در نتیجه خروجی شبکه منفرد، ایده آل نخواهد بود. اگر این شبکه با استفاده از یک سری الگوهای آموزشی یکسان به طور مکرر با پارامترهای اولیه تصادفی آموزش یابد، هر بار به کمینه متفاوتی همگرا می شود و بنابراین پارامترهای نهایی آن مقادیر خاصی را به خود می گیرند که با دیگر دفعات متفاوت است. انتخاب بهترین شبکه منفرد (NN_b)، اگر چه بهترین الگوی حاصل را تولید می کند اما باعث از دست رفتن اطلاعاتی می شود که در دیگر شبکه ها وجود داشته اند. بنابراین به جای این که یکی از شبکه ها را به عنوان بهترین شبکه انتخاب و از بقیه صرف نظر کنیم، با استفاده از روشی مناسب شبکه های آموزش دیده را با هم ترکیب می کنیم تا دقت و توان تعمیم را بهبود دهیم [۲۱]. با استفاده از ترکیب شبکه های منفرد، سیستم های چند شبکه ای را تولید می کنیم تا در مسائلی که یک شبکه قادر به حل آنها نیست و یا ممکن است با استفاده از سیستم های چند شبکه ای به طور مؤثرتری حل شوند، به نتیجه بهتری دست پیدا کنیم. دو

روش معمول برای ایجاد سیستم‌های چندشبکه‌ای وجود دارد: (۱) ترکیب آنسامبلی (۲) ترکیب مدوله‌ای. این دو نوع ترکیب با هم ناسازگار نیستند به این مفهوم که یک سیستم چندشبکه‌ای می‌تواند شامل هر دو نوع ترکیب باشد. های‌کین^۱ (۱۹۹۹) سیستم‌های چندشبکه‌ای را با عنوان ماشین کمیت‌های نام‌گذاری کرده است هر چند که این روش‌ها قبلاً شناخته شده و به کار رفته بود [۲۰]. در ادامه این فصل به شرح این دو نوع ترکیب می‌پردازیم.

۳-۲- ترکیب آنسامبلی

ترکیب آنسامبلی شبکه‌های آموزش دیده، روشی برای ساخت سیستم‌های چندشبکه‌ای است که می‌تواند به تلفیق اطلاعات شبکه‌های جزء و در نتیجه عملکرد بهتر منجر شود. در یک آنسامبل، هر یک از شبکه‌های جزء به تنهایی حلی را برای مسأله مورد نظر ارائه می‌دهند و نتایج به روشی مناسب با هم ترکیب می‌شوند. چون شبکه‌های مختلف به کمینه‌های مختلفی در سطح خطا همگرا می‌شوند انتظار داریم که ترکیب خروجی‌ها توان تعمیم را بهبود دهد. واضح است که ترکیب شبکه‌های یکسان فایده‌ای ندارد زیرا به طور یکسان تعمیم می‌یابند. تأکید می‌کنیم که شبکه‌ها بایستی به طور متفاوت و یا حداقل مشابه و نه یکسان تعمیم یابند. بنابراین به منظور ایجاد شبکه‌های سازنده ترکیب که این شرط را دارا باشند، می‌توانیم یکی از موارد زیر را تغییر دهیم: پارامترهای اولیه، الگوهای آموزشی، ساختار شبکه‌ها و یا الگوریتم آموزش [۲۰].

تغییر الگوهای آموزشی رهیافتی متداول برای ساخت شبکه‌های جزء ترکیب آنسامبلی است. به این منظور چندین روش مختلف پیشنهاد شده است. روش نمونه‌گیری Bagging روشی است که در آن سری‌های آموزشی جدید با استفاده از نمونه‌گیری به همراه جایگزینی از سری آموزشی اصلی انتخاب می‌شوند. نتیجه، مجموعه‌ای از سری‌های آموزشی است که در هر یک الگوهایی مختلف وجود دارد. وقتی که سیستم‌های پیش‌گو رشد می‌یابند و یا با استفاده از این سری‌ها آموزش می‌یابند و سپس با استفاده از رهیافت رأی‌گیری ترکیب می‌شوند، اغلب از یک سیستم پیش‌گوی منفرد که با استفاده از تمام الگوهای آموزشی مجموعه اصلی آموزش می‌یابد، کارایی بهتری دارند. در این روش می‌توان

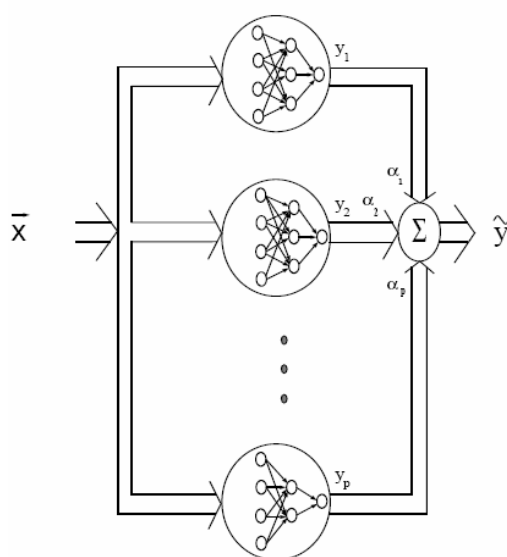
1- Haykin

سری‌های آموزشی جدید را که از سری آموزشی اصلی نمونه‌گیری می‌شوند مجزا از هم در نظر گرفت، یعنی نمونه‌گیری بدون جایگزینی یا تناوب [۳۶]. روش دیگر زمانی امکان‌پذیر است که داده‌های بیش از یک گیرنده یا اندازه‌گیر در دسترس باشد. در این شرایط یک آنسامبل می‌تواند از شبکه‌هایی که هر یک با استفاده از داده‌های یک گیرنده خاص آموزش یافته است، ایجاد شود [۳۶]. رهیافت دیگر آن است که داده‌ها را با استفاده از روش‌های مختلف، پیش‌پردازش کنیم و هر روش پیش‌پردازش را برای ساختن یکی از اجزاء آنسامبل به کار ببریم. همچنین روش‌هایی وجود دارد که می‌توانیم عبارت عمومی اعوجاج^۱ را در مورد آن‌ها به کار ببریم. در این روش‌ها اجزاء آنسامبل به وسیلهٔ اعوجاج ورودی‌ها یا خروجی‌های سری آموزش ایجاد می‌شوند [۳۶]. یک روش در حال حاضر رایج برای ایجاد آنسامبل، روش Adaboost است. آموزش با استفاده از این روش طی چندین مرحله انجام می‌شود. در مرحلهٔ اول، همهٔ الگوهای مجموعهٔ آموزشی اصلی شانس یکسانی برای این که در سری آموزش باشند خواهند داشت. در مرحلهٔ بعدی آموزش، آموزش مرحله اول برای همهٔ الگوهای مجموعهٔ آموزشی اصلی آزموده می‌شود و احتمال انتخاب الگوهایی که بد کلاس‌بندی شده‌اند اصلاح می‌شود (یا اگر الگوهای آموزش بتوانند وزن دار شوند، وزن آن الگوها افزایش می‌یابد) به نحوی که احتمال بیشتری دارند تا در سری آموزشی جدید وجود داشته باشند. این فرایند به تعداد دفعات معینی تکرار می‌شود. نشان داده شده است که آنسامبل‌هایی که به این روش ایجاد می‌شوند، نتایج بهتری را در مقایسه با روش نمونه‌گیری Bagging تولید می‌کنند [۳۶].

اگر چه برای ایجاد آنسامبل‌ها روش‌های زیادی پیشنهاد شده است، هنوز هیچ تصویر واضحی راجع به این که کدام روش بهترین است، وجود ندارد. این موضوع تا اندازه‌ای به خاطر محدودیت مطالعاتی است که در این زمینه انجام شده است. واضح است که مطالعات بیشتر، تصویر واضح‌تری را در این مورد ایجاد می‌کند. به عنوان مثال، احتمال دارد که فهم ترکیب آنسامبلی باعث شود ارزیابی‌های عملی شرایطی که تحت آن، روش خاصی مناسب‌تر است، بهبود یابد [۳۶].

۳-۲-۱- روش‌های ترکیب

وقتی که یک سری از شبکه‌ها برای ساخت ترکیبات آنسامبلی ایجاد شدند، بایستی نتایج این شبکه‌ها با روشی مناسب و مؤثر با هم ترکیب شود. چندین روش مختلف ترکیب وجود دارد که انتخاب یکی از آن‌ها به نوع مسأله‌ای که می‌خواهیم آن را حل کنیم بستگی دارد، یعنی مسأله، تخمین مقدار یک تابع، یا تشخیص طبقه‌بندی و یا دسته‌بندی داده‌ها است. در مسائل تخمین تابع، ترکیب آنسامبلی می‌تواند به طور خطی یا غیر خطی انجام شود. در شکل ۳-۱ نمایی از ترکیب خطی شبکه‌ها نشان داده شده است.



شکل ۳-۱: ترکیب خطی شبکه‌های آموزش دیده [۲۸].

چنان که در شکل دیده می‌شود ترکیب خروجی شبکه‌های مختلف، مشابه ایجاد یک شبکه بزرگ است که شبکه‌های سازنده آن به طور موازی با هم عمل می‌کنند و وزن‌های ترکیب، همان وزن‌های سیناپسی لایه خروجی هستند. برای ورودی \bar{x} ، خروجی ترکیب، \tilde{y} ، جمع وزنی خروجی‌های متناظر شبکه‌های جزء (y_j ها) است و α_j ها وزن‌های مربوط به آن‌ها است. اختلاف اصلی در این تشبیه آن است که در ترکیب، وزن‌های ارتباطی (α_j ها) ثابت هستند ولی در شبکه عصبی بزرگ معادل، این وزن‌ها که نقش اوزان سیناپسی لایه آخر را دارند در طول آموزش تغییر می‌کنند. مزیت دیگری که این ترکیب نسبت به شبکه هم‌ارز خود دارد آن است که در شبکه بزرگ معادلش، پارامترهای زیادی

وجود دارد که لازم است در فرایند آموزش تعیین شوند. بنابراین زمان آموزش زیادتر است و از طرفی به خاطر افزایش پارامترها، احتمال بیش‌برازش بیشتر می‌شود [۲۸].

ترکیب خطی خروجی‌های p شبکه به کار رفته در ساختار ترکیب، عبارت است از [۲۶]:

$$\tilde{y}(\bar{x}; \bar{\alpha}) = \sum_{j=1}^p \alpha_j y_j(\bar{x}) \quad (1-3)$$

که خطای آن:

$$\delta(\bar{x}; \bar{\alpha}) = r(\bar{x}) - \tilde{y}(\bar{x}; \bar{\alpha}) \quad (2-3)$$

در این رابطه $r(\bar{x})$ مقدار مطلوب برای ورودی \bar{x} است.

بنابراین لازم است که مقادیر مناسبی برای α_j ها پیدا کنیم.

یک روش همان انتخاب بهترین شبکه (NN_b) است که وزن متناظر آن (α_b) را برابر یک قرار دهیم و بقیه اوزان را صفر در نظر می‌گیریم. این روش اگر چه ساده است ولی چنان‌که قبلاً گفته شد اطلاعات مفید ($p-1$) شبکه دیگر را از دست می‌دهد.

روش دوم میانگین‌گیری ساده است که به طور گسترده‌ای مورد استفاده قرار می‌گیرد. در این روش فرض می‌شود که همه شبکه‌ها به یک میزان مفید هستند و دقت و توانایی یکسانی در حل مسئله دارند [۲۸]. با استفاده از این روش کاهشی در بایاس ایجاد نمی‌شود ولی واریانس می‌تواند کاهش یابد. روش سوم یافتن وزن‌های بهینه است^۱ به طوری که تابع عملکرد را که معمولاً میانگین مربعات خطا در نظر گرفته می‌شود کمینه کند. به این منظور هاشم و همکارانش روشی تحلیلی را پیشنهاد کرده‌اند که آن را MSE-OLC می‌نامند. در این روش، به منظور تصحیح بایاس y_j ها، جمله دیگری نیز به معادله ۱-۳ افزوده می‌شود. این جمله $\alpha_0 y_0(\bar{x})$ است که در آن $y_0(\bar{x}) = 1$. بنابراین خروجی ترکیب به صورت زیر در خواهد آمد:

$$\tilde{y}(\bar{x}; \bar{\alpha}) = \sum_{j=0}^p \alpha_j y_j(\bar{x}) = \bar{\alpha}' \bar{y}(\bar{x}) \quad (3-3)$$

که $\bar{y}(\bar{x})$ و $\bar{\alpha}$ بردارهایی $1 \times (p+1)$ هستند. برای رابطه ۳-۳ که حالت کلی را نشان می‌دهد سه حالت خاص نیز می‌توان در نظر گرفت. بنابراین در مجموع ۴ حالت خواهیم داشت. تفاوت این ۴

1- Optimal linear combination (OLC)

حالت به وجود یا عدم وجود جمله ثابت α_0 و شرط $\sum_{j=1}^p \alpha_j = 1$ مربوط می‌شود. اگر مجموع ضرایب شبکه‌ها ۱ باشد ترکیب را مقید^۱ و در غیر این صورت ترکیب را نامقید^۲ می‌نامند. پس این ۴ حالت عبارتند از ۱- نامقید با جمله ثابت (حالت کلی) ۲- مقید با جمله ثابت ۳- نامقید بدون جمله ثابت ۴- مقید بدون جمله ثابت [۲۶].

برای ۴ حالت MSE-OLC می‌توان رگرسیون حداقل مربعات معمولی هم‌ارز نیز تعریف کرد که مطالعه آن‌ها را به خصوص در مورد ۳ حالت خاص بسیار آسان‌تر می‌کند. در ادامه به بررسی تفصیلی این ۴ حالت می‌پردازیم.

۳-۲-۱-۱- حالت نامقید با جمله ثابت

مسئله با عبارت زیر بیان می‌شود [۲۶]:

$$P1: \min_{\vec{\alpha}} MSE(\tilde{y}(\vec{x}; \vec{\alpha})).$$

با مشتق‌گیری از MSE نسبت به $\vec{\alpha}$ ، بردار وزن‌های بهینه که آن را با $\vec{\alpha}_{(1)}^* = (\alpha_0^*, \alpha_1^*, \dots, \alpha_p^*)$ نشان می‌دهیم به صورت زیر به دست می‌آید [۲۶]:

$$\vec{\alpha}_{(1)}^* = \Psi^{-1} \vec{U} \quad (۴-۳)$$

که در آن:

$$\Psi = [\Psi_{ij}] = [E \langle y_i(\vec{x}) y_j(\vec{x}) \rangle] \quad (۵-۳)$$

$$\vec{U} = [u_i] = [E \langle r(\vec{x}) y_i(\vec{x}) \rangle] \quad (۶-۳)$$

E عملگر امید ریاضی، Ψ یک ماتریس $(p+1) \times (p+1)$ و \vec{U} یک بردار $(p+1) \times 1$ است.

بنابراین کمینه MSE نظیر برابر است با:

$$MSE^{(1)} = E(r^2(\vec{x})) - \vec{U}' \Psi^{-1} \vec{U} \quad (۷-۳)$$

حالت نامقید با جمله ثابت هم‌ارز رگرسیون $r(\vec{x})$ بر حسب $y_i(\vec{x})$: $i=1,2,\dots,p$ با یک جمله ثابت است، یعنی:

1- Constrained
2- Unconstrained

$$r(\vec{x}) = \alpha_0 + \sum_{i=1}^p \alpha_i y_i(\vec{x}) + \varepsilon(\vec{x}) \quad (8-3)$$

که $\varepsilon(\vec{x})$ خطای مدل و میانگین آن صفر است. بنابراین وزن‌های بهینه ترکیب همان ضرایب رگرسیون OLS هستند. با توجه به این موضوع چنانچه $\kappa = \{ \kappa_n : \kappa_n = (\vec{x}_n, r(\vec{x}_n), \bar{y}(\vec{x}_n)), n = 1, \dots, k \}$ یک مجموعه از الگوهای تحقق‌یافته باشد که به ازاء ورودی \vec{x}_n مقدار تابع هدف و خروجی هر یک از شبکه‌ها مشخص است آن گاه روابط ۳-۴، ۳-۵ و ۳-۶ به شکل زیر در می‌آید [۲۶]:

$$\widehat{\alpha}_{(1)}^* = \widehat{\Psi}^{-1} \widehat{U} \quad (9-3)$$

$$\widehat{\Psi} = [\widehat{\Psi}_{ij}] = \left[\sum_{n=1}^k (y_i(\vec{x}_n) \times y_j(\vec{x}_n)) / k \right] \quad (10-3)$$

$$\widehat{U} = [\widehat{u}_i] = \left[\sum_{n=1}^k (r(\vec{x}_n) y_i(\vec{x}_n)) / k \right] \quad (11-3)$$

که در آن‌ها $y_0(\vec{x}_n) = 1$.

به مجموعه الگوهای κ ، سری الگوهای ترکیب^۱ گفته می‌شود. به طور معمول κ همان مجموعه الگوهای آموزشی شبکه‌ها انتخاب می‌شود.

۳-۲-۱-۲-۲- حالت مقید با جمله ثابت

در این حالت مجموع ضرایب شبکه‌های جزء آنسامبل برابر ۱ است.

مسئله با عبارت زیر بیان می‌شود [۲۶]:

$$\vec{\alpha}' \vec{1}_z = 1 \quad \text{به طوری که } P2: \min_{\vec{\alpha}} MSE(\tilde{y}(\vec{x}; \vec{\alpha}))$$

که $\vec{1}_z$ برداری $1 \times (p+1)$ است که مؤلفه اول آن صفر و بقیه مؤلفه‌های آن ۱ است. با حل لاگرانژین

هم‌ارز $P2$ ، بردار وزن‌های بهینه به صورت زیر به دست می‌آید [۲۶]:

$$\vec{\alpha}_{(2)}^* = \vec{\alpha}_{(1)}^* - \beta_{(2)} \Psi^{-1} \vec{1}_z \quad (12-3)$$

که در آن:

$$\beta_{(2)} = (-1 + \bar{1}'_z \Psi^{-1} \bar{U}) / (\bar{1}'_z \Psi^{-1} \bar{1}_z) \quad (13-3)$$

بنابراین کمینه MSE نظیر برابر است با:

$$MSE^{(2)} = MSE^{(1)} + \beta_{(2)}^2 \bar{1}'_z \Psi^{-1} \bar{1}_z \quad (14-3)$$

جمله دوم در سمت راست رابطه (14-3) نامنفی است، بنابراین $MSE^{(2)} \geq MSE^{(1)}$.

این حالت هم‌ارز رگرسیون $r'(\bar{x}) = (r(\bar{x}) - y_c(\bar{x}))$ بر حسب $y'_i(\bar{x}) = (y_i(\bar{x}) - y_c(\bar{x}))$ با یک جمله ثابت است که $c \in \{1, 2, \dots, p\}$ و $i = 1, \dots, p$ و $i \neq c$. این معادله رگرسیون را می‌توان به صورت زیر نوشت:

$$r'(\bar{x}) = \alpha_0 + \sum_{i=1, i \neq c}^p \alpha_i y'_i(\bar{x}) + \varepsilon(\bar{x}) \quad (15-3)$$

بنابراین معادله (12-3) به صورت زیر ساده می‌شود [26]:

$$\bar{\alpha}_{(2)}^* = \Psi'^{-1} \bar{U}' \quad (16-3)$$

که در آن:

$$\Psi' = [\Psi'_{ij}] = [E \langle (y_i(\bar{x}) - \tau_{(i)0}) y_c(\bar{x}) (y_j(\bar{x}) - \tau_{(j)0}) y_c(\bar{x}) \rangle] \quad (17-3)$$

یک ماتریس $(p \times p)$ است و $\tau_{(i)0}$ متغیر شاخصی است که برای $i > 0$ برابر ۱ است و در غیر این صورت صفر می‌باشد. همچنین:

$$\bar{U}' = [u'_i] = [E \langle (r(\bar{x}) - y_c(\bar{x})) (y_i(\bar{x}) - \tau_{(i)0}) y_c(\bar{x}) \rangle] \quad (18-3)$$

یک بردار $(p \times 1)$ است. با استفاده از سری الگوهای ترکیب K ، سه رابطه اخیر به صورت زیر در می‌آیند [26]:

$$\widehat{\alpha}_{(2)}^* = \widehat{\Psi}'^{-1} \widehat{U}' \quad (19-3)$$

$$\widehat{\Psi}' = [\widehat{\Psi}'_{ij}] = \left[\sum_{n=1}^k (y'_i(\bar{x}_n) y'_j(\bar{x}_n)) / k \right] \quad (20-3)$$

$$\widehat{U}' = [\widehat{u}'_i] = \left[\sum_{n=1}^k (r'(\bar{x}_n) y'_i(\bar{x}_n)) / k \right] \quad (21-3)$$

که $y'_0(\bar{x}_n) = 1$.

روابطی که تا کنون برای این حالت مقید نوشتیم تابعی از خروجی شبکه‌های جزء (y_j ها) هستند. برای حالات مقید (این حالت و حالت ۴) می‌توانیم به جای خروجی شبکه‌ها ضرایب بهینه را با استفاده از خطای تخمین آن‌ها به دست بیاوریم. بنابراین [۲۶]:

$$\vec{\alpha}' \vec{1}_z = 1 \text{ که } P2' : \min_{\vec{\alpha}} E \left[\left(\tilde{\delta}(\vec{x}; \vec{\alpha}) \right)^2 \right]$$

که در آن:

$$\begin{aligned} \tilde{\delta}(\vec{x}; \vec{\alpha}) &= r(\vec{x}) - \tilde{y}(\vec{x}; \vec{\alpha}) = r(\vec{x}) \vec{\alpha}' \vec{1}_z - \vec{\alpha}' \vec{y}(\vec{x}) \\ &= \vec{\alpha}' \left(r(\vec{x}) \vec{1}_z - \vec{y}(\vec{x}) \right) = \vec{\alpha}' \vec{\delta}(\vec{x}) \end{aligned}$$

$\vec{\delta}(\vec{x})$ یک بردار $(p+1) \times 1$ است که $\delta_0(\vec{x}) = -y_0(\vec{x}) = -1$ و $\delta_i(\vec{x}) = r(\vec{x}) - y_i(\vec{x}); i = 1, 2, \dots, p$ با حل لاگرانژین هم‌ارز $P2'$ بردار وزن‌های بهینه برابر خواهد بود با [۲۶]:

$$\vec{\alpha}_{(2)}^* = \frac{\Omega^{-1} \vec{1}_z}{\vec{1}_z' \Omega^{-1} \vec{1}_z} \quad (۲۲-۳)$$

که در آن:

$$\Omega = [\omega_{ij}] = \left[E \left\langle \delta_i(\vec{x}) \delta_j(\vec{x}) \right\rangle \right] \quad (۲۳-۳)$$

یک ماتریس $(p+1) \times (p+1)$ است. بنابراین کمینه MSE نظیر با استفاده از رابطه زیر به دست می‌آید:

$$MSE^{(2)} = \frac{1}{\vec{1}_z' \Omega^{-1} \vec{1}_z} \quad (۲۴-۳)$$

بنابراین با استفاده از سری الگوهای ترکیب κ ، مقدار $\widehat{\alpha}_{(2)}^*$ را به دست می‌آوریم که [۲۶]:

$$\widehat{\alpha}_{(2)}^* = \frac{\widehat{\Omega}^{-1} \vec{1}_z}{\vec{1}_z' \widehat{\Omega}^{-1} \vec{1}_z} \quad (۲۵-۳)$$

$$\widehat{\Omega} = [\widehat{\omega}_{ij}] = \left[\frac{\sum_{n=1}^k (\delta_i(\vec{x}_n) \delta_j(\vec{x}_n))}{k} \right] \quad (۲۶-۳)$$

۳-۲-۱-۳- حالت نامقید بدون جمله ثابت

مسئله با عبارت زیر بیان می‌شود [۲۶]:

$$\bar{\alpha}' \bar{g}_z = 0 \text{ که } P3: \min_{\bar{\alpha}} MSE(\tilde{y}(\bar{x}; \bar{\alpha}))$$

که \bar{g}_z برداری $1 \times (p+1)$ که مؤلفه اول آن ۱ و بقیه مؤلفه‌های آن صفر است. با حل لاگرانژین هم‌ارز $P3$ بردار وزن‌های بهینه به صورت زیر به دست خواهد آمد [۲۶]:

$$\bar{\alpha}_{(3)}^* = \bar{\alpha}_{(1)}^* - \beta_{(3)} \Psi^{-1} \bar{g}_z \quad (۲۷-۳)$$

که در آن:

$$\beta_{(3)} = (\bar{g}_z' \Psi^{-1} \bar{U}) / (\bar{g}_z' \Psi^{-1} \bar{g}_z) \quad (۲۸-۳)$$

بنابراین کمینه MSE نظیر برابر است با:

$$MSE^{(3)} = MSE^{(1)} + \beta_{(3)}^2 \bar{g}_z' \Psi^{-1} \bar{g}_z \quad (۲۹-۳)$$

جمله دوم در سمت راست رابطه (۲۹-۳) نامنفی است، بنابراین $MSE^{(3)} \geq MSE^{(1)}$.

رگرسیون هم‌ارز $P3$ همان معادله رگرسیون $P1$ (معادله ۳-۸) است با این تفاوت که جمله ثابت α_0 را ندارد.

$$r(\bar{x}) = \sum_{i=1}^p \alpha_i y_i(\bar{x}) + \varepsilon(\bar{x}) \quad (۳۰-۳)$$

بنابراین معادله (۲۷-۳) به صورت زیر ساده می‌شود [۲۶]:

$$\bar{\alpha}_{(3)}^* = \Psi'^{-1} \bar{U}' \quad (۳۱-۳)$$

که در آن:

$$\Psi' = [\Psi'_{ij}] = [E \langle y_i(\bar{x}) y_j(\bar{x}) \rangle], (i, j > 0) \quad (۳۲-۳)$$

$$\bar{U}' = [u'_i] = [E \langle r(\bar{x}) y_i(\bar{x}) \rangle], (i > 0) \quad (۳۳-۳)$$

Ψ' یک ماتریس $(p \times p)$ و \bar{U}' یک بردار $(p \times 1)$ است. اگر از سری الگوهای ترکیب κ برای به دست آوردن وزن‌های بهینه ترکیب استفاده کنیم آن‌گاه [۲۶]:

$$\widehat{\alpha}_{(3)}^* = \widehat{\Psi}'^{-1} \widehat{U}' \quad (۳۴-۳)$$

$$\widehat{\Psi}' = [\widehat{\Psi}'_{ij}] = \left[\sum_{n=1}^k (y_i(\bar{x}_n) \times y_j(\bar{x}_n)) / k \right], (i, j > 0) \quad (۳۵-۳)$$

$$\widehat{U}' = [\widehat{u}'_i] = \left[\sum_{n=1}^k (r(\bar{x}_n) y_i(\bar{x}_n)) / k \right], (i > 0) \quad (36-3)$$

۳-۲-۱-۴- حالت مقید بدون جمله ثابت

در این حالت مسأله عبارت است از [۲۶]:

$$\min_{\vec{\alpha}} MSE(\tilde{y}(\vec{x}; \vec{\alpha})) \text{ به طوری که } \vec{\alpha}' \vec{g}_z = 0 \text{ و } \vec{\alpha}' \vec{1}_z = 1$$

از حل لاگرانژین هم‌ارز $P4$ بردار وزن‌های بهینه به صورت زیر به دست خواهد آمد [۲۶]:

$$\vec{\alpha}_{(4)}^* = \vec{\alpha}_{(1)}^* - \beta_{(4a)} \Psi^{-1} \vec{g}_z - \beta_{(4b)} \Psi^{-1} \vec{1}_z \quad (37-3)$$

که در آن:

$$\beta_{(4a)} = \left(\vec{g}_z' \Psi^{-1} \vec{U} - \beta_{(4b)} \vec{g}_z' \Psi^{-1} \vec{1}_z \right) / \left(\vec{g}_z' \Psi^{-1} \vec{g}_z \right) \quad (38-3)$$

$$\beta_{(4b)} = \frac{(\vec{1}_z' \Psi^{-1} \vec{U} - 1)(\vec{g}_z' \Psi^{-1} \vec{g}_z) - (\vec{g}_z' \Psi^{-1} \vec{U})(\vec{1}_z' \Psi^{-1} \vec{g}_z)}{(\vec{1}_z' \Psi^{-1} \vec{1}_z)(\vec{g}_z' \Psi^{-1} \vec{g}_z) - (\vec{g}_z' \Psi^{-1} \vec{1}_z)^2} \quad (39-3)$$

کمینه MSE نظیر از رابطه زیر به دست می‌آید:

$$MSE^{(4)} = MSE^{(1)} + \beta_{(4a)}^2 (\vec{g}_z' \Psi^{-1} \vec{g}_z) + \beta_{(4b)}^2 (\vec{1}_z' \Psi^{-1} \vec{1}_z) + 2\beta_{(4a)}\beta_{(4b)} (\vec{g}_z' \Psi^{-1} \vec{1}_z) \quad (40-3)$$

سه جمله اضافی که در سمت راست معادله فوق ظاهر شده است همگی نامنفی هستند

$$\text{بنابراین } MSE^{(4)} \geq MSE^{(1)}.$$

رگرسیون هم‌ارز در این حالت همان معادله ۳-۱۵ است با این تفاوت که فاقد جمله ثابت α_0 می‌باشد،

بنابراین:

$$r'(\vec{x}) = \sum_{i=1, i \neq c}^p \alpha_i y'_i(\vec{x}) + \varepsilon(\vec{x}) \quad (41-3)$$

بنابراین معادله (۳۷-۳) به صورت زیر ساده می‌شود [۲۶]:

$$\vec{\alpha}_{(4)}^* = \Psi^{-1} \vec{U}' \quad (42-3)$$

که در آن:

$$\Psi' = [\Psi'_{ij}] = \left[E \langle (y_i(\vec{x}) - y_c(\vec{x}))(y_j(\vec{x}) - y_c(\vec{x})) \rangle \right], (i, j > 0) \quad (43-3)$$

$$\vec{U}' = [u'_i] = \left[E \langle (r(\vec{x}) - y_c(\vec{x}))(y_i(\vec{x}) - y_c(\vec{x})) \rangle \right], (i > 0) \quad (44-3)$$

Ψ' یک ماتریس $(p-1) \times (p-1)$ و \vec{U}' یک بردار $1 \times (p-1)$ است.

اگر از سری الگوهای ترکیب κ برای محاسبه ضرایب بهینه ترکیب استفاده شود، در این صورت [۲۶]:

$$\widehat{\alpha}_{(4)}^* = \widehat{\Psi}'^{-1} \widehat{U}' \quad (۴۵-۳)$$

$$\widehat{\Psi}' = [\widehat{\Psi}'_{ij}] = \left[\sum_{n=1}^k (y'_i(\bar{x}_n) y'_j(\bar{x}_n)) / k \right], (i, j > 0) \quad (۴۶-۳)$$

$$\widehat{U}' = [\widehat{u}'_i] = \left[\sum_{n=1}^k (r'(\bar{x}_n) y'_i(\bar{x}_n)) / k \right], (i > 0) \quad (۴۷-۳)$$

مانند حالت ۲ اگر بخواهیم ضرایب بهینه را با استفاده از خطای تخمین شبکه‌های جزء به دست بیاوریم مسأله به صورت زیر خواهد بود [۲۶]:

$$\vec{\alpha}' \vec{1}_z = 1 \text{ و } \vec{\alpha}' \vec{g}_z = 0 \text{ که به طوری به } P4': \min_{\vec{\alpha}} E \left[(\tilde{\delta}(\bar{x}; \vec{\alpha}))^2 \right]$$

با حل لاگرانژین هم‌ارز $P4'$ بردار ضرایب بهینه ترکیب به صورت زیر به دست خواهد آمد [۲۶]:

$$\overline{\alpha}_{(4)}^* = (0, \overline{\alpha}_{(4)}^{*t})^t \quad (۴۸-۳)$$

$$\overline{\alpha}_{(4)}^{*t} = \frac{\Omega^{n-1} \vec{1}}{\vec{1}' \Omega^{n-1} \vec{1}} \quad (۴۹-۳)$$

$\vec{1}$ برداری $(p \times 1)$ که همه عضوهای آن ۱ است و

$$\Omega^n = [\omega^n_{ij}] = [E \langle \delta_i(\bar{x}) \delta_j(\bar{x}) \rangle], (i, j > 0) \quad (۵۰-۳)$$

یک ماتریس $(p \times p)$ است. کمینه MSE نظیر عبارت است از:

$$MSE^{(4)} = \frac{1}{\vec{1}' \Omega^{n-1} \vec{1}} \quad (۵۱-۳)$$

با استفاده از مجموعه الگوهای ترکیب κ روابط ۴۸-۳، ۴۹-۳ و ۵۰-۳ به صورت زیر به دست می‌آیند [۲۶]:

$$\widehat{\alpha}_{(4)}^* = (0, \widehat{\alpha}_{(4)}^{*t})^t \quad (۵۲-۳)$$

$$\widehat{\alpha}_{(4)}^{*t} = \frac{\widehat{\Omega}^{n-1} \vec{1}}{\vec{1}' \widehat{\Omega}^{n-1} \vec{1}} \quad (۵۳-۳)$$

$$\widehat{\Omega}^n = [\widehat{\omega}^n_{ij}] = \left[\sum_{n=1}^k (\delta_i(\bar{x}_n) \delta_j(\bar{x}_n)) / k \right], (i, j > 0) \quad (۵۴-۳)$$

اثبات روابط مربوط به این ۴ حالت توسط هاشم (۱۹۹۳) بیان شده است.

هاشم درباره روش تحلیلی MSE-OLC خود چنین نتیجه‌گیری می‌کند که اگر شبکه‌هایی که در ساختار ترکیب به کار رفته‌اند خوب آموزش دیده باشند^۱، به طور اتوماتیک مجموع ضرایب بهینه شبکه‌های ترکیب به سمت یک و جمله ثابت به سمت صفر میل می‌کند، در حالی که برای آنسامبلی متشکل از شبکه‌هایی که خوب آموزش ندیده‌اند^۲، مجموع ضرایب بهینه از یک دور خواهد بود و جمله ثابت به طور واضحی با صفر تفاوت دارد. منظور از شبکه خوب آموزش دیده آن است که برای بردارهای ورودی مختلف، خروجی‌ها به مقادیر مطلوب خیلی نزدیک باشند به طوری که خطاهای تخمین کوچک شود [۲۶، ۲۷ و ۲۸].

۳-۲-۲- همبستگی خطی مضر^۳

به طور مشابه با شبکه‌های عصبی منفرد، توان تعمیم MSE-OLC عبارت است از کارایی آن برای الگوهایی که در سری ترکیب نبوده‌اند و در برآورد ضرایب بهینه نقشی نداشته‌اند. روش MSE-OLC بدون شک خطای تخمین را برای الگوهای ترکیب (κ) نسبت به NN_b و میانگین‌گیری ساده کاهش می‌دهد اما این امر، طبیعی است. مهم آن است که در مرحله تعمیم، نتایج را نسبت به این دو روش بهبود بخشد. اگر این روش ترکیب برای الگوهای تعمیم نتایج بهتری به دست ندهد استفاده از آن مزیتی ندارد و این زمانی است که بین خروجی شبکه‌ها همبستگی خطی مضر وجود داشته باشد که به شدت قابلیت تعمیم را کاهش می‌دهد. زمانی دو متغیر، همبستگی خطی دارند که بردارهای آن دو متغیر بر روی یک خط قرار بگیرند و یا به عبارت دیگر، آن دو در یک زیرفضای یک‌بعدی باشند. در حالت کلی‌تر، k متغیر همبسته خطی‌اند اگر بتوان یکی از آن‌ها را به صورت ترکیب خطی بقیه به دست آورد. در این صورت این k متغیر در زیرفضایی با بعد کم‌تر از k قرار دارند [۲۶]. چون همه شبکه‌های جزء برای تخمین تابع واحدی، $r(\bar{x})$ ، به کار می‌روند، بنابراین وجود همبستگی خطی دور از انتظار نیست اما هر همبستگی خطی بر توان تعمیم اثر نامطلوب ندارد، بلکه بعضی از آن‌ها اثر نامطلوب دارد که از آن به همبستگی خطی مضر تعبیر می‌شود. راحت‌ترین روش برای تشخیص این

1- Well-trained component networks
2- Poorly trained networks
3- Harmful collinearity

نوع همبستگی، استفاده از یک سری الگوی آزمون مجزا از الگوهای ترکیب، برای ارزیابی توانایی تعمیم ترکیب است. البته هر چه سری آزمون بزرگتر باشد ارزیابی ما از توان تعمیم واقع بینانه تر خواهد بود. در مواقعی که با همبستگی خطی مضر مواجه می شویم روش هایی برای رفع اثرات نامطلوب آن وجود دارد. در صورتی که بتوانیم الگوهای بیشتری را در تعیین ضرایب به کار ببریم (κ بزرگتر باشد) ممکن است این اثرات تا حدودی رفع شود اما تضمینی وجود ندارد [۲۶]. اگر امکان افزایش الگوهای κ نباشد و یا با این کار بهبودی در تعمیم حاصل نشود می توانیم با استفاده از **رهیافت آزمون و انتخاب**، از میان شبکه های آموزش دیده، شبکه های مناسبی را برای ساخت ترکیب انتخاب کنیم و به کار ببریم به طوری که ترکیب آن ها توان تعمیم بهتری نسبت به بهترین شبکه منفرد حاصل و میانگین گیری ساده داشته باشد. برای اجرای رهیافت آزمون و انتخاب از یک سری الگوی مجزا از الگوهای ترکیب استفاده می کنیم تا آنسامبلی با عملکرد خوب را تشخیص دهیم [۳۶]. نکته مهمی که باید به آن اشاره کنیم آن است که اگر در حالت نامقید با جمله ثابت که از نظر تئوری کمترین خطا را دارد با همبستگی خطی مضر مواجه شویم احتمال دارد که با استفاده از ۳ روش خاص دیگر، تعمیم بهتری به دست آید [۲۶].

۳-۲-۳- ترکیب آنسامبلی به روش MSE-OLC برای شبکه های با چند خروجی

برای این کار لازم است تا هر خروجی را به طور جداگانه در نظر بگیریم و برای هر یک از آن ها بردار وزن های بهینه را پیدا کنیم. این روش میانگین مربعات خطای کل را کاهش می دهد و بنابراین در بسیاری از مسائل کاربرد دارد [۲۶].

۳-۲-۴- ترکیب آنسامبلی با استفاده از روش های هوش مصنوعی

علاوه بر روش هایی که به آن ها اشاره شد مطالعاتی نیز صورت گرفته است که در آن ها از روش الگوریتم ژنتیک برای یافتن ضرایب بهینه ترکیب خطی معادله ۱-۳ استفاده شده است [۲۲].

همچنین از خروجی شبکه‌های مختلف می‌توانیم به عنوان ورودی یک شبکه عصبی مصنوعی دیگر استفاده کنیم. واضح است که در این حالت، ترکیب غیرخطی است. شبکه عصبی ترکیب کننده، آموزش می‌بیند تا شبکه‌ها را با وزن‌هایی که در فضای ورودی تغییر می‌کنند ترکیب کند [۲۱].

۳-۳- ترکیب مدوله‌ای

اگر چه ممکن است بتوانیم مسأله را با طراحی یک شبکه عصبی یکپارچه حل کنیم، اما در سیستم‌های چندشبکه‌ای مدوله‌ای، آن را به تعدادی زیرمسأله تفکیک می‌کنیم که هر زیرمسأله توسط یک شبکه عصبی یا آنسامبلی از آن‌ها حل می‌شود و در نهایت با تلفیق پاسخ آن‌ها کل مسأله حل می‌شود. در اغلب موارد، استفاده از ترکیب مدوله‌ای که در واقع همان رهیافت تفرقه بینداز و حکومت کن^۱ است بهبود چشمگیری در حل مسأله ایجاد می‌کند [۲۰]. چنانچه بخواهیم یک شبکه منفرد را با شبکه‌هایی که مجموعاً به حل مسأله می‌پردازند جایگزین کنیم چون تعداد پارامترهای متغیر این شبکه زیاد است احتمال بیش‌برازش افزایش می‌یابد. همچنین مدت زمان آموزش این شبکه یکپارچه از شبکه‌هایی که به طور مدوله‌ای و موازی هم به حل مسأله می‌پردازند بیشتر است. در سیستم مدوله‌ای ساختار و الگوریتم آموزشی که برای حل هر بخش به کار می‌رود می‌تواند متفاوت از سایر بخش‌ها باشد. همچنین لزومی ندارد که هر بخش مسأله منحصراً با روش محاسبات عصبی حل شود مثلاً پیش‌پردازش داده‌های آموزشی می‌تواند با روش‌های غیرعصبی حل شود. دلیل دیگری که به رهیافت مدوله‌ای رهنمون می‌شود کاهش پیچیدگی نسبت به سیستم یکپارچه است به گونه‌ای که فهم، تعدیل و توسعه سیستم را سهل‌تر می‌کند. مهمترین فاکتور مؤثر در کارایی سیستم‌های مدوله‌ای کیفیت تفکیک مسأله است. مثلاً در روش تفکیک کلاسی^۲ که به وسیله آناند^۳ و همکارانش (۱۹۹۵) پیشنهاد شده است طبقه‌بندی به k کلاس به k طبقه‌بندی دوکلاسه تبدیل می‌شود و در عین حال همان تعداد الگوهای آموزش که در مسأله اصلی به کار می‌رود برای هر یک از k زیرمسأله به کار برده می‌شود.

1- Divide and conquer approach
2- Class decomposition
3- Anand

فصل چهارم:

تخمین تخلخل سنگ مخزن گازی کنگان

در میدان پارس جنوبی

با استفاده از شبکه‌های عصبی مصنوعی کمیته‌ای

۴-۱- مقدمه

حوضه رسوبی آب‌های فلات قاره ایران (خلیج فارس) بی‌شک به عنوان یکی از غنی‌ترین مناطق جهان از نظر تجمع مواد هیدروکربوری شناخته شده است و فعالیت‌های شرکت‌های مختلف نفتی در گذشته که منجر به کشف میادین متعدد در این منطقه گردیده است مؤید این ادعا است. عوامل متعددی در به وجود آوردن وضعیت مناسب نفتی - گازی فوق سهمیه بوده‌اند که از آن جمله نهشته‌های رسوبی منسجم با ضخامت کافی، وجود سنگهای مادر، مخزن و پوششی با کیفیت خوب و مفید و بالاخره چند لایه بودن افق‌های هیدروکربوری را می‌توان نام برد [۱۱].

مرحله جدید اکتشاف جهت کشف ذخایر جدید و مشخص نمودن بیشتر خصوصیات ساختمانی میادین مکشوفه موجود از سال ۱۳۶۷ توسط طرح اکتشاف و حفاری خلیج فارس در مدیریت اکتشاف شرکت ملی نفت ایران با عملیات لرزه نگاری در یک سری مناطق از پیش تعیین شده شروع گردید. مرحله اول برداشتهای لرزه نگاری منحصر به محدوده آبهای میانی و به موازات خط مرزی در خلیج فارس اختصاص یافت. نتیجه برداشتهای اولیه کشف پدیده‌های مثبت زمین شناسی متعدد در مناطق مختلف بوده که از جمله آنها ساختمان زمین شناسی پارس جنوبی که ادامه یال شمالی میدان عظیم گازی شمال قطر است می‌باشد [۱۱]. میدان پارس جنوبی در موقعیت ۲۵ تا ۲۵/۵ درجه شرقی و ۲۶/۵ تا ۲۷ درجه شمالی، در ۱۰۵ کیلومتری جنوب غربی بندر عسلویه واقع شده و بزرگ‌ترین میدان دریایی در جهان است. سازندهای کنگان و دالان به سن پرموتریاس مخازن اصلی این میدان را تشکیل می‌دهند [۱۴]. شکل ۴-۱ موقعیت جغرافیایی این میدان را نشان می‌دهد.

مرحله دوم عملیات لرزه نگاری که در سالهای ۶۹-۱۳۶۸ در مناطق مختلف دیگر از جمله بخش شمال شرقی ساختمان پارس جنوبی انجام پذیرفت، ابعاد گسترده‌تر از وسعت این میدان را در اختیار گذاشت.



شکل ۴-۱: موقعیت جغرافیایی میدان پارس جنوبی [۱۱].

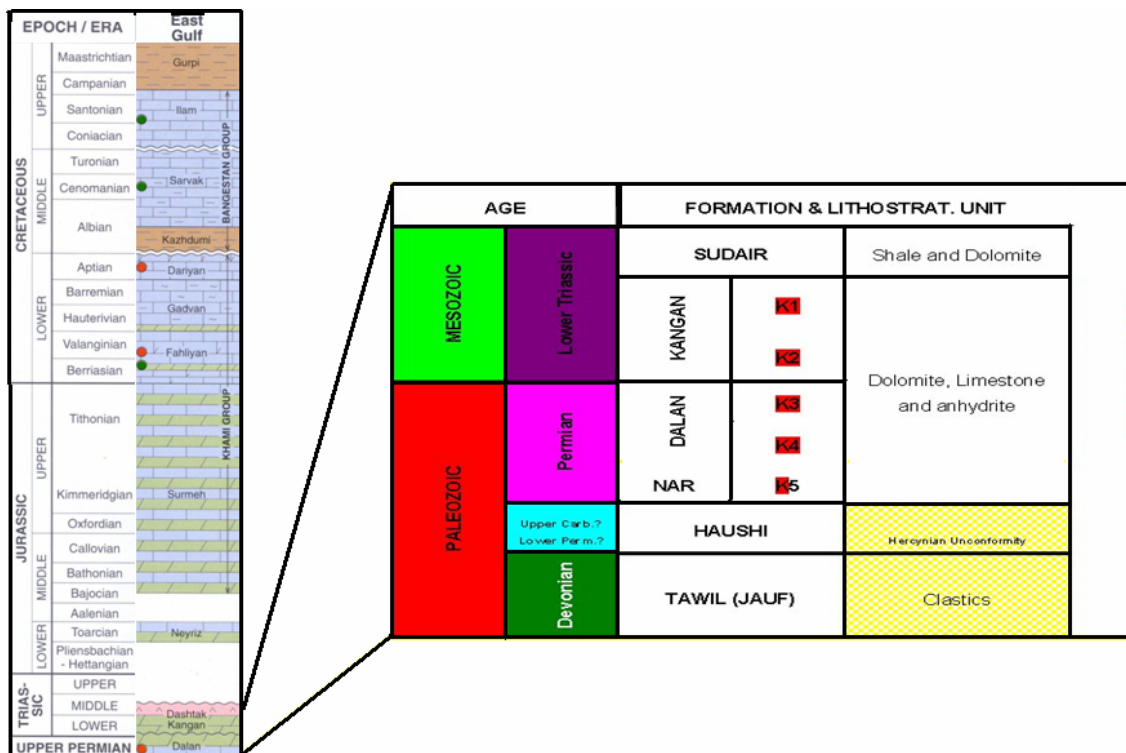
بر اساس اطلاعات حاصله از برداشت‌های لرزه نگاری، اولین چاه اکتشافی در منطقه قلعه‌ای ساختمان، نزدیک به خط میانی آب‌های خلیج فارس در سال ۷۰-۱۳۶۹ حفاری و در عمق ۳۵۲۲ متری در رسوبات ماسه سنگی فراقون به پایان رسید. نتیجه بررسی‌های حفاری و آزمایشات بهره دهی وجود مواد هیدروکربوری در طبقات جهرم (نفت آسفالتی-قیر)، داریان، گدوان و فهلیان (نفت)، سورمه بالائی (نفت سنگین) و مقادیر عظیمی از هیدروکربورهای گازی در سازندهای کنگان و دالان را مشخص نمود [۱۱].

به منظور بررسی قابلیت و توانایی شبکه‌های عصبی مصنوعی کمپته‌ای در بهبود نتایج حاصل از شبکه‌های عصبی مصنوعی منفرد، مطالعه موردی در میدان عظیم هیدروکربوری پارس جنوبی در بخش‌های گازدار k-1 و k-2 سازند کنگان انجام شد و تخلخل مؤثر ابتدا با استفاده از شبکه‌های عصبی مصنوعی منفرد و سپس با استفاده از ترکیب آنسامبلی نتایج آن‌ها تخمین زده شد.

۲-۴- زمین‌شناسی سازند کنگان

نام این سازند از منطقه کنگان واقع در جنوب ایران گرفته شده است. با توجه به اطلاعات به دست آمده از چاه‌های میدان پارس جنوبی و بر اساس نمونه‌های سنگی برداشت شده به هنگام حفاری و نمودارهای گرفته شده، رسوبات این سازند متعلق به دوران تریاس زیرین گزارش شده است. این سازند با یک ناپیوستگی موازی بر روی سازند دالان به سن پرمین قرار گرفته و در بالا توسط شیل‌های سازند دشتک پوشیده شده است [۹]. از نظر زمین‌شناسی ساختمانی، کانسار گازی کنگان تاقدیس نامتقارنی با امتداد شمال غربی- جنوب شرقی است که دامنه جنوب شرقی آن برگشته است [۱]. با ظاهر شدن اولین لایه دولومیت نخودی مایل به خاکستری رس‌دار در زیر شیل‌های آغار، شروع سازند کنگان مشخص می‌شود. بخش فوقانی شامل دولومیت خاکستری- نخودی مایل به خاکستری- بلورین و تخریبی آلی و لایه‌های نازکی از شیل به رنگ‌های سبز و سبز مایل به خاکستری می‌باشد. این بخش در چاه‌های مختلف پارس جنوبی ۲۵ تا ۳۰ متر ضخامت دارد. بعد از این بخش، ناحیه گازدار سازند کنگان قرار دارد که به طور کلی از سنگ آهک و آهک دولومیتی سفید و سفید مایل به نخودی تا قهوه‌ای روشن بلورین شبه تخمکی متخلخل و تخریبی آلی تشکیل شده است. این قسمت دارای تخلخل مناسبی می‌باشد و یکی از بخش‌های مهم گازدار میدان پارس جنوبی است. ضخامت این بخش در چاه‌های مختلف به طور متوسط ۱۵۵ متر می‌باشد [۱۱]. بر اساس توصیف مغزه و اطلاعات پتروفیزیکی و تلفیق مطالعات میکروسکوپی و ماکروسکوپی و با به کار بردن خصوصیات از قبیل سنگ‌شناسی، رخساره و خصوصیات مخزنی و با توجه به تقسیم‌بندی توالی کربنات خوف^۱ در سپر عربستان این ناحیه خود به دو واحد مخزنی مشخص و مجزا شامل k-1 و k-2 تقسیم شده است [۱۴]. در شکل ۲-۴ ستون چینه‌شناسی میدان پارس جنوبی نمایش داده شده است.

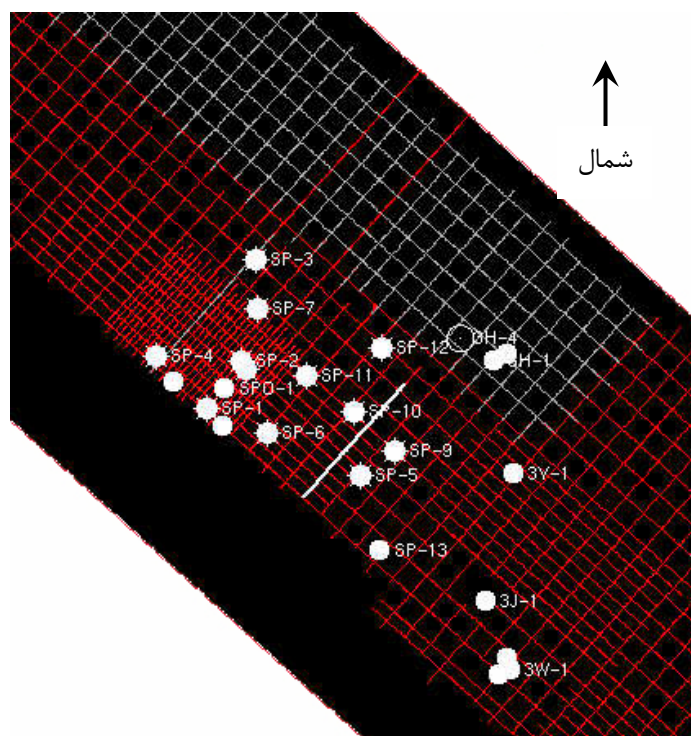
1- Khuff



شکل ۴-۲: ستون چینہ‌شناسی میدان پارس جنوبی [۱۱].

۳-۴ - چاه‌های مورد مطالعه

شکل ۴-۳ موقعیت جغرافیایی چاه‌های میدان پارس جنوبی را نشان می‌دهد.



شکل ۴-۳: موقعیت جغرافیایی چاه‌های میدان پارس جنوبی [۱۱].

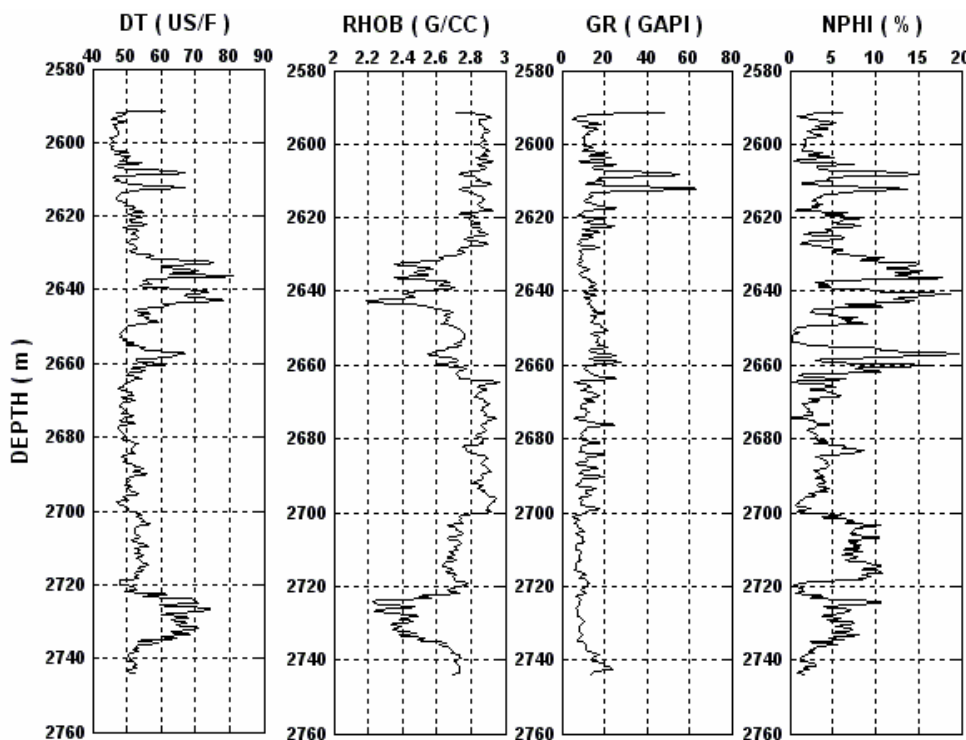
از داده‌های چاه‌های SP1، SP3، SP6 و SP13 در بازه عمقی سازند کنگان به منظور طراحی، آموزش و آزمون شبکه‌ها استفاده شد. جدول ۴-۱ عمق سرسازندهای چاه‌های مورد استفاده در این تحقیق را نشان می‌دهد.

جدول ۴-۱: عمق سرسازندهای چاه‌های مورد استفاده (بر حسب متر حفار) [۱۱].

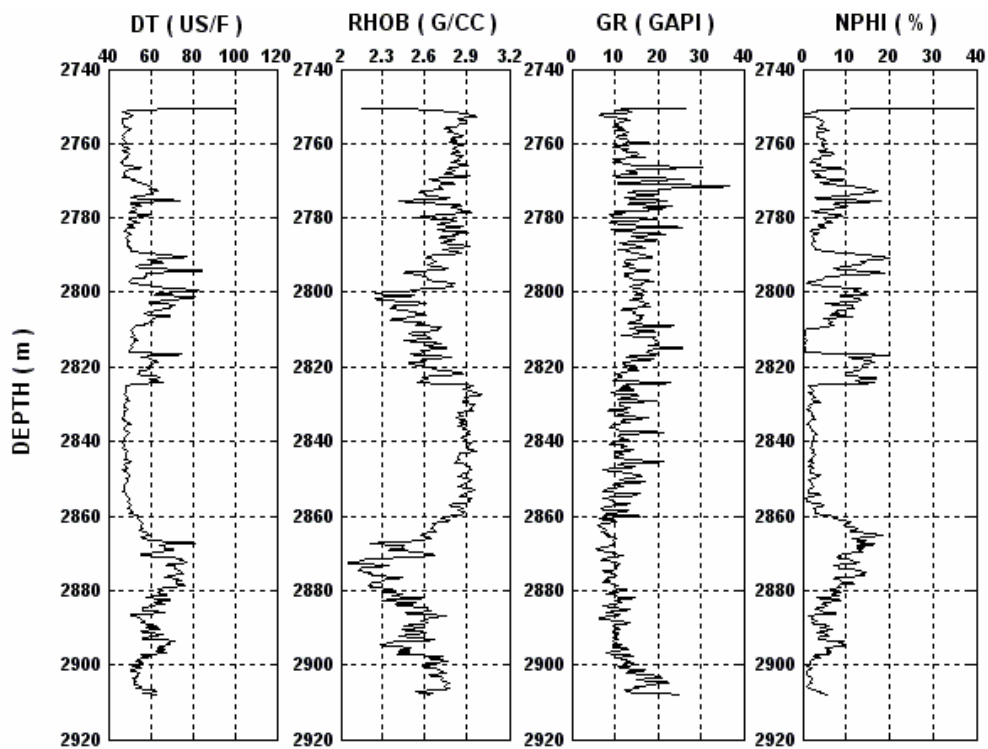
SP13	SP6	SP3	SP1	چاه سازند
۲۸۶۰/۵	۲۶۴۳	۲۷۵۰	۲۵۹۱	کنگان (K1)
۲۹۶۶	۲۷۴۷	۲۸۶۱	۲۷۰۰	کنگان (K2)
۳۰۱۳	۲۷۹۹	۲۹۰۸	۲۷۴۴	دالان (K3)

این داده‌ها شامل داده‌های نگارهای صوتی (DT)، چگالی (RHOB)، اشعه گاما (GR) و تخلخل نوترونی (NPHI) است. همچنین برای هر نقطه مقدار تخلخل مؤثر (PHIE) را در اختیار داریم. مقدار تخلخل مؤثر از روی داده‌های نگارها و با استفاده از مطالعات مغزه و اطلاعات زمین‌شناسی عمقی پس از اعمال تصحیحات مورد نیاز محاسبه شده است.

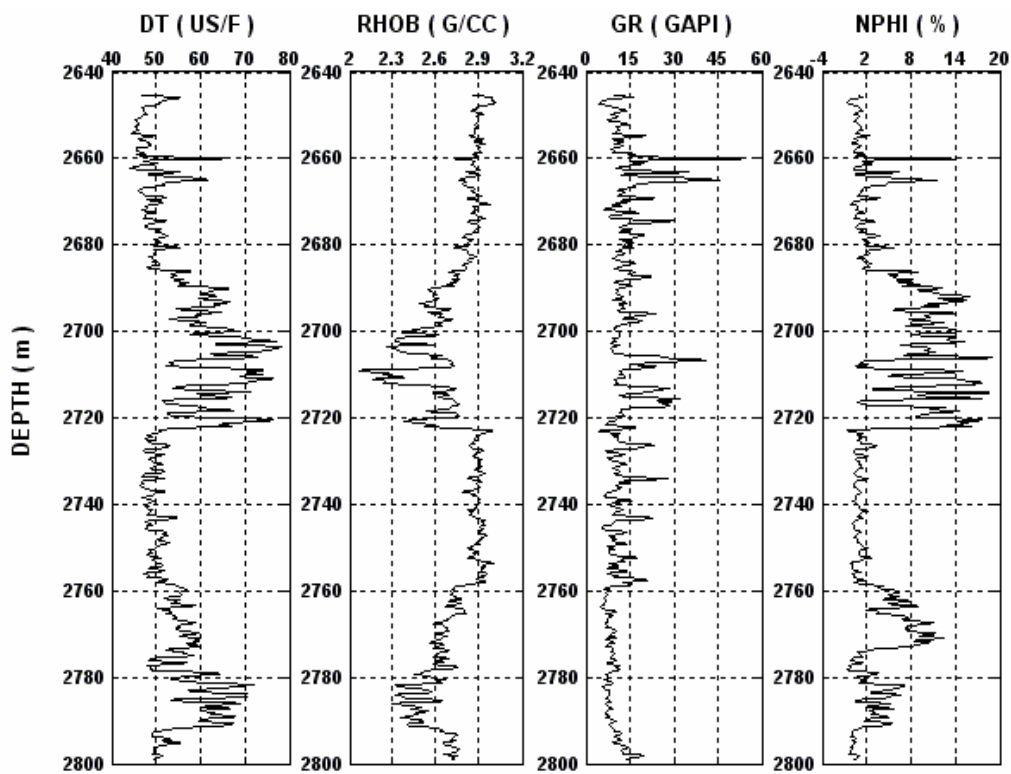
در شکل‌های ۴-۴ تا ۴-۷ نگارها بر حسب عمق برای چاه مورد استفاده نمایش داده شده‌اند.



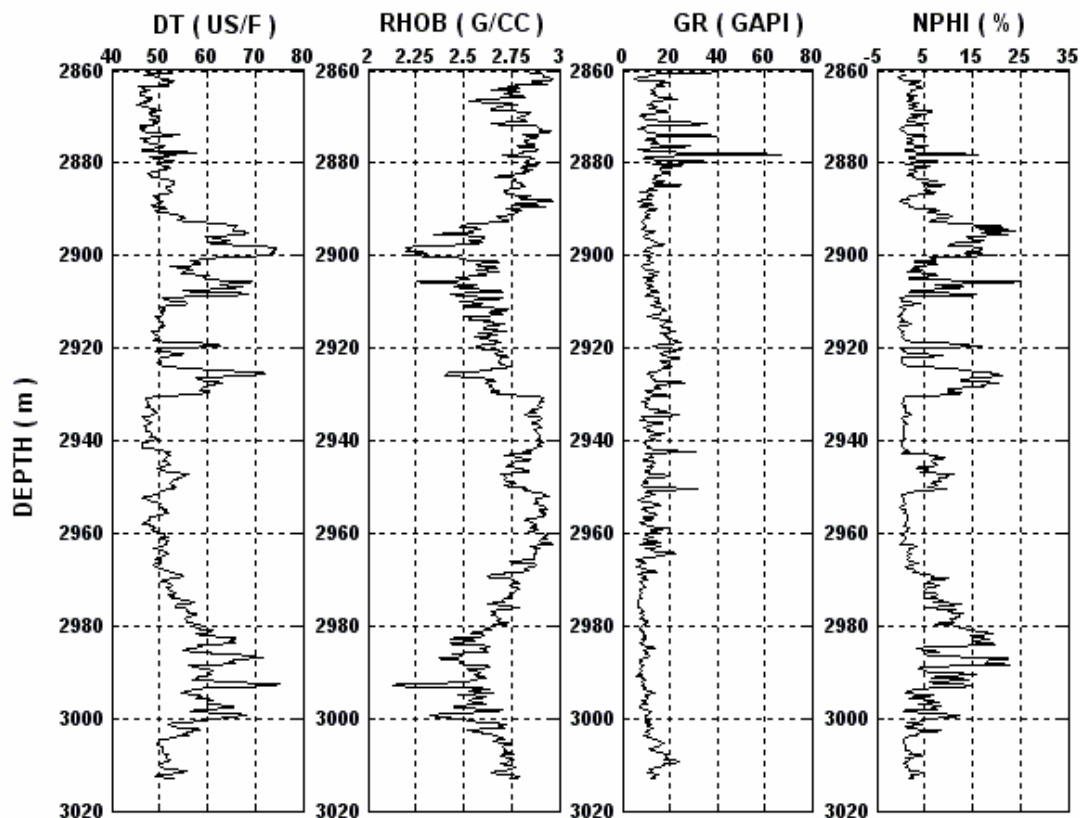
شکل ۴-۴: نگارهای چاه SP1.



شکل ۴-۵: نگارهای چاه SP3.



شکل ۴-۶: نگارهای چاه SP6.



شکل ۴-۷: نگارهای چاه SP13.

۴-۴- انتخاب الگوهای آموزش و آزمون

نگارهای صوتی، چگالی، اشعه گاما و تخلخل نوترونی به عنوان ورودی و تخلخل مؤثر به عنوان خروجی شبکه‌ها به کار رفتند. جدول ۲-۴ ضریب همبستگی خطی بین تخلخل مؤثر و هر یک از چهار نگار ورودی که از تمام داده‌های ۴ چاه مورد استفاده (۴۰۴۵ الگوی داده) به دست آمده است را نشان می‌دهد. در این جدول r نشان دهنده ضریب همبستگی خطی است.

جدول ۲-۴: ضریب همبستگی خطی بین خروجی و ورودی‌های شبکه‌های عصبی مصنوعی.

PHIE-NPHI	PHIE-GR	PHIE-RHOB	PHIE-DT
$r = ۰/۷۴۱$	$r = -۰/۱۷۴$	$r = -۰/۹۲۸$	$r = ۰/۹۲۹$

از الگوهای چاه‌های SP1، SP3 و SP13 به منظور آموزش شبکه‌ها و از الگوهای چاه SP6 برای آزمون توان تعمیم شبکه‌ها استفاده شد. علت این نحوه انتخاب آن بود که چاه SP6 تقریباً در قسمت مرکزی

منطقه مورد مطالعه قرار دارد و طبیعی است که میزان خطای تخمین کمتر باشد. در صورتی که انتخاب به شکل دیگری انجام می‌شد، مثلاً از الگوهای چاه SP3 یا SP13 برای آزمون استفاده می‌شد، چون این چاه‌ها در حاشیه منطقه مورد نظر قرار دارد این احتمال وجود داشت که نتایج با خطای بیشتری همراه باشد. الگوهای سه چاه SP1، SP3 و SP13 در یک فایل صفحه گسترده (مانند Excel) قرار داده شد و در مجموع ۳۰۳۶ الگوی داده (۱۰۰۱ الگو متعلق به چاه SP1، ۱۰۳۴ الگو متعلق به چاه SP3 و ۱۰۰۱ الگو متعلق به چاه SP13) به دست آمد که هر یک از این الگوها شامل ۶ ستون عمق، نگار صوتی، چگالی، اشعه گاما، تخلخل نوترونی و تخلخل مؤثر است. در جدول ۳-۴ قسمتی از این فایل که بر حسب نگار صوتی مرتب^۱ شده به نمایش درآمده است.

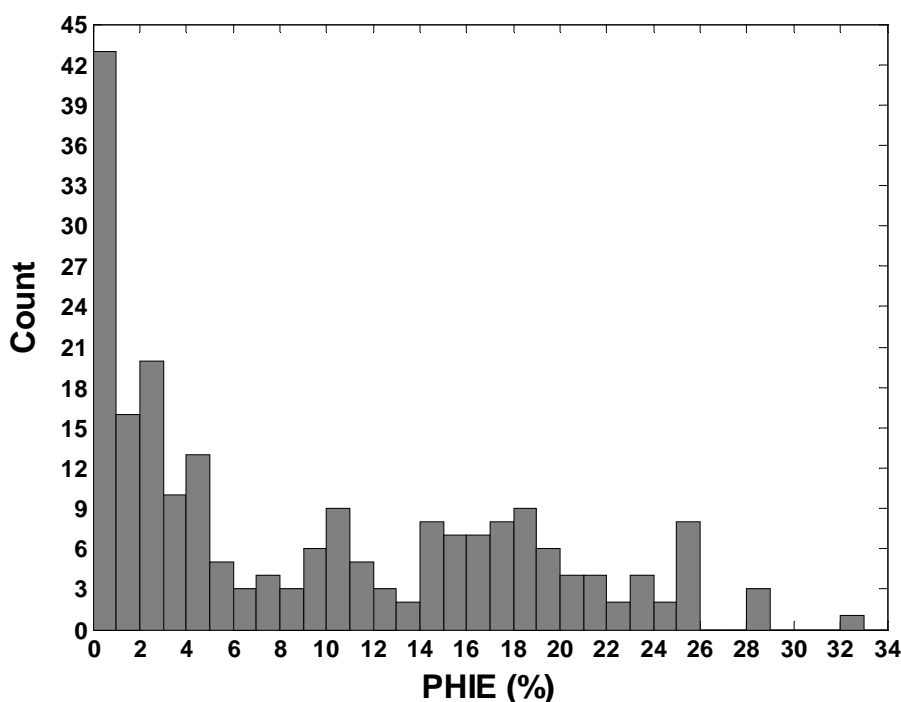
جدول ۳-۴: نمایش بخشی از فایل Excel شامل الگوهای چاه‌های SP1،

SP3 و SP13 که الگوهای آموزش از میان آن‌ها انتخاب شده است.

DEPTH (m)	DT (US/F)	RHOB (G/CC)	GR (GAPI)	NPHI	PHIE
2638.04	57.0529	2.6235	15.7539	0.0438	0.0755
2779.01	57.0620	2.7394	9.4563	0.0733	0.0557
2977.74	57.0625	2.6550	8.0582	0.1133	0.0992
2888.13	57.0681	2.5703	10.4832	0.0316	0.0669
3002.74	57.0795	2.5325	9.4581	0.0842	0.1278
2630.88	57.0806	2.6162	8.0390	0.1089	0.0991
2974.85	57.0883	2.7174	6.4070	0.0690	0.0676
2865.42	57.0900	2.5992	8.2380	0.1606	0.0905
2824.43	57.1103	2.6753	20.9055	0.1410	0.0675
2795.47	57.1175	2.5619	12.7981	0.0947	0.1117
2984.91	57.1261	2.5888	13.6866	0.0321	0.0983
2660.75	57.1285	2.7695	10.4954	0.0820	0.0485
2975.00	57.1402	2.6766	6.3553	0.0882	0.0815
2805.68	57.1429	2.5549	14.1446	0.0804	0.1072
2993.90	57.1646	2.5570	10.1976	0.0285	0.0931
2723.39	57.1760	2.5074	9.1027	0.0370	0.0871
2901.39	57.1887	2.6800	11.7884	0.0279	0.0930
2984.75	57.2043	2.6035	13.3712	0.0364	0.1073
2722.93	57.2044	2.5349	8.0952	0.0276	0.0750

همچنین به طور مشابه از چاه SP6 نیز تعداد ۱۰۰۹ الگوی داده موجود است که الگوهای آزمون تعمیم از میان آن‌ها انتخاب می‌شود. اگر چه مشخص است که محدودیتی از نظر تعداد الگوها وجود ندارد و مقادیر تخلخل مؤثر توسط مهندسین شرکت نفت محاسبه شده است، اما در شرایط واقعی که معمولاً یک کارشناس مخزن با آن مواجه است محاسبه مقادیر تخلخل با محدودیت همراه می‌باشد و مستلزم محاسبات فراوان، استفاده از اطلاعات زمین‌شناسی و مغزه‌ها و انجام تصحیحات مورد نیاز و به طور کلی صرف هزینه و وقت زیاد است. به همین دلیل، به منظور صرفه‌جویی در هزینه و زمان در تعیین مقادیر تخلخل از روی نگارها از روش شبکه‌های عصبی استفاده می‌شود.

همان طور که قبلاً اشاره شد انتخاب یک مجموعه آموزشی مناسب که دامنه وسیعی از الگوهای تحقق‌یافته را در بر گیرد از اهمیت زیادی برخوردار است، زیرا کیفیت عملکرد نهایی شبکه‌های عصبی مستقیماً تابع الگوهای آموزشی به کار رفته در آموزش آن است. مهم‌ترین شرط لازم برای مجموعه آموزشی آن است که تمام دامنه مقادیر ورودی‌ها و خروجی‌ها را پوشش دهد. در بسیاری از مسائل، پوشش مناسب ورودی‌ها به دامنه تغییراتشان موجب پوشش مناسب به خروجی‌ها می‌شود. به منظور پوشش مناسب ورودی‌ها و در نتیجه خروجی، ۳۰۳۶ الگوی داده ۴ مرتبه، هر بار بر حسب یکی از نگارهای ورودی به صورت افزایشی مرتب شدند و انتخاب به گونه‌ای انجام شد که تمام دامنه مقادیر نگار مورد نظر را در بر گیرد. در مرتب‌سازی بر حسب نگار صوتی ۶۰ الگوی داده، بر حسب نگار چگالی ۵۳ الگوی داده، بر حسب نگار گاما ۶۵ الگوی داده و بر حسب نگار نوترون نیز ۴۹ الگوی داده و در مجموع ۲۲۷ الگوی داده انتخاب شد که پس از حذف الگوهای تکراری، ۲۱۵ الگو برای آموزش شبکه‌ها باقی ماند. پس از انتخاب این ۲۱۵ الگو مشخص شد که ۶۷ الگوی آن‌ها متعلق به چاه SP1، ۷۷ الگو متعلق به چاه SP3 و ۷۱ الگو متعلق به چاه SP13 می‌باشد. علت تفاوت در تعداد الگوهای انتخابی در مرتب‌سازی بر حسب نگارهای مختلف عبارت است از تفاوت در دامنه تغییرات نگارهای مختلف، تفاوت فراوانی تعداد الگوها در بازه‌های مختلف یک نگار خاص و تأکید بر انتخاب الگوها به گونه‌ای که از همه محدوددها به خصوص محدوددهای با فراوانی کم نیز انتخاب صورت گیرد. شکل ۴-۸ نمودار توزیع فراوانی مقادیر تخلخل مؤثر در الگوهای آموزشی را نشان می‌دهد. با توجه به این شکل مشخص است که فراوانی الگوهای با تخلخل بزرگ نسبت به الگوهای با تخلخل کوچک، بسیار کمتر است.



شکل ۴-۸: نمودار توزیع فراوانی مقادیر تخلخل مؤثر در ۲۱۵ الگوی آموزشی.

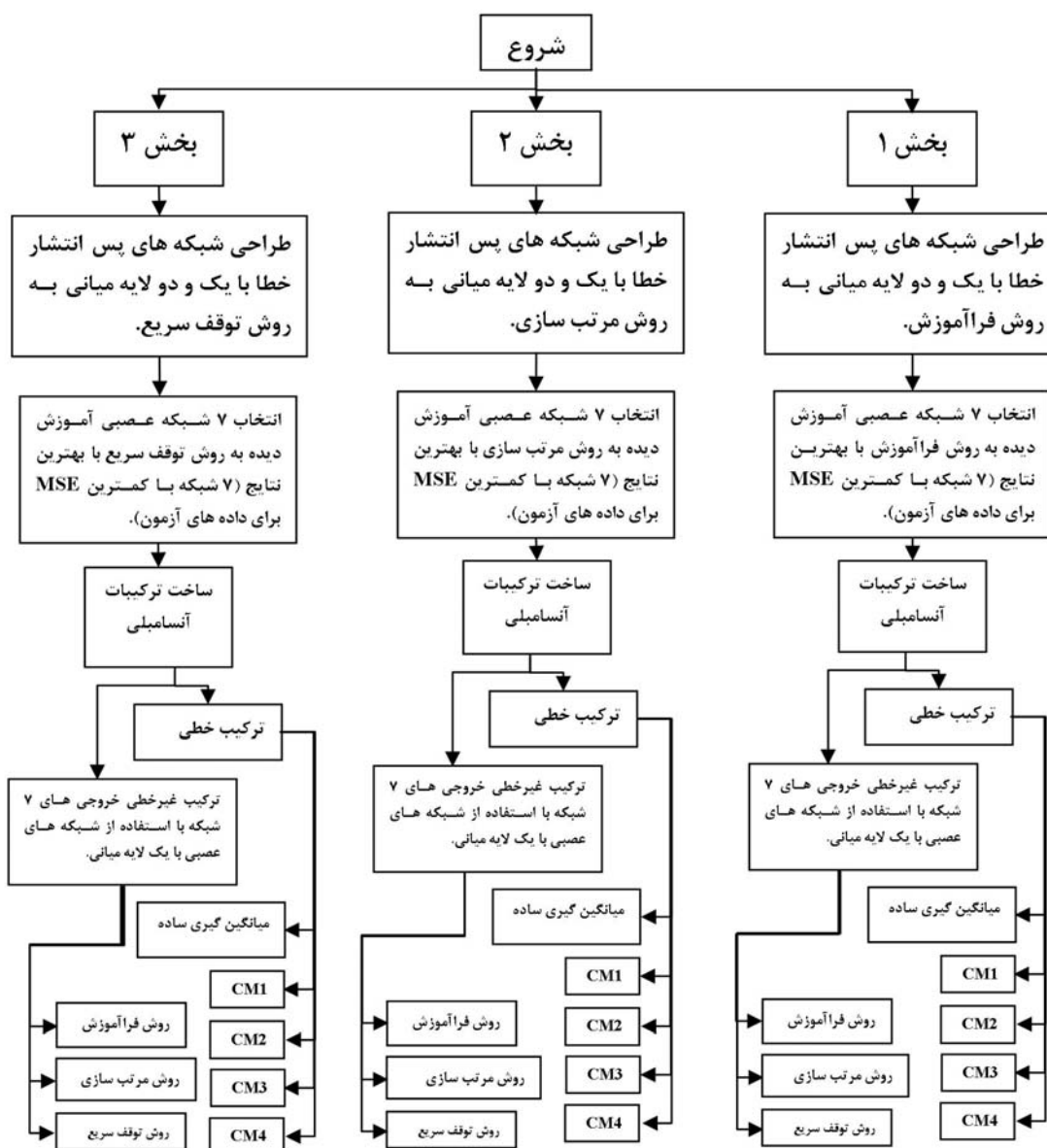
برای انتخاب الگوهای آزمون، ۱۰۰۹ الگوی داده چاه SP6 بر حسب عمق مرتب شدند و الگوها تقریباً از تمام بازه‌های عمقی انتخاب شدند. در نهایت ۸۹ الگوی آزمون به دست آمد.

۴-۵- برنامه‌نویسی کامپیوتری

کلیه برنامه‌های مورد استفاده در این تحقیق و همچنین کارهای محاسباتی در محیط نرم‌افزار MATLAB تهیه و انجام شده است. طراحی شبکه‌های پس‌انتشار خطا به سه روش فرآآموزش، مرتب‌سازی و توقف سریع انجام شد. در هر سه مورد به منظور پیش‌پردازش داده‌ها از تابع `Prestd` که مقادیر داده‌ها را به توزیع گوسی استاندارد تبدیل می‌کند استفاده کردیم. تعداد دوره‌های تکرار برای آموزش شبکه‌ها در هر سه مورد ۴۰۰ مرتبه انتخاب شد. تابع انتقال یا فعال‌سازی نرون‌های لایه‌های میانی تابع تانژانت هیپربولیک و تابع انتقال نرون‌های لایه خروجی تابع خطی انتخاب شد. در شکل ۴-۹ مجموعه کارهایی که انجام شده و در ادامه این فصل به شرح آن‌ها می‌پردازیم، نمایش داده شده است. این کارها در سه بخش کلی انجام شده‌اند: ۱- طراحی شبکه‌های پس‌انتشار خطا به روش فرآآموزش و ترکیب نتایج، ۲- طراحی شبکه‌های پس‌انتشار خطا به روش مرتب‌سازی و ترکیب نتایج و

۳- طراحی شبکه‌های پس‌انتشار خطا به روش توقف سریع و ترکیب نتایج. در هر بخش، شبکه‌هایی با یک و دو لایه میانی با تعداد نرون‌های مختلف در این لایه‌ها، در نظر گرفته شده و در مورد هر ساختار خاص شبکه (نرون‌های میانی مشخص)، یک مرحله آزمون و خطای طولانی با شرایط اولیه مختلف انجام شده است و بهترین نتیجه که کمترین میانگین مربعات خطای تخمین را برای الگوهای آزمون داشت ثبت گردید. سپس از هر یک از سه روش آموزش، ۷ شبکه با ساختارهای مختلف که بهترین نتایج را داشتند (کمترین MSE تخمین را برای الگوهای آزمون داشتند) برای ساخت ترکیبات آنسامبلی انتخاب شدند. ترکیبات آنسامبلی این ۷ شبکه به دو صورت کلی ساخته شدند: ۱- ترکیب خطی شبکه‌ها و ۲- ترکیب غیرخطی شبکه‌ها با استفاده از شبکه‌های عصبی با یک لایه میانی. ترکیب خطی شبکه‌ها با ۵ روش انجام شد. این ۵ روش عبارتند از میانگین‌گیری ساده، ترکیب خطی بهینه نامقید با جمله ثابت (CM1)، ترکیب خطی بهینه مقید با جمله ثابت (CM2)، ترکیب خطی بهینه نامقید بدون جمله ثابت (CM3) و ترکیب خطی بهینه مقید بدون جمله ثابت (CM4). در هر یک از این ۵ روش ۱۲۰ ترکیب آنسامبلی ممکن این ۷ شبکه، یعنی ۲۱ ترکیب دوشبکه‌ای، ۳۵ ترکیب سه‌شبکه‌ای، ۳۵ ترکیب چهارشبکه‌ای، ۲۱ ترکیب پنج‌شبکه‌ای، ۷ ترکیب شش‌شبکه‌ای و ۱ ترکیب هفت‌شبکه‌ای ساخته شد و بهترین ترکیب انتخاب گردید.

در هر یک از سه بخش اصلی، به منظور ترکیب آنسامبلی غیرخطی ۷ شبکه منتخب، از خروجی این ۷ شبکه به عنوان ورودی شبکه‌های پس‌انتشارخطا با یک لایه میانی که تابع انتقال نرون‌های آن، تابع تانژانت هیپربولیک و تابع انتقال نرون‌های لایه خروجی تابع خطی بود استفاده شد. این بار نیز تعداد دوره‌های تکرار ۴۰۰ بار انتخاب شد. برای آموزش شبکه‌های ترکیب‌کننده، هر سه روش آموزشی فراآموزش، مرتب‌سازی و توقف سریع به کار رفتند و در هر روش آموزش از تعداد نرون‌های مختلف در لایه میانی استفاده گردید. برای هر ساختار خاص یک مرحله طولانی آزمون و خطا با شرایط اولیه مختلف انجام شد. در نهایت بهترین نتیجه حاصل از هر یک از سه روش آموزش، در هر یک از سه بخش اصلی انتخاب شد.



شکل ۴-۹: نمای شماتیک مجموعه کارهایی که در رابطه با طراحی شبکه های عصبی منفرد و ترکیب آن ها در فصل چهارم انجام شده است.

۴-۶- طراحی شبکه های پس انتشار خطا به روش فراآموزش و ترکیب نتایج

برای آموزش به روش فراآموزش از تابع آموزش لوببرگ- مارکوارت استفاده شد. شبکه های پس انتشار خطا با یک و دو لایه میانی با تعداد نرون های متعدد با شرایط اولیه مختلف طی یک مرحله طولانی آزمون و خطا مورد آزمون قرار گرفت و برای هر ساختار خاص بهترین نتیجه ثبت شد. معیار توقف روند افزایش نرون در لایه های میانی، افزایش احتمال بیش برآزش بوده است. در جدول ۴-۴ نتایج کامل این مرحله نمایش داده شده است. ستون های این جدول از سمت راست عبارتند از شماره

ردیف، ساختار شبکه، میانگین مربعات خطای تخمین شبکه برای الگوهای آموزش، ضریب همبستگی بین مقادیر تخلخل و تخمین شبکه برای الگوهای آموزش، میانگین مربعات خطای تخمین شبکه برای الگوهای آزمون، ضریب همبستگی بین مقادیر تخلخل و تخمین شبکه برای الگوهای آزمون و تعداد کل پارامترهای شبکه. در بالای این جدول و جدول‌های مشابه این فصل، مقادیر به کار رفته پارامترهای تابع آموزشی ذکر شده‌اند. این پارامترها در پیوست الف معرفی شده است.

جدول ۴-۴: نتایج حاصل از شبکه‌های عصبی پس‌انتشار خطا با یک و دو لایه میانی به روش فراآموزش.

epochs = 400, show = 20, goal = 0, min_grad = 1.00×10^{-10} , mem_reduce = 1, mu = 0.0010, mu_dec = 0.100, mu_inc = 10, mu_max = $1.00 \times 10^{+010}$.

تعداد کل پارامترهای شبکه	ضریب همبستگی (الگوهای آزمون)	MSE (الگوهای آزمون)	ضریب همبستگی (الگوهای آموزش)	MSE (الگوهای آموزش)	ساختار شبکه	ردیف
۱۹	۰/۹۸۲۹	۲/۶۳۵۹	۰/۹۹۱۴	۱/۲۳۸۱	۴-۳-۱	۱
۲۵	۰/۹۸۵۵	۲/۵۳۵۹	۰/۹۹۲۳	۱/۱۰۱	۴-۴-۱	۲
۳۱	۰/۹۸۳	۲/۵۴۱۷	۰/۹۹۲۵	۱/۰۷۰۵	۴-۵-۱	۳
۳۷	۰/۹۸۴۴	۲/۴۴۶۷	۰/۹۹۳۲	۰/۹۷۷۱	۴-۶-۱	۴
۴۳	۰/۹۸۲	۲/۵۳۲۶	۰/۹۹۴۷	۰/۷۵۹۵	۴-۷-۱	۵
۴۹	۰/۹۸۵۹	۲/۷۵۱۳	۰/۹۹۴۴	۰/۷۹۷	۴-۸-۱	۶
۵۵	۰/۹۸۳۷	۲/۸۹۲۵	۰/۹۹۴۸	۰/۷۵۳۸	۴-۹-۱	۷
۶۱	۰/۹۸۲	۲/۵۶۲۲	۰/۹۹۵۳	۰/۶۷۹۴	۴-۱۰-۱	۸
۶۷	۰/۹۸۱۳	۲/۸۸۲	۰/۹۹۶۱	۰/۵۶۱۶	۴-۱۱-۱	۹
۷۳	۰/۹۸۲۹	۲/۹۱۴۷	۰/۹۹۴۷	۰/۷۶۳۷	۴-۱۲-۱	۱۰
۷۹	۰/۹۷۴۹	۳/۴۸۵۲	۰/۹۹۶۲	۰/۵۴۰۱	۴-۱۳-۱	۱۱
۱۹	۰/۹۸۴۲	۲/۴۹۷۷	۰/۹۸۹۹	۱/۴۵۱۳	۴-۲-۲-۱	۱۲
۲۳	۰/۹۸۶۲	۲/۴۷۸۲	۰/۹۹۱۱	۱/۲۶۸۸	۴-۲-۳-۱	۱۳
۲۷	۰/۹۸۵۴	۲/۴۱۳۶	۰/۹۹۱۷	۱/۱۹۲۵	۴-۲-۴-۱	۱۴
۳۱	۰/۹۸۶۵	۲/۳۶۹۵	۰/۹۹۱	۱/۲۹۷	۴-۲-۵-۱	۱۵
۳۵	۰/۹۸۶۸	۲/۴۵۹۲	۰/۹۹۱۶	۱/۲۰۱۸	۴-۲-۶-۱	۱۶
۳۹	۰/۹۸۵۴	۲/۵۰۷۴	۰/۹۹۲۴	۱/۰۹۲۴	۴-۲-۷-۱	۱۷
۴۳	۰/۹۸۶	۲/۴۶۷۵	۰/۹۹۲۸	۱/۰۳۲۶	۴-۲-۸-۱	۱۸
۴۷	۰/۹۸۵۵	۲/۵۳۴۳	۰/۹۹۳۴	۰/۹۴۷۱	۴-۲-۹-۱	۱۹
۵۱	۰/۹۸۴۱	۲/۵۲۳	۰/۹۹۲۳	۱/۱۰۲۱	۴-۲-۱۰-۱	۲۰
۵۵	۰/۹۸۶۸	۲/۲۸۷۷	۰/۹۹۲۶	۱/۰۶۵۹	۴-۲-۱۱-۱	۲۱
۵۹	۰/۹۸۴۳	۲/۴۲۴۸	۰/۹۹۲۷	۱/۰۴۶۸	۴-۲-۱۲-۱	۲۲
۶۳	۰/۹۸۷۴	۲/۴۲۹۴	۰/۹۹۳۵	۰/۹۲۹۲	۴-۲-۱۳-۱	۲۳
۶۷	۰/۹۸۵	۲/۴۶۸۴	۰/۹۹۱۷	۱/۱۹۱۴	۴-۲-۱۴-۱	۲۴
۷۱	۰/۹۸۶۳	۲/۵۸۷۲	۰/۹۹۳۳	۰/۹۵۷۱	۴-۲-۱۵-۱	۲۵
۷۵	۰/۹۸۳۳	۲/۴۲۶۴	۰/۹۹۱۸	۱/۱۷۱۷	۴-۲-۱۶-۱	۲۶
۷۹	۰/۹۸۴۲	۲/۳۷۴	۰/۹۹۲۶	۱/۰۵۴۷	۴-۲-۱۷-۱	۲۷

ادامه جدول ۴-۴

تعداد کل پارامترهای شبکه	ضریب همبستگی (الگوهای آزمون)	MSE (الگوهای آزمون)	ضریب همبستگی (الگوهای آموزش)	MSE (الگوهای آموزش)	ساختار شبکه	ردیف
۸۳	۰/۹۸۰۵	۲/۷۲۳	۰/۹۹۵۶	۰/۶۳۷۳	۴-۲-۱۸-۱	۲۸
۸۷	۰/۹۸۸۱	۲/۲۴۳۴	۰/۹۹۲۱	۱/۱۳۶۳	۴-۲-۱۹-۱	۲۹
۹۱	۰/۹۸۴۵	۲/۷۰۴	۰/۹۹۱۵	۱/۲۱۲۹	۴-۲-۲۰-۱	۳۰
۲۶	۰/۹۸۷۳	۲/۲۵۸۲	۰/۹۹۰۸	۱/۳۱۹۹	۴-۳-۲-۱	۳۱
۳۱	۰/۹۸۷۱	۲/۳۳۳۸	۰/۹۹۲۳	۱/۱۰۴	۴-۳-۳-۱	۳۲
۳۶	۰/۹۸۶۹	۲/۴۱۲۳	۰/۹۹۲۶	۱/۰۶۱۷	۴-۳-۴-۱	۳۳
۴۱	۰/۹۸۶۳	۲/۳۷۵۴	۰/۹۹۳۲	۰/۹۷۵	۴-۳-۵-۱	۳۴
۴۶	۰/۹۸۴۱	۲/۵۹۰۶	۰/۹۹۲۸	۱/۰۳۱۶	۴-۳-۶-۱	۳۵
۵۱	۰/۹۸۳۵	۲/۵۰۸۶	۰/۹۹۴۲	۰/۸۳۱	۴-۳-۷-۱	۳۶
۵۶	۰/۹۸۵۸	۲/۴۴۹۲	۰/۹۹۲۵	۱/۰۸۰۳	۴-۳-۸-۱	۳۷
۶۱	۰/۹۸۳۸	۲/۷۷۶۹	۰/۹۹۴۵	۰/۷۸۸۳	۴-۳-۹-۱	۳۸
۶۶	۰/۹۸۴۷	۲/۶۹۵۷	۰/۹۹۹۵	۰/۷۱۷	۴-۳-۱۰-۱	۳۹
۷۱	۰/۹۸۴۵	۲/۶۷۵۹	۰/۹۹۹۳	۱/۰۰۳۷	۴-۳-۱۱-۱	۴۰
۷۶	۰/۹۸۴۹	۲/۴۱۰۵	۰/۹۹۹۵	۰/۷۲۴۵	۴-۳-۱۲-۱	۴۱
۸۱	۰/۹۸۲	۳/۰۷۰۳	۰/۹۹۴۱	۰/۸۴۶۷	۴-۳-۱۳-۱	۴۲
۸۶	۰/۹۸۱۲	۳/۰۹۲۱	۰/۹۹۳۸	۰/۸۹۵۷	۴-۳-۱۴-۱	۴۳
۳۳	۰/۹۸۵۲	۲/۴۹۲۶	۰/۹۹۲۶	۱/۰۶۶۲	۴-۴-۲-۱	۴۴
۳۹	۰/۹۸۵۳	۲/۶۴۰۲	۰/۹۹۲۷	۱/۰۴۳۱	۴-۴-۳-۱	۴۵
۴۵	۰/۹۸۴	۲/۷۳۷۶	۰/۹۹۲۵	۱/۰۷۱۲	۴-۴-۴-۱	۴۶
۵۱	۰/۹۸۷۶	۲/۵۰۳۴	۰/۹۹۹۳	۱	۴-۴-۵-۱	۴۷
۵۷	۰/۹۸۲۶	۲/۷۶۲۹	۰/۹۹۳۹	۰/۸۶۹	۴-۴-۶-۱	۴۸
۶۳	۰/۹۸۲۹	۲/۶۶۷۶	۰/۹۹۹۷	۰/۴۳۵۴	۴-۴-۷-۱	۴۹
۶۹	۰/۹۸۳۴	۲/۵۰۷۵	۰/۹۸۷۶	۰/۸۹۲۸	۴-۴-۸-۱	۵۰
۷۵	۰/۹۷۸۹	۲/۹۱۳۴	۰/۹۹۵۹	۰/۵۸۸۵	۴-۴-۹-۱	۵۱
۸۱	۰/۹۷۹۷	۲/۹۴۴۷	۰/۹۹۹۶	۰/۵۶۹۶	۴-۴-۱۰-۱	۵۲
۸۷	۰/۹۷۹۹	۲/۹۶۸۳	۰/۹۹۴۷	۰/۷۵۴۸	۴-۴-۱۱-۱	۵۳
۹۳	۰/۹۷۷۴	۲/۹۱۰۵	۰/۹۹۵۶	۰/۶۳۷۳	۴-۴-۱۲-۱	۵۴
۹۹	۰/۹۸۳۹	۲/۴۳۲۸	۰/۹۹۶۹	۰/۴۴۱۲	۴-۴-۱۳-۱	۵۵
۱۰۵	۰/۹۷۹۸	۳/۸۷۹	۰/۹۹۷۸	۰/۳۱۹۵	۴-۴-۱۴-۱	۵۶
۴۰	۰/۹۸۴۵	۲/۶۳۹۶	۰/۹۹۳۸	۰/۸۹۶۶	۴-۵-۲-۱	۵۷
۴۷	۰/۹۸۲۹	۲/۵۹۲۷	۰/۹۹۳۲	۰/۹۸۲۲	۴-۵-۳-۱	۵۸
۵۴	۰/۹۸۴۵	۲/۵۰۶۳	۰/۹۹۴۴	۰/۸۰۱۴	۴-۵-۴-۱	۵۹
۶۱	۰/۹۸۴۱	۲/۹۱۵۶	۰/۹۹۹۵	۰/۷۱۲۳	۴-۵-۵-۱	۶۰
۶۸	۰/۹۷۹۴	۳/۰۲۰۳	۰/۹۹۹۵	۰/۷۲۱۶	۴-۵-۶-۱	۶۱
۷۵	۰/۹۸۱	۲/۶۸۲۸	۰/۹۹۵۹	۰/۵۸۷۷	۴-۵-۷-۱	۶۲
۸۲	۰/۹۸۲۷	۲/۸۴۹۹	۰/۹۹۹۷	۰/۴۳۷۱	۴-۵-۸-۱	۶۳
۸۹	۰/۹۸۳۱	۲/۸۶۵۳	۰/۹۹۳۷	۰/۸۹۹۲	۴-۵-۹-۱	۶۴
۹۶	۰/۹۷۸۱	۲/۹۴۷۶	۰/۹۹۹۷	۰/۴۳۱۳	۴-۵-۱۰-۱	۶۵

ادامه جدول ۴-۴

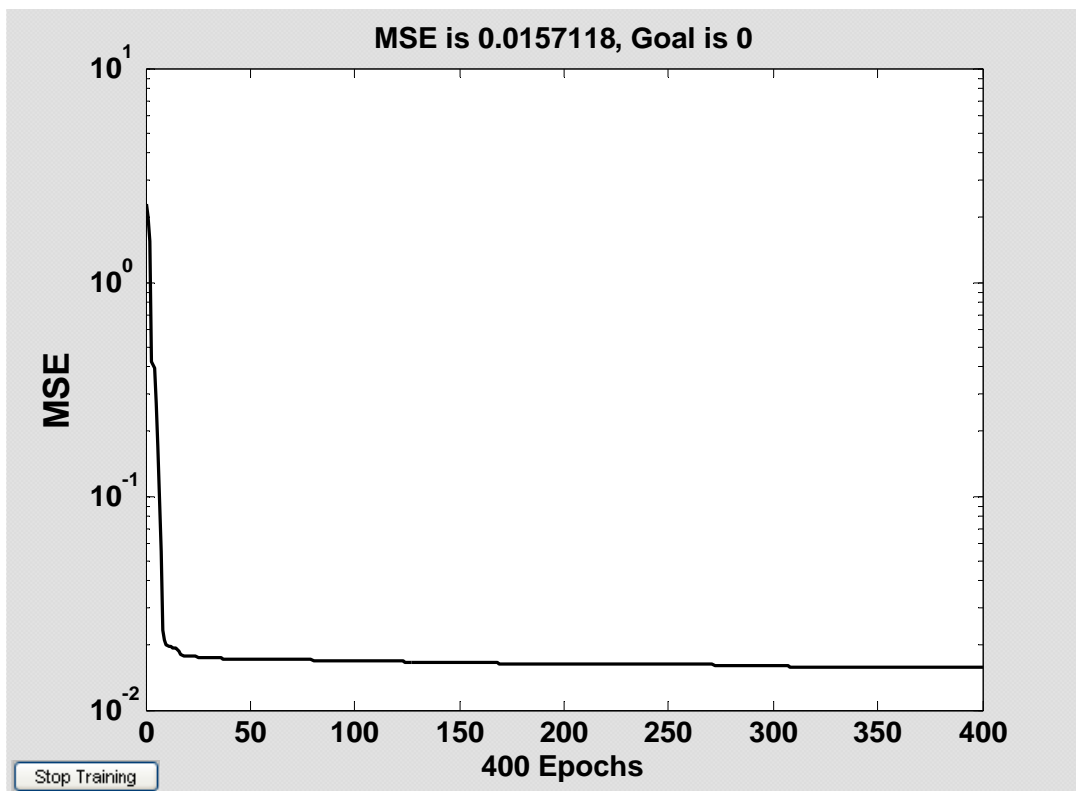
ردیف	ساختار شبکه	MSE (الگوهای آموزش)	ضریب همبستگی (الگوهای آموزش)	MSE (الگوهای آزمون)	ضریب همبستگی (الگوهای آزمون)	تعداد کل پارامترهای شبکه
۶۶	۴-۶-۲-۱	۰/۸۵۵۷	۰/۹۹۴	۲/۶۰۸۵	۰/۹۸۶۱	۴۷
۶۷	۴-۶-۳-۱	۰/۸۱۸۴	۰/۹۹۴۳	۲/۹۲۶۱	۰/۹۸۰۹	۵۵
۶۸	۴-۶-۴-۱	۰/۷۷۷۳	۰/۹۹۴۶	۲/۴۵۱۲	۰/۹۸۳۹	۶۳
۶۹	۴-۶-۵-۱	۰/۶۵۱۸	۰/۹۹۵۵	۳/۲۳۵۵	۰/۹۷۷۶	۷۱
۷۰	۴-۶-۶-۱	۰/۴۵۴۴	۰/۹۹۶۸	۳/۲۷۸۳	۰/۹۷۵۷	۷۹
۷۱	۴-۶-۷-۱	۰/۴۱۸۳	۰/۹۹۷۱	۳/۱۵۶۸	۰/۹۷۴۳	۸۷
۷۲	۴-۶-۸-۱	۰/۸۱۰۵	۰/۹۹۴۴	۳/۱۹۴۷	۰/۹۷۷۸	۹۵
۷۳	۴-۷-۲-۱	۰/۸۵۹۲	۰/۹۹۴	۳/۰۹۵۴	۰/۹۸۲۶	۵۴
۷۴	۴-۷-۳-۱	۰/۹۱۹۱	۰/۹۹۳۶	۲/۷۴۴۹	۰/۹۸۲۲	۶۳
۷۵	۴-۷-۴-۱	۰/۵۵۴۵	۰/۹۹۶۱	۲/۸۶۶۲	۰/۹۷۹۱	۷۲
۷۶	۴-۷-۵-۱	۰/۴۷۷۱	۰/۹۹۶۷	۳/۳۳۸۲	۰/۹۷۹۹	۸۱
۷۷	۴-۷-۶-۱	۰/۳۷۹۷	۰/۹۹۷۴	۳/۴۱۳۴	۰/۹۷۹۱	۹۰
۷۸	۴-۷-۷-۱	۰/۳۵۲۱	۰/۹۹۷۶	۳/۶۸۹۱	۰/۹۷۶۶	۹۹
۷۹	۴-۷-۸-۱	۰/۲۳۵۵	۰/۹۹۸۴	۳/۶۴۸۹	۰/۹۷۳۴	۱۰۸
۸۰	۴-۸-۲-۱	۰/۶۳۰۱	۰/۹۹۵۶	۲/۷۹۵۳	۰/۹۸۳۳	۶۱
۸۱	۴-۸-۳-۱	۰/۶۳۲۷	۰/۹۹۵۶	۲/۷۸۰۳	۰/۹۸۲۵	۷۱
۸۲	۴-۸-۴-۱	۰/۹۷۱۷	۰/۹۹۳۲	۲/۵۵۸۴	۰/۹۸۴۱	۸۱
۸۳	۴-۸-۵-۱	۰/۵۹۹۵	۰/۹۹۵۸	۳/۶۶۱۵	۰/۹۷۵۴	۹۱
۸۴	۴-۸-۶-۱	۰/۳۴۶۸	۰/۹۹۷۶	۳/۶۰۳۵	۰/۹۷۴۵	۱۰۱
۸۵	۴-۹-۲-۱	۰/۷۶۳۵	۰/۹۹۴۷	۲/۵۸۹۱	۰/۹۸۳۸	۶۸
۸۶	۴-۹-۳-۱	۰/۵۰۳۳	۰/۹۹۶۵	۳/۵۷۶۵	۰/۹۷۷۹	۷۹
۸۷	۴-۹-۴-۱	۰/۳۳۹۴	۰/۹۹۷۶	۳/۶۶۸۸	۰/۹۷۲۵	۹۰
۸۸	۴-۱۰-۲-۱	۰/۳۹۸۱	۰/۹۹۷۲	۳/۵۴۱۸	۰/۹۷۷۵	۷۵
۸۹	۴-۱۰-۳-۱	۰/۶۵۱۳	۰/۹۹۵۵	۳/۲۵۷۷	۰/۹۷۴۲	۸۷
۹۰	۴-۱۰-۴-۱	۰/۲۷۱۳	۰/۹۹۸۱	۳/۸۱۶۸	۰/۹۷۰۱	۹۹

ردیف ۲۹ جدول ۴-۴ که با زمینه تیره مشخص شده است، ساختار و نتایج شبکه‌ای را نشان می‌دهد که کمترین میانگین مربعات خطای تخمین الگوهای آزمون را داشته است (بهترین شبکه). در ادامه این فصل، موارد مشابه به همین روش مشخص شده‌اند. نتایج مذکور در جدول ۴-۴ بر اساس میانگین مربعات خطای تخمین الگوهای آزمون مرتب شدند و در نهایت ۷ شبکه که کمترین میانگین مربعات خطای تخمین الگوهای آزمون را داشتند برای انجام مراحل بعدی انتخاب شدند. در جدول ۴-۵ نتایج این ۷ شبکه آمده است.

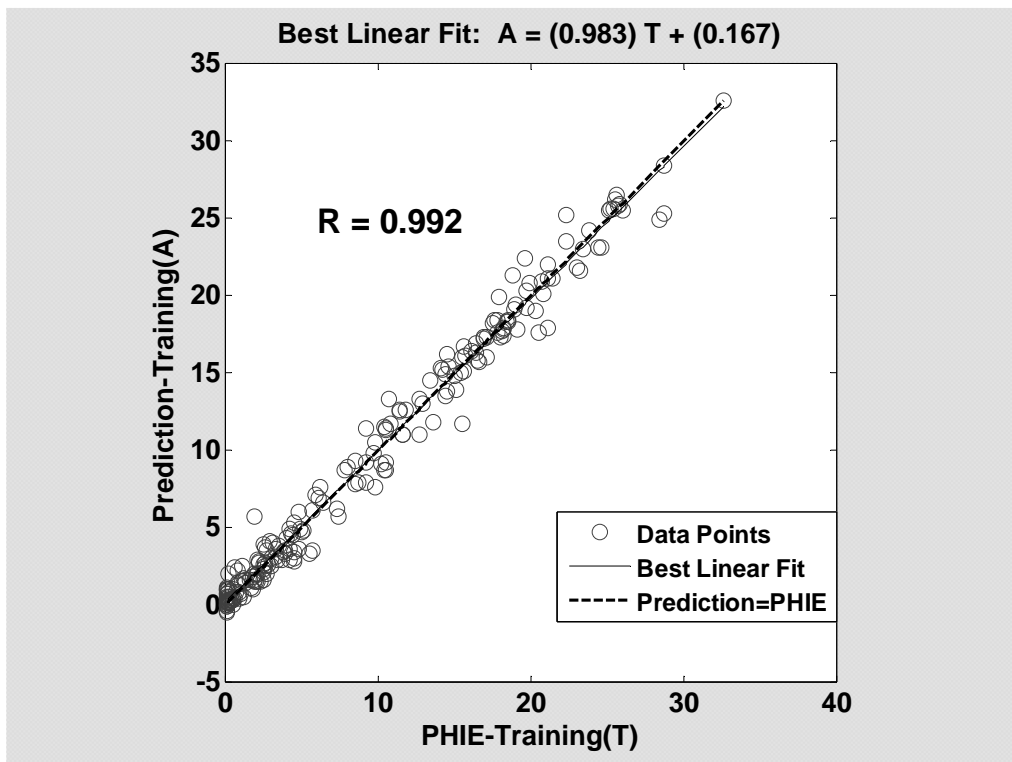
جدول ۴-۵: ۷ شبکه منفرد آموزش دیده به روش فراآموزش با بهترین نتایج.

تعداد کل پارامترهای شبکه	ضریب همبستگی (الگوهای آزمون)	MSE (الگوهای آزمون)	ضریب همبستگی (الگوهای آموزش)	MSE (الگوهای آموزش)	ساختار شبکه	ردیف
۸۷	۰/۹۸۸	۲/۲۴۳	۰/۹۹۲	۱/۱۳۶	۴-۲-۱۹-۱	۱
۲۶	۰/۹۸۷	۲/۲۵۸	۰/۹۹۱	۱/۳۲	۴-۳-۲-۱	۲
۵۵	۰/۹۸۷	۲/۲۸۸	۰/۹۹۳	۱/۰۶۶	۴-۲-۱۱-۱	۳
۳۱	۰/۹۸۷	۲/۳۳۴	۰/۹۹۲	۱/۱۰۴	۴-۳-۳-۱	۴
۳۱	۰/۹۸۶	۲/۳۷	۰/۹۹۱	۱/۲۹۷	۴-۲-۵-۱	۵
۷۹	۰/۹۸۴	۲/۳۷۴	۰/۹۹۳	۱/۰۵۵	۴-۲-۱۷-۱	۶
۴۱	۰/۹۸۶	۲/۳۷۵	۰/۹۹۳	۰/۹۷۵	۴-۳-۵-۱	۷

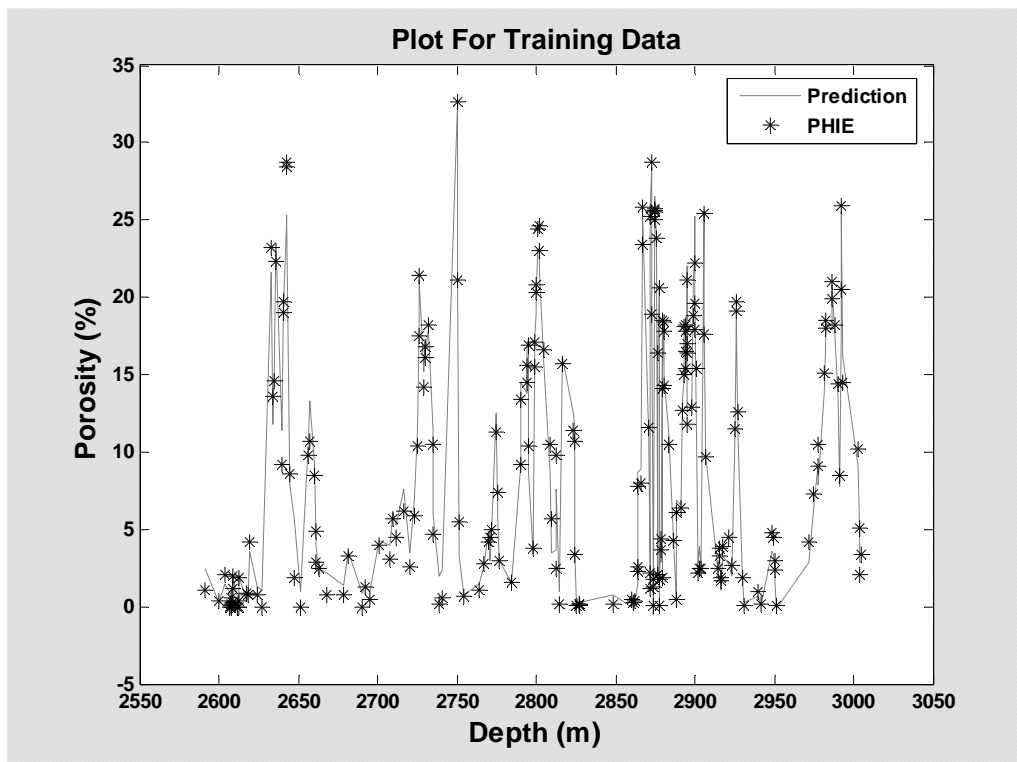
در شکل ۴-۱۰ نتایج بهترین شبکه منفرد روش فراآموزش (شبکه شماره ۱ جدول ۴-۵) نمایش داده شده است.



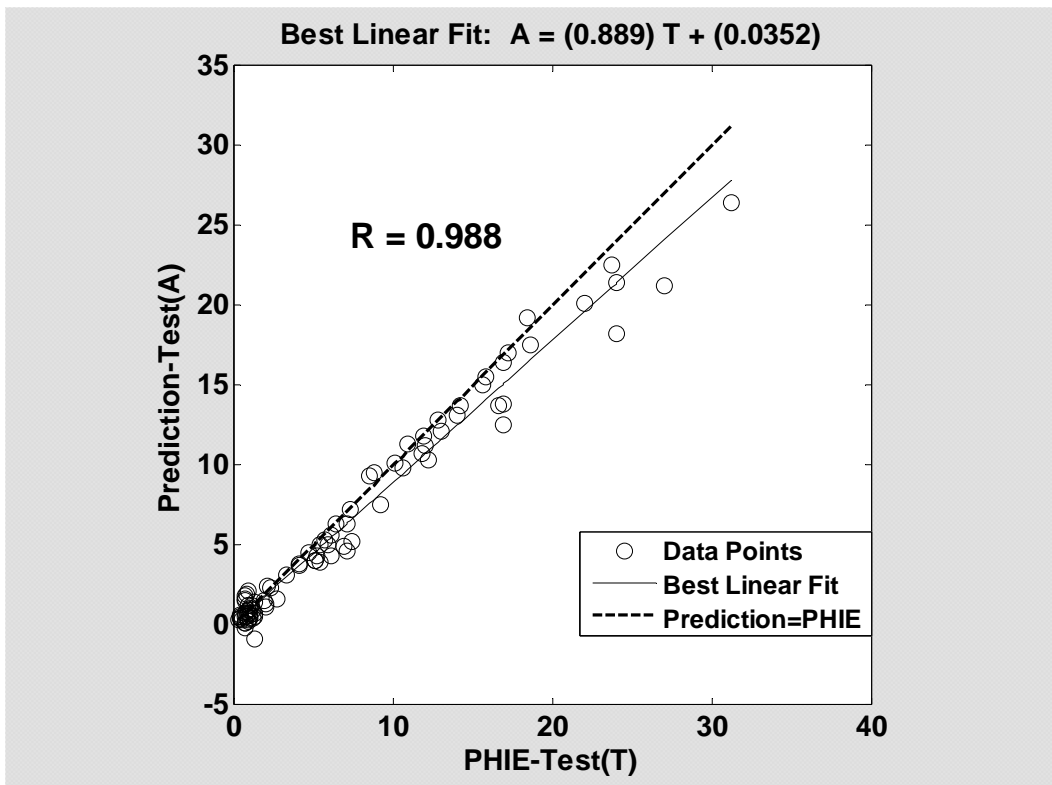
شکل ۴-۱۰-الف: نمایش پیشرفت روند آموزش و کاهش میانگین مربعات خطا برای الگوهای آموزش برای شبکه شماره ۱ جدول ۴-۵.



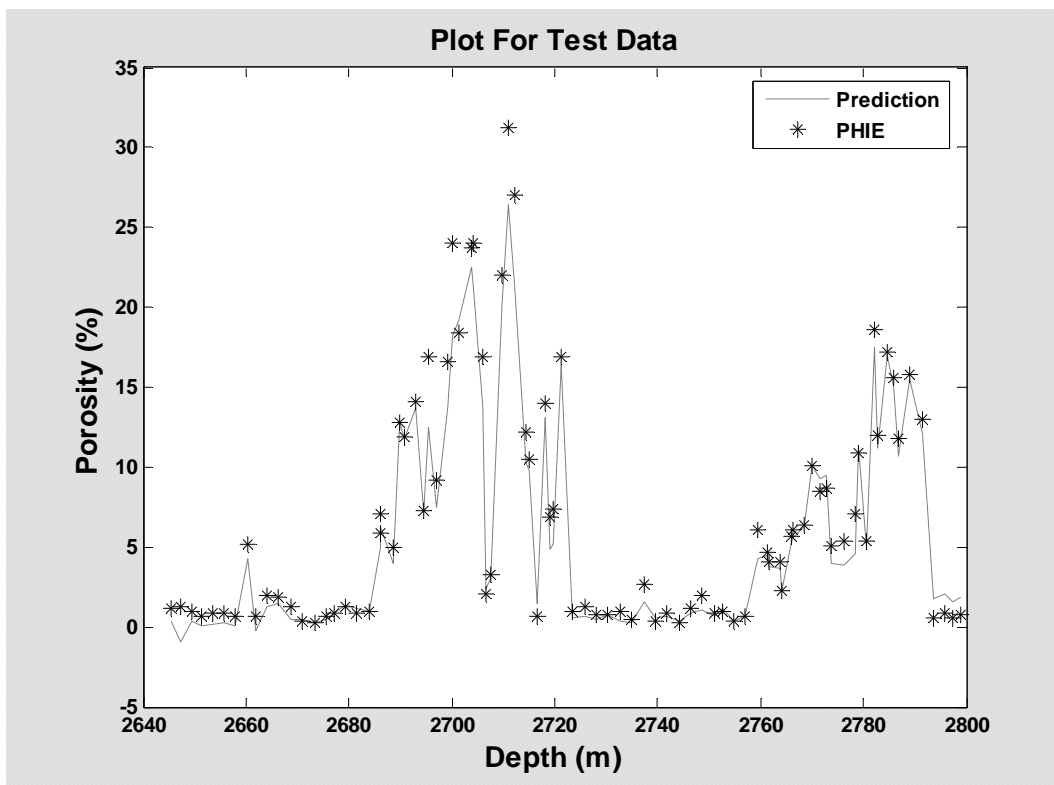
شکل ۴-۱۰-ب: نمودار همبستگی (Cross plot) بین تخلخل و تخمین شبکه شماره ۱ جدول ۴-۵ برای الگوهای آموزش.



شکل ۴-۱۰-ج: مقادیر تخلخل و تخمین شبکه شماره ۱ جدول ۴-۵ برای الگوهای آموزش.



شکل ۴-۱۰-۵: نمودار همبستگی بین تخلخل و تخمین شبکه شماره ۱ جدول ۴-۵ برای الگوهای آزمون.



شکل ۴-۱۰-۵: مقادیر تخلخل و تخمین شبکه شماره ۱ جدول ۴-۵ برای الگوهای آزمون.

شکل ۴-۱۰-د نشان می‌دهد که دقت تخمین شبکه برای مقادیر بزرگ تخلخل کمتر از دقت تخمین آن برای مقادیر کوچک تخلخل است. علت این موضوع آن است که همان طور که قبلاً نیز در مورد شکل ۴-۸ اشاره شد فراوانی الگوهای آموزشی با مقادیر بزرگ تخلخل از الگوهای آموزشی با مقادیر کوچک تخلخل به مراتب کمتر است.

۴-۶-۱- ترکیبات آنسامبلی خطی

چنان که قبلاً اشاره شد ۱۲۰ ترکیب آنسامبلی ممکن ۷ شبکه جدول ۴-۵ با استفاده از ۵ روش ترکیب خطی شامل میانگین‌گیری ساده و ۴ روش ترکیب خطی بهینه (ترکیب خطی نامقید با جمله ثابت (CM1)، ترکیب خطی مقید با جمله ثابت (CM2)، ترکیب خطی نامقید بدون جمله ثابت (CM3) و ترکیب خطی مقید بدون جمله ثابت (CM4))، ساخته شد. نتایج این ۱۲۰ ترکیب آنسامبلی به پنج روش ترکیب خطی در جدول ۴-۶ آورده شده است. ستون اول از سمت چپ این جدول (NO.) شماره ردیف و ستون دوم (Ensemble) ترکیب آنسامبلی را مشخص می‌کند، بقیه ستون‌ها میانگین مربعات خطای تخمین را برای الگوهای آموزش و آزمون در هر یک از ۵ روش ترکیب خطی نشان می‌دهد. به عنوان مثال ردیف ۹۵ میانگین مربعات خطای تخمین ترکیب آنسامبلی ۵ شبکه ۲، ۳، ۴، ۶ و ۷ جدول ۴-۵ را به ۵ روش ترکیب خطی نشان می‌دهد. نتایج بهترین ترکیب آنسامبلی از بین ۱۲۰ ترکیب در هر روش با زمینه تیره مشخص شده است. در مورد ۲ جدول دیگر در این فصل که نتایجی مشابه جدول ۴-۶ را نشان می‌دهند نیز همین کار انجام شده است.

جدول ۴-۶: نتایج ۱۲۰ ترکیب آنسامبلی شبکه‌های عصبی آموزش‌دیده به روش فراآموزش (۷ شبکه جدول ۴-۵) به ۵ روش ترکیب خطی (میانگین‌گیری ساده، CM1، CM2، CM3 و CM4).

NO.	Ensemble	Simple averaging		CM1		CM2		CM3		CM4	
		MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)
1	1,2	1.103	2.156	1.086	2.160	1.086	2.172	1.086	2.150	1.086	2.166
2	1,3	0.996	2.207	0.993	2.210	0.993	2.217	0.993	2.200	0.993	2.213
3	1,4	1.041	2.241	1.040	2.243	1.040	2.250	1.040	2.235	1.040	2.246
4	1,5	1.168	2.265	1.135	2.243	1.135	2.250	1.135	2.234	1.135	2.241
5	1,6	0.963	2.201	0.959	2.208	0.959	2.217	0.959	2.198	0.959	2.213
6	1,7	0.990	2.251	0.965	2.302	0.965	2.307	0.965	2.298	0.965	2.314
7	2,3	1.089	2.204	1.050	2.237	1.050	2.240	1.050	2.233	1.050	2.239
8	2,4	1.119	2.254	1.087	2.286	1.087	2.290	1.087	2.283	1.087	2.290
9	2,5	1.239	2.243	1.238	2.244	1.238	2.248	1.238	2.240	1.238	2.248
10	2,6	1.095	2.263	1.047	2.330	1.047	2.332	1.047	2.328	1.047	2.332
11	2,7	1.028	2.265	0.966	2.320	0.966	2.325	0.966	2.317	0.966	2.334
12	3,4	1.013	2.270	1.012	2.264	1.012	2.268	1.012	2.260	1.012	2.267
13	3,5	1.086	2.284	1.051	2.274	1.051	2.276	1.051	2.270	1.051	2.276
14	3,6	0.985	2.274	0.985	2.273	0.985	2.276	0.985	2.268	0.985	2.275
15	3,7	0.938	2.288	0.932	2.292	0.932	2.296	0.932	2.288	0.932	2.303
16	4,5	1.129	2.305	1.096	2.312	1.096	2.314	1.096	2.310	1.096	2.314
17	4,6	1.021	2.291	1.019	2.296	1.019	2.299	1.019	2.293	1.019	2.298
18	4,7	0.988	2.346	0.968	2.350	0.968	2.354	0.968	2.349	0.968	2.363
19	5,6	1.068	2.283	1.034	2.309	1.034	2.313	1.034	2.305	1.034	2.312
20	5,7	1.042	2.339	0.973	2.353	0.973	2.356	0.973	2.352	0.973	2.366
21	6,7	0.933	2.306	0.928	2.299	0.928	2.303	0.928	2.295	0.928	2.310
22	1,2,3	1.026	2.165	0.992	2.202	0.993	2.210	0.993	2.193	0.993	2.206
23	1,2,4	1.054	2.196	1.036	2.219	1.036	2.229	1.036	2.211	1.036	2.225
24	1,2,5	1.143	2.198	1.063	2.143	1.064	2.157	1.064	2.133	1.064	2.149
25	1,2,6	1.015	2.178	0.955	2.228	0.956	2.236	0.956	2.218	0.956	2.231

ادامه جدول ۴-۶

NO.	Ensemble	Simple averaging		CM1		CM2		CM3		CM4	
		MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)
26	1,2,7	1.006	2.201	0.961	2.284	0.961	2.290	0.961	2.279	0.961	2.297
27	1,3,4	0.988	2.223	0.983	2.215	0.984	2.223	0.983	2.206	0.984	2.220
28	1,3,5	1.055	2.236	0.977	2.204	0.977	2.210	0.977	2.195	0.977	2.204
29	1,3,6	0.946	2.202	0.944	2.202	0.944	2.210	0.944	2.191	0.944	2.207
30	1,3,7	0.947	2.231	0.930	2.272	0.931	2.277	0.930	2.266	0.931	2.283
31	1,4,5	1.090	2.255	1.027	2.252	1.027	2.259	1.027	2.245	1.027	2.253
32	1,4,6	0.978	2.220	0.959	2.209	0.959	2.217	0.959	2.198	0.959	2.213
33	1,4,7	0.985	2.267	0.962	2.305	0.962	2.310	0.962	2.301	0.962	2.317
34	2,3,4	1.044	2.226	1.012	2.258	1.012	2.262	1.012	2.254	1.012	2.262
35	2,3,5	1.108	2.223	1.047	2.246	1.047	2.249	1.047	2.241	1.047	2.249
36	2,3,6	1.026	2.227	0.984	2.280	0.984	2.282	0.984	2.276	0.984	2.282
37	2,3,7	0.985	2.234	0.932	2.294	0.932	2.299	0.932	2.290	0.932	2.306
38	2,4,5	1.136	2.249	1.087	2.284	1.087	2.288	1.087	2.281	1.087	2.288
39	2,4,6	1.051	2.252	1.019	2.292	1.019	2.295	1.019	2.289	1.019	2.295
40	2,4,7	1.016	2.277	0.964	2.323	0.964	2.328	0.964	2.320	0.964	2.337
41	2,5,6	1.104	2.239	1.034	2.313	1.034	2.316	1.034	2.309	1.034	2.316
42	2,5,7	1.072	2.264	0.965	2.319	0.965	2.324	0.965	2.316	0.965	2.333
43	3,4,5	1.049	2.272	1.012	2.266	1.012	2.269	1.012	2.262	1.012	2.269
44	3,4,6	0.984	2.260	0.979	2.259	0.979	2.263	0.979	2.254	0.979	2.263
45	3,4,7	0.957	2.291	0.932	2.292	0.932	2.297	0.932	2.288	0.932	2.304
46	3,5,6	1.015	2.259	0.984	2.269	0.984	2.272	0.984	2.264	0.984	2.272
47	3,5,7	0.992	2.290	0.929	2.296	0.929	2.300	0.929	2.294	0.929	2.307
48	4,5,6	1.046	2.271	1.017	2.287	1.017	2.291	1.017	2.283	1.017	2.291
49	4,5,7	1.029	2.320	0.968	2.350	0.968	2.354	0.968	2.349	0.968	2.363
50	1,5,6	1.034	2.223	0.929	2.218	0.930	2.226	0.930	2.208	0.930	2.219
51	1,5,7	1.044	2.270	0.964	2.293	0.964	2.298	0.964	2.290	0.964	2.304
52	1,6,7	0.931	2.226	0.922	2.254	0.922	2.261	0.922	2.247	0.922	2.265

ادامه جدول ۴-۶

NO.	Ensemble	Simple averaging		CM1		CM2		CM3		CM4	
		MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)
53	2,6,7	0.986	2.258	0.927	2.308	0.927	2.312	0.927	2.305	0.928	2.319
54	3,6,7	0.926	2.270	0.916	2.274	0.917	2.280	0.917	2.269	0.917	2.285
55	4,6,7	0.960	2.299	0.926	2.303	0.926	2.307	0.926	2.299	0.926	2.314
56	5,6,7	0.983	2.288	0.926	2.307	0.926	2.311	0.926	2.304	0.926	2.318
57	4,5,6,7	0.992	2.286	0.925	2.307	0.925	2.310	0.925	2.304	0.925	2.318
58	3,5,6,7	0.964	2.267	0.910	2.283	0.910	2.287	0.910	2.279	0.910	2.293
59	3,4,6,7	0.945	2.272	0.910	2.276	0.910	2.281	0.910	2.271	0.910	2.287
60	3,4,5,7	0.994	2.287	0.929	2.297	0.929	2.300	0.929	2.294	0.929	2.307
61	3,4,5,6	1.010	2.256	0.979	2.261	0.979	2.265	0.979	2.257	0.979	2.265
62	2,5,6,7	1.021	2.252	0.926	2.307	0.926	2.311	0.926	2.305	0.926	2.318
63	2,4,6,7	0.989	2.263	0.925	2.308	0.925	2.312	0.925	2.304	0.925	2.319
64	2,4,5,7	1.049	2.271	0.962	2.323	0.962	2.328	0.962	2.321	0.962	2.337
65	2,4,5,6	1.071	2.243	1.017	2.290	1.017	2.294	1.017	2.287	1.017	2.293
66	2,3,6,7	0.965	2.238	0.913	2.293	0.913	2.297	0.913	2.289	0.913	2.303
67	2,3,5,7	1.023	2.244	0.928	2.285	0.928	2.289	0.928	2.282	0.928	2.297
68	2,3,5,6	1.048	2.227	0.983	2.282	0.983	2.285	0.983	2.277	0.983	2.284
69	1,5,6,7	0.983	2.241	0.901	2.225	0.901	2.232	0.901	2.218	0.901	2.233
70	1,4,6,7	0.950	2.243	0.915	2.246	0.916	2.253	0.916	2.239	0.916	2.257
71	1,4,5,7	1.026	2.271	0.958	2.291	0.958	2.296	0.958	2.287	0.958	2.301
72	1,4,5,6	1.023	2.231	0.929	2.217	0.929	2.225	0.929	2.207	0.929	2.219
73	1,3,6,7	0.922	2.221	0.914	2.249	0.914	2.256	0.914	2.242	0.914	2.259
74	1,3,5,7	0.996	2.249	0.917	2.240	0.918	2.245	0.918	2.234	0.918	2.248
75	1,3,5,6	0.997	2.219	0.907	2.201	0.907	2.208	0.907	2.191	0.907	2.203
76	1,3,4,7	0.957	2.246	0.930	2.269	0.930	2.275	0.930	2.264	0.930	2.281
77	1,3,4,6	0.960	2.216	0.943	2.201	0.943	2.209	0.943	2.190	0.943	2.206
78	1,2,6,7	0.967	2.204	0.919	2.266	0.919	2.272	0.919	2.259	0.919	2.275
79	1,2,5,7	1.052	2.224	0.949	2.233	0.950	2.241	0.949	2.227	0.950	2.245

ادامه جدول ۴-۶

NO.	Ensemble	Simple averaging		CM1		CM2		CM3		CM4	
		MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)
80	1,2,5,6	1.058	2.197	0.927	2.206	0.928	2.215	0.928	2.196	0.928	2.209
81	1,2,4,7	1.000	2.227	0.960	2.289	0.960	2.296	0.960	2.284	0.960	2.302
82	1,2,4,6	1.008	2.200	0.955	2.228	0.956	2.236	0.956	2.218	0.956	2.231
83	1,2,3,7	0.974	2.197	0.930	2.276	0.930	2.282	0.930	2.271	0.930	2.287
84	1,2,3,6	0.986	2.181	0.936	2.231	0.937	2.237	0.937	2.221	0.937	2.234
85	2,3,4,7	0.986	2.250	0.932	2.293	0.932	2.297	0.932	2.289	0.932	2.304
86	2,3,4,6	1.013	2.232	0.978	2.272	0.978	2.275	0.978	2.268	0.978	2.275
87	2,3,4,5	1.070	2.234	1.011	2.256	1.011	2.260	1.011	2.252	1.011	2.260
88	1,3,4,5	1.033	2.239	0.958	2.213	0.958	2.220	0.958	2.205	0.958	2.215
89	1,2,4,5	1.092	2.215	1.003	2.206	1.004	2.217	1.004	2.198	1.004	2.212
90	1,2,3,5	1.068	2.195	0.964	2.158	0.965	2.167	0.964	2.148	0.965	2.161
91	1,2,3,4	1.012	2.193	0.983	2.221	0.983	2.228	0.983	2.212	0.983	2.225
92	3,4,5,6,7	0.973	2.269	0.907	2.281	0.907	2.285	0.907	2.278	0.907	2.291
93	2,4,5,6,7	1.016	2.258	0.925	2.307	0.925	2.311	0.925	2.304	0.925	2.318
94	2,3,5,6,7	0.995	2.240	0.910	2.287	0.910	2.291	0.910	2.284	0.910	2.297
95	2,3,4,6,7	0.971	2.246	0.908	2.291	0.908	2.295	0.908	2.287	0.908	2.301
96	2,3,4,5,7	1.016	2.253	0.928	2.285	0.928	2.290	0.928	2.283	0.928	2.297
97	2,3,4,5,6	1.034	2.233	0.978	2.272	0.978	2.275	0.978	2.268	0.978	2.275
98	1,4,5,6,7	0.987	2.249	0.899	2.222	0.899	2.229	0.899	2.215	0.899	2.230
99	1,3,5,6,7	0.964	2.233	0.888	2.212	0.888	2.218	0.888	2.204	0.888	2.218
100	1,3,4,6,7	0.939	2.234	0.905	2.238	0.905	2.245	0.905	2.231	0.905	2.249
101	1,3,4,5,7	0.993	2.254	0.917	2.240	0.917	2.246	0.917	2.235	0.917	2.249
102	1,3,4,5,6	0.996	2.226	0.907	2.201	0.907	2.208	0.907	2.191	0.907	2.203
103	1,2,5,6,7	1.007	2.217	0.900	2.217	0.901	2.225	0.900	2.210	0.901	2.225
104	1,2,4,6,7	0.974	2.222	0.914	2.257	0.914	2.263	0.914	2.250	0.914	2.267
105	1,2,4,5,7	1.036	2.236	0.946	2.241	0.946	2.249	0.946	2.236	0.946	2.253
106	1,2,4,5,6	1.042	2.211	0.927	2.206	0.927	2.216	0.927	2.196	0.927	2.209

ادامه جدول ۴-۶

NO.	Ensemble	Simple averaging		CM1		CM2		CM3		CM4	
		MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)	MSE (Tr.)	MSE (Test.)
107	1,2,3,6,7	0.953	2.202	0.908	2.265	0.908	2.271	0.908	2.259	0.908	2.274
108	1,2,3,5,7	1.013	2.216	0.913	2.211	0.914	2.218	0.914	2.205	0.914	2.221
109	1,2,3,5,6	1.022	2.197	0.907	2.194	0.907	2.203	0.907	2.185	0.907	2.197
110	1,2,3,4,7	0.977	2.217	0.930	2.272	0.930	2.278	0.930	2.267	0.930	2.283
111	1,2,3,4,6	0.986	2.198	0.936	2.230	0.936	2.237	0.936	2.220	0.936	2.233
112	1,2,3,4,5	1.046	2.209	0.951	2.183	0.951	2.192	0.951	2.174	0.951	2.186
113	1,2,3,4,5,6	1.015	2.208	0.907	2.194	0.907	2.203	0.907	2.184	0.907	2.197
114	1,2,3,4,5,7	1.008	2.228	0.913	2.212	0.914	2.219	0.914	2.206	0.914	2.222
115	1,2,3,4,6,7	0.961	2.216	0.900	2.255	0.901	2.261	0.901	2.248	0.901	2.264
116	1,2,3,5,6,7	0.986	2.213	0.887	2.208	0.888	2.215	0.888	2.200	0.888	2.215
117	1,2,4,5,6,7	1.004	2.228	0.898	2.213	0.899	2.220	0.898	2.205	0.899	2.221
118	1,3,4,5,6,7	0.970	2.241	0.884	2.207	0.884	2.213	0.884	2.199	0.884	2.214
119	2,3,4,5,6,7	0.995	2.246	0.907	2.286	0.907	2.290	0.907	2.283	0.907	2.296
120	1,2,3,4,5,6,7	0.987	2.223	0.884	2.202	0.884	2.209	0.884	2.194	0.884	2.209

نتایج بهترین آنسامبل‌های حاصل از ۵ روش ترکیب خطی از بین ۱۲۰ ترکیب آنسامبلی، که بیشترین کاهش را در میانگین مربعات خطای تخمین الگوهای آزمون ایجاد کرده‌اند، در جدول ۴-۷ آمده است. این نتایج با دقت سه رقم اعشار گرد شده‌اند.

جدول ۴-۷: نتایج بهترین آنسامبل‌های حاصل از ۵ روش ترکیب خطی ۷ شبکه جدول ۴-۵.

روش ترکیب خطی نتایج	میانگین‌گیری ساده	ترکیب خطی بهینه نامقید با جمله ثابت (CM1)	ترکیب خطی بهینه مقید با جمله ثابت (CM2)	ترکیب خطی بهینه نامقید بدون جمله ثابت (CM3)	ترکیب خطی بهینه مقید بدون جمله ثابت (CM4)
ترکیب آنسامبلی	۲۱	۵ و ۲، ۱	۵ و ۲، ۱	۵ و ۲، ۱	۵ و ۲، ۱
ضریب همبستگی (الگوهای آموزش)	۰/۹۹۲	۰/۹۹۳	۰/۹۹۳	۰/۹۹۳	۰/۹۹۳
MSE (الگوهای آموزش)	۱/۱۰۳	۱/۰۶۳	۱/۰۶۴	۱/۰۶۴	۱/۰۶۴
کاهش در MSE الگوهای آموزش.	٪ ۲/۹	٪ ۶/۴	٪ ۶/۳	٪ ۶/۳	٪ ۶/۳
ضریب همبستگی (الگوهای آزمون)	۰/۹۸۹	۰/۹۸۹	۰/۹۸۹	۰/۹۸۹	۰/۹۸۹
MSE (الگوهای آزمون)	۲/۱۵۶	۲/۱۴۳	۲/۱۵۷	۲/۱۳۳	۲/۱۴۹
کاهش در MSE الگوهای آزمون.	٪ ۳/۹	٪ ۴/۵	٪ ۳/۸	٪ ۴/۹	٪ ۴/۲
α_0	۰	-۰/۰۳۲	-۰/۰۰۶	۰	۰
α_1	۰/۵	۰/۹۵۱	۰/۹۴۹	۰/۹۴۹	۰/۹۴۹
α_2	۰/۵	۰/۵۰۸	۰/۵۰۷	۰/۵۰۷	۰/۵۰۷
α_3	۰	۰	۰	۰	۰
α_4	۰	۰	۰	۰	۰
α_5	۰	-۰/۴۵۶	-۰/۴۵۶	-۰/۴۵۵	-۰/۴۵۶
α_6	۰	۰	۰	۰	۰
α_7	۰	۰	۰	۰	۰
$\sum \alpha_i$	۱	۱/۰۰۳	۱	۱/۰۰۱	۱

در این جدول ردیف دوم بهترین ترکیب آنسامبلی از بین ۱۲۰ ترکیب برای هر یک از روش‌های ترکیب خطی، ردیف سوم ضریب همبستگی بین مقادیر تخلخل و تخمین ترکیب آنسامبلی برای الگوهای آموزش، ردیف چهارم میانگین مربعات خطای تخمین ترکیب برای الگوهای آموزش و ردیف پنجم درصد کاهش میانگین مربعات خطای تخمین الگوهای آموزش با استفاده از ترکیب نسبت به بهترین شبکه منفرد (شبکه شماره ۱ جدول ۴-۵) را نشان می‌دهند. درصد کاهش میانگین مربعات خطای تخمین الگوهای آموزش از تقسیم میانگین مربعات خطای تخمین ترکیب بر میانگین مربعات خطای تخمین بهترین شبکه منفرد برای الگوهای آموزش به دست آمده است. به طور مشابه،

ردیف‌های ششم، هفتم و هشتم به ترتیب، ضریب همبستگی بین مقادیر تخلخل و تخمین ترکیب آنسامبلی، میانگین مربعات خطای تخمین ترکیب آنسامبلی و درصد کاهش میانگین مربعات خطای تخمین با استفاده از ترکیب نسبت به شبکه شماره ۱ جدول ۴-۵ را برای الگوهای آزمون نشان می‌دهند. α_0 مقدار جمله ثابت، α_1 ضریب رگرسیون خطی شبکه ۱، α_2 ضریب رگرسیون خطی شبکه ۲ و را نشان می‌دهند. $\sum \alpha_i$ مجموع ضرایب شبکه‌هایی می‌باشد که آنسامبل از آن‌ها تشکیل شده است. به عنوان مثال ستون چهارم جدول از سمت راست نشان می‌دهد که در روش ترکیب خطی بهینه مقید با جمله ثابت، از بین ۱۲۰ ترکیب آنسامبلی، ترکیب سه شبکه‌ای شبکه‌های ۱، ۲ و ۵ بهترین بوده است زیرا در بین ۱۲۰ ترکیب، برای الگوهای آزمون، بیشترین کاهش را در میانگین مربعات خطای تخمین نسبت به شبکه شماره ۱ جدول ۴-۵ ایجاد کرده و کمترین میانگین مربعات خطا را داشته است. در این ترکیب مقدار جمله ثابت ۰/۰۰۶-، ضریب رگرسیون خطی شبکه‌های ۱، ۲ و ۵ به ترتیب ۰/۹۴۹، ۰/۵۰۷ و ۰/۴۵۶- است و ضریب رگرسیون خطی بقیه شبکه‌ها (شبکه‌های ۳، ۴، ۶ و ۷) صفر می‌باشد، زیرا این شبکه‌ها در این ترکیب استفاده نشده‌اند. همان طور که انتظار می‌رود در این ترکیب مجموع ضرایب شبکه‌های جزء آنسامبل ۱ است زیرا ترکیب مقید است.

اطلاعات جدول ۴-۷ نشان می‌دهد که در ترکیبات نامقید، مجموع ضرایب شبکه‌ها به ۱ خیلی نزدیک است و در ترکیبات با جمله ثابت، مقدار این جمله به صفر خیلی نزدیک است که بنابر نتایج هاشم، نشان می‌دهد که شبکه‌های جزء آنسامبل خوب آموزش دیده هستند [۲۶، ۲۷ و ۲۸].

همچنین اعداد این جدول مشخص می‌کند که از بین ۵ روش ترکیب خطی (یعنی مجموعاً از بین ۶۰۰ ترکیب خطی)، ترکیب خطی بهینه نامقید بدون جمله ثابت شبکه‌های ۱، ۲ و ۵ کمترین میانگین مربعات خطای تخمین الگوهای آزمون را داشته و بیشترین کاهش را در میانگین مربعات خطا نسبت به شبکه شماره ۱ جدول ۴-۵ (۴/۹٪) برای الگوهای آزمون ایجاد کرده است.

۴-۶-۲- ترکیبات آنسامبلی غیرخطی با استفاده از شبکه‌های عصبی مصنوعی منفرد

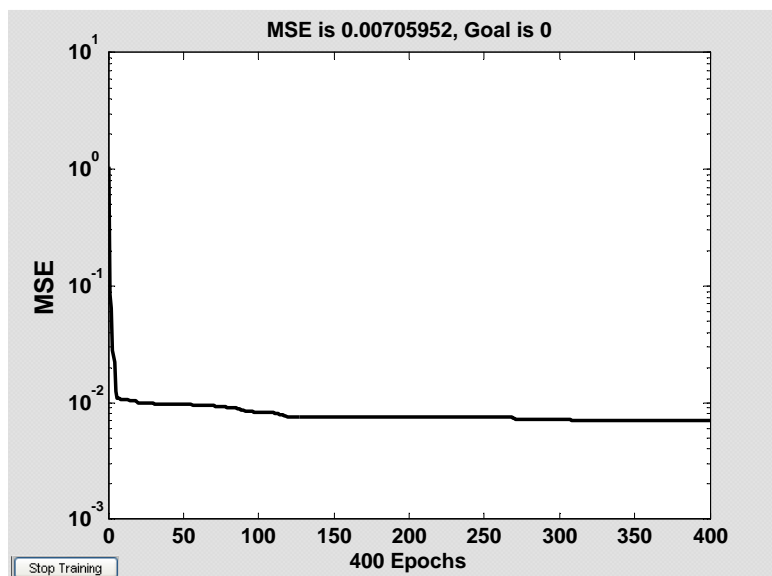
در این قسمت از خروجی ۷ شبکه منفرد جدول ۴-۵ به عنوان ورودی یک شبکه عصبی منفرد دیگر استفاده شد. واضح است که ترکیب به این روش غیرخطی است. طراحی این شبکه پس‌انتشار خطا به

هر سه روش فراآموزش، مرتب‌سازی و توقف سریع انجام شد. از شبکه‌هایی با یک لایه میانی که تابع انتقال نرون‌های آن تانژانت هیپربولیک و تابع انتقال نرون‌های لایه خروجی تابع خطی است استفاده گردید. تابع پیش‌پردازش Prestd و دوره‌های تکرار ۴۰۰ مرتبه انتخاب شد. آزمون و خطا با تعداد نرون‌های مختلف در لایه میانی و شرایط اولیه گوناگون انجام شد و در نهایت بهترین نتیجه از هر یک از سه روش آموزش پس‌انتشار خطا انتخاب گردید. جدول ۴-۸ بهترین نتیجه هر یک از سه روش پس‌انتشار خطا را نشان می‌دهد.

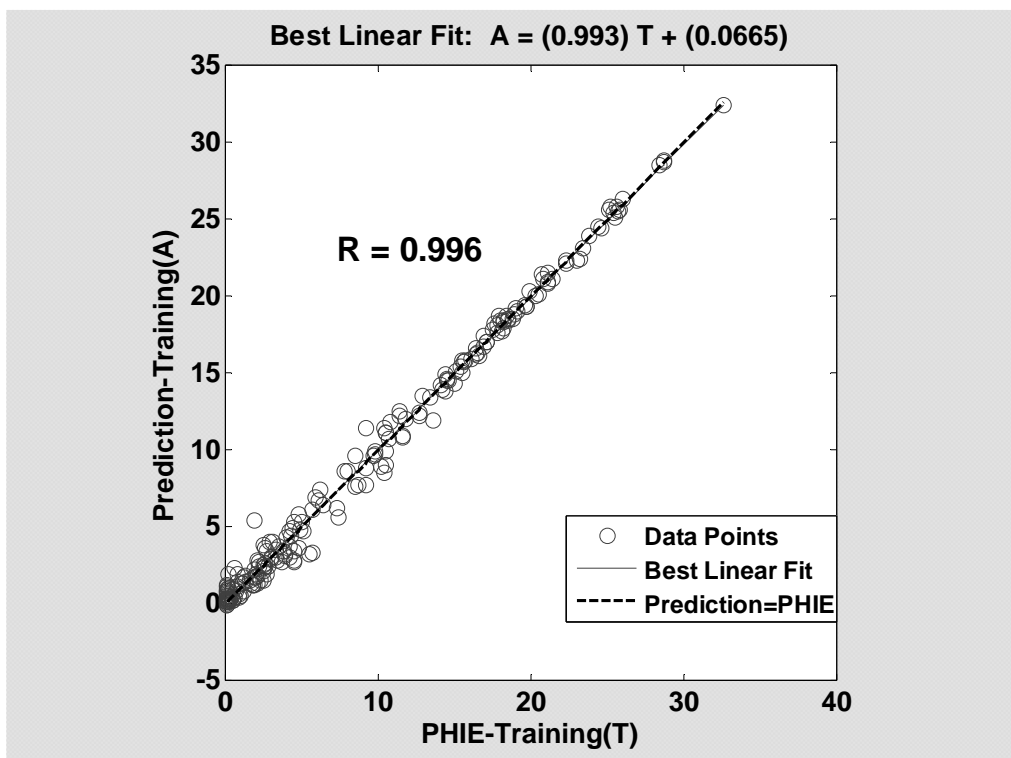
جدول ۴-۸: بهترین نتایج حاصل از ترکیب غیرخطی خروجی‌های ۷ شبکه جدول ۴-۵ با استفاده از شبکه عصبی ترکیب‌کننده آموزش دیده به سه روش فراآموزش، مرتب‌سازی و توقف سریع.

کاهش در MSE الگوهای آزمون.	MSE (الگوهای آزمون)	ضریب همبستگی (الگوهای آزمون)	کاهش در MSE الگوهای آموزش.	MSE (الگوهای آموزش)	ضریب همبستگی (الگوهای آموزش)	ساختار شبکه	نتایج روش آموزش شبکه ترکیب‌کننده
٪ ۱۹	۱/۸۱۶	۰/۹۸۸	٪ ۵۵	۰/۵۱۱	۰/۹۹۶	۷-۷-۱	فراآموزش
٪ ۵/۶	۲/۱۱۸	۰/۹۸۷	٪ ۲۶/۸	۰/۸۳۲	۰/۹۹۴	۷-۶-۱	مرتب‌سازی
٪ ۹/۴	۲/۰۳۲	۰/۹۸۸	٪ ۱/۹	۱/۱۱۴	۰/۹۹۳	۷-۶-۱	توقف سریع

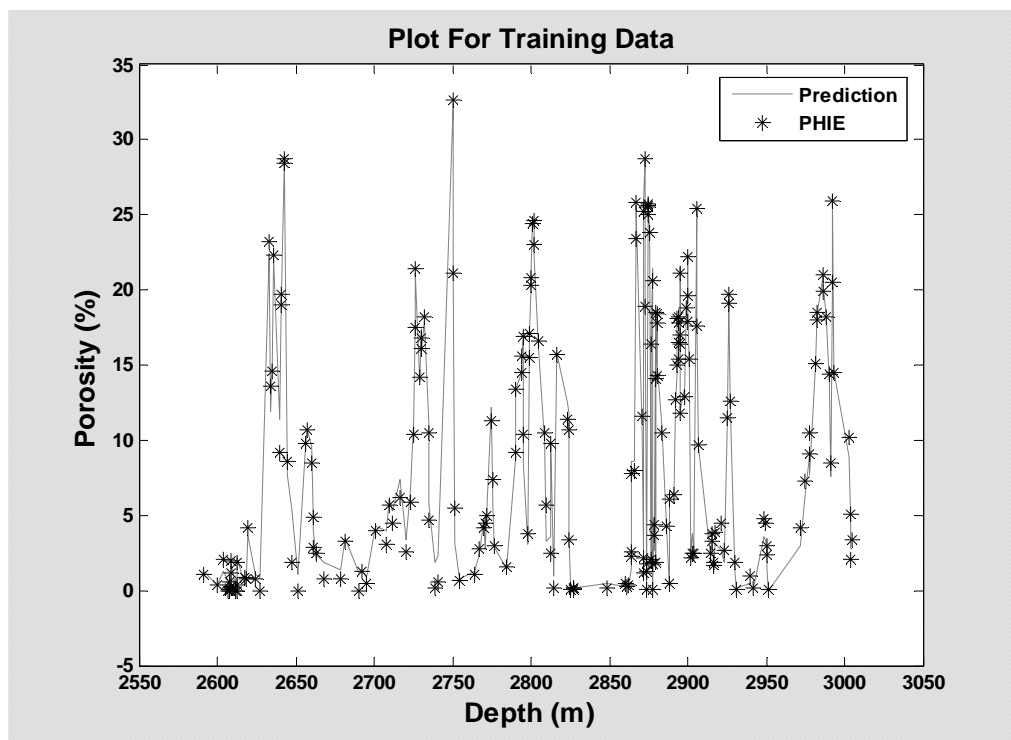
با توجه به جدول فوق مشاهده می‌شود که ترکیب تخمین ۷ شبکه منفرد جدول ۴-۵ با استفاده از یک شبکه پس‌انتشارخطا با ساختار ۷-۷-۱ که به روش فراآموزش آموزش دیده است، بیشترین کاهش را در میانگین مربعات خطای تخمین الگوهای آزمون به میزان ٪ ۱۹ نسبت به شبکه ۱ جدول ۴-۵ موجب شده است. در شکل ۴-۱۱ نتایج حاصل از این ترکیب نمایش داده شده است.



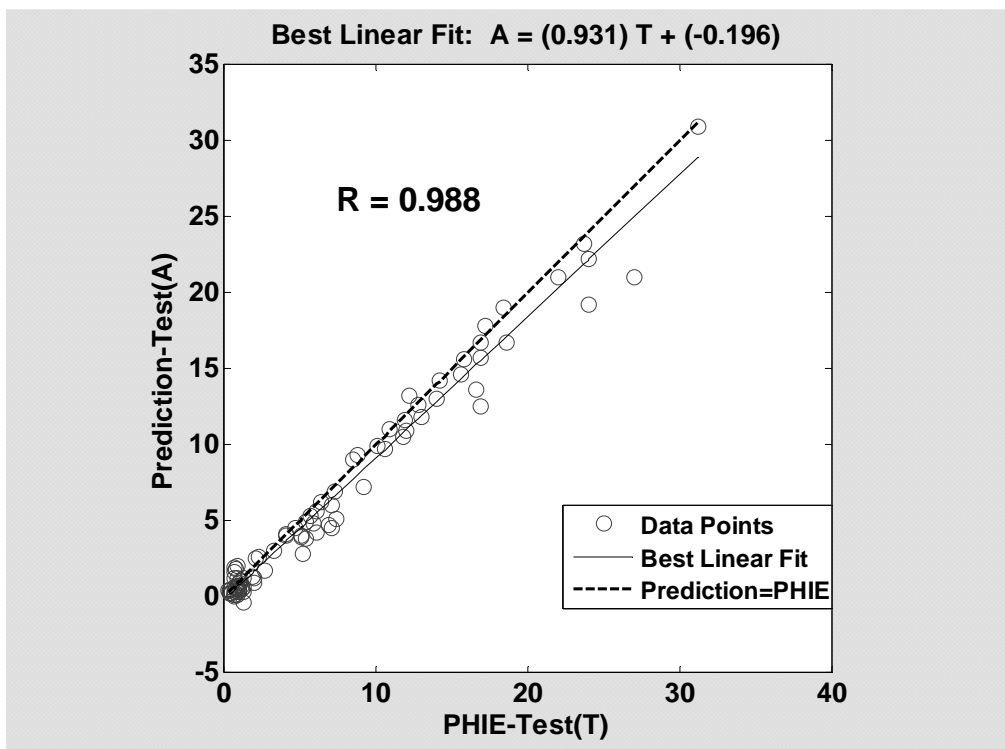
شکل ۴-۱۱-الف: نمایش پیشرفت روند آموزش و کاهش MSE برای الگوهای آموزش در ترکیب تخمین شبکه‌های منفرد جدول ۴-۵ توسط شبکه‌ای با ساختار ۷-۷-۱ که به روش فراآموزش آموزش دیده است.



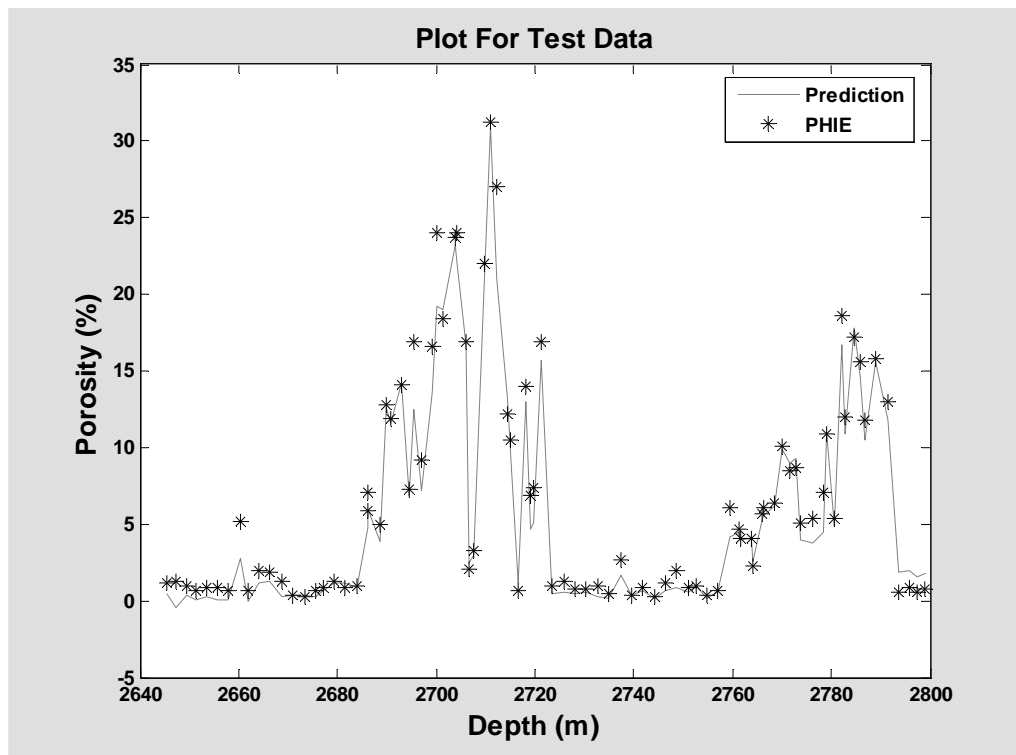
شکل ۴-۱۱-ب: نمودار همبستگی بین تخلخل و تخمین شبکه ترکیب کننده برای الگوهای آموزش در ترکیب تخمین شبکه های منفرد جدول ۴-۵ توسط شبکه آموزش دیده با روش فراآموزش.



شکل ۴-۱۱-ج: مقادیر تخلخل و تخمین شبکه ترکیب کننده برای الگوهای آموزش در ترکیب تخمین شبکه های منفرد جدول ۴-۵ توسط شبکه آموزش دیده با روش فراآموزش.



شکل ۴-۱۱-۵: نمودار همبستگی بین تخلخل و تخمین شبکه ترکیب کننده برای الگوهای آزمون در ترکیب تخمین شبکه های منفرد جدول ۴-۵ توسط شبکه آموزش دیده با روش فراآموزش.



شکل ۴-۱۱-۵: مقادیر تخلخل و تخمین شبکه ترکیب کننده برای الگوهای آزمون در ترکیب تخمین شبکه های منفرد جدول ۴-۵ توسط شبکه آموزش دیده با روش فراآموزش.

مقایسه اشکال ۴-۱۰ و ۴-۱۱ با هم نشان می‌دهد که با استفاده از ترکیب به وسیله یک شبکه عصبی، شیب بهترین خط برازش داده شده در نمودارهای همبستگی برای الگوهای آموزش از ۰/۹۸۳ به ۰/۹۹۳ و برای الگوهای آزمون از ۰/۸۸۹ به ۰/۹۳۱ افزایش یافته که از بهبود نتایج و افزایش دقت حکایت دارد.

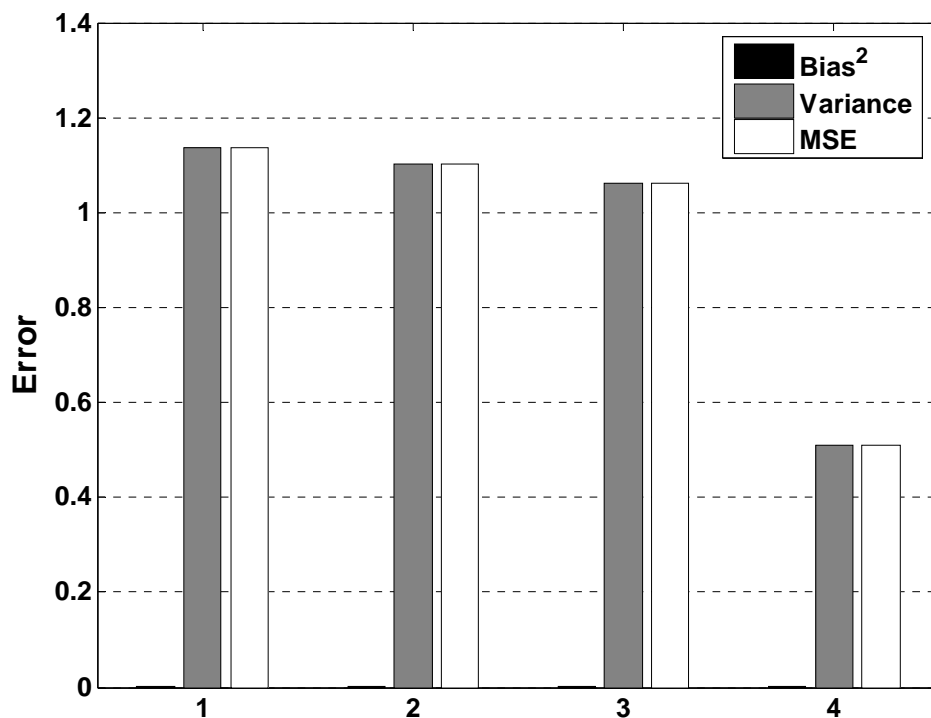
۴-۶-۳- واریانس و بایاس خطا

در جدول ۴-۹ واریانس، بایاس و میانگین مربعات خطای تخمین الگوهای آموزش و آزمون برای بهترین شبکه منفرد یعنی شبکه شماره ۱ جدول ۴-۵ (۱) و بهترین نتایج ترکیب یعنی ترکیب شبکه‌های ۱ و ۲ جدول ۴-۵ به روش میانگین‌گیری ساده (۲)، ترکیب خطی بهینه شبکه‌های ۱، ۲ و ۵ جدول ۴-۵ به روش نامقید بدون جمله ثابت (۳) و ترکیب ۷ شبکه جدول ۴-۵ با استفاده از شبکه عصبی پس‌انتشار خطا ۱-۷-۷ آموزش‌دیده به روش فراآموزش (۴) آورده شده است.

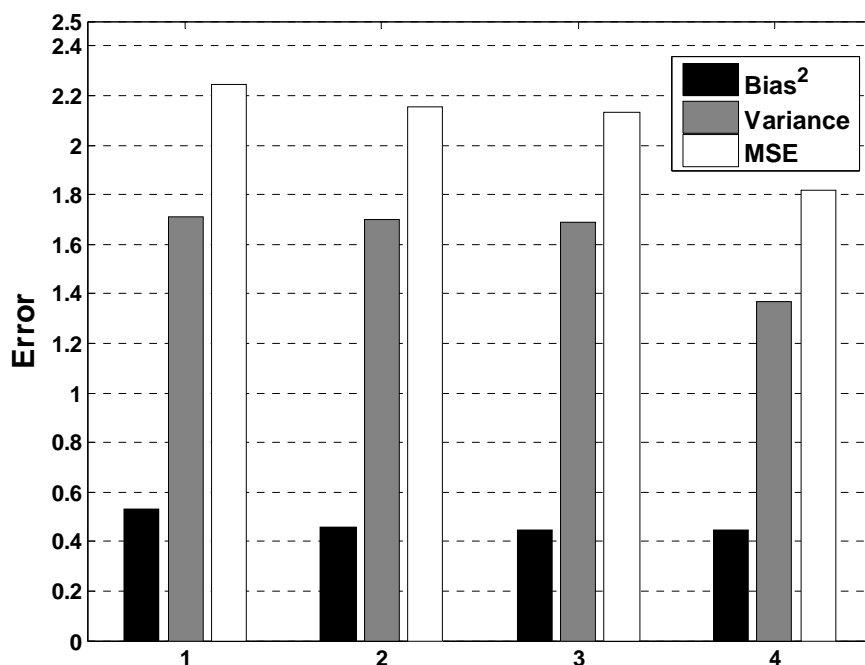
جدول ۴-۹: واریانس، بایاس و میانگین مربعات خطای تخمین شبکه شماره ۱ جدول ۴-۵ و بهترین نتایج حاصل از ترکیب شبکه‌های آن جدول.

نتایج ردیف	بایاس خطای تخمین (الگوهای آموزش)	واریانس خطای تخمین (الگوهای آموزش)	میانگین مربعات خطای تخمین (الگوهای آموزش)	بایاس خطای تخمین (الگوهای آموزش)	واریانس خطای تخمین (الگوهای آموزش)	میانگین مربعات خطای تخمین (الگوهای آموزش)
۱	۰/۰۰۶	۱/۱۳۶	۱/۱۳۶	-۰/۷۲۹	۱/۷۱۳	۲/۲۴۳
۲	۰/۰۰۳	۱/۱۰۳	۱/۱۰۳	-۰/۶۷۷	۱/۶۹۷	۲/۱۵۶
۳	-۰/۰۱۴	۱/۰۶۴	۱/۰۶۴	-۰/۶۶۶	۱/۶۸۹	۲/۱۳۳
۴	4×10^{-4}	۰/۵۱۱	۰/۵۱۱	-۰/۶۶۸	۱/۳۷	۱/۸۱۶

شکل‌های ۴-۱۲ نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطای تخمین را برای الگوهای آموزش و آزمون برای ۴ حالت جدول ۴-۹ نشان می‌دهد. مطابق شکل ۴-۱۲ الف، برای الگوهای آموزش، واریانس مؤلفه اصلی خطا است و بایاس تقریباً نقشی در خطا ندارد ولی مطابق شکل ۴-۱۲ ب، برای الگوهای آزمون هر دو مؤلفه بایاس و واریانس مؤثرند اگر چه نقش واریانس بیشتر است.



شکل ۴-۱۲-الف: نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطا در بهترین حالات برای الگوهای آموزش (۱: شبکه شماره ۱ جدول ۴-۵، ۲: بهترین ترکیب شبکه‌های جدول ۴-۵ به روش میانگین‌گیری ساده، ۳: بهترین ترکیب خطی بهینه شبکه‌های جدول ۴-۵، ۴: بهترین ترکیب ۷ شبکه جدول ۴-۵ با استفاده از شبکه عصبی).



شکل ۴-۱۲-ب: نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطا در بهترین حالات برای الگوهای آزمون (۱: شبکه شماره ۱ جدول ۴-۵، ۲: بهترین ترکیب شبکه‌های جدول ۴-۵ به روش میانگین‌گیری ساده، ۳: بهترین ترکیب خطی بهینه شبکه‌های جدول ۴-۵، ۴: بهترین ترکیب ۷ شبکه جدول ۴-۵ با استفاده از شبکه عصبی).

۴-۷- طراحی شبکه‌های پس‌انتشار خطا به روش مرتب‌سازی و ترکیب نتایج

برای آموزش به روش مرتب‌سازی از تابع آموزش مرتب‌سازی به روش بیزین استفاده شد. شبکه‌های پس‌انتشار خطا با یک و دو لایه میانی با تعداد نرون‌های متعدد با شرایط اولیه مختلف مورد آزمون قرار گرفت و پس از یک مرحله طولانی آزمون و خطا، برای هر ساختار خاص بهترین نتیجه ثبت شد. ساختارهای مورد استفاده همان ساختارهایی است که در روش فراآموزش مورد استفاده قرار گرفت. در جدول ۴-۱۰ نتایج حاصل از هر یک از این ساختارها آورده شده است. کمیت‌های مذکور در ستون‌های این جدول مشابه ستون‌های جدول ۴-۴ می‌باشد که قبلاً توضیح داده شده‌اند و تنها تفاوت، ستون اضافی پارامترهای مؤثر شبکه است.

جدول ۴-۱۰: نتایج حاصل از شبکه‌های عصبی پس‌انتشار خطا با یک و دو لایه میانی به روش مرتب‌سازی.

epochs = 400, show = 20, goal = 0, min_grad = 1.00×10^{-10} , mem_reduce = 1,
mu = 0.0050, mu_dec = 0.100, mu_inc = 10, mu_max = 1.00×10^{10} .

ردیف	ساختار شبکه	MSE (الگوهای آموزش)	ضریب همبستگی (الگوهای آموزش)	MSE (الگوهای آزمون)	ضریب همبستگی (الگوهای آزمون)	تعداد کل پارامترهای شبکه	تعداد پارامترهای مؤثر شبکه
۱	۴-۳-۱	۱/۲۵۵۶	۰/۹۹۱۲	۲/۴۸۰۹	۰/۹۸۵۱	۱۹	۱۷
۲	۴-۴-۱	۱/۲۰۱۷	۰/۹۹۱۶	۲/۴۲۶۱	۰/۹۸۵۶	۲۵	۲۱
۳	۴-۵-۱	۱/۱۴۴۳	۰/۹۹۹۲	۲/۵۱۶۹	۰/۹۸۵۴	۳۱	۲۶
۴	۴-۶-۱	۱/۱۰۰۶	۰/۹۹۲۳	۲/۵۹۷۸	۰/۹۸۵۵	۳۷	۲۸
۵	۴-۷-۱	۱/۰۶۸	۰/۹۹۲۶	۲/۵۸۱۸	۰/۹۸۵۵	۴۳	۳۱
۶	۴-۸-۱	۱/۰۹۲۳	۰/۹۹۲۴	۲/۶۰۵	۰/۹۸۵۴	۴۹	۲۹
۷	۴-۹-۱	۱/۰۹۳۳	۰/۹۹۲۴	۲/۶۰۲۹	۰/۹۸۵۵	۵۵	۲۹
۸	۴-۱۰-۱	۱/۰۹۸۲	۰/۹۹۲۳	۲/۵۹۳۴	۰/۹۸۵۳	۶۱	۲۸
۹	۴-۱۱-۱	۱/۱۱۷	۰/۹۹۲۲	۲/۶۷۵۸	۰/۹۸۵۴	۶۷	۲۷
۱۰	۴-۱۲-۱	۱/۰۹۷۳	۰/۹۹۲۳	۲/۵۸۸۲	۰/۹۸۵۳	۷۳	۲۷
۱۱	۴-۱۳-۱	۱/۰۹۷۴	۰/۹۹۲۳	۲/۵۸۷۶	۰/۹۸۵۲	۷۹	۲۷
۱۲	۴-۱۴-۱	۱/۱۱۸۷	۰/۹۹۲۲	۲/۶۶	۰/۹۸۵۴	۸۵	۲۷
۱۳	۴-۱۵-۱	۱/۰۹۷۷	۰/۹۹۲۳	۲/۵۸۷۱	۰/۹۸۵۲	۹۱	۲۷
۱۴	۴-۲-۲-۱	۱/۵۶۸۶	۰/۹۸۹	۲/۴۲۲۶	۰/۹۸۴۹	۱۹	۱۳
۱۵	۴-۲-۳-۱	۱/۲۸۰۹	۰/۹۹۱۱	۲/۵۷۲۶	۰/۹۸۴۳	۲۳	۱۵
۱۶	۴-۲-۴-۱	۱/۲۷۹۵	۰/۹۹۱۱	۲/۵۷۴۴	۰/۹۸۴۴	۲۷	۱۵
۱۷	۴-۲-۵-۱	۱/۲۷۸۸	۰/۹۹۱۱	۲/۵۷۵۸	۰/۹۸۴۴	۳۱	۱۵
۱۸	۴-۲-۶-۱	۱/۵۶۶۸	۰/۹۸۹۱	۲/۴۳	۰/۹۸۴۹	۳۵	۱۲

ادامه جدول ۴-۱۰

ردیف	ساختار شبکه	MSE (الگوهای آموزش)	ضریب همبستگی (الگوهای آموزش)	MSE (الگوهای آزمون)	ضریب همبستگی (الگوهای آزمون)	تعداد کل پارامترهای شبکه	تعداد پارامترهای مؤثر شبکه
۱۹	۴-۲-۷-۱	۱/۲۶۳۵	۰/۹۹۱۲	۲/۵۴۲۴	۰/۹۸۴۵	۳۹	۱۷
۲۰	۴-۲-۸-۱	۱/۱۹۲۸	۰/۹۹۱۷	۲/۴۳۱	۰/۹۸۵۳	۴۳	۲۱
۲۱	۴-۲-۹-۱	۱/۲۶۷۶	۰/۹۹۱۲	۲/۵۴۷۲	۰/۹۸۴۵	۴۷	۱۶
۲۲	۴-۲-۱۰-۱	۱/۵۶۴۲	۰/۹۸۹۱	۲/۴۳۱۶	۰/۹۸۴۹	۵۱	۱۲
۲۳	۴-۲-۱۱-۱	۱/۱۸۰۹	۰/۹۹۱۸	۲/۴۲۸۸	۰/۹۸۵۴	۵۵	۲۱
۲۴	۴-۲-۱۲-۱	۱/۵۶۳۵	۰/۹۸۹۱	۲/۴۳۲۳	۰/۹۸۴۹	۵۹	۱۲
۲۵	۴-۲-۱۳-۱	۱/۲۷۶۵	۰/۹۹۱۱	۲/۵۷۶۳	۰/۹۸۴۴	۶۳	۱۵
۲۶	۴-۲-۱۴-۱	۱/۲۲۸۱	۰/۹۹۱۴	۲/۴۰۵۳	۰/۹۸۶۶	۶۷	۲۳
۲۷	۴-۲-۱۵-۱	۱/۵۶۲۶	۰/۹۸۹۱	۲/۴۳۳۳	۰/۹۸۴۹	۷۱	۱۲
۲۸	۴-۲-۱۶-۱	۱/۱۳۹۲	۰/۹۹۲۱	۲/۳۹۴۴	۰/۹۸۵۴	۷۵	۲۵
۲۹	۴-۲-۱۷-۱	۱/۵۶۲۱	۰/۹۸۹۱	۲/۴۳۴	۰/۹۸۴۹	۷۹	۱۲
۳۰	۴-۲-۱۸-۱	۱/۱۶۲۲	۰/۹۹۱۹	۲/۴۱۹۴	۰/۹۸۵۵	۸۳	۲۳
۳۱	۴-۲-۱۹-۱	۱/۵۶۱۷	۰/۹۸۹۱	۲/۴۳۵	۰/۹۸۴۹	۸۷	۱۲
۳۲	۴-۲-۲۰-۱	۱/۲۷۶۲	۰/۹۹۱۱	۲/۵۷۷۶	۰/۹۸۴۴	۹۱	۱۵
۳۳	۴-۳-۲-۱	۱/۲۰۰۶	۰/۹۹۱۶	۲/۴۲۷۱	۰/۹۸۵۴	۲۶	۲۱
۳۴	۴-۳-۳-۱	۱/۲۰۰۵	۰/۹۹۱۶	۲/۴۲۹۲	۰/۹۸۵۵	۳۱	۲۱
۳۵	۴-۳-۴-۱	۱/۲۰۰۲	۰/۹۹۱۶	۲/۴۲۵۳	۰/۹۸۵۵	۳۶	۲۱
۳۶	۴-۳-۵-۱	۱/۱۹۹۲	۰/۹۹۱۶	۲/۴۲۱۲	۰/۹۸۵۶	۴۱	۲۱
۳۷	۴-۳-۶-۱	۱/۱۹۸۴	۰/۹۹۱۶	۲/۴۱۷۵	۰/۹۸۵۶	۴۶	۲۰
۳۸	۴-۳-۷-۱	۱/۱۹۷۸	۰/۹۹۱۶	۲/۴۱۴۳	۰/۹۸۵۶	۵۱	۲۰
۳۹	۴-۳-۸-۱	۱/۱۹۷۸	۰/۹۹۱۶	۲/۴۱۹	۰/۹۸۵۶	۵۶	۲۱
۴۰	۴-۳-۹-۱	۱/۱۹۷۵	۰/۹۹۱۶	۲/۴۱۹۳	۰/۹۸۵۶	۶۱	۲۱
۴۱	۴-۳-۱۰-۱	۱/۱۹۷۲	۰/۹۹۱۶	۲/۴۱۹۳	۰/۹۸۵۶	۶۶	۲۱
۴۲	۴-۳-۱۱-۱	۱/۱۹۶۹	۰/۹۹۱۷	۲/۴۱۹۳	۰/۹۸۵۶	۷۱	۲۱
۴۳	۴-۳-۱۲-۱	۱/۱۹۶۷	۰/۹۹۱۷	۲/۴۱۹۳	۰/۹۸۵۶	۷۶	۲۱
۴۴	۴-۳-۱۳-۱	۱/۱۹۶۵	۰/۹۹۱۷	۲/۴۱۹۲	۰/۹۸۵۶	۸۱	۲۱
۴۵	۴-۳-۱۴-۱	۱/۱۹۶۴	۰/۹۹۱۷	۲/۴۱۹۲	۰/۹۸۵۶	۸۶	۲۱
۴۶	۴-۴-۲-۱	۱/۱۷۴۵	۰/۹۹۱۸	۲/۴۸۹۵	۰/۹۸۵۸	۳۳	۲۵
۴۷	۴-۴-۳-۱	۱/۰۰۷	۰/۹۹۳	۲/۴۳۰۱	۰/۹۸۵۴	۳۹	۳۱
۴۸	۴-۴-۴-۱	۱/۱۲۹۹	۰/۹۹۲۱	۲/۵۵۲۷	۰/۹۸۴	۴۵	۲۴
۴۹	۴-۴-۵-۱	۱/۱۳	۰/۹۹۲۱	۲/۵۵۲۸	۰/۹۸۴	۵۱	۲۴

ادامه جدول ۴-۱۰

تعداد پارامترهای مؤثر شبکه	تعداد کل پارامترهای شبکه	ضریب همبستگی (الگوهای آزمون)	MSE (الگوهای آزمون)	ضریب همبستگی (الگوهای آموزش)	MSE (الگوهای آموزش)	ساختار شبکه	ردیف
۲۴	۵۷	۰/۹۸۴	۲/۵۵۲۸	۰/۹۹۲۱	۱/۱۳	۴-۴-۶-۱	۵۰
۳۱	۶۳	۰/۹۸۴۱	۲/۵۴۰۹	۰/۹۹۲۹	۱/۰۱۱۷	۴-۴-۷-۱	۵۱
۲۴	۶۹	۰/۹۸۵۶	۲/۴۵۵	۰/۹۹۱۹	۱/۱۶۶۱	۴-۴-۸-۱	۵۲
۲۲	۷۵	۰/۹۸۵۸	۲/۴۸۹۴	۰/۹۹۱۷	۱/۱۸۵۷	۴-۴-۹-۱	۵۳
۳۱	۸۱	۰/۹۸۴۲	۲/۵۲۶	۰/۹۹۳	۱/۰۰۸	۴-۴-۱۰-۱	۵۴
۲۵	۸۷	۰/۹۸۵۱	۲/۵۲۴۳	۰/۹۹۲	۱/۱۴۸۳	۴-۴-۱۱-۱	۵۵
۲۲	۹۳	۰/۹۸۵۸	۲/۴۸۹۶	۰/۹۹۱۷	۱/۱۸۵۸	۴-۴-۱۲-۱	۵۶
۲۲	۹۹	۰/۹۸۵۸	۲/۴۸۹۷	۰/۹۹۱۷	۱/۱۸۵۸	۴-۴-۱۳-۱	۵۷
۲۲	۱۰۵	۰/۹۸۵۸	۲/۴۸۹۷	۰/۹۹۱۷	۱/۱۸۵۸	۴-۴-۱۴-۱	۵۸
۳۱	۴۰	۰/۹۸۴۵	۲/۵۷۰۴	۰/۹۹۲۵	۱/۰۷۹۴	۴-۵-۲-۱	۵۹
۳۰	۴۷	۰/۹۸۴۸	۲/۶۵۱۳	۰/۹۹۲۶	۱/۰۶۱۲	۴-۵-۳-۱	۶۰
۲۹	۵۴	۰/۹۸۳۹	۲/۶۸۱۵	۰/۹۹۲۷	۱/۰۴۷۹	۴-۵-۴-۱	۶۱
۲۹	۶۱	۰/۹۸۳۹	۲/۶۸۲۵	۰/۹۹۲۷	۱/۰۴۸	۴-۵-۵-۱	۶۲
۲۹	۶۸	۰/۹۸۵۸	۲/۵۱۵۶	۰/۹۹۲۶	۱/۰۵۶۶	۴-۵-۶-۱	۶۳
۲۹	۷۵	۰/۹۸۵۸	۲/۵۲۷۵	۰/۹۹۲۶	۱/۰۵۹۱	۴-۵-۷-۱	۶۴
۲۹	۸۲	۰/۹۸۵۸	۲/۵۴۶۴	۰/۹۹۲۶	۱/۰۶۱	۴-۵-۸-۱	۶۵
۲۹	۸۹	۰/۹۸۵۸	۲/۵۵۴۴	۰/۹۹۲۶	۱/۰۶۲۲	۴-۵-۹-۱	۶۶
۲۹	۹۶	۰/۹۸۴	۲/۶۸۵۸	۰/۹۹۲۷	۱/۰۴۸۱	۴-۵-۱۰-۱	۶۷
۳۴	۴۷	۰/۹۸۴۷	۲/۶۲۳۷	۰/۹۹۲۸	۱/۰۳۲۵	۴-۶-۲-۱	۶۸
۳۲	۵۵	۰/۹۸۴۴	۲/۶۶۵۸	۰/۹۹۲۷	۱/۰۴۲۴	۴-۶-۳-۱	۶۹
۳۲	۶۳	۰/۹۸۴۵	۲/۶۶۵۲	۰/۹۹۲۷	۱/۰۴۳۹	۴-۶-۴-۱	۷۰
۳۲	۷۱	۰/۹۸۴۵	۲/۶۶۵۵	۰/۹۹۲۷	۱/۰۴۵۲	۴-۶-۵-۱	۷۱
۳۲	۷۹	۰/۹۸۴۶	۲/۶۶۵۶	۰/۹۹۲۷	۱/۰۴۶	۴-۶-۶-۱	۷۲
۳۲	۸۷	۰/۹۸۴۶	۲/۶۶۵۶	۰/۹۹۲۷	۱/۰۴۶۶	۴-۶-۷-۱	۷۳
۳۳	۹۵	۰/۹۸۵	۲/۶۰۷۱	۰/۹۹۲۷	۱/۰۴۲۲	۴-۶-۸-۱	۷۴
۴۷	۵۴	۰/۹۸۲۱	۲/۷۴۸۳	۰/۹۹۴۸	۰/۷۴۶۲	۴-۷-۲-۱	۷۵
۴۸	۶۳	۰/۹۸	۳/۲۵۷۶	۰/۹۹۴۱	۰/۸۵۱۷	۴-۷-۳-۱	۷۶
۴۹	۷۲	۰/۹۷۹۶	۳/۰۶۸۳	۰/۹۹۴۱	۰/۸۴۴۹	۴-۷-۴-۱	۷۷
۳۹	۸۱	۰/۹۸۱	۳/۱۰۸۴	۰/۹۹۳۲	۰/۹۷	۴-۷-۵-۱	۷۸
۳۸	۹۰	۰/۹۸۱۳	۳/۰۸۴۱	۰/۹۹۳۲	۰/۹۷۳۸	۴-۷-۶-۱	۷۹
۳۸	۹۹	۰/۹۸۱۳	۳/۰۸۱۶	۰/۹۹۳۲	۰/۹۷۴۵	۴-۷-۷-۱	۸۰

ادامه جدول ۴-۱۰

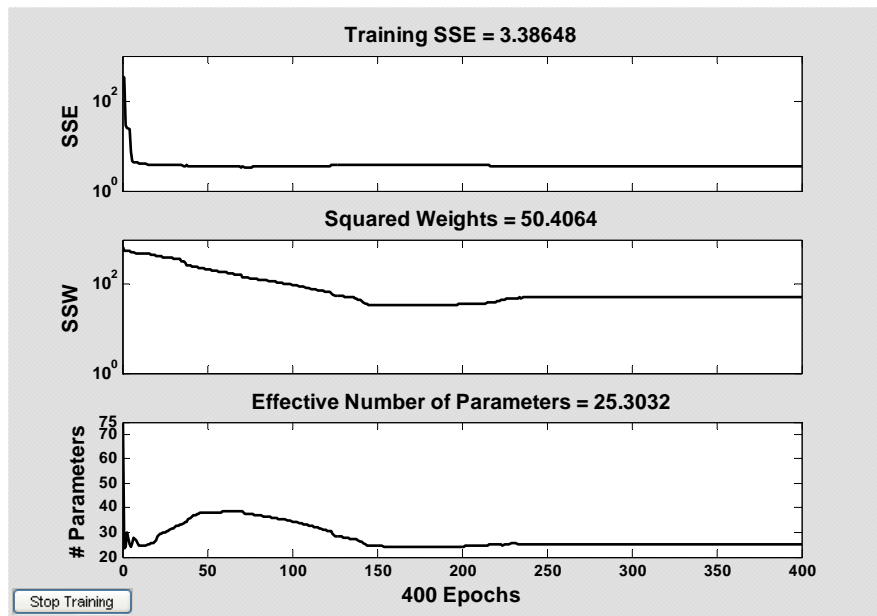
ردیف	ساختار شبکه	MSE (الگوهای آموزش)	ضریب همبستگی (الگوهای آموزش)	MSE (الگوهای آزمون)	ضریب همبستگی (الگوهای آزمون)	تعداد کل پارامترهای شبکه	تعداد پارامترهای مؤثر شبکه
۸۱	۴-۷-۸-۱	۱/۰۴۲۶	۰/۹۹۲۷	۲/۷۴۹	۰/۹۸۳۳	۱۰۸	۳۴
۸۲	۴-۸-۲-۱	۰/۸۹۷۴	۰/۹۹۳۷	۳/۰۳۳۵	۰/۹۸۲۱	۶۱	۴۳
۸۳	۴-۸-۳-۱	۰/۶۰۹	۰/۹۹۵۸	۳/۲۹۴۳	۰/۹۷۸۲	۷۱	۶۲
۸۴	۴-۸-۴-۱	۰/۵۹۳۶	۰/۹۹۵۹	۳/۵۰۷	۰/۹۷۸۲	۸۱	۶۵
۸۵	۴-۸-۵-۱	۰/۹۵۵۵	۰/۹۹۳۳	۳/۱۵۰۵	۰/۹۸۰۸	۹۱	۴۱
۸۶	۴-۸-۶-۱	۰/۹۷۱۲	۰/۹۹۳۲	۳/۰۹۶۵	۰/۹۸۱۲	۱۰۱	۳۹
۸۷	۴-۹-۲-۱	۰/۹۰۱۳	۰/۹۹۳۷	۲/۹۷۸۴	۰/۹۸۰۱	۶۸	۴۵
۸۸	۴-۹-۳-۱	۰/۹۳۳۹	۰/۹۹۳۵	۳/۱۶۲۶	۰/۹۷۹۸	۷۹	۴۳
۸۹	۴-۹-۴-۱	۰/۹۴۶۸	۰/۹۹۳۴	۳/۱۷۶۹	۰/۹۸۰۱	۹۰	۴۲
۹۰	۴-۱۰-۲-۱	۰/۹۱۶۴	۰/۹۹۳۶	۳/۰۰۱۷	۰/۹۸۱۳	۷۵	۴۴
۹۱	۴-۱۰-۳-۱	۰/۶۷۶۶	۰/۹۹۵۳	۳/۷۴۲۵	۰/۹۷۶۹	۸۷	۶۵
۹۲	۴-۱۰-۴-۱	۰/۹۵۴	۰/۹۹۳۴	۳/۱۳۰۲	۰/۹۸۰۴	۹۹	۴۲

نتایج جدول ۴-۱۰ بر اساس میانگین مربعات خطای تخمین الگوهای آزمون مرتب شدند و در نهایت ۷ شبکه که کمترین میانگین مربعات خطا را داشتند برای انجام مراحل بعدی انتخاب شدند. در جدول ۴-۱۱ نتایج این ۷ شبکه آمده است.

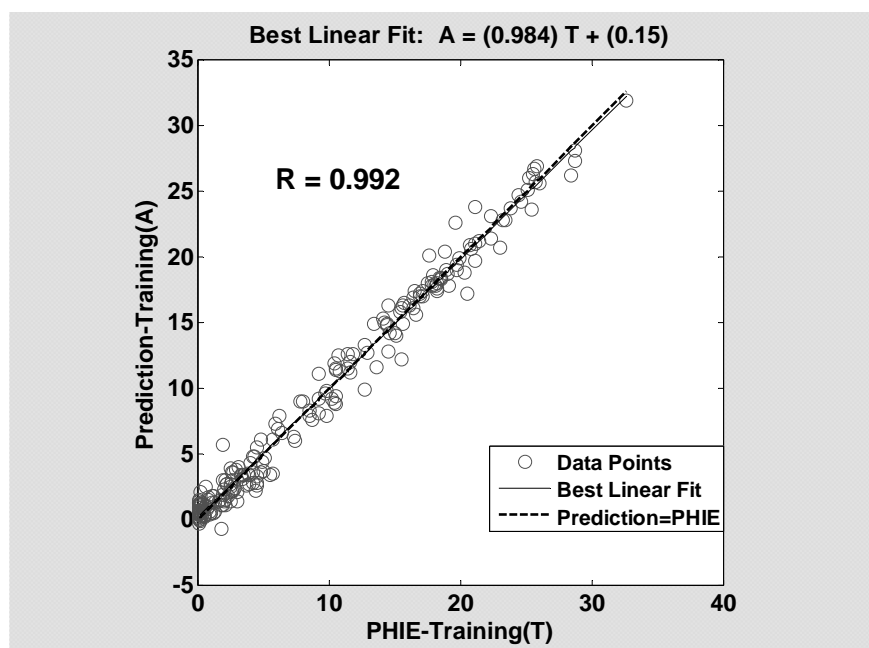
جدول ۴-۱۱: ۷ شبکه منفرد آموزش دیده به روش مرتب سازی با بهترین نتایج.

ردیف	ساختار شبکه	MSE (الگوهای آموزش)	ضریب همبستگی (الگوهای آموزش)	MSE (الگوهای آزمون)	ضریب همبستگی (الگوهای آزمون)	تعداد کل پارامترهای شبکه	تعداد پارامترهای مؤثر شبکه
۱	۴-۲-۱۶-۱	۱/۱۳۹	۰/۹۹۲	۲/۳۹۴	۰/۹۸۵	۷۵	۲۵
۲	۴-۲-۱۴-۱	۱/۲۲۸	۰/۹۹۱	۲/۴۰۵	۰/۹۸۷	۶۷	۲۳
۳	۴-۳-۷-۱	۱/۱۹۸	۰/۹۹۲	۲/۴۱۴	۰/۹۸۶	۵۱	۲۰
۴	۴-۳-۶-۱	۱/۱۹۸	۰/۹۹۲	۲/۴۱۸	۰/۹۸۶	۴۶	۲۰
۵	۴-۳-۸-۱	۱/۱۹۸	۰/۹۹۲	۲/۴۱۹	۰/۹۸۶	۵۶	۲۱
۶	۴-۳-۱۳-۱	۱/۱۹۷	۰/۹۹۲	۲/۴۱۹	۰/۹۸۶	۸۱	۲۱
۷	۴-۳-۱۴-۱	۱/۱۹۶	۰/۹۹۲	۲/۴۱۹	۰/۹۸۶	۸۶	۲۱

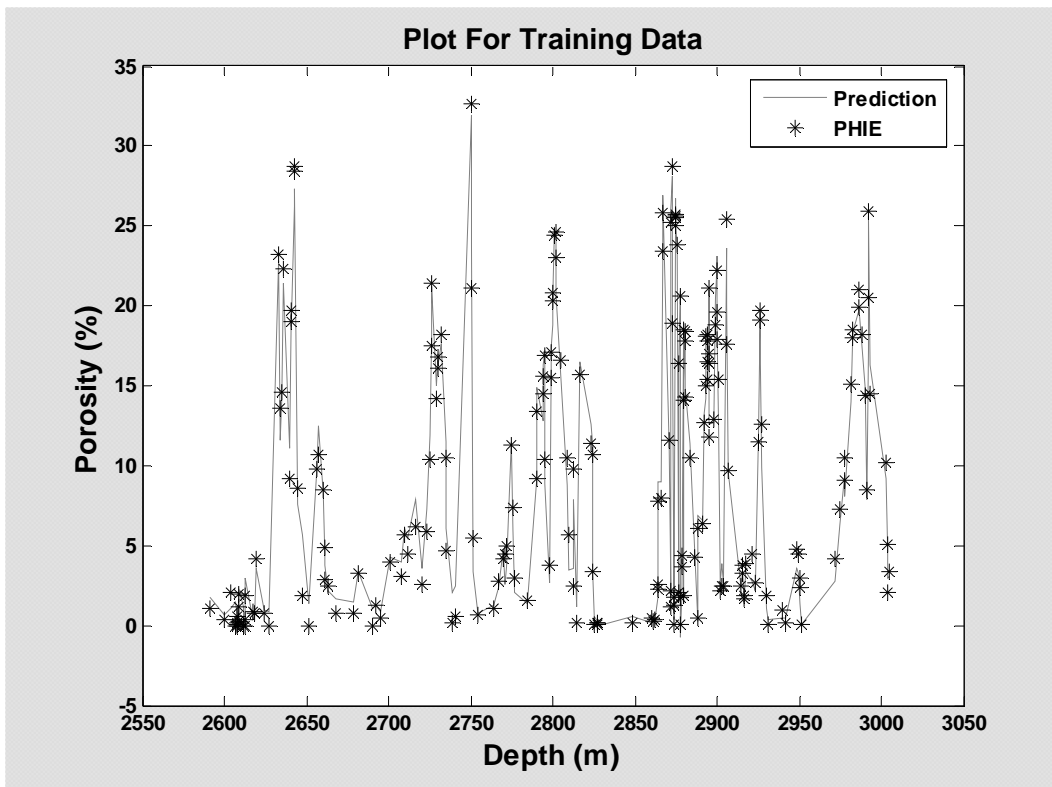
در شکل ۴-۱۳ نتایج بهترین شبکه منفرد حاصل از روش مرتب‌سازی که کمترین MSE تخمین را برای الگوهای آزمون دارد (شبکه شماره ۱ جدول ۴-۱۱)، نمایش داده شده است. در شکل ۴-۱۳ الف، نمودار بالایی مجموع مربعات خطا (SSE)، نمودار وسطی مجموع مربعات پارامترهای شبکه (SSW) و نمودار پایینی تعداد پارامترهای مؤثر شبکه را نشان می‌دهد.



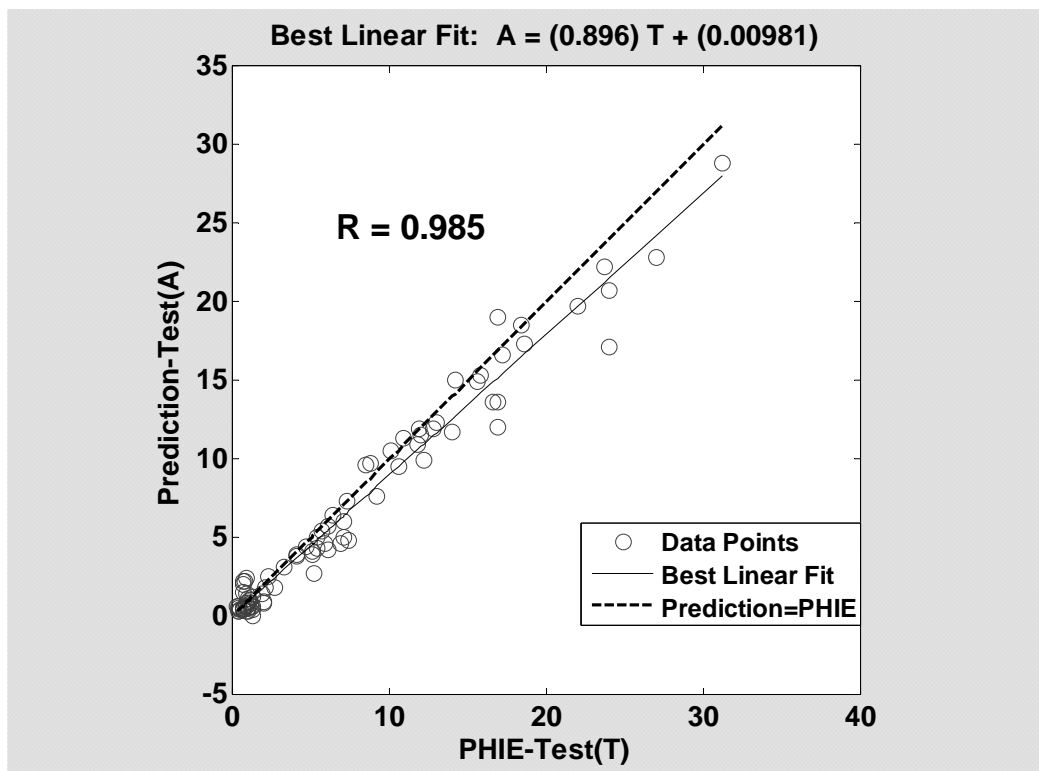
شکل ۴-۱۳ الف: نمایش پیشرفت روند آموزش و همگرایی SSE، SSW و تعداد پارامترهای مؤثر شبکه شماره ۱ جدول ۴-۱۱.



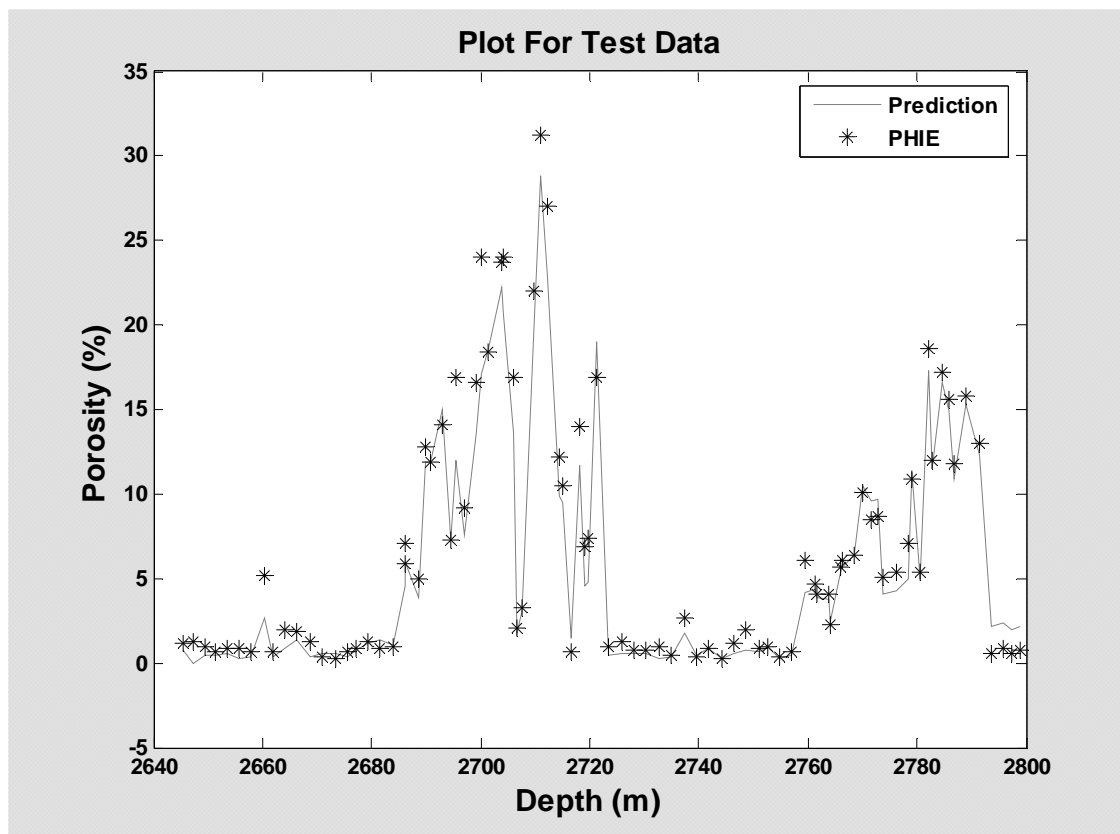
شکل ۴-۱۳ ب: نمودار همبستگی بین تخلخل و تخمین شبکه شماره ۱ جدول ۴-۱۱ برای الگوهای آموزش.



شکل ۴-۱۳-ج: مقادیر تخلخل و تخمین شبکه شماره ۱ جدول ۴-۱۱ برای الگوهای آموزش.



شکل ۴-۱۳-د: نمودار همبستگی بین تخلخل و تخمین شبکه شماره ۱ جدول ۴-۱۱ برای الگوهای آزمون.



شکل ۴-۱۳-۵: مقادیر تخلخل و تخمین شبکه شماره ۱ جدول ۴-۱۱ برای الگوهای آزمون.

۴-۷-۱- ترکیبات آنسامبلی خطی

۱۲۰ ترکیب آنسامبلی ممکن ۷ شبکه جدول ۴-۱۱ با استفاده از ۵ روش ترکیب خطی، روش میانگین‌گیری ساده و ۴ روش ترکیب خطی بهینه (CM1، CM2، CM3 و CM4)، ساخته شد. نتایج این ۱۲۰ ترکیب آنسامبلی به ۵ روش ترکیب در جدول ۴-۱۲ آورده شده است. ستون اول این جدول از سمت چپ (NO.) شماره ترکیب و ستون دوم (Ensemble) ترکیب آنسامبلی، یعنی شبکه‌هایی از جدول ۴-۱۱ که با هم ترکیب شده‌اند، را مشخص می‌کند. بقیه ستون‌های این جدول میانگین مربعات خطای هر یک از ترکیبات آنسامبلی را برای الگوهای آموزش و آزمون در هر یک از ۵ روش ترکیب خطی مشخص می‌کند. به عنوان مثال ردیف شماره ۶۰ این جدول، میانگین مربعات خطای تخمین ترکیب آنسامبلی شبکه‌های ۳، ۴، ۵ و ۷ جدول ۴-۱۱ را به ۵ روش ترکیب خطی برای الگوهای آموزش و آزمون نشان می‌دهد.

جدول ۴-۱۲: نتایج ۱۲۰ ترکیب آنسامبلی شبکه‌های عصبی آموزش‌دیده به روش مرتب‌سازی (۷ شبکه جدول ۴-۱۱) به ۵ روش ترکیب خطی (میانگین‌گیری ساده، CM1، CM2، CM3 و CM4).

NO.	Ensemble	Simple averaging		CM1		CM2		CM3		CM4	
		Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)
1	1,2	1.101	2.329	1.094	2.326	1.095	2.332	1.095	2.320	1.095	2.332
2	1,3	1.093	2.313	1.090	2.309	1.090	2.314	1.090	2.303	1.090	2.314
3	1,4	1.093	2.314	1.090	2.310	1.090	2.315	1.090	2.304	1.090	2.315
4	1,5	1.093	2.315	1.090	2.310	1.090	2.315	1.090	2.304	1.090	2.316
5	1,6	1.092	2.315	1.089	2.311	1.090	2.315	1.090	2.305	1.090	2.316
6	1,7	1.092	2.315	1.089	2.311	1.090	2.315	1.090	2.305	1.090	2.316
7	2,3	1.138	2.390	1.137	2.384	1.138	2.390	1.137	2.378	1.138	2.391
8	2,4	1.138	2.391	1.138	2.386	1.138	2.391	1.138	2.380	1.138	2.392
9	2,5	1.138	2.392	1.137	2.387	1.137	2.392	1.137	2.380	1.137	2.393
10	2,6	1.138	2.392	1.137	2.387	1.137	2.392	1.137	2.381	1.137	2.393
11	2,7	1.138	2.392	1.137	2.387	1.137	2.392	1.137	2.381	1.137	2.393
12	3,4	1.198	2.416	1.197	2.395	1.197	2.397	1.197	2.394	1.197	2.397
13	3,5	1.198	2.417	1.198	2.419	1.198	2.420	1.198	2.417	1.198	2.421
14	3,6	1.197	2.417	1.191	2.470	1.191	2.472	1.191	2.469	1.191	2.472
15	3,7	1.197	2.417	1.190	2.471	1.190	2.472	1.190	2.470	1.190	2.472
16	4,5	1.198	2.418	1.192	2.455	1.192	2.456	1.192	2.454	1.192	2.456
17	4,6	1.197	2.418	1.187	2.444	1.187	2.445	1.187	2.443	1.187	2.444
18	4,7	1.197	2.418	1.187	2.442	1.187	2.442	1.187	2.441	1.187	2.441
19	5,6	1.197	2.419	1.185	2.430	1.185	2.431	1.185	2.429	1.185	2.430
20	5,7	1.197	2.419	1.185	2.428	1.185	2.428	1.185	2.427	1.185	2.427
21	6,7	1.197	2.419	1.189	2.416	1.189	2.415	1.189	2.416	1.189	2.410
22	1,2,3	1.085	2.324	1.078	2.308	1.078	2.314	1.078	2.301	1.078	2.315
23	1,2,4	1.085	2.324	1.078	2.309	1.078	2.315	1.078	2.301	1.078	2.316
24	1,2,5	1.085	2.325	1.078	2.309	1.078	2.316	1.078	2.302	1.078	2.316
25	1,2,6	1.084	2.325	1.077	2.309	1.078	2.316	1.077	2.302	1.078	2.316
26	1,2,7	1.084	2.325	1.077	2.309	1.078	2.316	1.077	2.302	1.078	2.316

ادامه جدول ۴-۱۲

NO.	Ensemble	Simple averaging		CM1		CM2		CM3		CM4	
		Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)
27	1,3,4	1.111	2.327	1.090	2.305	1.090	2.310	1.090	2.300	1.090	2.311
28	1,3,5	1.111	2.328	1.089	2.333	1.089	2.338	1.089	2.327	1.089	2.338
29	1,3,6	1.111	2.328	1.082	2.367	1.082	2.372	1.082	2.362	1.082	2.372
30	1,3,7	1.111	2.328	1.082	2.366	1.082	2.371	1.082	2.361	1.082	2.370
31	1,4,5	1.111	2.329	1.084	2.347	1.084	2.351	1.084	2.342	1.084	2.351
32	1,4,6	1.111	2.329	1.082	2.336	1.083	2.340	1.083	2.331	1.083	2.339
33	1,4,7	1.111	2.329	1.083	2.334	1.083	2.338	1.083	2.329	1.083	2.337
34	2,3,4	1.142	2.395	1.137	2.376	1.137	2.381	1.137	2.370	1.137	2.381
35	2,3,5	1.141	2.395	1.137	2.400	1.137	2.405	1.137	2.393	1.137	2.406
36	2,3,6	1.141	2.395	1.131	2.438	1.131	2.443	1.131	2.432	1.131	2.444
37	2,3,7	1.141	2.395	1.130	2.436	1.131	2.442	1.130	2.431	1.131	2.442
38	2,4,5	1.142	2.396	1.131	2.422	1.132	2.427	1.132	2.416	1.132	2.426
39	2,4,6	1.141	2.396	1.129	2.410	1.129	2.415	1.129	2.405	1.129	2.415
40	2,4,7	1.141	2.396	1.129	2.408	1.130	2.413	1.130	2.403	1.130	2.412
41	2,5,6	1.141	2.397	1.128	2.398	1.128	2.403	1.128	2.393	1.128	2.402
42	2,5,7	1.141	2.397	1.129	2.396	1.129	2.401	1.129	2.391	1.129	2.400
43	3,4,5	1.198	2.417	1.191	2.442	1.191	2.443	1.191	2.441	1.191	2.443
44	3,4,6	1.198	2.417	1.187	2.452	1.187	2.453	1.187	2.451	1.187	2.452
45	3,4,7	1.198	2.417	1.187	2.451	1.187	2.452	1.187	2.451	1.187	2.451
46	3,5,6	1.197	2.418	1.185	2.441	1.185	2.442	1.185	2.440	1.185	2.441
47	3,5,7	1.197	2.418	1.185	2.440	1.185	2.441	1.185	2.440	1.185	2.439
48	4,5,6	1.198	2.419	1.182	2.386	1.182	2.387	1.182	2.385	1.182	2.386
49	4,5,7	1.198	2.419	1.183	2.389	1.183	2.389	1.183	2.388	1.183	2.387
50	1,5,6	1.111	2.329	1.082	2.325	1.082	2.329	1.082	2.320	1.082	2.328
51	1,5,7	1.111	2.329	1.083	2.323	1.083	2.327	1.083	2.318	1.083	2.326
52	1,6,7	1.110	2.329	1.086	2.314	1.086	2.318	1.086	2.310	1.086	2.315
53	2,6,7	1.141	2.397	1.134	2.386	1.134	2.390	1.134	2.382	1.134	2.387

ادامه جدول ۴-۱۲

NO.	Ensemble	Simple averaging		CM1		CM2		CM3		CM4	
		Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)
54	3,6,7	1.197	2.418	1.187	2.449	1.187	2.448	1.187	2.448	1.187	2.444
55	4,6,7	1.197	2.419	1.187	2.441	1.187	2.442	1.187	2.441	1.187	2.440
56	5,6,7	1.197	2.419	1.184	2.440	1.184	2.442	1.184	2.438	1.184	2.444
57	4,5,6,7	1.197	2.419	1.180	2.392	1.180	2.396	1.180	2.390	1.181	2.400
58	3,5,6,7	1.197	2.418	1.184	2.444	1.184	2.446	1.184	2.442	1.184	2.448
59	3,4,6,7	1.197	2.418	1.187	2.449	1.187	2.449	1.187	2.449	1.187	2.448
60	3,4,5,7	1.198	2.418	1.177	2.411	1.177	2.411	1.177	2.410	1.177	2.409
61	3,4,5,6	1.198	2.418	1.175	2.410	1.175	2.412	1.175	2.409	1.175	2.412
62	2,5,6,7	1.149	2.401	1.122	2.423	1.123	2.431	1.123	2.415	1.123	2.439
63	2,4,6,7	1.149	2.400	1.127	2.436	1.127	2.443	1.127	2.430	1.127	2.447
64	2,4,5,7	1.149	2.400	1.129	2.387	1.129	2.391	1.129	2.382	1.129	2.390
65	2,4,5,6	1.149	2.400	1.128	2.382	1.128	2.387	1.128	2.377	1.128	2.386
66	2,3,6,7	1.149	2.400	1.130	2.430	1.130	2.435	1.130	2.425	1.130	2.434
67	2,3,5,7	1.149	2.400	1.128	2.418	1.128	2.423	1.128	2.413	1.128	2.421
68	2,3,5,6	1.149	2.400	1.127	2.419	1.127	2.424	1.127	2.414	1.127	2.423
69	1,5,6,7	1.126	2.344	1.079	2.337	1.079	2.344	1.079	2.331	1.079	2.349
70	1,4,6,7	1.126	2.344	1.081	2.352	1.081	2.358	1.081	2.346	1.081	2.361
71	1,4,5,7	1.126	2.344	1.083	2.328	1.083	2.332	1.083	2.323	1.083	2.331
72	1,4,5,6	1.126	2.344	1.082	2.324	1.082	2.329	1.082	2.319	1.082	2.328
73	1,3,6,7	1.126	2.343	1.082	2.362	1.082	2.366	1.082	2.357	1.082	2.365
74	1,3,5,7	1.126	2.343	1.081	2.355	1.081	2.359	1.081	2.350	1.081	2.358
75	1,3,5,6	1.126	2.343	1.080	2.355	1.081	2.360	1.081	2.350	1.081	2.359
76	1,3,4,7	1.126	2.343	1.081	2.358	1.081	2.363	1.081	2.353	1.081	2.362
77	1,3,4,6	1.126	2.343	1.081	2.359	1.081	2.364	1.081	2.354	1.081	2.363
78	1,2,6,7	1.094	2.336	1.075	2.312	1.075	2.317	1.075	2.305	1.075	2.314
79	1,2,5,7	1.094	2.336	1.071	2.321	1.072	2.326	1.071	2.314	1.072	2.325
80	1,2,5,6	1.094	2.336	1.071	2.322	1.071	2.328	1.071	2.316	1.071	2.327

ادامه جدول ۴-۱۲

NO.	Ensemble	Simple averaging		CM1		CM2		CM3		CM4	
		Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)
81	1,2,4,7	1.095	2.335	1.071	2.331	1.071	2.337	1.071	2.325	1.071	2.336
82	1,2,4,6	1.095	2.335	1.071	2.333	1.071	2.339	1.071	2.326	1.071	2.338
83	1,2,3,7	1.095	2.335	1.070	2.362	1.070	2.368	1.070	2.355	1.070	2.368
84	1,2,3,6	1.095	2.335	1.070	2.364	1.070	2.370	1.070	2.357	1.070	2.370
85	2,3,4,7	1.149	2.399	1.129	2.424	1.129	2.429	1.129	2.419	1.129	2.428
86	2,3,4,6	1.149	2.399	1.129	2.426	1.129	2.431	1.129	2.421	1.129	2.430
87	2,3,4,5	1.149	2.399	1.131	2.421	1.132	2.426	1.132	2.416	1.132	2.426
88	1,3,4,5	1.127	2.343	1.084	2.355	1.084	2.360	1.084	2.350	1.084	2.359
89	1,2,4,5	1.095	2.335	1.072	2.344	1.072	2.350	1.072	2.338	1.072	2.350
90	1,2,3,5	1.095	2.335	1.076	2.333	1.076	2.339	1.076	2.325	1.076	2.340
91	1,2,3,4	1.095	2.334	1.078	2.306	1.078	2.312	1.078	2.298	1.078	2.313
92	3,4,5,6,7	1.197	2.418	1.174	2.412	1.174	2.415	1.174	2.410	1.174	2.419
93	2,4,5,6,7	1.156	2.403	1.121	2.396	1.122	2.406	1.121	2.389	1.122	2.415
94	2,3,5,6,7	1.156	2.403	1.122	2.428	1.123	2.437	1.123	2.421	1.123	2.445
95	2,3,4,6,7	1.156	2.402	1.127	2.436	1.127	2.443	1.127	2.430	1.127	2.447
96	2,3,4,5,7	1.156	2.402	1.125	2.400	1.125	2.404	1.125	2.395	1.125	2.403
97	2,3,4,5,6	1.156	2.402	1.123	2.398	1.123	2.404	1.123	2.393	1.123	2.404
98	1,4,5,6,7	1.137	2.355	1.079	2.329	1.079	2.336	1.079	2.322	1.079	2.342
99	1,3,5,6,7	1.137	2.355	1.078	2.356	1.079	2.363	1.078	2.350	1.079	2.368
100	1,3,4,6,7	1.137	2.354	1.081	2.362	1.081	2.368	1.081	2.357	1.081	2.370
101	1,3,4,5,7	1.138	2.354	1.079	2.344	1.079	2.348	1.079	2.339	1.079	2.346
102	1,3,4,5,6	1.138	2.354	1.078	2.343	1.078	2.347	1.078	2.338	1.078	2.347
103	1,2,5,6,7	1.106	2.346	1.065	2.341	1.066	2.350	1.065	2.332	1.066	2.358
104	1,2,4,6,7	1.106	2.346	1.067	2.359	1.068	2.367	1.068	2.351	1.068	2.372
105	1,2,4,5,7	1.106	2.346	1.071	2.331	1.071	2.337	1.071	2.325	1.071	2.336
106	1,2,4,5,6	1.107	2.346	1.071	2.327	1.071	2.333	1.071	2.321	1.071	2.332
107	1,2,3,6,7	1.106	2.346	1.070	2.362	1.070	2.368	1.070	2.355	1.070	2.367

ادامه جدول ۴-۱۲

NO.	Ensemble	Simple averaging		CM1		CM2		CM3		CM4	
		Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)
108	1,2,3,5,7	1.106	2.345	1.069	2.353	1.069	2.358	1.069	2.346	1.069	2.357
109	1,2,3,5,6	1.106	2.346	1.069	2.354	1.069	2.359	1.069	2.347	1.069	2.359
110	1,2,3,4,7	1.107	2.345	1.070	2.356	1.070	2.361	1.070	2.349	1.070	2.361
111	1,2,3,4,6	1.107	2.345	1.069	2.357	1.070	2.363	1.069	2.351	1.070	2.363
112	1,2,3,4,5	1.107	2.345	1.071	2.354	1.072	2.360	1.071	2.348	1.072	2.360
113	1,2,3,4,5,6	1.117	2.354	1.067	2.343	1.067	2.349	1.067	2.336	1.067	2.349
114	1,2,3,4,5,7	1.117	2.354	1.068	2.344	1.068	2.350	1.068	2.338	1.068	2.348
115	1,2,3,4,6,7	1.117	2.354	1.067	2.364	1.068	2.372	1.067	2.356	1.068	2.376
116	1,2,3,5,6,7	1.117	2.354	1.065	2.357	1.065	2.367	1.065	2.348	1.065	2.374
117	1,2,4,5,6,7	1.117	2.355	1.065	2.335	1.066	2.345	1.065	2.327	1.066	2.354
118	1,3,4,5,6,7	1.146	2.363	1.076	2.343	1.076	2.351	1.076	2.337	1.076	2.356
119	2,3,4,5,6,7	1.161	2.405	1.118	2.406	1.118	2.416	1.118	2.399	1.118	2.424
120	1,2,3,4,5,6,7	1.125	2.361	1.063	2.346	1.063	2.356	1.063	2.337	1.063	2.363

نتایج بهترین آنسامبل‌های حاصل از هر یک از این ۵ روش که نسبت به شبکه ۱ جدول ۴-۱۱ بیشترین کاهش را در میانگین مربعات خطای تخمین الگوهای آزمون ایجاد کرده‌اند، در جدول ۴-۱۳ با دقت سه رقم اعشار آمده است. ردیف‌های این جدول مشابه جدول ۴-۷ است و مقایسه نسبت به بهترین شبکه منفرد (شبکه ۱ جدول ۴-۱۱) انجام شده است.

جدول ۴-۱۳: نتایج بهترین ترکیبات آنسامبلی حاصل از ۵ روش ترکیب خطی (میانگین‌گیری ساده و ۴ روش ترکیب خطی بهینه) ۷ شبکه جدول ۴-۱۱.

روش ترکیب خطی نتایج	میانگین‌گیری ساده	ترکیب خطی بهینه نامقید با جمله ثابت (CM1)	ترکیب خطی بهینه مقید با جمله ثابت (CM2)	ترکیب خطی بهینه نامقید بدون جمله ثابت (CM3)	ترکیب خطی بهینه مقید بدون جمله ثابت (CM4)
ترکیب آنسامبلی	۱ و ۳	۱، ۳ و ۴	۱، ۳ و ۴	۱، ۲، ۳ و ۴	۱، ۳ و ۴
ضریب همبستگی (الگوهای آموزش)	۰/۹۹۲	۰/۹۹۲	۰/۹۹۲	۰/۹۹۳	۰/۹۹۲
MSE (الگوهای آموزش)	۱/۰۹۳	۱/۰۹۰	۱/۰۹۰	۱/۰۷۸	۱/۰۹۰
کاهش در MSE الگوهای آموزش.	٪ ۴	٪ ۴/۳	٪ ۴/۳	٪ ۵/۴	٪ ۴/۳
ضریب همبستگی (الگوهای آزمون)	۰/۹۸۷	۰/۹۸۷	۰/۹۸۷	۰/۹۸۷	۰/۹۸۷
MSE (الگوهای آزمون)	۲/۳۱۳	۲/۳۰۵	۲/۳۱۰	۲/۲۹۸	۲/۳۱۱
کاهش در MSE الگوهای آزمون.	٪ ۳/۴	٪ ۳/۷	٪ ۳/۵	٪ ۴	٪ ۳/۵
α_0	۰	-۰/۰۱۳	۰/۰۰۰	۰	۰
α_1	۰/۵	۰/۵۹۶	۰/۵۹۶	۰/۴۹۶	۰/۵۹۶
α_2	۰	۰	۰	۰/۲۲۶	۰
α_3	۰/۵	۱/۷۶۸	۱/۸۰۹	۱/۲۲۱	۱/۸۰۹
α_4	۰	-۱/۳۶۳	-۱/۴۰۵	-۰/۹۴۲	-۱/۴۰۵
α_5	۰	۰	۰	۰	۰
α_6	۰	۰	۰	۰	۰
α_7	۰	۰	۰	۰	۰
$\sum \alpha_i$	۱	۱/۰۰۱	۱	۱/۰۰۱	۱

به عنوان مثال، ستون پنجم از سمت راست جدول ۴-۱۳ نشان می‌دهد که در روش ترکیب خطی بهینه نامقید بدون جمله ثابت (CM3)، از بین ۱۲۰ ترکیب آنسامبلی، ترکیب شبکه‌های شماره ۱، ۲، ۳ و ۴ جدول ۴-۱۱ بهترین نتیجه را داشته است و میانگین مربعات خطای تخمین آن برای الگوهای

آزمون ۲/۲۹۸ است که ۴٪ از نتیجه بهترین شبکه منفرد (۲/۳۹۴) کمتر است. این ترکیب جمله ثابت ندارد و ضریب رگرسیون شبکه‌های ۱، ۲، ۳ و ۴ در این ترکیب به ترتیب عبارتند از ۰/۴۹۶، ۰/۲۲۶، ۱/۲۲۱۱ و ۰/۹۴۲- و مجموع این ضرایب ۱/۰۰۱ است که خیلی به ۱ نزدیک است. ضرایب بقیه شبکه‌ها در این ترکیب صفر است زیرا از آن‌ها در این ترکیب استفاده نشده است. بنابر اطلاعات این جدول، از بین ۶۰۰ ترکیب آنسامبلی خطی شبکه‌های جدول ۴-۱۱، این ترکیب کمترین میانگین مربعات خطای تخمین الگوهای آزمون را داشته است.

در این جا نیز در مورد ترکیبات نامقید، مجموع ضرایب شبکه‌های جزء آنسامبل به ۱ خیلی نزدیک است و در مورد ترکیبات دارای جمله ثابت، مقدار این جمله به صفر خیلی نزدیک است که بنابر نظر هاشم بیانگر آن است که شبکه‌ها خوب آموزش دیده‌اند.

۴-۷-۲- ترکیبات آنسامبلی غیرخطی با استفاده از شبکه‌های عصبی مصنوعی منفرد

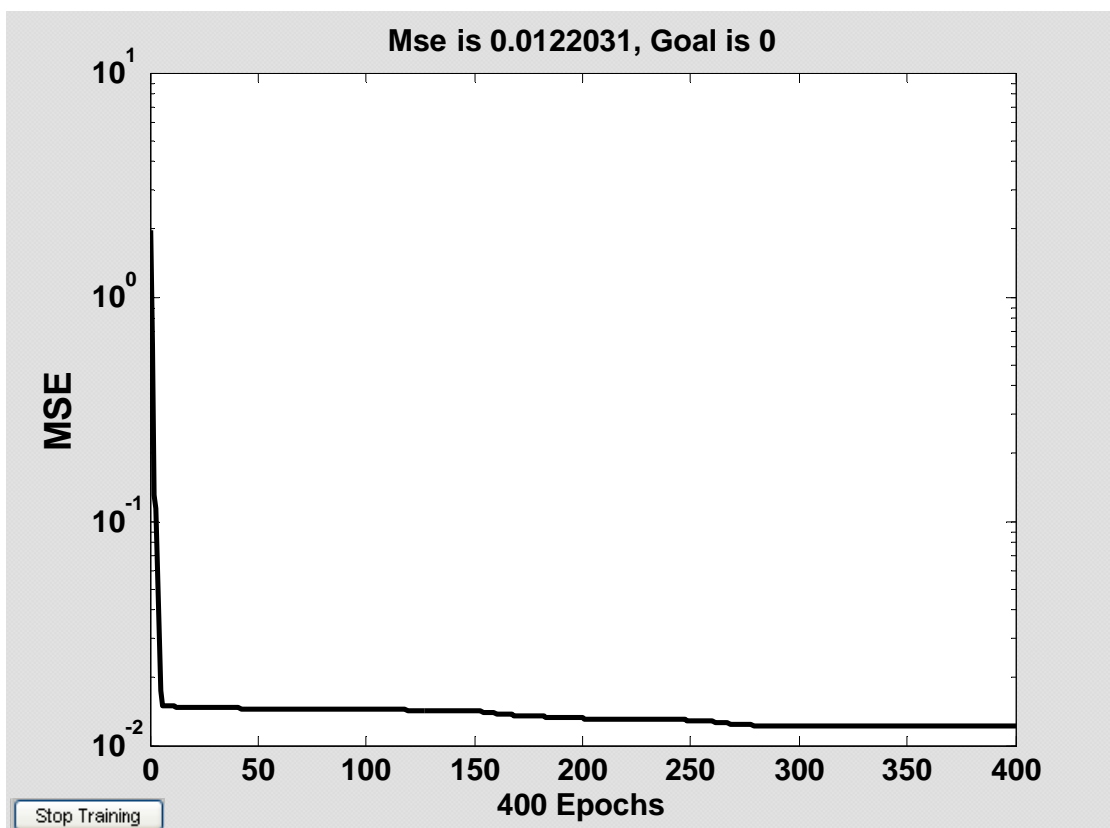
در این قسمت نیز مانند بخش ۴-۶-۲ از خروجی ۷ شبکه منفرد مذکور در جدول ۴-۱۱ به عنوان ورودی یک شبکه عصبی منفرد دیگر استفاده شد. از شبکه‌هایی با یک لایه میانی که تابع انتقال نرون‌های آن تانژانت هیپربولیک و تابع انتقال نرون‌های لایه خروجی تابع خطی است استفاده گردید. برای آموزش شبکه ترکیب‌کننده هر سه روش فرآیند آموزش، مرتب‌سازی و توقف سریع به کار رفتند. تابع پیش‌پردازش Prestd و دوره‌های تکرار ۴۰۰ مرتبه انتخاب شد. برای هر یک از روش‌های آموزش، یک مرحله طولانی آزمون و خطا با تعداد نرون‌های مختلف در لایه میانی و شرایط اولیه گوناگون انجام شد و در نهایت بهترین نتایج انتخاب گردید. جدول ۴-۱۴ بهترین نتیجه هر یک از سه روش پس‌انتشار خطا را نشان می‌دهد.

جدول ۴-۱۴: بهترین نتایج حاصل از ترکیب غیرخطی خروجی‌های ۷ شبکه منفرد مذکور در جدول ۴-۱۱ با استفاده از شبکه عصبی ترکیب‌کننده آموزش دیده به سه روش فرآیند آموزش، مرتب‌سازی و توقف سریع.

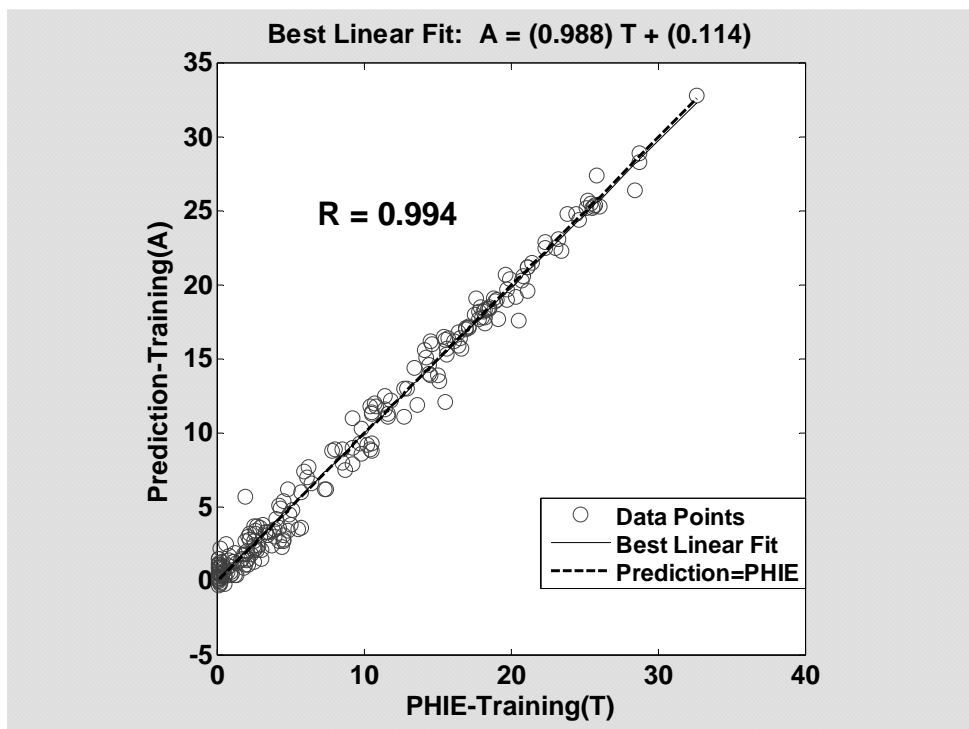
کاهش در MSE الگوهای آزمون.	MSE (الگوهای آزمون)	ضریب همبستگی (الگوهای آزمون)	کاهش در MSE الگوهای آموزش.	MSE (الگوهای آموزش)	ضریب همبستگی (الگوهای آموزش)	ساختار شبکه	نتایج روش آموزش شبکه ترکیب‌کننده
٪ ۲۱/۶	۱/۸۷۷	۰/۹۸۷	٪ ۲۲/۵	۰/۸۸۳	۰/۹۹۴	۷-۴-۱	فرآیند آموزش
٪ ۴/۳	۲/۲۹۲	۰/۹۸۷	٪ ۵/۴	۱/۰۷۷	۰/۹۹۳	۷-۳-۱	مرتب‌سازی
٪ ۵/۶	۲/۲۵۹	۰/۹۸۶	---	۱/۴۰۲	۰/۹۹۱	۷-۷-۱	توقف سریع

با توجه به اطلاعات جدول ۴-۱۴، میانگین مربعات خطای تخمین الگوهای آموزش با استفاده از ترکیب به وسیله شبکه آموزش دیده به روش توقف سریع نسبت به شبکه شماره ۱ جدول ۴-۱۱ نه تنها کاهشی پیدا نکرده، بلکه بیشتر نیز می‌باشد، به همین دلیل در ردیف چهارم ستون پنجم از سمت راست از علامت سه خط تیره (---) استفاده شده است تا این مطلب نشان داده شود. در ادامه این فصل نیز در موارد مشابه از همین علامت استفاده شده است.

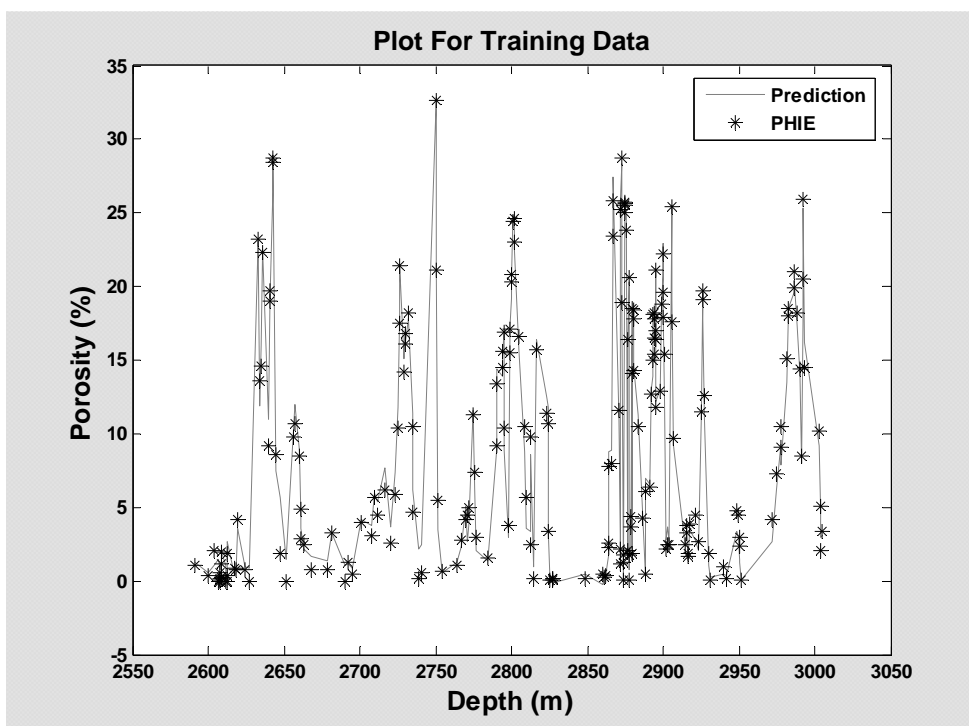
اطلاعات جدول ۴-۱۴ نشان می‌دهد که ترکیب تخمین شبکه‌های منفرد جدول ۴-۱۱ با استفاده از یک شبکه پس‌انتشارخطا با ساختار ۱-۴-۷ که به روش فرآیند آموزش آموزش دیده است، کمترین میانگین مربعات خطا را برای الگوهای آزمون (۱/۸۷۷) داشته و بیشترین کاهش در میانگین مربعات خطا به میزان ۲۱/۶٪ را نسبت به بهترین شبکه منفرد (۲/۳۹۴) ایجاد کرده است. در شکل ۴-۱۴ نتایج حاصل از این ترکیب نمایش داده شده است.



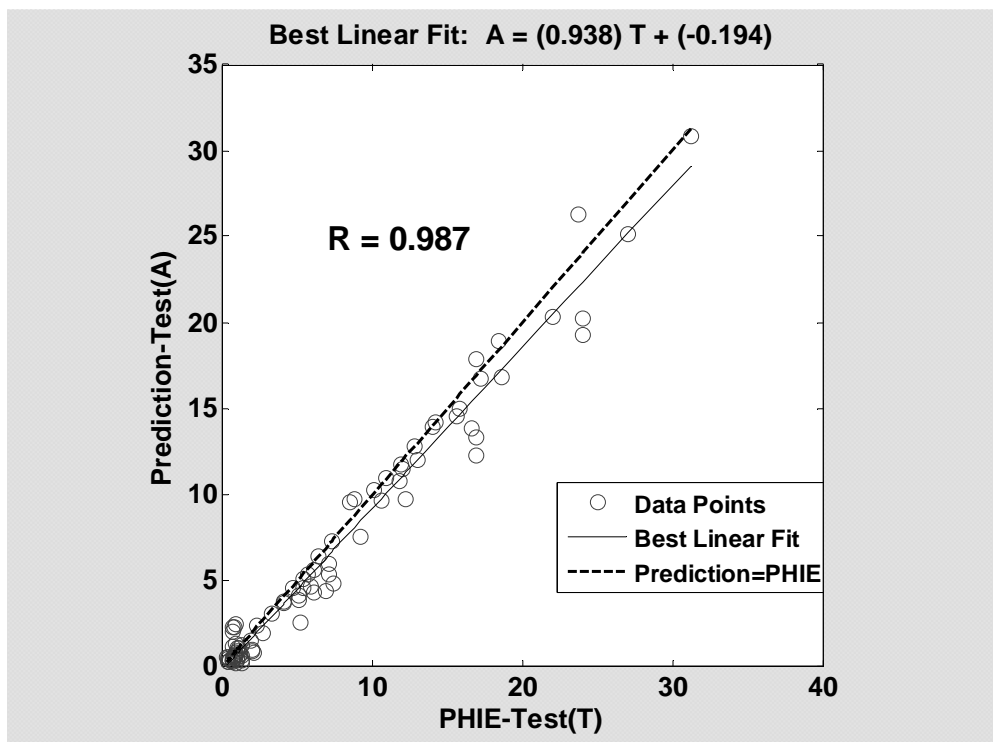
شکل ۴-۱۴ الف: نمایش پیشرفت روند آموزش و کاهش MSE برای الگوهای آموزش در ترکیب تخمین شبکه‌های جدول ۴-۱۱ با استفاده از شبکه‌ای با ساختار ۱-۴-۷ که به روش فرآیند آموزش آموزش دیده است.



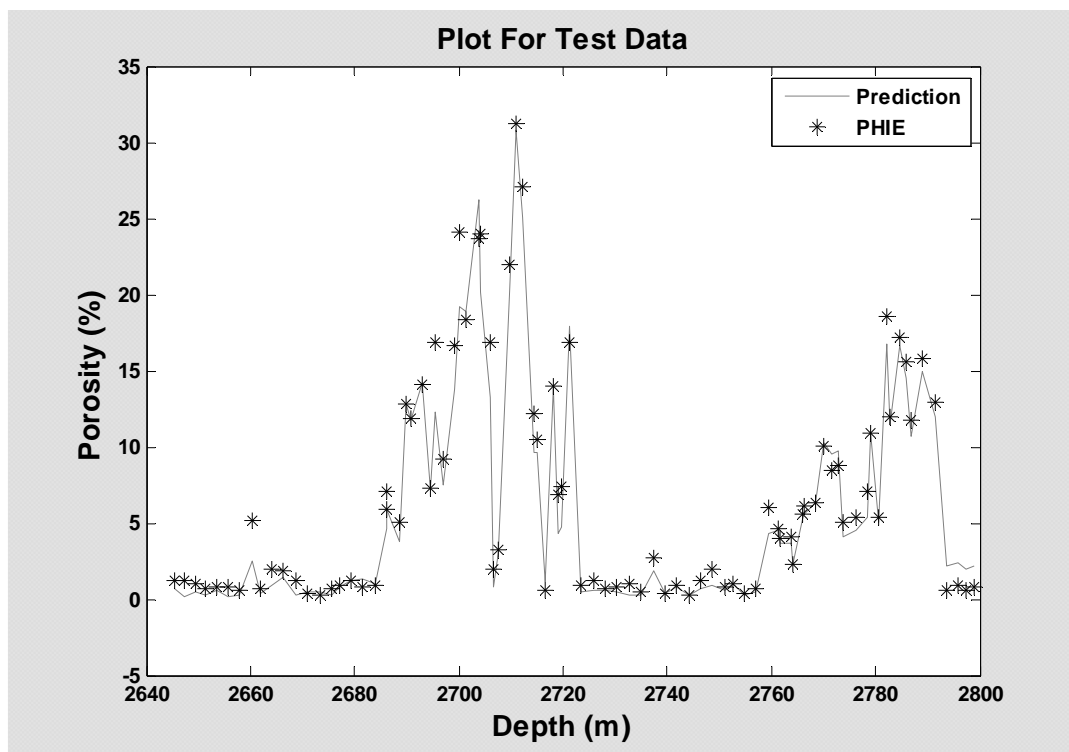
شکل ۴-۱۴ ب: نمودار همبستگی بین تخلخل و تخمین شبکه ترکیب کننده برای الگوهای آموزش در ترکیب تخمین شبکه های منفرد جدول ۴-۱۱ توسط شبکه آموزش دیده با روش فرآ آموزش.



شکل ۴-۱۴ ج: مقادیر تخلخل و تخمین شبکه ترکیب کننده برای الگوهای آموزش در ترکیب تخمین شبکه های منفرد جدول ۴-۱۱ توسط شبکه آموزش دیده با روش فرآ آموزش.



شکل ۴-۱۴-۵: نمودار همبستگی بین تخلخل و تخمین شبکه ترکیب کننده برای الگوهای آزمون در ترکیب تخمین شبکه‌های منفرد جدول ۴-۱۱ توسط شبکه آموزش دیده با روش فرآیند آموزش.



شکل ۴-۱۴-۵: مقادیر تخلخل و تخمین شبکه ترکیب کننده برای الگوهای آزمون در ترکیب تخمین شبکه‌های منفرد جدول ۴-۱۱ توسط شبکه آموزش دیده با روش فرآیند آموزش.

مقایسه اشکال ۴-۱۳-د و ۴-۱۴-د با هم نشان می‌دهد که با استفاده از ترکیب غیرخطی به وسیله یک شبکه عصبی، شیب بهترین خط برازش داده شده در نمودارهای همبستگی برای الگوهای آزمون از ۰/۸۹۶ به ۰/۹۳۸ افزایش یافته است.

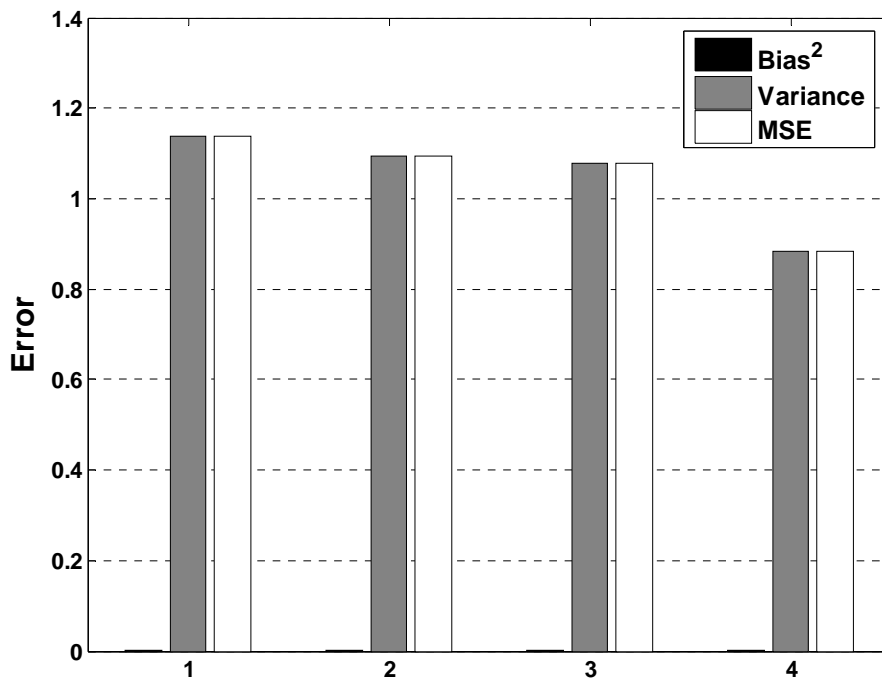
۴-۷-۳- واریانس و بایاس خطا

در جدول ۴-۱۵ واریانس، بایاس و میانگین مربعات خطای تخمین الگوهای آموزش و آزمون برای شبکه شماره ۱ جدول ۴-۱۱ (۱) و بهترین نتایج ترکیب شبکه‌های جدول ۴-۱۱ یعنی میانگین‌گیری ساده شبکه‌های ۱ و ۳ (۲)، ترکیب خطی بهینه نامقید بدون جمله ثابت شبکه‌های ۱، ۲، ۳ و ۴ (۳) و ترکیب غیرخطی تخمین ۷ شبکه با استفاده از شبکه عصبی آموزش دیده به روش فراآموزش با ساختار ۴-۷-۱ (۴) آورده شده است.

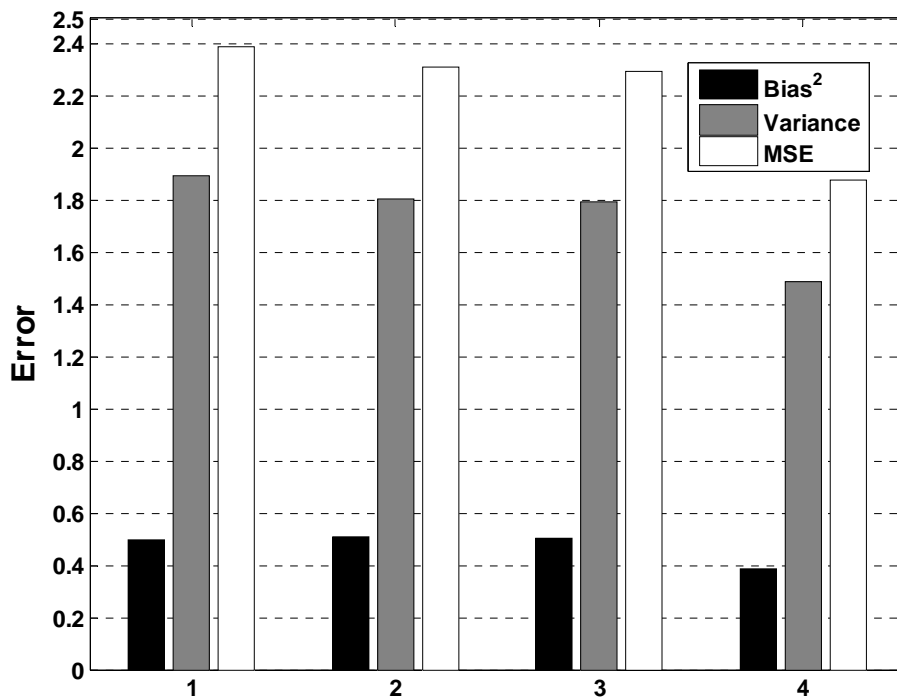
جدول ۴-۱۵: واریانس، بایاس و میانگین مربعات خطای تخمین شبکه شماره ۱ جدول ۴-۱۱ و بهترین نتایج حاصل از ترکیب شبکه‌های آن جدول.

نتایج ردیف	بایاس خطای تخمین (الگوهای آموزش)	واریانس خطای تخمین (الگوهای آموزش)	میانگین مربعات خطای تخمین (الگوهای آموزش)	بایاس خطای تخمین (الگوهای آموزش)	واریانس خطای تخمین (الگوهای آموزش)	میانگین مربعات خطای تخمین (الگوهای آزمون)
۱	-3×10^{-4}	۱/۱۳۹	۱/۱۳۹	-۰/۷۰۶	۱/۸۹۵	۲/۳۹۴
۲	-5×10^{-4}	۱/۰۹۳	۱/۰۹۳	-۰/۷۱۳	۱/۸۰۵	۲/۳۱۳
۳	۰/۰۰۸	۱/۰۷۸	۱/۰۷۸	-۰/۷۱۰	۱/۷۹۴	۲/۲۹۸
۴	2×10^{-4}	۰/۸۸۳	۰/۸۸۳	-۰/۶۲۱	۱/۴۹۱	۱/۸۷۷

شکل‌های ۴-۱۵ نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطای تخمین را برای الگوهای آموزش و آزمون برای ۴ حالت مذکور در جدول ۴-۱۵ نشان می‌دهد. مطابق شکل ۴-۱۵-الف، برای الگوهای آموزش واریانس مؤلفه اصلی خطا است و بایاس تقریباً نقشی در خطا ندارد ولی مطابق شکل ۴-۱۵-ب، برای الگوهای آزمون هر دو مؤلفه بایاس و واریانس مؤثرند اگر چه نقش واریانس بیشتر است.



شکل ۴-۱۵-الف: نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطا در بهترین حالات برای الگوهای آموزش (۱: شبکه شماره ۱ جدول ۴-۱۱، ۲: بهترین ترکیب شبکه‌های جدول ۴-۱۱ به روش میانگین‌گیری ساده، ۳: بهترین ترکیب خطی بهینه شبکه‌های جدول ۴-۱۱، ۴: بهترین ترکیب ۷ شبکه جدول ۴-۱۱ با استفاده از شبکه عصبی).



شکل ۴-۱۵-ب: نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطا در بهترین حالات برای الگوهای آموزش (۱: شبکه شماره ۱ جدول ۴-۱۱، ۲: بهترین ترکیب شبکه‌های جدول ۴-۱۱ به روش میانگین‌گیری ساده، ۳: بهترین ترکیب خطی بهینه شبکه‌های جدول ۴-۱۱، ۴: بهترین ترکیب ۷ شبکه جدول ۴-۱۱ با استفاده از شبکه عصبی).

۴-۸- طراحی شبکه‌های پس‌انتشار خطا به روش توقف سریع و ترکیب نتایج

برای آموزش به روش توقف سریع از تابع آموزش لونیبرگ- مارکوارت استفاده شد. با توجه به توضیحات پیوست الف، برای این که همگرایی خیلی سریع اتفاق نیفتد، پارامترهای تابع آموزش به گونه‌ای مناسب انتخاب شدند ($\mu_{inc}=1.2$, $\mu_{dec}=0.95$, $\mu=1$). الگوهای آموزش به دو دسته، ۷۰ درصد آن‌ها برای آموزش شبکه‌ها و ۳۰ درصد باقی‌مانده به منظور اعتبارسنجی حین آموزش، تقسیم شدند. شبکه‌های پس‌انتشار خطا با یک و دو لایه میانی با تعداد نرون‌های متعدد طی یک مرحله طولانی آزمون و خطا با شرایط اولیه مختلف مورد آزمون قرار گرفت و برای هر ساختار خاص بهترین نتیجه ثبت شد. ساختارهای مورد استفاده همان ساختارهایی است که در روش فرآیند آموزش مورد استفاده قرار گرفت. نتایج حاصل از این مرحله در جدول ۴-۱۶ نمایش داده شده است. کمیت‌هایی که هر یک از ستون‌های این جدول نشان می‌دهند مشابه ستون‌های جداول ۴-۴ و ۴-۱۰ است و تنها تفاوت، ستون اضافی دوره توقف آموزش است.

جدول ۴-۱۶: نتایج حاصل از شبکه‌های عصبی پس‌انتشار خطا با یک و دو لایه میانی به روش توقف سریع.

Epochs = 400, show = 20, goal = 0, min_grad = $1.00 e^{-10}$, max_fail = 5, mem_reduce = 1, $\mu=1$, $\mu_{dec} = 0.95$, $\mu_{inc} = 1.2$, $\mu_{max} = 1.00 e^{+010}$.

ردیف	ساختار شبکه	MSE (الگوهای آموزش)	ضریب همبستگی (الگوهای آموزش)	MSE (الگوهای آزمون)	ضریب همبستگی (الگوهای آزمون)	تعداد کل پارامترهای شبکه	دوره توقف آموزش
۱	۴-۳-۱	۲/۹۰۲۹	۰/۹۸۳۸	۳/۷۳۷۳	۰/۹۷۷۵	۱۹	۷
۲	۴-۴-۱	۲/۰۳۱۴	۰/۹۸۵۹	۲/۶۵۸۸	۰/۹۸۴۹	۲۵	۲۲
۳	۴-۵-۱	۱/۶۸۴۷	۰/۹۸۸۶	۲/۵۳۹۵	۰/۹۸۵۳	۳۱	۸
۴	۴-۶-۱	۱/۵۷۳۷	۰/۹۸۹۱	۲/۸۱۷۵	۰/۹۸۳۳	۳۷	۸
۵	۴-۷-۱	۱/۹۴۵۹	۰/۹۸۶۵	۲/۷۹۲۱	۰/۹۸۵۱	۴۳	۱۵
۶	۴-۸-۱	۱/۸۹۹۲	۰/۹۸۶۸	۲/۴۸۱۷	۰/۹۸۶۵	۴۹	۸
۷	۴-۹-۱	۱/۳۵۴۶	۰/۹۹۰۶	۲/۶۹۷۳	۰/۹۸۵۲	۵۵	۱۲
۸	۴-۱۰-۱	۱/۴۲۰۷	۰/۹۹۰۲	۲/۸۴۲۳	۰/۹۸۵۸	۶۱	۸
۹	۴-۱۱-۱	۱/۴۸۷	۰/۹۸۹۶	۲/۴۶۹۷	۰/۹۸۵۴	۶۷	۱۱
۱۰	۴-۱۲-۱	۱/۹۹۰۴	۰/۹۸۶۳	۲/۴۴۶۹	۰/۹۸۴۴	۷۳	۸
۱۱	۴-۱۳-۱	۱/۲۸۶۳	۰/۹۹۱۱	۲/۵۱۷۷	۰/۹۸۵۶	۷۹	۴۲
۱۲	۴-۱۴-۱	۱/۴۲۸۶	۰/۹۹۰۱	۲/۶۸۰۶	۰/۹۸۵۴	۸۵	۹

1- Stop epoch

ادامه جدول ۴-۱۶

ردیف	ساختار شبکه	MSE (الگوهای آموزش)	ضریب همبستگی (الگوهای آموزش)	MSE (الگوهای آزمون)	ضریب همبستگی (الگوهای آزمون)	تعداد کل پارامترهای شبکه	دوره توقف آموزش
۱۳	۴-۱۵-۱	۱/۵۱۹	۰/۹۸۹۴	۲/۵۹۳۱	۰/۹۸۷۳	۹۱	۱۴
۱۴	۴-۲-۲-۱	۴/۵۶۳۷	۰/۹۸۱۷	۴/۳۶۰۹	۰/۹۸۰۳	۱۹	۸
۱۵	۴-۲-۳-۱	۲/۱۷۲۷	۰/۹۸۵۷	۳/۰۹۶۹	۰/۹۷۶۹	۲۳	۱۰
۱۶	۴-۲-۴-۱	۲/۰۰۲۵	۰/۹۸۶	۳/۲۳۱	۰/۹۷۵۶	۲۷	۱۰
۱۷	۴-۲-۵-۱	۲/۲۷۹۲	۰/۹۸۵۷	۳/۱۹۱۲	۰/۹۸۲۴	۳۱	۹
۱۸	۴-۲-۶-۱	۲/۶۴۷۴	۰/۹۸۳۹	۳/۳۸۹۵	۰/۹۷۸۷	۳۵	۱۰
۱۹	۴-۲-۷-۱	۱/۷۱۶۵	۰/۹۸۸	۲/۶۸۶۶	۰/۹۸۳۸	۳۹	۲۱
۲۰	۴-۲-۸-۱	۱/۹۱۸۱	۰/۹۸۶۸	۲/۵۱۶۴	۰/۹۸۳۴	۴۳	۱۰
۲۱	۴-۲-۹-۱	۱/۹۶۵۴	۰/۹۸۶۸	۲/۶۴۱۱	۰/۹۸۴۲	۴۷	۸
۲۲	۴-۲-۱۰-۱	۲/۰۶۹۷	۰/۹۸۶۱	۳/۰۰۷۹	۰/۹۸۰۸	۵۱	۱۰
۲۳	۴-۲-۱۱-۱	۲/۳۶۵۵	۰/۹۸۴۴	۲/۸۶۴	۰/۹۷۷	۵۵	۸
۲۴	۴-۲-۱۲-۱	۱/۸۰۰۱	۰/۹۸۷۴	۲/۹۷۵۷	۰/۹۷۷۷	۵۹	۸
۲۵	۴-۲-۱۳-۱	۱/۶۶۰۱	۰/۹۸۸۴	۲/۷۰۰۵	۰/۹۸۴۱	۶۳	۱۹
۲۶	۴-۲-۱۴-۱	۱/۹۱۹۸	۰/۹۸۷۹	۲/۶۰۲۳	۰/۹۸۲۸	۶۷	۸
۲۷	۴-۲-۱۵-۱	۲/۷۵۲۳	۰/۹۸۳۲	۲/۷۹۸۴	۰/۹۸۳۶	۷۱	۸
۲۸	۴-۲-۱۶-۱	۱/۹۹۰۱	۰/۹۸۶۹	۲/۷۷۳۴	۰/۹۸۵۹	۷۵	۱۰
۲۹	۴-۲-۱۷-۱	۱/۸۱۳۴	۰/۹۸۸۴	۲/۲۵۷۴	۰/۹۸۶۴	۷۹	۹
۳۰	۴-۲-۱۸-۱	۲/۳۱۴۳	۰/۹۸۵۳	۲/۶۸۶۱	۰/۹۸۴۴	۸۳	۸
۳۱	۴-۲-۱۹-۱	۲/۳۰۲۹	۰/۹۸۴۸	۲/۳۳۴۶	۰/۹۸۴۸	۸۷	۸
۳۲	۴-۲-۲۰-۱	۱/۸۶۹۷	۰/۹۸۷۱	۲/۷۷۰۹	۰/۹۸۳۳	۹۱	۱۰
۳۳	۴-۳-۲-۱	۳/۴۰۴۲	۰/۹۸۱۵	۳/۱۳۱۵	۰/۹۸۱۸	۲۶	۹
۳۴	۴-۳-۳-۱	۲/۱۴۶۱	۰/۹۸۵۵	۳/۰۵۶۷	۰/۹۸۵	۳۱	۹
۳۵	۴-۳-۴-۱	۲/۳۹۷	۰/۹۸۳۵	۲/۹۷۵۸	۰/۹۸۱۱	۳۶	۸
۳۶	۴-۳-۵-۱	۳/۴۷۹۳	۰/۹۸۳۴	۳/۵۷۶۱	۰/۹۷۸۷	۴۱	۷
۳۷	۴-۳-۶-۱	۳/۰۸۵۴	۰/۹۸۵۱	۳/۵۰۵۷	۰/۹۸۴۱	۴۶	۸
۳۸	۴-۳-۷-۱	۲/۰۰۶۸	۰/۹۸۶۱	۲/۵۳۱۴	۰/۹۸۲۵	۵۱	۹
۳۹	۴-۳-۸-۱	۲/۳۱۲۹	۰/۹۸۶۵	۲/۷۲۱	۰/۹۸۵۳	۵۶	۹
۴۰	۴-۳-۹-۱	۱/۶۹۱	۰/۹۸۸۲	۲/۶۸۷۵	۰/۹۸۳۸	۶۱	۹
۴۱	۴-۳-۱۰-۱	۱/۵۵۸۵	۰/۹۸۹۲	۲/۹۳۳۷	۰/۹۸۳۳	۶۶	۸
۴۲	۴-۳-۱۱-۱	۱/۷۶۶	۰/۹۸۸۶	۳/۱۲۱۴	۰/۹۸۱	۷۱	۸
۴۳	۴-۳-۱۲-۱	۲/۱۵۷۴	۰/۹۸۵۹	۲/۷۹۰۱	۰/۹۸۰۲	۷۶	۹

ادامه جدول ۴-۱۶

ردیف	ساختار شبکه	MSE (الگوهای آموزش)	ضریب همبستگی (الگوهای آموزش)	MSE (الگوهای آزمون)	ضریب همبستگی (الگوهای آزمون)	تعداد کل پارامترهای شبکه	دوره توقف آموزش
۴۴	۴-۳-۱۳-۱	۱/۱۵۶۳	۰/۹۸۹۵	۲/۶۵۰۷	۰/۹۸۴۳	۸۱	۲۰
۴۵	۴-۳-۱۴-۱	۱/۸۴۸۸	۰/۹۸۷۲	۲/۴۲۹۸	۰/۹۸۵۳	۸۶	۸
۴۶	۴-۴-۲-۱	۳/۵۸۷۲	۰/۹۸۱۱	۴/۳۱۴۶	۰/۹۶۹	۳۳	۸
۴۷	۴-۴-۳-۱	۱/۶۹۸۸	۰/۹۸۸۴	۲/۵۹۴۶	۰/۹۸۲۷	۳۹	۸
۴۸	۴-۴-۴-۱	۱/۹۲۴۴	۰/۹۸۸۶	۲/۳۰۴۹	۰/۹۸۶	۴۵	۸
۴۹	۴-۴-۵-۱	۲/۱۲۱۲	۰/۹۸۵۵	۲/۶۰۹۱	۰/۹۸۶۴	۵۱	۹
۵۰	۴-۴-۶-۱	۲/۲۰۵۳	۰/۹۸۶۲	۳/۵۵۸۱	۰/۹۸۵۷	۵۷	۷
۵۱	۴-۴-۷-۱	۲/۰۳۵۹	۰/۹۸۶۳	۲/۹۵۴۲	۰/۹۸۲۷	۶۳	۸
۵۲	۴-۴-۸-۱	۲/۰۵۱۵	۰/۹۸۵۷	۲/۵۰۹۹	۰/۹۸۷۹	۶۹	۱۴
۵۳	۴-۴-۹-۱	۲/۶۷۸۴	۰/۹۸۱۳	۲/۶۹۵۷	۰/۹۸۴۹	۷۵	۱۷
۵۴	۴-۴-۱۰-۱	۱/۸۶۴۴	۰/۹۸۷	۲/۵۱۳۶	۰/۹۸۶	۸۱	۱۰
۵۵	۴-۴-۱۱-۱	۲/۳۸۵۶	۰/۹۸۶۸	۲/۹۶۹۴	۰/۹۸۴۳	۸۷	۸
۵۶	۴-۴-۱۲-۱	۱/۸۸۵۵	۰/۹۸۷	۲/۷۵۶۱	۰/۹۸۴۹	۹۳	۸
۵۷	۴-۴-۱۳-۱	۱/۴۵۰۶	۰/۹۸۹۹	۲/۶۵۲	۰/۹۸۵	۹۹	۱۰
۵۸	۴-۴-۱۴-۱	۲/۲۷۴۲	۰/۹۸۴۱	۲/۴۸۷۸	۰/۹۸۶۳	۱۰۵	۲۲
۵۹	۴-۵-۲-۱	۲/۲۳۴۴	۰/۹۸۵۴	۲/۶۶۳۹	۰/۹۸۵۷	۴۰	۹
۶۰	۴-۵-۳-۱	۲/۱۸۱۹	۰/۹۸۷۳	۳/۲۹۵۱	۰/۹۸۱۶	۴۷	۹
۶۱	۴-۵-۴-۱	۲/۰۷۵۳	۰/۹۸۶۹	۳/۰۱۸۴	۰/۹۸۵۳	۵۴	۹
۶۲	۴-۵-۵-۱	۱/۹۲۷۹	۰/۹۸۶۹	۲/۵۲۸۷	۰/۹۸۵۷	۶۱	۸
۶۳	۴-۵-۶-۱	۱/۷۸۹۱	۰/۹۸۷۶	۲/۴۳۸۸	۰/۹۸۴۶	۶۸	۱۰
۶۴	۴-۵-۷-۱	۱/۷۲۰۵	۰/۹۸۸۲	۲/۶۲۰۷	۰/۹۸۶۱	۷۵	۹
۶۵	۴-۵-۸-۱	۱/۹۰۹۸	۰/۹۸۷	۲/۸۵۰۶	۰/۹۸۴	۸۲	۹
۶۶	۴-۵-۹-۱	۱/۸۵۵۷	۰/۹۸۷۲	۲/۹۵۴۲	۰/۹۸۳۲	۸۹	۱۸
۶۷	۴-۵-۱۰-۱	۲/۰۰۰۸	۰/۹۸۶۴	۲/۳۴۸۸	۰/۹۸۴۹	۹۶	۹
۶۸	۴-۶-۲-۱	۳/۴۸۴۸	۰/۹۸۲۵	۴/۲۳۲۵	۰/۹۸۲۵	۴۷	۹
۶۹	۴-۶-۳-۱	۳/۲۹۴۳	۰/۹۸۶۲	۴/۷۷۹۸	۰/۹۸۳۱	۵۵	۹
۷۰	۴-۶-۴-۱	۱/۶۶۵	۰/۹۸۸۵	۲/۴۶۸۷	۰/۹۸۶	۶۳	۱۱
۷۱	۴-۶-۵-۱	۲/۱۴۷۲	۰/۹۸۵۱	۲/۸۸۲۳	۰/۹۸۳۸	۷۱	۱۰
۷۲	۴-۶-۶-۱	۱/۲۶۴۵	۰/۹۹۱۲	۲/۵۴۳۷	۰/۹۸۵۷	۷۹	۹
۷۳	۴-۶-۷-۱	۲/۶۵۳۶	۰/۹۸۴۱	۲/۹۱۴۷	۰/۹۸۳۱	۸۷	۷
۷۴	۴-۶-۸-۱	۱/۶۸۱۶	۰/۹۸۸۳	۲/۶۵۵	۰/۹۸۴۷	۹۵	۹

ادامه جدول ۴-۱۶

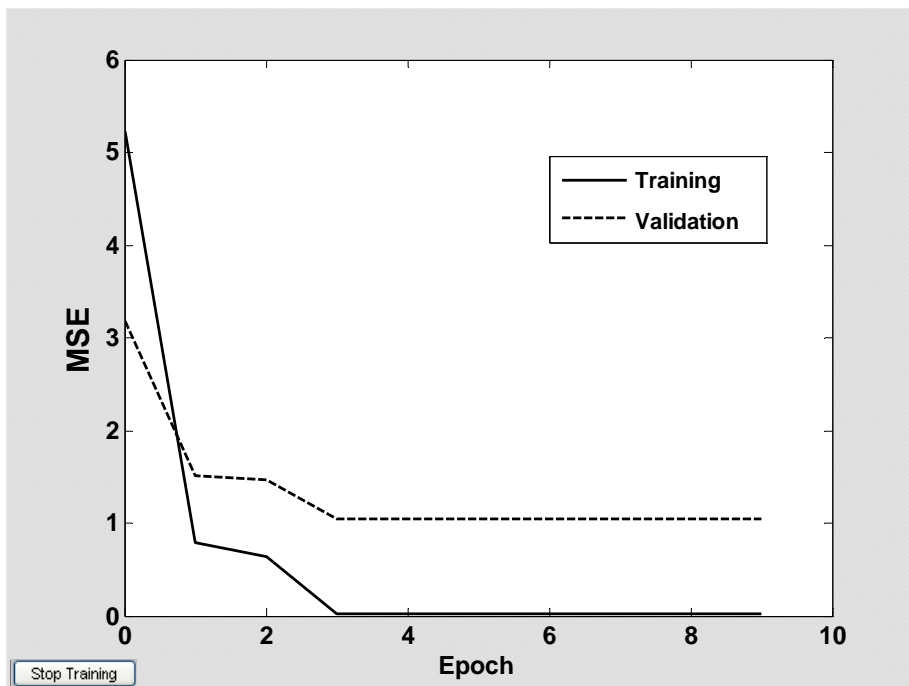
ردیف	ساختار شبکه	MSE (الگوهای آموزش)	ضریب همبستگی (الگوهای آموزش)	MSE (الگوهای آزمون)	ضریب همبستگی (الگوهای آزمون)	تعداد کل پارامترهای شبکه	دوره توقف آموزش
۷۵	۴-۷-۲-۱	۱/۴۴۴۶	۰/۹۹	۳/۰۴۸۶	۰/۹۸۳۳	۵۴	۸
۷۶	۴-۷-۳-۱	۱/۵۲۰۳	۰/۹۸۹۴	۲/۳۶۸۳	۰/۹۸۷۷	۶۳	۱۲
۷۷	۴-۷-۴-۱	۲/۶۲۳	۰/۹۸۱۸	۲/۶۶۴۳	۰/۹۸۶۳	۷۲	۱۳
۷۸	۴-۷-۵-۱	۲/۰۲۹	۰/۹۸۶۴	۲/۸۰۴۵	۰/۹۸۳۷	۸۱	۹
۷۹	۴-۷-۶-۱	۲/۱۱۶۹	۰/۹۸۵۲	۲/۵۸۲	۰/۹۸۶۴	۹۰	۱۵
۸۰	۴-۷-۷-۱	۱/۵۹۷۶	۰/۹۸۸۹	۲/۴۹۲۶	۰/۹۸۶۷	۹۹	۹
۸۱	۴-۷-۸-۱	۲/۳۳۶۱	۰/۹۸۵۶	۲/۹۰۴۲	۰/۹۸۰۶	۱۰۸	۸
۸۲	۴-۸-۲-۱	۳/۰۹۹۳	۰/۹۸۴۶	۴/۱۵۱۱	۰/۹۷۷۴	۶۱	۱۰
۸۳	۴-۸-۳-۱	۴/۰۰۳۲	۰/۹۸۸۴	۴/۴۳۶۲	۰/۹۷۸۸	۷۱	۸
۸۴	۴-۸-۴-۱	۲/۸۱۳۹	۰/۹۸۴۶	۳/۱۰۰۹	۰/۹۸۱۹	۸۱	۸
۸۵	۴-۸-۵-۱	۱/۹۹۳۴	۰/۹۸۷۳	۲/۳۲۳۹	۰/۹۸۲۹	۹۱	۸
۸۶	۴-۸-۶-۱	۱/۸۹۰۱	۰/۹۸۷	۲/۸۶۳۸	۰/۹۸۲	۱۰۱	۹
۸۷	۴-۹-۲-۱	۳/۷۲۲	۰/۹۸۳۵	۴/۰۵۳۶	۰/۹۸۲	۶۸	۸
۸۸	۴-۹-۳-۱	۳/۲۸۴۹	۰/۹۸۱۳	۳/۷۷۶	۰/۹۸۳۱	۷۹	۱۱
۸۹	۴-۹-۴-۱	۳/۴۸۱۸	۰/۹۸۳	۳/۹۷۹	۰/۹۷۵۳	۹۰	۹
۹۰	۴-۱۰-۲-۱	۳/۹۷۳۸	۰/۹۷۵۲	۴/۷۳۲۷	۰/۹۶۵۴	۷۵	۸
۹۱	۴-۱۰-۳-۱	۳/۲۹۵۲	۰/۹۷۷۴	۳/۰۸۰۷	۰/۹۷۴۵	۸۷	۷
۹۲	۴-۱۰-۴-۱	۳/۰۶۸	۰/۹۸۶	۴/۲۹۶۲	۰/۹۸۲۲	۹۹	۸

این نتایج بر اساس میانگین مربعات خطای تخمین الگوهای آزمون مرتب شدند و در نهایت ۷ شبکه که کمترین میانگین مربعات خطا را داشتند برای انجام مراحل بعدی انتخاب شدند. در جدول ۴-۱۷ نتایج این ۷ شبکه آمده است.

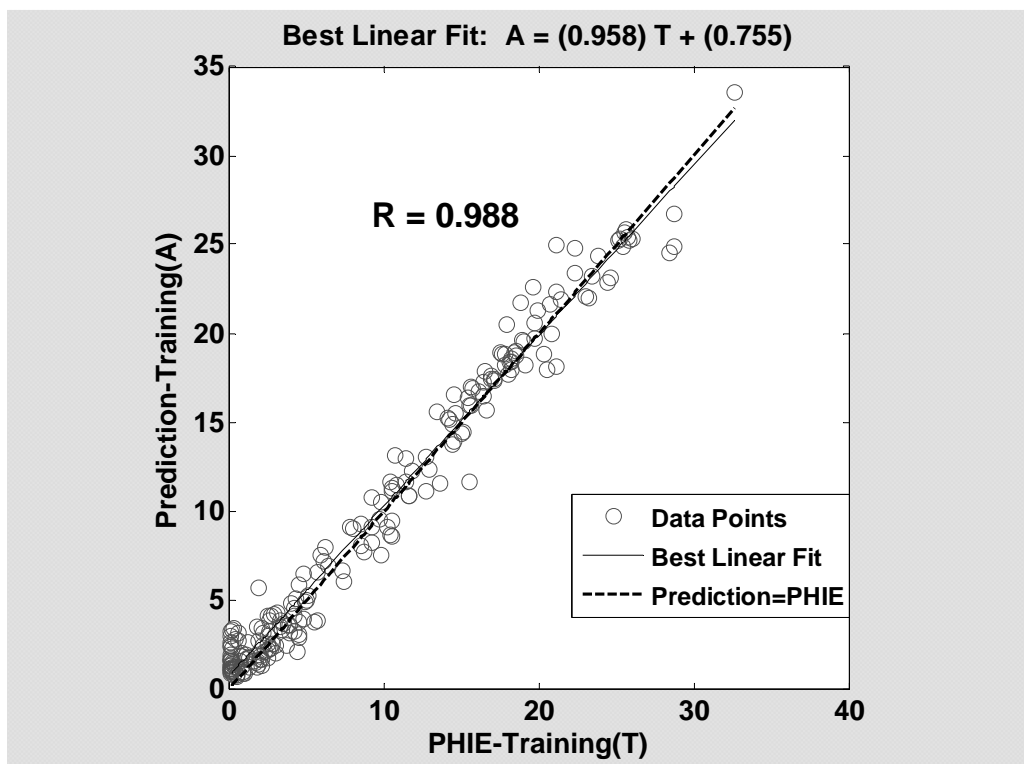
جدول ۱۷-۴: ۷ شبکه منفرد آموزش دیده به روش توقف سریع با بهترین نتایج.

ردیف	ساختار شبکه	MSE (الگوهای آموزش)	ضریب همبستگی (الگوهای آموزش)	MSE (الگوهای آزمون)	ضریب همبستگی (الگوهای آزمون)	تعداد کل پارامترهای شبکه	دوره توقف آموزش
۱	۴-۲-۱۷-۱	۱/۸۱۳	۰/۹۸۸	۲/۲۵۷	۰/۹۸۶	۷۹	۹
۲	۴-۴-۴-۱	۱/۹۲۴	۰/۹۸۹	۲/۳۰۵	۰/۹۸۶	۴۵	۸
۳	۴-۸-۵-۱	۱/۹۹۳	۰/۹۸۷	۲/۳۲۴	۰/۹۸۳	۹۱	۸
۴	۴-۲-۱۹-۱	۲/۳۰۳	۰/۹۸۵	۲/۳۳۵	۰/۹۸۵	۸۷	۸
۵	۴-۵-۱۰-۱	۲/۰۰۱	۰/۹۸۶	۲/۳۴۹	۰/۹۸۵	۹۶	۹
۶	۴-۷-۳-۱	۱/۵۲۰	۰/۹۸۹	۲/۳۶۸	۰/۹۸۸	۶۳	۱۲
۷	۴-۳-۱۴-۱	۱/۸۴۹	۰/۹۸۷	۲/۴۳	۰/۹۸۵	۸۶	۸

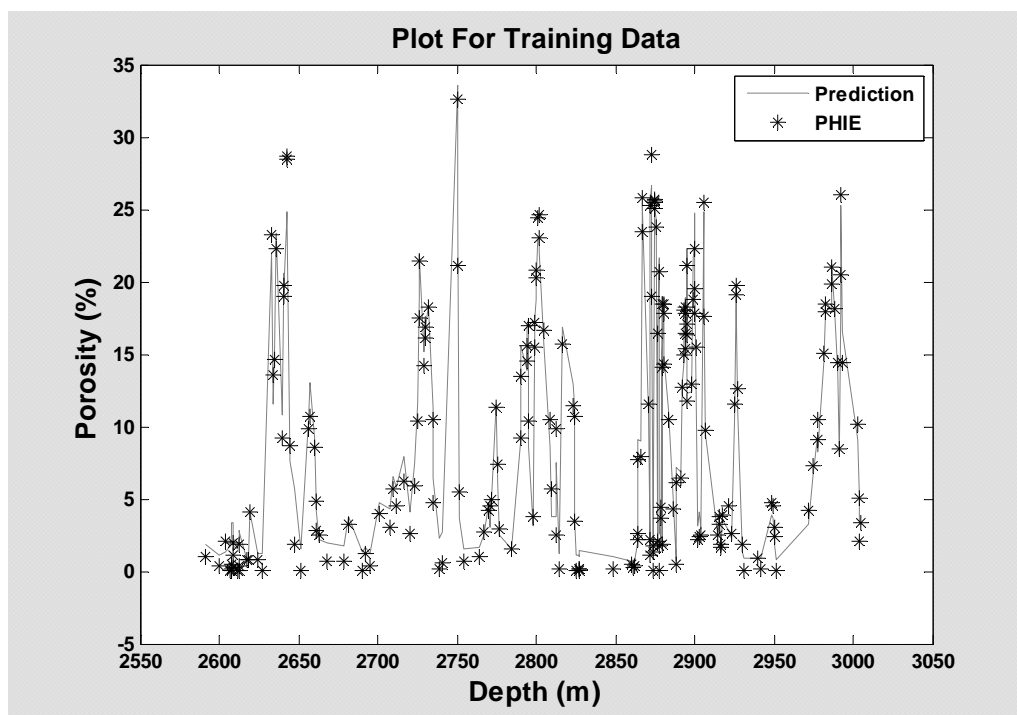
در شکل ۱۶-۴ نتایج بهترین شبکه منفرد آموزش دیده به روش توقف سریع (شبکه شماره ۱ جدول ۱۷-۴) نمایش داده شده است.



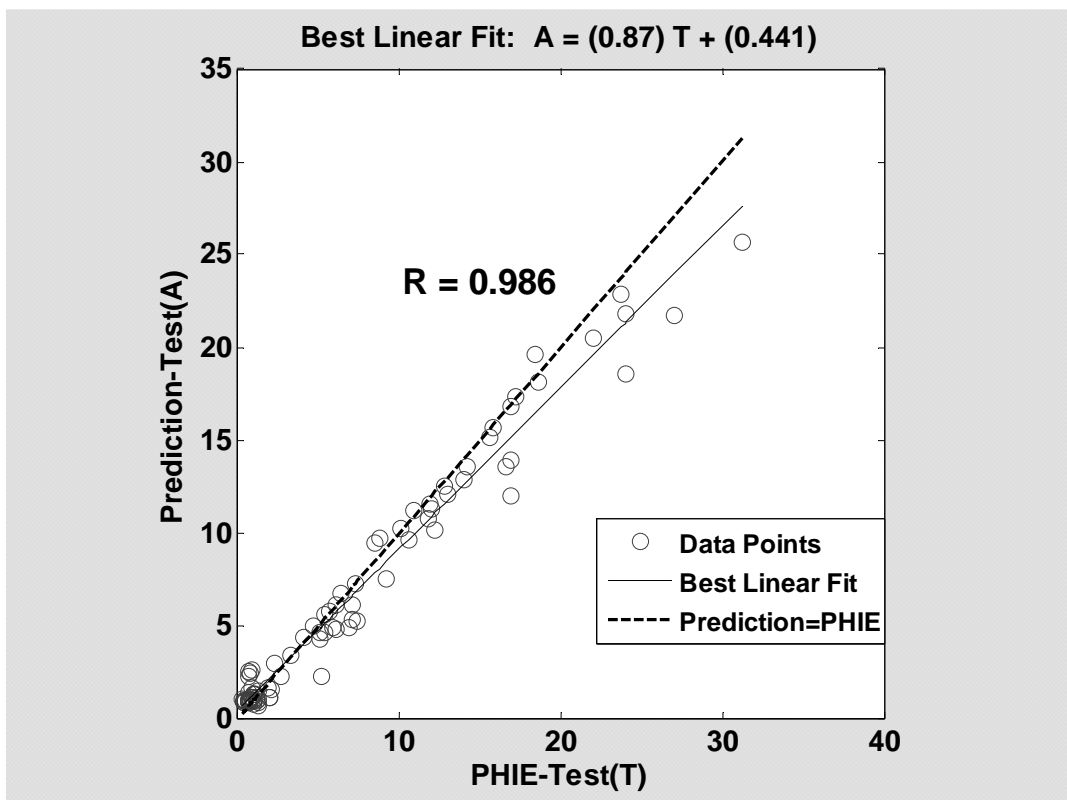
شکل ۱۶-۴ الف: نمایش پیشرفت روند آموزش و کاهش میانگین مربعات خطا برای الگوهای آموزش و اعتبارسنجی در شبکه شماره ۱ جدول ۱۷-۴.



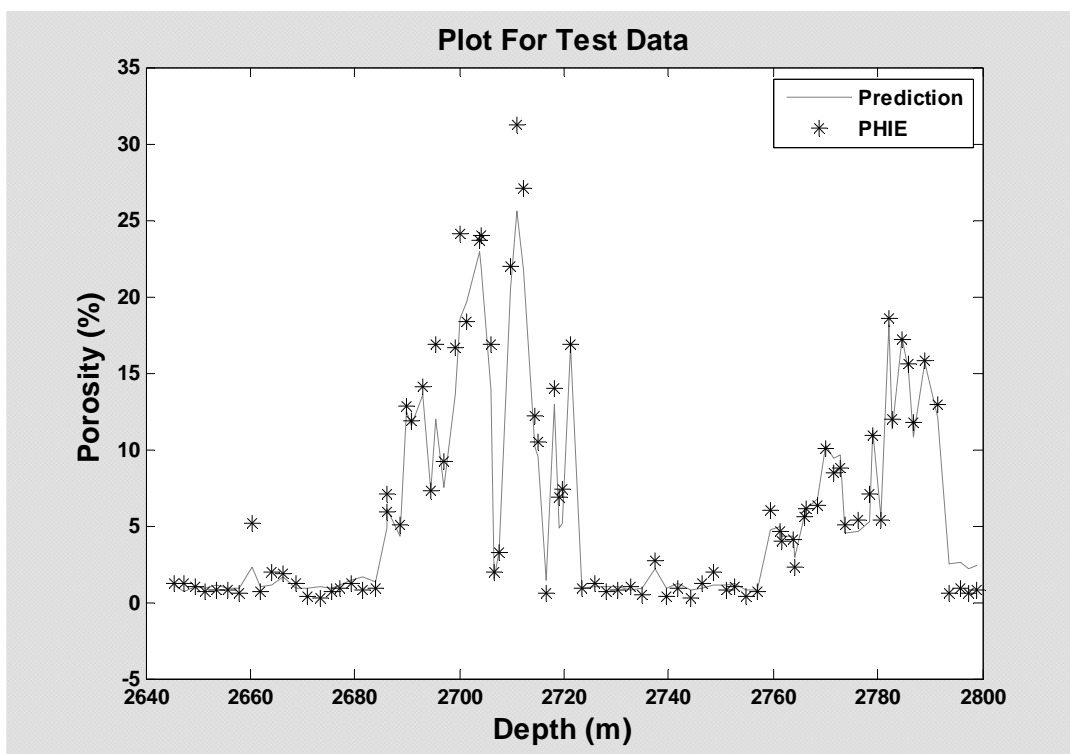
شکل ۴-۱۶ ب: نمودار همبستگی بین تخلخل و تخمین شبکه شماره ۱ جدول ۴-۱۷ برای مجموعه الگوهای آموزش و اعتبارسنجی.



شکل ۴-۱۶ ج: مقادیر تخلخل و تخمین شبکه شماره ۱ جدول ۴-۱۷ برای مجموعه الگوهای آموزش و اعتبارسنجی.



شکل ۴-۱۶-۵: نمودار همبستگی بین تخلخل و تخمین شبکه شماره ۱ جدول ۴-۱۷ برای الگوهای آزمون.



شکل ۴-۱۶-۵: مقادیر تخلخل و تخمین شبکه شماره ۱ جدول ۴-۱۷ برای الگوهای آزمون.

۴-۸-۱- ترکیبات آنسامبلی خطی

۱۲۰ ترکیب آنسامبلی خطی ممکن ۷ شبکه جدول ۴-۱۷ با استفاده از روش میانگین‌گیری ساده و ۴ روش ترکیب خطی بهینه (CM1، CM2، CM3 و CM4) ساخته شد. نتایج این ۱۲۰ ترکیب آنسامبلی به هر ۵ روش در جدول ۴-۱۸ آورده شده است. در این جدول ستون اول از سمت چپ شماره ردیف و ستون دوم شبکه‌هایی از جدول ۴-۱۷ که در ساخت ترکیب آنسامبلی خطی به کار رفته‌اند را نشان می‌دهد. بقیه ستون‌های جدول میانگین مربعات خطای تخمین الگوهای آموزش و آزمون را برای هر یک از ۵ روش ترکیب خطی نشان می‌دهد. به عنوان مثال، در ردیف ۱۱۰ این جدول نتایج ترکیب شبکه‌های شماره ۱، ۲، ۳، ۴ و ۷ جدول ۴-۱۷ آورده شده است.

جدول ۴-۱۸: نتایج ۱۲۰ ترکیب آنسامبلی شبکه‌های عصبی آموزش‌دیده به روش توقف سریع (۷ شبکه جدول ۴-۱۷) به ۵ روش ترکیب خطی (میانگین‌گیری ساده، CM1، CM2، CM3 و CM4).

NO.	Ensemble	Simple averaging		CM1		CM2		CM3		CM4	
		Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)
1	1,2	1.731	2.247	1.507	2.598	1.583	2.707	1.710	2.403	1.725	2.243
2	1,3	1.710	2.238	1.526	2.547	1.567	2.637	1.678	2.412	1.699	2.234
3	1,4	1.891	2.251	1.644	2.618	1.676	2.698	1.778	2.453	1.801	2.247
4	1,5	1.644	2.244	1.514	2.583	1.548	2.640	1.625	2.369	1.635	2.237
5	1,6	1.509	2.232	1.455	2.420	1.464	2.440	1.475	2.292	1.475	2.275
6	1,7	1.633	2.283	1.548	2.539	1.572	2.605	1.626	2.390	1.633	2.279
7	2,3	1.857	2.275	1.603	2.582	1.705	2.685	1.844	2.391	1.854	2.274
8	2,4	1.987	2.258	1.635	2.650	1.764	2.763	1.911	2.377	1.917	2.282
9	2,5	1.790	2.235	1.592	2.591	1.688	2.660	1.785	2.300	1.788	2.234
10	2,6	1.568	2.221	1.478	2.403	1.497	2.407	1.502	2.265	1.502	2.291
11	2,7	1.711	2.260	1.573	2.515	1.647	2.593	1.708	2.310	1.709	2.268
12	3,4	2.039	2.252	1.813	2.540	1.855	2.627	1.964	2.442	1.984	2.287
13	3,5	1.780	2.216	1.659	2.494	1.699	2.550	1.772	2.315	1.780	2.216
14	3,6	1.649	2.215	1.518	2.346	1.518	2.350	1.518	2.347	1.519	2.396
15	3,7	1.819	2.276	1.741	2.473	1.764	2.537	1.803	2.377	1.806	2.308
16	4,5	1.936	2.257	1.822	2.567	1.856	2.597	1.906	2.337	1.909	2.269
17	4,6	1.768	2.252	1.496	2.349	1.497	2.346	1.499	2.397	1.502	2.459
18	4,7	1.962	2.298	1.829	2.516	1.837	2.561	1.848	2.458	1.849	2.429
19	5,6	1.582	2.283	1.496	2.355	1.500	2.366	1.500	2.289	1.501	2.324
20	5,7	1.686	2.308	1.642	2.477	1.661	2.515	1.680	2.338	1.680	2.316
21	6,7	1.612	2.348	1.518	2.347	1.519	2.353	1.518	2.332	1.519	2.379
22	1,2,3	1.717	2.239	1.492	2.562	1.555	2.654	1.674	2.399	1.692	2.231
23	1,2,4	1.822	2.236	1.486	2.633	1.567	2.745	1.703	2.424	1.718	2.257
24	1,2,5	1.658	2.221	1.478	2.572	1.535	2.646	1.623	2.357	1.632	2.229
25	1,2,6	1.552	2.208	1.440	2.444	1.461	2.455	1.474	2.286	1.474	2.269
26	1,2,7	1.635	2.241	1.489	2.541	1.544	2.620	1.616	2.358	1.622	2.255

ادامه جدول ۱۸-۴

NO.	Ensemble	Simple averaging		CM1		CM2		CM3		CM4	
		Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)
27	1,3,4	1.828	2.228	1.485	2.585	1.520	2.682	1.638	2.490	1.667	2.279
28	1,3,5	1.636	2.207	1.479	2.534	1.519	2.606	1.606	2.346	1.617	2.214
29	1,3,6	1.571	2.199	1.455	2.418	1.462	2.430	1.466	2.328	1.466	2.333
30	1,3,7	1.666	2.242	1.520	2.527	1.555	2.602	1.624	2.379	1.632	2.266
31	1,4,5	1.752	2.229	1.504	2.602	1.534	2.660	1.608	2.404	1.620	2.259
32	1,4,6	1.670	2.220	1.405	2.426	1.406	2.434	1.407	2.388	1.407	2.376
33	1,4,7	1.776	2.256	1.515	2.575	1.527	2.637	1.562	2.502	1.569	2.390
34	2,3,4	1.924	2.242	1.588	2.608	1.693	2.717	1.839	2.417	1.850	2.293
35	2,3,5	1.754	2.214	1.571	2.545	1.651	2.605	1.747	2.300	1.752	2.212
36	2,3,6	1.651	2.205	1.466	2.404	1.480	2.407	1.478	2.305	1.481	2.375
37	2,3,7	1.753	2.243	1.573	2.514	1.646	2.592	1.706	2.312	1.707	2.283
38	2,4,5	1.847	2.224	1.587	2.606	1.687	2.669	1.785	2.302	1.788	2.236
39	2,4,6	1.727	2.213	1.420	2.420	1.432	2.427	1.431	2.328	1.433	2.379
40	2,4,7	1.841	2.244	1.527	2.534	1.592	2.633	1.640	2.389	1.640	2.367
41	2,5,6	1.591	2.215	1.471	2.399	1.490	2.400	1.495	2.261	1.495	2.287
42	2,5,7	1.664	2.236	1.547	2.498	1.607	2.550	1.653	2.288	1.654	2.257
43	3,4,5	1.858	2.210	1.655	2.502	1.692	2.559	1.766	2.335	1.774	2.229
44	3,4,6	1.778	2.205	1.496	2.350	1.496	2.349	1.497	2.378	1.498	2.421
45	3,4,7	1.904	2.246	1.737	2.484	1.755	2.550	1.789	2.421	1.793	2.349
46	3,5,6	1.614	2.202	1.494	2.353	1.496	2.359	1.495	2.327	1.496	2.378
47	3,5,7	1.700	2.233	1.621	2.456	1.647	2.505	1.677	2.320	1.678	2.289
48	4,5,6	1.702	2.235	1.461	2.351	1.461	2.353	1.461	2.371	1.463	2.430
49	4,5,7	1.798	2.260	1.627	2.500	1.638	2.537	1.646	2.422	1.646	2.419
50	1,5,6	1.511	2.229	1.438	2.433	1.452	2.451	1.466	2.277	1.466	2.260
51	1,5,7	1.577	2.256	1.484	2.528	1.514	2.584	1.568	2.338	1.572	2.251
52	1,6,7	1.537	2.266	1.453	2.418	1.461	2.434	1.470	2.299	1.470	2.283
53	2,6,7	1.586	2.246	1.475	2.406	1.493	2.407	1.497	2.270	1.497	2.300

ادامه جدول ۱۸-۴

NO.	Ensemble	Simple averaging		CM1		CM2		CM3		CM4	
		Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)
54	3,6,7	1.662	2.248	1.517	2.346	1.518	2.350	1.517	2.346	1.519	2.393
55	4,6,7	1.744	2.273	1.495	2.352	1.495	2.349	1.498	2.399	1.500	2.462
56	5,6,7	1.572	2.290	1.495	2.354	1.499	2.362	1.498	2.296	1.499	2.333
57	4,5,6,7	1.677	2.251	1.457	2.360	1.457	2.362	1.457	2.378	1.459	2.437
58	3,5,6,7	1.611	2.228	1.494	2.352	1.496	2.358	1.495	2.326	1.496	2.376
59	3,4,6,7	1.754	2.228	1.495	2.353	1.495	2.351	1.496	2.382	1.497	2.426
60	3,4,5,7	1.787	2.222	1.591	2.474	1.608	2.523	1.628	2.391	1.628	2.360
61	3,4,5,6	1.711	2.197	1.460	2.353	1.460	2.355	1.460	2.359	1.461	2.403
62	2,5,6,7	1.576	2.232	1.469	2.401	1.487	2.400	1.490	2.265	1.491	2.294
63	2,4,6,7	1.703	2.229	1.416	2.422	1.429	2.433	1.428	2.333	1.430	2.384
64	2,4,5,7	1.758	2.227	1.489	2.517	1.539	2.588	1.570	2.371	1.570	2.364
65	2,4,5,6	1.689	2.207	1.410	2.412	1.421	2.416	1.421	2.322	1.422	2.373
66	2,3,6,7	1.643	2.220	1.466	2.405	1.480	2.406	1.478	2.305	1.481	2.374
67	2,3,5,7	1.690	2.217	1.547	2.497	1.607	2.550	1.651	2.293	1.651	2.276
68	2,3,5,6	1.626	2.193	1.461	2.401	1.475	2.401	1.473	2.301	1.475	2.369
69	1,5,6,7	1.518	2.249	1.436	2.431	1.449	2.445	1.461	2.282	1.461	2.267
70	1,4,6,7	1.658	2.242	1.400	2.438	1.401	2.447	1.403	2.400	1.403	2.386
71	1,4,5,7	1.692	2.239	1.416	2.573	1.429	2.625	1.457	2.477	1.460	2.400
72	1,4,5,6	1.627	2.216	1.372	2.440	1.375	2.449	1.378	2.376	1.378	2.365
73	1,3,6,7	1.586	2.226	1.453	2.419	1.460	2.429	1.465	2.325	1.465	2.328
74	1,3,5,7	1.611	2.221	1.476	2.521	1.511	2.582	1.567	2.345	1.571	2.267
75	1,3,5,6	1.552	2.195	1.437	2.430	1.447	2.438	1.452	2.319	1.452	2.330
76	1,3,4,7	1.769	2.231	1.459	2.563	1.480	2.641	1.536	2.492	1.547	2.358
77	1,3,4,6	1.687	2.200	1.397	2.443	1.400	2.454	1.404	2.378	1.405	2.353
78	1,2,6,7	1.553	2.227	1.437	2.446	1.456	2.452	1.469	2.287	1.469	2.274
79	1,2,5,7	1.600	2.226	1.467	2.531	1.511	2.590	1.568	2.344	1.572	2.255
80	1,2,5,6	1.548	2.205	1.432	2.446	1.451	2.454	1.465	2.286	1.465	2.267

ادامه جدول ۱۸-۴

NO.	Ensemble	Simple averaging		CM1		CM2		CM3		CM4	
		Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)
81	1,2,4,7	1.743	2.233	1.414	2.570	1.458	2.669	1.514	2.461	1.518	2.373
82	1,2,4,6	1.668	2.207	1.366	2.465	1.378	2.479	1.385	2.354	1.385	2.352
83	1,2,3,7	1.668	2.230	1.487	2.539	1.542	2.617	1.616	2.358	1.621	2.264
84	1,2,3,6	1.599	2.200	1.436	2.440	1.453	2.446	1.457	2.310	1.458	2.337
85	2,3,4,7	1.835	2.231	1.524	2.532	1.589	2.631	1.640	2.389	1.640	2.368
86	2,3,4,6	1.749	2.202	1.418	2.418	1.429	2.423	1.426	2.339	1.429	2.407
87	2,3,4,5	1.820	2.209	1.547	2.568	1.628	2.633	1.730	2.331	1.736	2.235
88	1,3,4,5	1.736	2.206	1.414	2.573	1.447	2.649	1.537	2.424	1.555	2.260
89	1,2,4,5	1.739	2.217	1.443	2.611	1.503	2.687	1.596	2.387	1.607	2.250
90	1,2,3,5	1.661	2.209	1.469	2.545	1.519	2.610	1.605	2.353	1.617	2.216
91	1,2,3,4	1.799	2.228	1.435	2.606	1.496	2.713	1.629	2.477	1.653	2.282
92	3,4,5,6,7	1.693	2.216	1.456	2.360	1.457	2.362	1.457	2.370	1.458	2.417
93	2,4,5,6,7	1.665	2.221	1.405	2.416	1.417	2.425	1.416	2.330	1.417	2.381
94	2,3,5,6,7	1.614	2.209	1.461	2.401	1.475	2.400	1.473	2.301	1.475	2.367
95	2,3,4,6,7	1.725	2.213	1.413	2.420	1.423	2.430	1.420	2.350	1.423	2.423
96	2,3,4,5,7	1.762	2.213	1.487	2.514	1.537	2.584	1.570	2.371	1.570	2.364
97	2,3,4,5,6	1.704	2.193	1.409	2.411	1.419	2.413	1.417	2.333	1.419	2.397
98	1,4,5,6,7	1.617	2.232	1.364	2.460	1.367	2.472	1.370	2.396	1.370	2.383
99	1,3,5,6,7	1.558	2.216	1.436	2.429	1.447	2.437	1.452	2.316	1.452	2.325
100	1,3,4,6,7	1.677	2.217	1.395	2.448	1.397	2.460	1.401	2.391	1.401	2.368
101	1,3,4,5,7	1.701	2.216	1.389	2.559	1.408	2.622	1.449	2.465	1.454	2.374
102	1,3,4,5,6	1.645	2.194	1.365	2.458	1.370	2.468	1.377	2.369	1.377	2.349
103	1,2,5,6,7	1.542	2.220	1.430	2.446	1.448	2.450	1.460	2.286	1.461	2.270
104	1,2,4,6,7	1.650	2.220	1.360	2.474	1.373	2.493	1.380	2.367	1.380	2.363
105	1,2,4,5,7	1.689	2.221	1.378	2.565	1.410	2.639	1.448	2.455	1.451	2.384
106	1,2,4,5,6	1.637	2.203	1.352	2.465	1.363	2.475	1.370	2.354	1.370	2.349
107	1,2,3,6,7	1.596	2.213	1.435	2.441	1.452	2.446	1.457	2.309	1.457	2.333

ادامه جدول ۱۸-۴

NO.	Ensemble	Simple averaging		CM1		CM2		CM3		CM4	
		Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)	Mse(Tr.)	Mse(Test.)
108	1,2,3,5,7	1.628	2.213	1.465	2.529	1.510	2.587	1.567	2.346	1.571	2.267
109	1,2,3,5,6	1.581	2.193	1.428	2.442	1.445	2.445	1.450	2.309	1.450	2.331
110	1,2,3,4,7	1.748	2.224	1.404	2.570	1.447	2.667	1.511	2.465	1.517	2.366
111	1,2,3,4,6	1.687	2.200	1.365	2.469	1.378	2.482	1.384	2.356	1.384	2.360
112	1,2,3,4,5	1.734	2.207	1.398	2.586	1.443	2.661	1.537	2.426	1.555	2.260
113	1,2,3,4,5,6	1.654	2.193	1.350	2.472	1.363	2.480	1.370	2.355	1.370	2.352
114	1,2,3,4,5,7	1.699	2.211	1.369	2.562	1.401	2.635	1.446	2.455	1.450	2.374
115	1,2,3,4,6,7	1.671	2.210	1.360	2.474	1.373	2.492	1.378	2.371	1.378	2.382
116	1,2,3,5,6,7	1.575	2.205	1.428	2.442	1.444	2.444	1.450	2.308	1.450	2.328
117	1,2,4,5,6,7	1.622	2.215	1.343	2.479	1.355	2.495	1.362	2.374	1.362	2.367
118	1,3,4,5,6,7	1.638	2.210	1.360	2.469	1.365	2.482	1.370	2.392	1.370	2.375
119	2,3,4,5,6,7	1.684	2.205	1.402	2.415	1.412	2.423	1.409	2.347	1.412	2.417
120	1,2,3,4,5,6,7	1.641	2.203	1.343	2.481	1.355	2.496	1.361	2.376	1.361	2.380

نتایج بهترین آنسامبل‌های حاصل از هر یک از این ۵ روش که کمترین میانگین مربعات خطای تخمین الگوهای آزمون را داشته‌اند، در جدول ۴-۱۹ آمده است. ردیف‌های این جدول مشابه جداول ۴-۷ و ۴-۱۳ است و مقایسه نسبت به بهترین شبکه منفرد یعنی شبکه ۱ جدول ۴-۱۷ انجام شده است. اعداد این جدول سه رقم اعشار دقت دارند.

جدول ۴-۱۹: نتایج بهترین آنسامبل‌های حاصل از ۵ روش ترکیب خطی ۷ شبکه جدول ۴-۱۷.

روش ترکیب خطی نتایج	میانگین‌گیری ساده	ترکیب خطی بهینه نامقید با جمله ثابت (CM1)	ترکیب خطی بهینه مقید با جمله ثابت (CM2)	ترکیب خطی بهینه نامقید بدون جمله ثابت (CM3)	ترکیب خطی بهینه مقید بدون جمله ثابت (CM4)
ترکیب آنسامبلی	۱، ۲، ۳، ۵ و ۶	۳، ۶ و ۷	۴ و ۶	۲، ۵ و ۶	۲، ۳ و ۵
ضریب همبستگی (الگوهای آموزش)	۰/۹۹۰	۰/۹۸۹	۰/۹۹۰	۰/۹۹۰	۰/۹۸۹
MSE (الگوهای آموزش)	۱/۵۸۱	۱/۵۱۷	۱/۴۹۷	۱/۴۹۵	۱/۷۵۲
کاهش در MSE الگوهای آموزش.	٪ ۱۲/۸	٪ ۱۶/۳	٪ ۱۷/۴	٪ ۱۷/۵	٪ ۳/۴
ضریب همبستگی (الگوهای آزمون)	۰/۹۸۶	۰/۹۸۸	۰/۹۸۸	۰/۹۸۸	۰/۹۸۵
MSE (الگوهای آزمون)	۲/۱۹۳	۲/۳۴۶	۲/۳۴۶	۲/۲۶۱	۲/۲۱۲
کاهش در MSE الگوهای آزمون.	٪ ۲/۸	---	---	---	٪ ۲
α_0	۰	۰/۰۰۰	۰/۰۷۷	۰	۰
α_1	۰/۲	۰	۰	۰	۰
α_2	۰/۲	۰	۰	۰/۱۱۱	۰/۳۰۱
α_3	۰/۲	-۰/۰۴۳	۰	۰	۰/۳۰۵
α_4	۰	۰	-۰/۲۲۲	۰	۰
α_5	۰/۲	۰	۰	۰/۱۱۴	۰/۳۹۴
α_6	۰/۲	۱/۰۹۱۵	۱/۲۲۲	۰/۷۷۷	۰
α_7	۰	-۰/۰۴۶	۰	۰	۰
$\sum \alpha_i$	۱	۱/۰۰۳	۱	۱/۰۰۲	۱

با توجه به جدول ۴-۱۹ مشخص است که از ۴ روش ترکیب خطی بهینه، سه روش CM1، CM2 و CM3 اگر چه نسبت به بهترین شبکه منفرد به خوبی میانگین مربعات خطای تخمین الگوهای آموزش را کاهش داده‌اند، اما در مورد الگوهای آزمون کاهشی در MSE ایجاد نکرده‌اند و میانگین مربعات خطای تخمین بزرگتری دارند که این در نتیجه اثرات همبستگی خطی مضر است. فقط

ترکیب خطی بهینه به روش CM4 توانسته است MSE تخمین الگوهای آزمون را به میزان ۲٪ کاهش دهد. همچنین میانگین مربعات خطای تخمین الگوهای آزمون با استفاده از میانگین‌گیری ساده ۲/۸٪ کاهش یافته است.

۴-۸-۲- ترکیبات آنسامبلی غیرخطی با استفاده از شبکه‌های عصبی مصنوعی منفرد

در این قسمت نیز مانند بخش ۴-۶-۲ و ۴-۷-۲ به منظور ترکیب غیرخطی ۷ شبکه منفرد جدول ۴-۱۷ از خروجی این ۷ شبکه به عنوان ورودی یک شبکه عصبی منفرد استفاده شد. برای آموزش شبکه عصبی منفرد ترکیب‌کننده از هر سه روش آموزش فرآیند آموزش، مرتب‌سازی و توقف سریع استفاده شد. در هر روش آموزش، از شبکه‌هایی با یک لایه میانی که تابع انتقال نرون‌های آن تانژانت هیپربولیک و تابع انتقال لایه خروجی تابع خطی است استفاده گردید. تابع پیش‌پردازش Prestd و دوره‌های تکرار ۴۰۰ مرتبه انتخاب شد. یک مرحله طولانی آزمون و خطا با تعداد نرون‌های مختلف در لایه میانی و شرایط اولیه گوناگون انجام شد و در نهایت بهترین نتایج انتخاب گردید. جدول ۴-۲۰ بهترین نتیجه حاصل از هر یک از سه روش آموزش پس‌انتشار خطا را نشان می‌دهد.

جدول ۴-۲۰: بهترین نتایج حاصل از ترکیب غیرخطی خروجی‌های ۷ شبکه جدول ۴-۱۷ با استفاده از شبکه عصبی مصنوعی ترکیب‌کننده آموزش‌دیده به سه روش فرآیند آموزش، مرتب‌سازی و توقف سریع.

کاهش در MSE الگوهای آزمون.	MSE (الگوهای آزمون)	ضریب همبستگی (الگوهای آزمون)	کاهش در MSE الگوهای آموزش.	MSE (الگوهای آموزش)	ضریب همبستگی (الگوهای آموزش)	ساختار شبکه	نتایج روش آموزش شبکه ترکیب‌کننده
۲/۳٪	۲/۲۰۶	۰/۹۸۸	۵۲/۵٪	۰/۸۶۲	۰/۹۹۴	۷-۵-۱	فرآیند آموزش
---	۲/۴۶۳	۰/۹۸۶	۳۱/۴٪	۱/۲۴۴	۰/۹۹۱	۷-۱۰-۱	مرتب‌سازی
---	۲/۴۰۵	۰/۹۸۵	۱/۹٪	۱/۷۷۸	۰/۹۸۸	۷-۶-۱	توقف سریع

با توجه به جدول فوق مشاهده می‌شود که ترکیب تخمین ۷ شبکه منفرد جدول ۴-۱۷ با استفاده از یک شبکه منفرد با ساختار ۷-۵-۱ که با روش فرآیند آموزش آموزش دیده، توانسته است میانگین مربعات خطای تخمین الگوهای آموزش و آزمون را به ترتیب ۵۲/۵٪ و ۲/۳٪ کاهش دهد. همچنین شبکه‌های عصبی ترکیب‌کننده که با دو روش مرتب‌سازی و توقف سریع آموزش دیده‌اند، اگر چه توانسته‌اند MSE تخمین الگوهای آموزش را کاهش دهند اما کاهش در MSE تخمین الگوهای

آزمون ایجاد نکرده‌اند و میانگین مربعات خطای تخمین آن‌ها برای الگوهای آزمون از شبکه شماره ۱ جدول ۴-۱۷ بزرگ‌تر است.

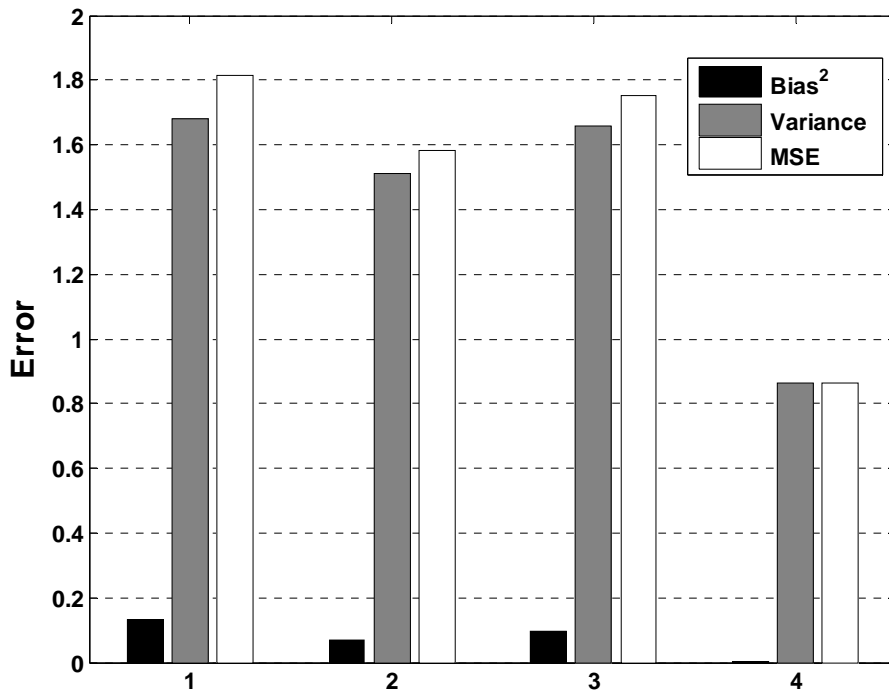
۴-۸-۳- واریانس و بایاس خطا

در جدول ۴-۲۱ واریانس، بایاس و میانگین مربعات خطای تخمین الگوهای آموزش و آزمون برای بهترین شبکه منفرد یعنی شبکه شماره ۱ جدول ۴-۱۷ (۱) و بهترین نتایج ترکیب شبکه‌های جدول ۴-۱۷ یعنی ترکیب شبکه‌های ۱، ۲، ۳، ۵ و ۶ به روش میانگین‌گیری ساده (۲)، ترکیب خطی بهینه شبکه‌های ۲، ۳ و ۵ به روش مقید بدون جمله ثابت (۳) و ترکیب غیرخطی خروجی‌های ۷ شبکه با استفاده از شبکه عصبی پس‌انتشار خطا آموزش‌دیده به روش فراآموزش (۴) آمده است.

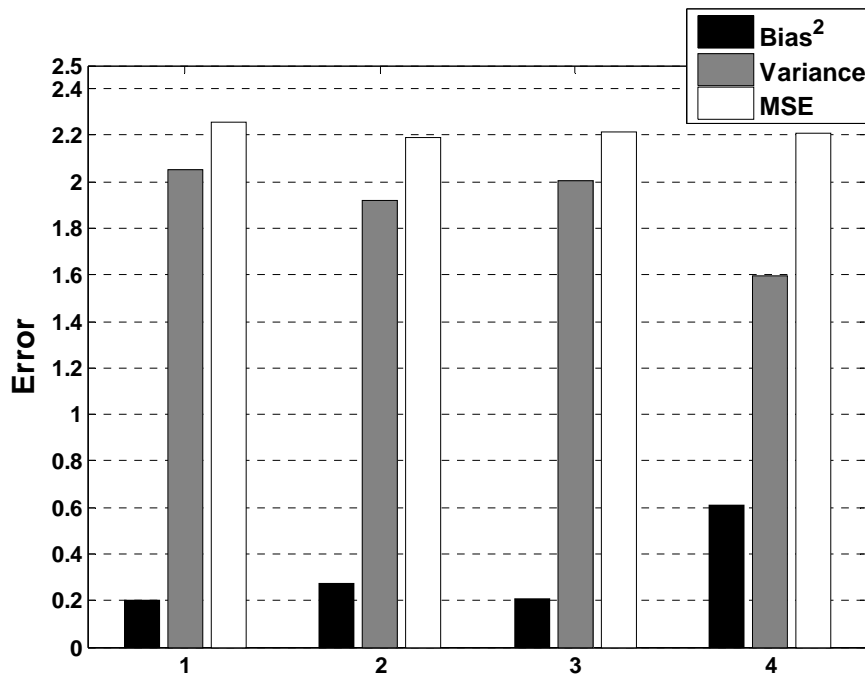
جدول ۴-۲۱: واریانس، بایاس و میانگین مربعات خطای تخمین شبکه شماره ۱ جدول ۴-۱۷ و بهترین نتایج حاصل از ترکیب شبکه‌های آن جدول.

نتایج ردیف	بایاس خطای تخمین (الگوهای آموزش)	واریانس خطای تخمین (الگوهای آموزش)	میانگین مربعات خطای تخمین (الگو- های آموزش)	بایاس خطای تخمین (الگوهای آموزش)	واریانس خطای تخمین (الگوهای آموزش)	میانگین مربعات خطای تخمین (الگو- های آموزش)
۱	۰/۳۶۱	۱/۶۸۳	۱/۸۱۳	-۰/۴۵۰	۲/۰۵۵	۲/۲۵۷
۲	۰/۲۶۲	۱/۵۱۲	۱/۵۸۱	-۰/۵۲۱	۱/۹۲۱	۲/۱۹۳
۳	۰/۳۰۷	۱/۶۵۷	۱/۷۵۲	-۰/۴۵۴	۲/۰۰۶	۲/۲۱۲
۴	-5×10^{-5}	۰/۸۶۲	۰/۸۶۲	-۰/۷۸۰	۱/۵۹۷	۲/۲۰۶

شکل‌های ۴-۱۷ نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطای تخمین الگوهای آموزش و آزمون برای ۴ حالت مذکور در جدول ۴-۲۱ را نشان می‌دهد. مطابق این دو شکل، در این ۴ حالت، هم برای الگوهای آموزش و هم برای الگوهای آزمون، واریانس مؤلفه اصلی خطا است.



شکل ۴-۱۷-الف: نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطا در بهترین حالات برای الگوهای آموزش (۱: شبکه شماره ۱ جدول ۴-۱۷، ۲: بهترین ترکیب شبکه‌های جدول ۴-۱۷ به روش میانگین‌گیری ساده، ۳: بهترین ترکیب خطی بهینه شبکه‌های جدول ۴-۱۷، ۴: بهترین ترکیب ۷ شبکه جدول ۴-۱۷ با استفاده از شبکه عصبی).



شکل ۴-۱۷-ب: نمودار میله‌ای مجذور بایاس، واریانس و میانگین مربعات خطا در بهترین حالات برای الگوهای آموزش (۱: شبکه شماره ۱ جدول ۴-۱۷، ۲: بهترین ترکیب شبکه‌های جدول ۴-۱۷ به روش میانگین‌گیری ساده، ۳: بهترین ترکیب خطی بهینه شبکه‌های جدول ۴-۱۷، ۴: بهترین ترکیب ۷ شبکه جدول ۴-۱۷ با استفاده از شبکه عصبی).

فصل پنجم:

نتایج و پیشنهادات

۵-۱- جمع‌بندی و نتیجه‌گیری

برای کاهش خطا و بهبود نتایج روش شبکه‌های عصبی مصنوعی منفرد، به جای استفاده از بهترین شبکه حاصل از فرایند آزمون و خطا (NN_b)، ترکیب نتایج چندین شبکه عصبی مصنوعی پیشنهاد شده است. ترکیب نتایج شبکه‌های عصبی به دو صورت می‌تواند انجام شود: (۱) ترکیب آنسامبلی که در آن نتایج چندین شبکه که همگی به یک منظور طراحی شده‌اند و یک کار را انجام می‌دهند با روشی مناسب ترکیب می‌شود. (۲) ترکیب مدوله‌ای که در آن مسأله به چندین زیرمسأله تفکیک می‌شود و برای حل هر یک از آن‌ها شبکه‌ای مناسب طراحی می‌شود. حل کامل مسأله مستلزم حل هر یک از زیرمسأله‌ها و ترکیب مناسب آن‌هاست.

در این تحقیق از ترکیب آنسامبلی شبکه‌های عصبی مصنوعی منفرد به منظور تخمین تداخل سنگ مخزن بهره گرفته شد. نتایج نشان می‌دهد که در تخمین تداخل، ترکیب آنسامبلی توانسته است نتایج شبکه‌های عصبی منفرد را بهبود دهد.

در ترکیب ۷ شبکه منفرد آموزش دیده به روش فراآموزش که کمترین میانگین مربعات خطای تخمین را برای الگوهای آزمون داشتند، بهترین ترکیب هر روش یعنی بهترین ترکیب با استفاده از میانگین‌گیری ساده، بهترین ترکیب خطی بهینه و بهترین ترکیب غیرخطی با استفاده از یک شبکه سه‌لایه، MSE تخمین الگوهای آزمون را نسبت به بهترین شبکه منفرد به ترتیب ۳/۹٪، ۴/۹٪ و ۱۹٪ کاهش دادند. در ترکیب ۷ شبکه منفرد آموزش دیده به روش مرتب‌سازی که بهترین نتایج را نسبت به شبکه‌های منفرد دیگر داشتند (کمترین MSE تخمین را برای الگوهای آزمون داشتند)، بهترین ترکیب با استفاده از میانگین‌گیری ساده ۳/۴٪، بهترین ترکیب خطی بهینه ۴٪ و بهترین ترکیب غیرخطی با استفاده از شبکه عصبی سه‌لایه ۲۱/۶٪ میانگین مربعات خطای تخمین الگوهای آزمون را نسبت به بهترین شبکه منفرد کاهش دادند. در ترکیب ۷ شبکه منفردی که با روش

توقف سریع آموزش دیده‌اند نیز میانگین مربعات خطای تخمین الگوهای آزمون نسبت به بهترین شبکه منفرد با استفاده از بهترین ترکیب به روش میانگین‌گیری ساده $2/8\%$ ، با استفاده از بهترین ترکیب خطی بهینه 2% و با استفاده از بهترین ترکیب غیرخطی به وسیله شبکه عصبی سه‌لایه $2/3\%$ کاهش یافت. در نمودار شکل ۵-۱ میانگین مربعات خطای تخمین الگوهای آزمون بهترین شبکه عصبی منفرد و بهترین نتایج ترکیب با استفاده از میانگین‌گیری ساده، ترکیب خطی بهینه و ترکیب غیرخطی با استفاده از شبکه عصبی منفرد با یک لایه میانی برای سه بخش مجزای ترکیب شبکه‌های آموزش دیده به روش فراآموزش، مرتب‌سازی و توقف سریع آمده است. اعداد انتهایی این نمودار که به صورت درصد نوشته شده‌اند، درصد کاهش میانگین مربعات خطای تخمین الگوهای آزمون در ترکیب‌ها نسبت به بهترین شبکه منفرد هر بخش را نشان می‌دهند.

با توجه به این تحقیق، می‌توان گفت که به طور کلی میزان توانایی ترکیب آنسامبلی در بهبود نتایج شبکه‌های عصبی مصنوعی منفرد و افزایش دقت و کاهش خطا در حل مسائل تابع عواملی به شرح زیر است:

اولین عامل مؤثر نوع و ماهیت مسأله است، به این مفهوم که چه نوع کمیاتی ورودی شبکه‌هاست و کمیات خروجی چه ماهیتی دارند؟ شبکه‌های منفردی که از آن‌ها به عنوان اجزاء آنسامبل استفاده می‌شود نیز در کارایی ترکیب آنسامبلی تأثیر زیادی دارند به طوری که ممکن است در یک مسأله، آنسامبلی نتایج را به خوبی بهبود دهد در حالی که آنسامبل دیگری نه تنها بهبودی در نتایج ایجاد نکند، بلکه نسبت به شبکه‌های جزء نتایجی با خطای بیشتر به دست دهد. از این منظر، مجموعه آموزش، ساختار، پارامترهای اولیه و روش آموزش شبکه‌های جزء اهمیت پیدا می‌کند. روش مناسب برای انتخاب ترکیبات آنسامبلی مناسب آن است که دقت و توانایی تخمین هر ترکیب آنسامبلی با استفاده از مجموعه‌ای از الگوها تحت عنوان الگوهای آزمون که از مجموعه الگوهای ترکیب (الگوهای آموزش) مجزا است آزموده شود. همان طور که در فرآیند سعی و خطا برای انتخاب یک شبکه منفرد مناسب نمی‌توانیم ادعا کنیم که بهترین شبکه حاصل، بهترین شبکه ممکن است در مورد ترکیبات آنسامبلی نیز نمی‌توانیم ادعا کنیم که بهترین ترکیب حاصل، بهترین ترکیب ممکن است. ترکیبات زیادی را می‌توان ساخت و مورد آزمون قرار داد. با افزایش شبکه‌های جزء، تعداد ترکیبات ممکن به طور مضاعفی افزایش می‌یابد، مثلاً چنانچه از ۸ شبکه برای ساخت ترکیبات آنسامبلی استفاده شود،

تعداد کل ترکیبات آنسامبلی ممکن این ۸ شبکه ۲۴۷ ترکیب خواهد بود، یعنی بیش از دو برابر ترکیبات آنسامبلی که با ۷ شبکه می‌توان ساخت.

عامل سومی که می‌توانیم در این جا به آن اشاره کنیم، روش ترکیبی است که از آن برای ساخت ترکیبات بهره می‌گیریم. چنان که دیدیم ترکیب با روش‌های میانگین‌گیری ساده، روش‌های مختلف ترکیب خطی بهینه و ترکیب غیرخطی با استفاده از شبکه عصبی منفرد ترکیب‌کننده، نتایج متفاوتی دارند.

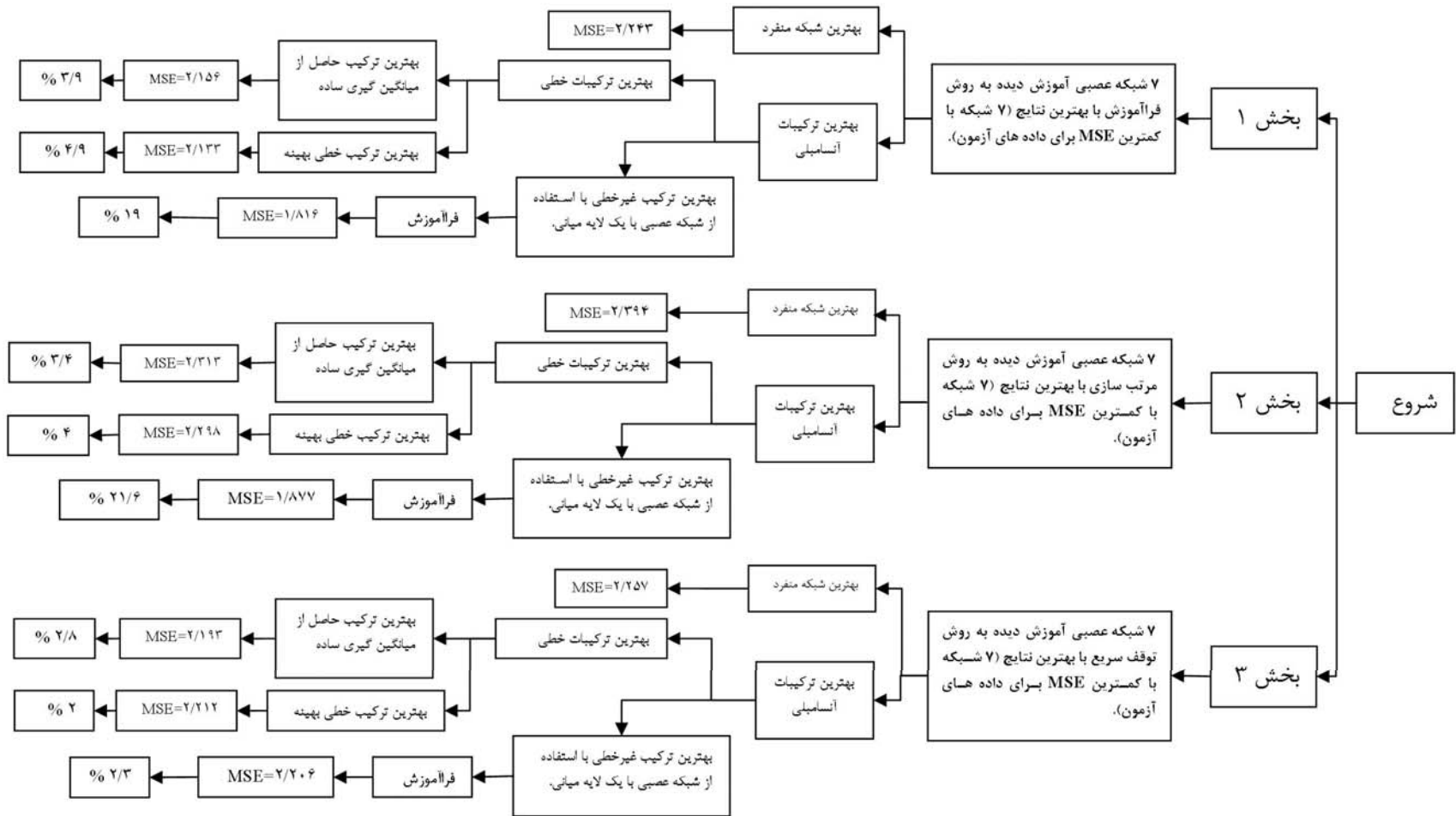
مشکلی که برای ساخت ترکیبات آنسامبلی مناسب به روش خطی وجود دارد آن است که انتخاب مناسب شبکه‌های منفرد جزء آنسامبل از بین تعداد زیادی شبکه منفرد مشکل است. همچنین در ترکیب نتایج شبکه‌ها با استفاده از یک شبکه عصبی ترکیب‌کننده دیگر، از آزمون و خطا استفاده می‌شود که وقت‌گیر است و برای طراحی بهینه شبکه ترکیب‌کننده روش دیگری وجود ندارد.

۵-۲- پیشنهادات

محدودیت مطالعات صورت گرفته درباره ماشین کمیته‌ای باعث شده است که نقاط قوت و ضعف آن‌ها و میزان توانایی‌شان در بهبود نتایج شبکه‌های منفرد مبهم باشد. اگر چه با مطالعه حاضر به میزان اندکی قابلیت این روش‌ها معرفی شد ولی با استفاده از آن نمی‌توان به زوایای مختلف عملکرد آن‌ها پی برد.

به نظر می‌رسد که توانایی ترکیب آنسامبلی بیش از آن است که در تخمین تخلخل دیده شد، لذا پیشنهاد می‌شود که تخمین تخلخل و سایر خصوصیات مخزنی از جمله تراوایی، اشباع سیالات مخزنی و و همچنین تخمین نگارهای چاه در میادین نفتی گوناگون، با استفاده از روش ترکیب آنسامبلی انجام گیرد.

در این مطالعه، به منظور تعیین ضرایب ترکیب خطی بهینه شبکه‌های جزء از روش‌های تحلیلی استفاده شد. روش الگوریتم ژنتیک نیز به عنوان یک روش هوش مصنوعی که بیشترین کاربرد را در بین روش‌های بهینه‌سازی دارد می‌تواند به این منظور استفاده شود و در نهایت نتایج این دو روش مورد مقایسه قرار گیرد.



شکل ۵-۱: میانگین مربعات خطای تخمین الگوهای آزمون برای بهترین شبکه عصبی منفرد حاصل و بهترین نتایج حاصل از ترکیب برای شبکه‌های آموزش دیده به سه روش فرآیند آموزش، مرتب‌سازی و توقف سریع.

تفکیک مسأله و حل هر قسمت آن توسط یک شبکه و یا ترکیبی از شبکه‌ها و در نهایت ترکیب مدوله‌ای حل‌ها می‌تواند به بهبود نتایج کمک کند. در این راستا، می‌توان دامنه تغییرات تخلخل را به چندین زیردامنه تقسیم کرد و برای تخمین در هر زیردامنه، از شبکه و یا ترکیب شبکه‌ها استفاده نمود. انتظار می‌رود با این کار دقت تخمین، مخصوصاً برای تخلخل‌های بالا که فراوانی کمتری دارند، افزایش یابد. به طور کلی ترکیب مدوله‌ای در مواردی که دامنه تغییرات کمیت‌ها به خصوص برای خروجی وسیع است می‌تواند دقت تخمین‌ها را بالا ببرد. بنابراین احتمال زیادی وجود دارد که تخمین کمیتی مانند تراوایی با استفاده از ماشین کمیت‌های نتایج بهتری بدهد.

موارد فوق نمونه‌هایی محدود از موضوعات فراوانی است که می‌تواند برای تحقیقات آینده در زمینه کاربرد ماشین کمیت‌های در مسائل مربوط به مخازن مورد توجه قرار گیرد.

فهرست منابع و مآخذ

منابع فارسی

۱. افشارحرب، ع.، ۱۳۸۱، زمین‌شناسی نفت، انتشارات دانشگاه پیام نور.
۲. انصاری، ع.، ۱۳۷۶، چاه‌پیمایی در هیدروژئولوژی، (مؤلف: دومینیکو کاپلییر)، انتشارات دانشگاه یزد.
۳. بولی بهمنی، ع.، ۱۳۸۲، اصول و چگونگی استخراج پارامترهای پتروفیزیکی سنگ مخزن نفت با استفاده از نگارهای چاه‌پیمایی، سمینار کارشناسی ارشد مهندسی اکتشاف معدن، دانشگاه صنعتی شاهرود.
۴. تدینی، م.، ربانی، ا.ر. و نبی‌بیدهندی، م.، ۱۳۸۳، تعیین نفوذپذیری سنگ مخزن میدان پارسی با استفاده از شبکه عصبی مصنوعی، کنفرانس مهندسی معدن ایران، دانشگاه تربیت مدرس.
۵. جلالی لیجائی، م.، نبی‌بیدهندی، م. و میرزایی، س.، ۱۳۸۳، تخمین تخلخل میدان نفتی گچساران با استفاده از تکنیک شبکه‌های عصبی مصنوعی و رگرسیون خطی چندگانه، کنفرانس مهندسی معدن ایران، دانشگاه تربیت مدرس.
۶. حبیبیان، ب.، نبی‌بیدهندی، م. و کاظم‌زاده، ع.، ۱۳۸۴، پیش‌بینی نفوذپذیری از روی داده‌های چاه‌نگاری با استفاده از شبکه‌های عصبی مصنوعی در یکی از مخازن کربناته جنوب ایران، مجله فیزیک زمین و فضا، جلد ۳۱، شماره ۱، ۸۶-۷۹.
۷. حسنی پاک، ع.ا. و شرف‌الدین، م.، ۱۳۸۰، تحلیل داده‌های اکتشافی، انتشارات دانشگاه تهران.
۸. خاکبازان، ا.ح.، فتحیان‌پور، ن. و باقری، ع.م.، ۱۳۸۳، تخمین نمودارهای خام چاه‌نگاری با استفاده از روش شبکه عصبی در چاه‌های قدیمی نفت، کنفرانس مهندسی معدن ایران، دانشگاه تربیت مدرس.
۹. خسروتهرانی، خ.، ۱۳۸۳، زمین‌شناسی ایران، انتشارات دانشگاه پیام نور، چاپ پنجم.
۱۰. ذاکری، م.، رحیمی، م.، کامکار روحانی، ا. و پویان، ع.ا.، ۱۳۸۶، تخمین تخلخل سنگ مخزن گازی کنگان در میدان پارس جنوبی با استفاده از ترکیب آنسامبلی شبکه‌های عصبی مصنوعی، مجموعه مقالات دومین کنگره ملی مهندسی نفت ایران، جلد دوم، ۴۳-۳۳.
۱۱. رحیمی، م.، ۱۳۸۱، تعبیر و تفسیر لرزه‌ای ناحیه پارس جنوبی در افق‌های آسماری، جهرم، سروک، داریان، فهلیان، سورمه، کنگان، نار و فراقون، مدیریت اکتشاف شرکت ملی نفت ایران.

۱۲. رضایی، م.ر.، ۱۳۸۰، زمین‌شناسی نفت، انتشارات علوی.
۱۳. روکی، ر.، ۱۳۸۶، تخمین نفوذپذیری از روی نگارهای چاه با استفاده از شبکه عصبی مصنوعی در یکی از میادین نفتی خلیج فارس، پایان‌نامه کارشناسی ارشد مهندسی اکتشاف معدن، دانشگاه صنعتی شاهرود.
۱۴. زارعی، س.، موحد، ب.، باقری، ع.م. و مردانی، ع.، ۱۳۸۶، ارزیابی تراوایی سازند کنگان با استفاده از داده‌های نگار CMR و مغزه در میدان پارس جنوبی، بیست و ششمین گردهمایی علوم زمین-شناسی.
۱۵. کامکار روحانی، ا.، ذاکری، م.، پویان، ع.ا. و رحیمی، م.، ۱۳۸۷، تخمین نگار نوترون با استفاده از ترکیب آنسامبلی شبکه‌های عصبی مصنوعی، مجموعه مقالات سیزدهمین کنفرانس ژئوفیزیک ایران، ۲۹۴-۲۹۸.
۱۶. مرادزاده، ع. و قوامی ریایی، ر.، ۱۳۸۰، چاه‌پیمایی برای مهندسین، انتشارات دانشگاه صنعتی شاهرود.
۱۷. منهای، م.ب.، ۱۳۷۹، مبانی شبکه‌های عصبی (جلد اول)، انتشارات دانشگاه صنعتی امیرکبیر.
۱۸. منهای، م.ب. و سیفی‌پور، ن.، ۱۳۷۷، کاربرد هوش محاسباتی در کنترل، انتشارات مرکز نشر پرفسور حسابی.
۱۹. نوروزی، غ.، ۱۳۷۸، چاه‌نگاری، (مؤلف: ابرتوسرا)، انتشارات دانشگاه صنعتی اصفهان.

منابع انگلیسی

20. Bhatt A., 2002, Reservoir properties from well logs using neural networks, PhD thesis, Department of Petroleum Engineering and Applied Geophysics, Norwegian University of Science and Technology.
21. Bhatt A. and Helle H.B., 2002, Committee neural networks for porosity and permeability prediction from well logs, *Geophysical Prospecting*, **50**, 645-660.
22. Chen Ch.H. and Lin Z.Sh., 2006, A committee machine with empirical formulas for permeability prediction, *Computer & Geosciences*, **32**, 485-496.
23. Demuth H. and Beale M., 2002, Neural network toolbox User's guide of MATLAB, Version 4, Online in www.mathworks.com.

24. Fung Ch.Ch., Wong K.W. and Eren H., 1997, Modular artificial neural network for prediction of petrophysical properties from well log data, *IEEE transaction on instrumentation and measurement*, **46**, 1295-1299.
25. Gopinath P. and Reddy N.P., 2000, Toward intelligent Web monitoring: Performance of committee neural networks vs single neural network, *Information technology applications in biomedicine, Proceedings of IEEE EMBs international conference*, 179-182.
26. Hashem S., 1993, Optimal linear combination of neural networks, PhD thesis, School of Industrial Engineering, Purdue University.
27. Hashem S., 1997, Optimal linear combination of neural networks, *Neural Networks*, **10**, 599-614.
28. Hashem S., Shmeiser B. and Yih Y., 1994, Optimal linear combination of neural networks: An overview, School of Industrial Engineering, Purdue University.
29. Hearst J.R., Nelson P.H. and Paillet F.L., 2000, Well logging for physical properties, A hand book for geophysicists, geologists and engineers, Jhon Wiley and sons Ltd.
30. Helle H.B., Bhatt A. and Ursin B., 2001, Porosity and permeability prediction from wireline logs using artificial neural networks: A North Sea case study, *Geophysical Prospecting*, **49**, 431-444.
31. Huang Z., Shimeld J., Williamson M. and Katsube J., 1996, Permeability prediction with artificial neural neural network modeling in the Venture gas field offshore eastern Canada, *Geophysics*, **61**, 422-436.
32. Huang Z. and Williamson M., 1997, Determination of porosity and permeability in reservoir intervals by artificial neural network modelling offshore eastern Canada, *Petroleum Geoscience*, **3**, 245-258.
33. Kaastra I. and Boyd M., 1996, Designing a neural network for forecasting financial and economic time series, *Neurocomputing*, **10**, 215-236.
34. Link C.A., Himmer Ph.A., 2003, Oil reservoir porosity prediction using a neural network ensemble approach, *In Geophysical application of artificial neural networks and fuzzy logic* (Eds W.Sandham and M.Leggett), 197-213, Kluwer academic publishers.
35. Selley R.C., 1998, *Elements of petroleum geology*, Second edition, Academic press.
36. Sharkey A., Sharkey N., Gerecke U. and Chandroth G.O., 2000, The "test and select" approach to ensemble combination, *In Multiple classifier systems* (Eds J.Kittler and F.Roli), 30-44, *Lecture notes in computer science*, Springer-Verlag, Inc.
37. Tiab D. and Donaldson E.C., 2004, *Petrophysics. Theory and practice of measuring reservoir rock and fluid properties*, Gulf publishing co.

پیوست

پیوست الف:

انواع الگوریتم‌های آموزشی پس‌انتشار خطا
و توابع مربوط به آن‌ها در نرم‌افزار MATLAB

انواع الگوریتم‌های آموزشی پس‌انتشار خطا و توابع مربوط به آن‌ها در نرم‌افزار

MATLAB

در آموزش با الگوریتم پس‌انتشار خطا دو روند آموزش مختلف وجود دارد:

۱- آموزش الگو به الگو^۱: در این روش به ازای هر بردار الگوی آموزشی ورودی به شبکه، مسیر رفت و برگشت الگوریتم پس‌انتشار خطا طی شده و تغییر پارامترهای شبکه محاسبه و تصحیح انجام می‌شود. این روند برای هر الگوی آموزشی تکرار می‌شود.

۲- آموزش گروهی^۲: در این روش به ازای هر بردار الگوی آموزشی ورودی مسیر رفت و برگشت طی شده و تغییر پارامترهای شبکه محاسبه می‌گردد، اما تصحیح پارامترها انجام نمی‌شود. به عبارت دیگر برای تمام الگوهای آموزشی تنها تغییرات پارامترها محاسبه می‌شوند. وقتی که تمام الگوهای آموزشی به شبکه عرضه شد و تغییرات پارامترها محاسبه گردید، تغییرات هر پارامتر مشخص به ازای تمام الگوهای آموزشی با هم جمع می‌شود تا تغییر آن پارامتر را مشخص کند. سپس تصحیح انجام می‌شود.

بعضی از توابع آموزش مانند تابع تفریق گرادیان با هر دو روند یاد شده قابل استفاده هستند ولی بعضی از آن‌ها مانند تابع لونبرگ-مارکوارت فقط با روند گروهی قابل استفاده می‌باشند [۱۸ و ۲۳].

حال به شرح انواع الگوریتم‌های آموزشی پس‌انتشار خطا می‌پردازیم.

همه الگوریتم‌های آموزشی پس‌انتشار خطا از گرادیان تابع عملکرد نسبت به بردار پارامترهای شبکه استفاده می‌کنند تا پارامترها را به گونه‌ای تغییر دهند که تابع عملکرد شبکه را کمینه کند.

1- Incremental mode

2- Batch mode

الف-۱- الگوریتم پس انتشار خطای تفریق گرادیان^۱

این الگوریتم در سال ۱۹۸۶ پیشنهاد شد. در این الگوریتم پارامترهای شبکه در جهت منفی گرادیان تغییر می‌کند و تابع عملکرد به سرعت کاهش می‌یابد، به همین دلیل به آن الگوریتم بیشترین شیب نیز گفته می‌شود. تغییر پارامترهای شبکه با رابطه زیر بدست می‌آید [۲۳]:

$$\vec{x}_{k+1} = \vec{x}_k - \alpha_k \vec{g}_k \quad (\text{الف-۱})$$

در این رابطه \vec{x}_k بردار پارامترهای شبکه، \vec{g}_k بردار گرادیان تابع عملکرد نسبت به پارامترهای شبکه و α_k سرعت یادگیری^۲ است. تابع مربوط به این الگوریتم در نرم‌افزار MATLAB، Traingd است. در رابطه با این تابع هفت پارامتر آموزشی وجود دارد که می‌توانند توسط کاربر تعیین شوند و در غیر این صورت مقادیر پیش فرض خود را حفظ خواهند کرد. این هفت پارامتر عبارتند از epochs, show, goal, time, min_grad, max_fail و lr. lr پارامتر سرعت یادگیری است. هر چه مقدار آن بزرگ‌تر باشد تغییرات پارامترهای شبکه در یک تصحیح، بزرگ‌تر خواهد بود اما باید توجه داشت که چنانچه مقدار آن به طور نامناسبی بزرگ باشد الگوریتم ناپایدار خواهد شد و اگر مقدار آن کوچک انتخاب شود زمان زیادی برای همگرایی الگوریتم نیاز خواهد بود. پارامتر show مشخص می‌کند که بعد از هر چند دوره آموزش، وضعیت آموزش نمایش داده شود. بقیه پارامترها مشخص می‌کنند که چه موقع آموزش متوقف شود. epochs تعداد دوره‌های تکرار را مشخص می‌کند. اگر تعداد تکرارها به آن برسد، یا اگر تابع عملکرد به مقداری کمتر از پارامتر goal برسد، یا اگر مقدار گرادیان کمتر از min_grad شود و یا اگر مدت زمان آموزش بر حسب ثانیه بیشتر از time شود، آموزش متوقف می‌شود. پارامتر max_fail که در آموزش به روش توقف سریع به کار می‌رود تعداد تکرارهایی را مشخص می‌کند که چنانچه خطا برای الگوهای اعتبارسنجی افزایش یابد آموزش متوقف می‌شود [۲۳].

ضعف عمده این الگوریتم، سرعت کم همگرایی آن و نیز توقف یادگیری شبکه قبل از فراگیری کامل مسأله به دلیل به دام افتادن پارامترهای شبکه در نقاط کمینه محلی می‌باشد [۱۸ و ۲۳].

1- Gradient descent back-propagation algorithm

2- Learning rate

الف-۲- الگوریتم پس انتشار خطای تفریق گرادیان با تکانه^۱

این الگوریتم نسبت به الگوریتم تفریق گرادیان اغلب همگرایی سریع تری را موجب می شود. تکانه به شبکه اجازه می دهد که تغییرات پارامترها را به گونه ای انجام دهد که علاوه بر گرادیان محلی، روند نهایی رویه خطا را نیز تأثیر دهد. بدون تکانه ممکن است شبکه در یک کمینه محلی به دام بیفتد ولی تکانه همانند یک فیلتر پایین گذر باعث می شود که شبکه ویژگی های کوچک رویه خطا را فراموش کند و از چنین کمینه هایی عبور کند [۲۳]. در این الگوریتم تغییرات پارامترهای شبکه مجموع دو جمله است. یک جمله کسری از حاصل ضرب گرادیان تابع عملکرد و سرعت یادگیری و جمله دیگر کسری از آخرین تغییرات پارامترهای شبکه است. میزان تأثیر آخرین تغییرات پارامترها به وسیله پارامتر ضریب تکانه^۲ که با mc نشان داده می شود و می تواند بین صفر و یک انتخاب شود تعیین می شود. وقتی که این ثابت صفر انتخاب می شود تغییرات پارامترها تنها بر اساس گرادیان تابع عملکرد صورت می گیرد و وقتی که یک انتخاب می شود تغییرات پارامترها برابر آخرین تغییرات پارامترها است. در یک تکرار معین چنانچه تابع عملکرد جدید از تابع عملکرد تکرار قبل به نسبتی بیش از پارامتر max_perf_inc که معمولاً $1/0.4$ تعیین می شود فزونی گیرد، پارامترهای حاصل غیرمفید هستند و رها می شوند و mc صفر انتخاب و تغییرات پارامترها تنها بر اساس جمله گرادیان تعیین می شود. در MATLAB برای این الگوریتم تابع $Traingdm$ پیش بینی شده است. با توجه به توضیحاتی که داده شد مشخص است که این تابع علاوه بر پارامترهایی که برای تابع $Traingd$ ذکر شد دو پارامتر اضافی mc و max_perf_inc نیز دارد. این تابع از روند آموزش گروهی استفاده می کند [۲۳].

الف-۳- الگوریتم های پس انتشار خطا با سرعت یادگیری متغیر^۳

در روش های تفریق گرادیان سرعت یادگیری در کل مرحله آموزش ثابت نگه داشته می شود و کارایی الگوریتم به تعیین مناسب سرعت یادگیری وابستگی زیادی دارد. اگر سرعت یادگیری بسیار بزرگ انتخاب شود ممکن است الگوریتم دچار نوسان و ناپایداری شود و چنانچه خیلی کوچک انتخاب شود زمان زیادی برای همگرایی الگوریتم نیاز است. در عمل ممکن نیست که قبل از آموزش، سرعت

1- Gradient descent with momentum back-propagation algorithm
2- Momentum coefficient
3- Variable learning rate back-propagation algorithms

یادگیری به طور بهینه تعیین شود و در واقع در طی آموزش با حرکت بر روی روبه خطا سرعت یادگیری بهینه تغییر می‌کند [۲۳].

اگر اجازه دهیم که سرعت یادگیری در طی آموزش تغییر کند کارایی الگوریتم بیشترین شیب بهبود خواهد یافت. سرعت یادگیری قابل تنظیم باعث می‌شود که تغییرات پارامترها تا آنجایی که ممکن است بزرگ باشد و در عین حال یادگیری پایدار بماند. در هر دوره با استفاده از سرعت یادگیری موجود، پارامترهای شبکه محاسبه و خروجی‌ها و خطاها تعیین می‌شوند. اگر خطای جدید نسبت به خطای تکرار قبل به نسبتی بیش از پارامتر max_perf_inc که معمولاً $1/0.4$ انتخاب می‌شود فزون‌ی گیرد، پارامترهای جدید غیر مفید هستند و رها می‌شوند. همچنین سرعت یادگیری با ضرب شدن در پارامتر lr_dec که معمولاً $0.7/10$ انتخاب می‌شود کاهش می‌یابد. در غیر این صورت پارامترهای جدید مفید هستند و نگه داشته می‌شوند. اگر خطای جدید از خطای قبلی کوچک‌تر باشد سرعت آموزش با ضرب شدن در پارامتر lr_inc که معمولاً $1/0.5$ انتخاب می‌شود افزایش می‌یابد. این روند، سرعت یادگیری را افزایش می‌دهد، البته فقط تا اندازه‌ای که شبکه بدون افزایش خطای زیادی آموزش ببیند. بنابراین یک سرعت یادگیری نزدیک به بهینه برای وضعیت محلی روبه خطا به دست می‌آید. وقتی که یک سرعت یادگیری بزرگ‌تر بتواند به یادگیری پایدار منجر شود سرعت یادگیری افزایش می‌یابد. همچنین وقتی که سرعت یادگیری به اندازه‌ای بالا است که کاهش خطا را تضمین نکند، کاهش می‌یابد تا پایداری حفظ شود [۲۳].

در نرم‌افزار MATLAB تابع Traingda به منظور آموزش به این روش طراحی شده است. در کاربرد این تابع علاوه بر پارامترهای آموزشی که در مورد تابع Traingd به آن‌ها اشاره شد، سه پارامتر اضافی دیگر نیز وجود دارد که عبارتند از max_perf_inc ، lr_dec و lr_inc که در مورد آن‌ها توضیحات لازم بیان شد. علاوه بر تابع Traingda ، تابع Traingdx نیز برای آموزش با سرعت یادگیری متغیر طراحی شده است. این تابع دو الگوریتم آموزش با سرعت یادگیری قابل تنظیم و آموزش با تکانه را با هم ترکیب می‌کند و واضح است که علاوه بر پارامترهای تابع Traingda ، پارامتر اضافی ضریب تکانه را نیز دارد [۲۳].

الف-۴- الگوریتم پس‌انتشار خطای بهبودپذیر^۱

معمولاً شبکه‌های چندلایه در لایه‌های میانی خود از توابع انتقال سیگموئیدی استفاده می‌کنند. این توابع یک دامنه ورودی گسترده و نامعین را به گستره‌ای مشخص و معین از خروجی‌ها می‌نگارند. شیب این توابع به ازای یک ورودی بزرگ به صفر نزدیک می‌شود. این موضوع باعث بروز مشکل به هنگام استفاده از روش بیشترین شیب می‌شود، زیرا گرادین مقدار بسیار کوچکی پیدا می‌کند و حتی اگر پارامترهای شبکه بسیار دورتر از مقادیر بهینه‌شان باشند تغییرات آن‌ها بسیار کوچک خواهد بود [۲۳]. هدف از الگوریتم پس‌انتشار خطای بهبودپذیر آن است که اثرات مقدار مشتقات جزئی حذف شود. در این الگوریتم تنها از علامت مشتق به منظور تعیین جهت تغییر پارامترها استفاده می‌شود و مقدار مشتق تاثیری بر تنظیم پارامترها ندارد. اندازه تغییر پارامترها به وسیله مقدار قابل تنظیم جداگانه‌ای تعیین می‌شود. هر گاه که مشتق تابع عملکرد نسبت به یک پارامتر خاص در دو تکرار متوالی هم‌علامت باشد این مقدار قابل تنظیم برای آن پارامتر با استفاده از فاکتور $delt_inc$ افزایش می‌یابد. همچنین هر وقت که این مشتق نسبت به تکرار قبلی تغییر علامت دهد مقدار قابل تنظیم به وسیله فاکتور $delt_dec$ کاهش می‌یابد. اگر مشتق صفر باشد مقدار قابل تنظیم ثابت باقی می‌ماند. هر گاه که پارامتر مورد نظر نوسان کند تغییرات آن کوچک می‌شود. اگر در چند تکرار متوالی تغییر پارامتر در یک جهت صورت گیرد بزرگی تغییرات آن پارامتر افزایش می‌یابد. در نرم‌افزار MATLAB تابع `Trainrp` برای آموزش با این الگوریتم پیش‌بینی شده است. پارامترهای آموزشی این تابع عبارتند از `show`, `epochs`, `goal`, `time`, `min_grad`, `max_fail`, `delt_inc`, `delt_dec`, `delta0` و `deltamax`. ۸ پارامتر اول قبلاً توضیح داده شده است، `delta0` اندازه اولیه تغییرات پارامترها و `deltamax` بیشینه تغییرات پارامترها در یک تصحیح می‌باشد. کارایی تابع `Trainrp` چندان متأثر از تنظیم پارامترهای آموزشی نیست و بیشتر این پارامترها می‌توانند مقدار پیش‌فرضشان باقی بمانند. به طور کلی این الگوریتم از الگوریتم‌هایی که پیش از این شرح داده شد خیلی سریع‌تر همگرا می‌شود [۲۳].

الف-۵- الگوریتم‌های گرادیان مزدوج^۱

این روش‌ها در سال ۱۹۹۲ جهت آموزش ساختارهای عصبی ارائه شد. در الگوریتم پس‌انتشار خطای بیشترین شیب، پارامترهای شبکه در جهت منفی گرادیان تغییر می‌کنند. چنان که قبلاً اشاره شد در این جهت تابع عملکرد سریعترین کاهش را خواهد داشت، اما این امر به این معنی نیست که سریع‌ترین همگرایی نیز در این جهت اتفاق می‌افتد. در روش گرادیان مزدوج تغییرات در جهت‌های مزدوج صورت می‌گیرد که به طور کلی نسبت به روش بیشترین شیب ده‌ها برابر سریعتر همگرا می‌شود.

دو مسیر \vec{p}_i و \vec{p}_j را نسبت به ماتریس هسین^۲ مزدوج گویند هر گاه داشته باشیم [۱۸]:

$$\vec{p}_i^T \cdot H \cdot \vec{p}_j = 0 \quad ; \quad \forall i, j, \quad i \neq j \quad (\text{الف-۲})$$

ماتریس هسین ماتریس مشتقات دوم تابع عملکرد است و با فرض این‌که تعداد پارامترهای شبکه N باشد یک ماتریس $N \times N$ خواهد بود [۱۸].

در بیشتر الگوریتم‌های گرادیان مزدوج، در هر تکرار، اندازه تغییرات تنظیم می‌شود. جستجویی در جهت گرادیان مزدوج انجام می‌شود تا اندازه تغییرات به گونه‌ای انتخاب شود که تابع عملکرد را در امتداد آن مسیر کمینه کند. در جعبه ابزار شبکه عصبی نرم‌افزار MATLAB ۵ تابع جستجوی خطی مختلف وجود دارد. بعضی از این توابع جستجو با برخی از توابع آموزشی روش گرادیان مزدوج، که در ادامه به آن‌ها می‌پردازیم، تناسب بیشتری دارند اگر چه در نهایت انتخاب بهینه تابع جستجو به ماهیت مسأله بستگی دارد [۲۳]. در این جا به منظور اختصار، از شرح توابع جستجو صرف‌نظر می‌کنیم و فقط توابع آموزشی روش گرادیان مزدوج را شرح می‌دهیم.

الف-۵-۱- الگوریتم فلچر-ریویس^۳

همه الگوریتم‌های گرادیان مزدوج، در نخستین تکرار در جهت منفی گرادیان جستجو می‌کنند [۲۳]:

$$\vec{p}_0 = -\vec{g}_0 \quad (\text{الف-۳})$$

-
- 1- Conjugate gradient algorithms
 - 2- Hessian matrix
 - 3- Fletcher-Reeves algorithm

یک جستجوی خطی انجام می‌شود تا اندازه تغییرات بهینه مشخص شود و سپس تغییرات در جهت جاری اعمال می‌شود [۲۳]:

$$\bar{x}_{k+1} = \bar{x}_k + \alpha_k \bar{p}_k \quad (\text{الف-۴})$$

جهت جستجوی جدید طوری انتخاب می‌شود که مزدوج جهت جستجوی قبلی باشد. روال کلی تعیین راستای جستجوی جدید آن است که جهت منفی گرادیان جاری را با جهت جستجوی قبلی ترکیب کنیم:

$$\bar{p}_k = -\bar{g}_k + \beta_k \bar{p}_{k-1} \quad (\text{الف-۵})$$

در روش‌های مختلف گرادیان مزدوج، ثابت β_k به روش‌های مختلفی محاسبه می‌شود. در روش فلچر-ریویس β_k به صورت زیر به دست می‌آید:

$$\beta_k = \frac{\bar{g}_k^T \bar{g}_k}{\bar{g}_{k-1}^T \bar{g}_{k-1}} \quad (\text{الف-۶})$$

در نرم‌افزار MATLAB تابع `Traincgf` برای آموزش با الگوریتم فلچر-ریویس تهیه شده است. پارامترهای آموزشی این تابع عبارتند از `max_fail`, `min_grad`, `time`, `goal`, `epochs`, `show`, `alpha`, `beta`, `gamma`, `delta`, `scalfch`, `low_lim`, `up_lim`, `maxstep`, `minstep` و `bmax`. شش پارامتر اول قبلاً توضیح داده شده‌اند. پارامتر `scalfch` نام تابع جستجوی خطی را مشخص می‌کند. بقیه پارامترها مربوط به تابع جستجو هستند. به طور کلی تابع `Traincgf` در مقایسه با `Trainrp` در تکرارهای کمتری همگرا می‌شود، اگر چه در هر تکرار محاسبات بیشتری نیاز است. معمولاً الگوریتم‌های گرادیان مزدوج از الگوریتم‌های پخش خطای سرعت یادگیری متغیر بسیار سریع‌تر هستند و گاهی اوقات نیز نسبت به الگوریتم پخش خطای بهبودپذیر سریع‌تر همگرا می‌شوند، اگر چه نتایج از مسأله‌ای به مسأله دیگر تفاوت دارد. این الگوریتم‌ها نسبت به الگوریتم‌های ساده‌تر فقط به میزان کمی حافظه بیشتر نیاز دارند و بنابراین انتخاب خوبی برای شبکه‌هایی با پارامترهای زیاد می‌باشند [۲۳].

الف-۵-۲- الگوریتم پولک-ریبیر^۱

نوع دیگری از الگوریتم گرادیان مزدوج است که مانند روش فلچر-ریویس جهت جستجو در هر تکرار با رابطه زیر مشخص می‌شود [۲۳]:

$$\vec{p}_k = -\vec{g}_k + \beta_k \vec{p}_{k-1} \quad (\text{الف-۷})$$

در این روش ثابت β_k به صورت زیر تعیین می‌شود:

$$\beta_k = \frac{\Delta \vec{g}_{k-1}^T \vec{g}_k}{\vec{g}_{k-1}^T \vec{g}_{k-1}} \quad (\text{الف-۸})$$

این عبارت برابر است با ضرب داخلی تغییرات تکرار قبلی گرادیان با گرادیان جاری تقسیم بر مجذور اندازه بردار گرادیان تکرار قبلی. در نرم‌افزار MATLAB تابع Traincgp برای آموزش با این روش گرادیان مزدوج پیش‌بینی شده است. پارامترهای آموزشی این تابع مشابه تابع Traincgf است و عملکرد مشابهی دارند به طوری که پیش‌بینی آن که کدام یک از آن‌ها در یک مسأله خاص بهتر عمل می‌کند، مشکل است. حجم حافظه‌ای که Traincgp نیاز دارد تا حدودی از Traincgf بیشتر است [۲۳].

الف-۵-۳- الگوریتم شروع پاول-بیل^۲

در یک الگوریتم گرادیان مزدوج، به طور دوره‌ای جهت جستجو به همان جهت اولیه منفی گرادیان برگشت داده می‌شود. نقطه استاندارد برگشت وقتی است که تعداد تکرارها برابر با تعداد پارامترهای شبکه شود، ولی معیارهای برگشت دیگری وجود دارد که کارایی آموزش را بهبود می‌دهند. یکی از این معیارها توسط پاول و بیل پیشنهاد شده است که بر اساس آن وقتی که نامساوی زیر برقرار شود جهت جستجو به منفی گرادیان تغییر می‌یابد [۲۳]:

$$|\vec{g}_{k-1}^T \vec{g}_k| \geq 0.2 \|\vec{g}_k\|^2 \quad (\text{الف-۹})$$

در MATLAB تابع Traincgb برای آموزش به این روش پیش‌بینی شده است که پارامترهای آموزشی آن مشابه تابع Traincgf است. برای برخی از مسائل، عملکرد تابع Traincgb تا اندازه‌ای از

1- Polak-Ribiere algorithm

2- Powell-Beale algorithm

تابع Traincgp بهتر است، اگر چه قضاوت نهایی به مسأله مورد نظر بستگی دارد. حجم حافظه مورد نیاز تابع Traincgb تا اندازه‌ای از Traincgp بیشتر است [۲۳].

الف-۵-۴- الگوریتم گرادیان مزدوج مقیاسی^۱

الگوریتم‌های گرادیان مزدوجی که تا این جا شرح داده شد در هر تکرار، به یک جستجوی خطی نیاز دارند. این جستجو از نظر محاسباتی سنگین است زیرا لازم است که در هر جستجو پاسخ شبکه به همه الگوهای آموزشی چندین بار محاسبه شود. گرادیان مزدوج مقیاسی طراحی شده است تا از جستجوی خطی که وقت‌گیر است اجتناب شود. توضیح این الگوریتم پیچیده در چند سطر مشکل است، ولی می‌توان گفت که این الگوریتم ترکیبی از رهیافتی که در بخشی از روش لوبنبرگ-مارکوارت به کار می‌رود با رهیافت گرادیان مزدوج است. در MATLAB تابع Trainscg برای آموزش به این روش پیش‌بینی شده است. این تابع نسبت به سایر توابع گرادیان مزدوج به تکرارهای بیشتری نیاز دارد تا همگرا شود، ولی تعداد محاسباتی که در هر تکرار نیاز است به میزان قابل توجهی کاهش می‌یابد، زیرا جستجوی خطی انجام نمی‌شود. حجم حافظه مورد نیاز این تابع مشابه تابع Traincgf است [۲۳].

الف-۶- الگوریتم‌های شبه نیوتنی^۲

الف-۶-۱- الگوریتم بی‌اف‌جی‌اس^۳

با هدف بهینه‌سازی سریع، روش نیوتن می‌تواند به عنوان جایگزینی برای روش‌های گرادیان مزدوج باشد. تغییر پارامترها در این الگوریتم مطابق رابطه زیر انجام می‌شود [۲۳]:

$$\vec{x}_{k+1} = \vec{x}_k - H_k^{-1} \vec{g}_k \quad (\text{الف-۱۰})$$

که در آن H_k ماتریس هسین تابع عملکرد به ازای مقادیر جاری پارامترهای شبکه است. اغلب، روش نیوتن سریع‌تر از روش‌های گرادیان مزدوج همگرا می‌شود. متأسفانه محاسبه ماتریس هسین برای شبکه‌های عصبی پیش‌رو پیچیده و مشکل است. گروهی از الگوریتم‌ها وجود دارند که بر اساس روش

1- Scaled conjugate gradient algorithm
2- Quasi-Newton algorithms
3- BFGS algorithm

نیوتن عمل می‌کنند اما به محاسبه مشتقات دوم نیازی ندارند و به آن‌ها روش‌های شبه‌نیوتنی یا متقاطع^۱ گفته می‌شود. این روش‌ها در هر تکرار الگوریتم، تخمین ماتریس هسین مرحله قبل را تصحیح می‌کنند و از تخمین جدید برای محاسبه پارامترهای شبکه استفاده می‌کنند. روش شبه‌نیوتنی که در مطالعات منتشر شده موفق‌ترین نتایج را نشان داده است، روش تصحیح بی‌اف‌جی‌اس بوده است. در MATLAB تابع آموزشی Trainbfg به منظور آموزش شبکه‌ها با این الگوریتم پیش‌بینی شده است. پارامترهای آموزشی این تابع همان پارامترهای تابع Traincgf است. این الگوریتم نسبت به روش‌های گرادیان مزدوج در هر تکرار به محاسبات و حافظه بیشتری نیاز دارد، زیرا ماتریس هسین که برای شبکه‌ای با N پارامتر یک ماتریس $N \times N$ است باید تخمین و ذخیره شود ولی به طور کلی در تکرارهای کمتری همگرا می‌شود. ممکن است برای شبکه‌های خیلی بزرگ که پارامترهای زیادی دارند استفاده از روش‌های گرادیان مزدوج و یا پس‌انتشار خطای بهبودپذیر بهتر باشد ولی برای شبکه‌های کوچک‌تر Trainbfg یک تابع کارآمد است [۲۳].

الف-۶-۲- الگوریتم متقاطع یک مرحله‌ای^۲

چون الگوریتم بی‌اف‌جی‌اس نسبت به الگوریتم‌های گرادیان مزدوج در هر تکرار به محاسبات و حافظه بیشتری نیازمند است، بنابراین مطلوب است که تخمین به گونه‌ای انجام شود که این محاسبات و حجم حافظه کاهش یابد. روش متقاطع یک مرحله‌ای، روشی حد واسط الگوریتم‌های گرادیان مزدوج و الگوریتم‌های شبه نیوتنی است. در این الگوریتم فرض می‌شود که ماتریس هسین تکرار قبلی ماتریس واحد باشد. این الگوریتم نسبت به الگوریتم بی‌اف‌جی‌اس در هر دوره به محاسبات و حافظه کمتری نیاز دارد ولی در مقایسه با الگوریتم‌های گرادیان مزدوج اندکی به حافظه و محاسبات بیشتر نیازمند است. در MATLAB تابع آموزشی Trainoss برای آموزش با این الگوریتم پیش‌بینی شده است. پارامترهای آموزشی این تابع همان پارامترهای تابع Traincgf است [۲۳].

1- Secant

2- One step secant algorithm

الف-۷- الگوریتم لونیبرگ-مارکوارت^۱

این روش در سال ۱۹۹۴ ارائه شد و تا به امروز به عنوان سریع‌ترین روش آموزش شبکه‌های عصبی برای شبکه‌های پیش‌رو با اندازه متوسط (شبکه‌هایی با چند صد پارامتر) شناخته شده است. این روش ده‌ها برابر از روش گرادیان مزدوج و صدها برابر از روش بیشترین شیب سریع‌تر است. از ویژگی‌های برجسته این روش، کاهش شدید در نرون‌های لایه مخفی می‌باشد و از محدودیت‌های آن این است که تنها با روند یادگیری گروهی قابل استفاده است و هیچ تضمینی ندارد که با آموزش الگو به الگو همگرا شود [۱۸ و ۲۳].

وقتی که تابع عملکرد به شکل مجموع مربعات (مثلاً تابع MSE) انتخاب شود می‌توان ماتریس هسین را به صورت زیر تخمین زد [۲۳]:

$$H = J'J \quad (\text{الف-۱۱})$$

همچنین گرادیان به صورت زیر محاسبه می‌شود:

$$\vec{g} = J' \vec{e} \quad (\text{الف-۱۲})$$

J ماتریس ژاکوبین است که مشتقات اول خطاهای تخمین شبکه نسبت به پارامترهای شبکه می‌باشد و \vec{e} یک بردار از خطاهای شبکه است. ماتریس ژاکوبین با استفاده از یک روش پسانتشار خطا با پیچیدگی بسیار کمتر نسبت به ماتریس هسین قابل محاسبه است. در الگوریتم لونیبرگ-مارکوارت تغییر پارامترها با استفاده از رابطه زیر انجام می‌شود [۲۳]:

$$\vec{x}_{k+1} = \vec{x}_k - [J^T J + \mu I]^{-1} J^T \vec{e} \quad (\text{الف-۱۳})$$

وقتی که مقدار کمیت اسکالر μ صفر در نظر گرفته شود الگوریتم لونیبرگ-مارکوارت به الگوریتم نیوتنی تبدیل می‌شود که از تخمین ماتریس هسین استفاده کرده است، ولی وقتی که مقدار این کمیت بزرگ است به روش بیشترین شیب با سرعت یادگیری کوچک تبدیل می‌شود. در نزدیکی یک کمینه خطا، روش نیوتن سریع‌تر و دقیق‌تر است، بنابراین در چنین موقعیتی هدف آن است که تا آن جا که ممکن است به سرعت به روش نیوتنی نزدیک شویم. بنابراین بعد از هر مرحله موفقیت‌آمیز کاهش تابع عملکرد، μ کاهش می‌یابد و فقط بعد از هر افزایش تابع عملکرد، μ افزایش می‌یابد. با این روش تابع عملکرد بعد از هر تکرار الگوریتم، همیشه کاهش می‌یابد. در MATLAB برای آموزش

1- Levenberg-Marquardt algorithm

با الگوریتم لونیبرگ- مارکوارت تابع Trainlm پیش‌بینی شده است. پارامترهای این تابع عبارتند از epochs, show, goal, time, min_grad, max_fail, mu, mu_dec, mu_inc, mu_max, mem_reduc. شش پارامتر اول قبلاً توضیح داده شده‌اند. μ مقدار اولیه μ است. هر گاه تابع عملکرد کاهش یابد، μ در پارامتر mu_dec ضرب می‌شود و هر گاه افزایش یابد، μ در mu_inc ضرب می‌شود. اگر μ از mu_max بزرگ‌تر شود الگوریتم متوقف می‌شود. پارامتر mem_reduc برای کنترل میزان حافظه‌ای که الگوریتم مورد استفاده قرار می‌دهد به کار می‌رود [۲۳].

باید توجه داشت که اگر در آموزش شبکه‌ها با روش توقف سریع از یک الگوریتم سریع استفاده می‌شود بایستی پارامترهای آموزش به گونه‌ای تعیین شوند که الگوریتم نسبتاً آهسته همگرا شود، مثلاً در صورت استفاده از الگوریتم لونیبرگ- مارکوارت، μ نسبتاً بزرگ مثلاً ۱ و μ_{inc} و μ_{dec} نیز نزدیک به ۱، مثلاً ۱/۵ و ۰/۸ انتخاب شوند [۲۳].

الف-۸- تابع مرتب‌سازی بیزین^۱

در آموزش شبکه‌ها به روش مرتب‌سازی، مطلوب است که پارامترهای مرتب‌سازی بهینه به طور اتوماتیک تعیین شود. بنیان بیزین مک‌کی^۲ رهیافتی برای این منظور است که در آن، پارامترهای شبکه متغیرهایی تصادفی با توزیع‌های معین در نظر گرفته می‌شوند. پارامترهای مرتب‌سازی با واریانس‌های نامعلوم وابسته به این توزیع‌ها مربوط هستند. این پارامترها با استفاده از روش‌های آماری تعیین می‌شوند. شرح مفصل این روش، از سطح این نوشتار خارج است.

در MATLAB تابع آموزشی Trainbr برای آموزش با این روش پیش‌بینی شده است. یکی از ویژگی‌های این روش آن است که تعداد پارامترهایی از شبکه را که به طور مؤثر توسط شبکه استفاده می‌شوند مشخص می‌کند. با فرض اطمینان از همگرایی، صرف‌نظر از این‌که تعداد کل پارامترهای شبکه چقدر است، تعداد مؤثر پارامترها بایستی تقریباً ثابت باشد.

به هنگام استفاده از تابع Trainbr، باید اجازه دهیم تا تعداد مؤثر پارامترهای شبکه، مجموع مربعات خطا و مجموع مربعات پارامترهای شبکه در چندین تکرار به مقداری نسبتاً ثابت همگرا شود [۲۳].

1- Bayesian

2- Bayesian framework of David MacKay