

تخمین تخلخل یک مخزن هیدروکربوری با استفاده از سیستم‌های چند شبکه‌ای

پژوهش‌نفت

سال بیست و سوم

شماره ۷۴

صفحه، ۱۰۹-۱۲۵ ۱۳۹۲

تاریخ دریافت مقاله: ۹۰/۹/۲۶

تاریخ پذیرش مقاله: ۹۱/۶/۱۴

محمود ذاکری^{۱*} و ابوالقاسم کامکار روحانی^۲

۱- سازمان زمین‌شناسی و اکتشافات معدنی کشور

۲- دانشگاه صنعتی شاهرود، دانشکده مهندسی معدن، نفت و ژئوفیزیک

mahmood_zakery2006@yahoo.com

که میانگین مربعات خطای تخمین تخلخل الگوهای آموزش و آزمون را نسبت به بهترین شبکه عصبی مصنوعی منفرد به ترتیب ۱۴/۷٪ و ۱۲/۵٪ کاهش داده است.

واژه‌های کلیدی: شبکه عصبی مصنوعی، سیستم‌های چند شبکه‌ای، ترکیب آنسامبلی، الگوریتم ژنتیک، تخلخل و نگار

مقدمه

الگو و الهام بخش شبکه‌های عصبی مصنوعی^۱، سیستم عصبی موجودات زنده است که به گونه‌ای کاملاً متفاوت از کامپیوترهای دیجیتالی مرسوم، محاسبات را انجام می‌دهند. این شبکه‌ها برخلاف کامپیوتر که نیازمند دستورهای کاملاً صریح و مشخص است، به مدل‌های ریاضی محض نیاز ندارند؛ بلکه مانند انسان تجربه کسب کرده و سپس نتیجه این تجربیات را تعمیم می‌دهند [۱].

در مواردی که روابط بین داده‌ها غیرخطی، مبهم و ناشناخته است و تصویر روشنی از آنها در دست نیست، روش شبکه عصبی مصنوعی در مقایسه با روش‌های متعارف،

1. Artificial Neural Networks (ANNs)

چکیده

در این مقاله سیستم‌های چندشبکه‌ای به منظور کاهش خطا و افزایش دقت نتایج حاصل از روش شبکه عصبی مصنوعی پیشنهاد شده است. در این سیستم‌ها نتایج چندین شبکه که به طور منفرد و مجزا آموزش دیده‌اند، به روشی مناسب با هم ترکیب می‌شود. در این مطالعه تخلخل مؤثر یکی از مخازن هیدروکربوری میدان عظیم پارس جنوبی با استفاده از سیستم‌های چندشبکه‌ای برآورد شده است. از شبکه‌های عصبی مصنوعی پس انتشار خطا که به روش اعتبارسنجی همزمان آموزش دیده‌اند، به عنوان اجزای سیستم‌های چندشبکه‌ای استفاده شد. داده نگارهای چاه از ۴ چاه این میدان در بازه عمقی سازند کنگان مورد استفاده قرار گرفت. نگارهای صوتی، چگالی، اشعه گاما و تخلخل نوترونی به عنوان ورودی شبکه‌ها و تخلخل مؤثر به عنوان خروجی شبکه‌ها انتخاب گردید. روش ترکیب آنسامبلی شبکه‌ها که دارای ساختاری موازی است، برای تشکیل سیستم‌های چند شبکه‌ای مورد استفاده قرار گرفت. نتایج نشان می‌دهد که ترکیب‌های آنسامبلی مناسب می‌تواند خروجی شبکه‌های عصبی مصنوعی منفرد آموزش دیده به روش اعتبارسنجی همزمان را بهبود بخشد. بهترین ترکیب آنسامبلی حاصل در این تحقیق، ترکیبی سه شبکه‌ای است

عبارتند از: فانگ و همکاران از ترکیب مدوله‌ای شبکه‌های عصبی برای تشخیص جنس و درصد کانی‌های سنگ‌های مجاور یک چاه از روی نگارهای چاه استفاده کرده‌اند و نتایج را با شبکه پس انتشار خطای^۱ منفرد مقایسه نموده‌اند. مقایسه نشان می‌دهد که ترکیب مدوله‌ای شبکه‌های عصبی، مدت زمان آموزش و میانگین مربعات خطا را نسبت به شبکه پس انتشار خطای منفرد به میزان زیادی کاهش داده است [۷].

بات در رساله دکتری خود، با استفاده از شبکه‌های عصبی کمیته‌ای به تخمین تخلخل، تراوایی و اشباع سیالات مخزنی در میدین نفتی دریای شمال پرداخته است. وی برای تخمین تخلخل از ترکیب آنسامبلی شبکه‌ها بهره گرفته است. 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 تخلخل را تخمین بزند. وی در تخمین تراوایی از روش مدوله‌ای استفاده کرده است. دامنه تراوایی به سه زیردامنه تقسیم و هر یک از سه مدول سیستم با استفاده از یکی از زیر دامنه‌ها 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 کمیته‌ای نشان می‌دهد که در تخمین تراوایی، ماشین کمیته‌ای دقت بیشتر و خطای کمتری داشته است. برای محاسبه اشباع هر یک از سه سیال مخزنی آب، نفت و گاز، از یک ماشین کمیته‌ای استفاده شده که ترکیب آنسامبلی ۹ شبکه پرسپترون با ساختار ۱-۴-۴ می‌باشد. هر یک از این ترکیب‌های 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 [۲].

ابزاری توانمند برای حل مسأله است. معمولاً این شرایط در داده‌های مربوط به علوم زمین وجود دارد، به همین دلیل از این روش به طور روزافزون در این علوم استفاده می‌شود [۲ و ۳].

در روش شبکه عصبی، تعدادی شبکه آموزش دیده تولید می‌شود، این شبکه‌ها با استفاده از یک معیار عملکرد مناسب مثلاً میانگین مربعات خطاها^۱ (MSE) سنجش شده و بر این اساس بهترین شبکه انتخاب می‌شود. انتخاب بهترین شبکه منفرد، اگر چه بهترین الگوی حاصل را تولید می‌کند، اما باعث از دست رفتن اطلاعاتی می‌شود که در دیگر شبکه‌ها وجود داشته است، بنابراین به جای این که یکی از شبکه‌ها به عنوان بهترین شبکه انتخاب و از بقیه صرف نظر شود، ترکیب نتایج شبکه‌های آموزش دیده با استفاده از روشی مناسب پیشنهاد شده است. این کار ممکن است به تلفیق اطلاعات شبکه‌های جزء ترکیب و در نتیجه افزایش دقت و توان تعمیم کمک کند. با استفاده از ترکیب نتایج شبکه‌های منفرد، سیستم‌های چندشبکه‌ای^۲ تولید می‌شود تا در مسائلی که یک شبکه قادر به حل آنها نیست و یا ممکن است با استفاده از سیستم‌های چندشبکه‌ای به طور مؤثرتری حل شوند، به نتیجه بهتری دست یافت [۳-۵].

دو روش معمول برای ایجاد سیستم‌های چندشبکه‌ای وجود دارد: ۱- ترکیب آنسامبلی^۳ که در این مقاله از آن استفاده شده است و توضیح داده خواهد شد. ۲- ترکیب مدوله‌ای^۴ که در آن مسأله به تعدادی زیرمسأله تفکیک و هر زیرمسأله توسط یک شبکه عصبی یا آنسامبلی از آنها حل شده و در نهایت با تلفیق پاسخ آنها کل مسأله حل می‌شود. این دو نوع ترکیب با هم ناسازگار نیستند به این مفهوم که یک سیستم چندشبکه‌ای می‌تواند شامل هر دو نوع ترکیب باشد. در مقالات ارائه شده در منابع ۲ و ۶، سیستم‌های چندشبکه‌ای را با عنوان «ماشین کمیته‌ای^۵» نام‌گذاری کرده، هر چند که این روش‌ها قبلاً شناخته شده و به کار رفته است [۲ و ۶].

میزان به‌کارگیری سیستم‌های چندشبکه‌ای در علوم مربوط به زمین و از جمله چاه‌نگاری، نسبت به شبکه‌های عصبی مصنوعی منفرد، بسیار محدودتر است. چند مورد از آنها

1. Mean Square Errors (MSE)
2. Multiple Networks Systems
3. Ensemble Combination
4. Modular Neural Network (MNN)
5. Committee Machine (CM)
6. Back Propagation (BP)

است که طی آن شبکه، الگوی موجود در ورودی‌ها و ارتباط بین ورودی‌ها و خروجی‌های مجموعه الگوهای آموزشی را فرا می‌گیرد. تعمیم، توانایی شبکه برای ارائه جواب قابل قبول در قبال ورودی‌هایی است که در مجموعه آموزشی نبوده‌اند. استفاده از شبکه برای انجام عملکردی که به آن منظور طراحی شده است را اجرا گویند [۱]. بدیهی است که شبکه‌ای که در مرحله کاربرد و اجرا مورد استفاده قرار می‌گیرد، باید به خوبی آموزش دیده باشد و از توان تعمیم مناسبی برخوردار باشد.

یکی از مشکلاتی که ممکن است در حل مسأله با استفاده از روش شبکه عصبی پیش بیاید، بیش‌برازش^۷ و یا همان توانایی تعمیم کم است، یعنی در مرحله آموزش، شبکه عملکرد خوبی دارد و خطا به مقدار بسیار کوچکی می‌رسد، اما وقتی که با الگوهایی به غیر از الگوهای آموزشی آزموده می‌شود، عملکرد ضعیفی داشته و خطای پاسخ زیاد است. یکی از روش‌هایی که به منظور کاهش احتمال بیش‌برازش پیشنهاد شده است، آموزش به روش اعتبارسنجی همزمان است [۹] که در این مقاله از آن به منظور آموزش شبکه‌های عصبی مصنوعی منفرد استفاده شده است.

آموزش به روش اعتبارسنجی همزمان

در این روش، الگوهای آموزش به دو دسته تقسیم می‌شوند که از دسته اول برای آموزش شبکه استفاده می‌شود. گروه دوم سری اعتبارسنجی^۸ است که در فرآیند تعیین پارامترهای شبکه نقشی ندارد ولی در هنگام آموزش، خطای شبکه برای آن 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 برای تعداد خاصی دوره آموزش افزایش یابد، آموزش متوقف شده

چون و لین نیز از یک ماشین کمیت‌های 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 را با استفاده از روش الگوریتم ژنتیک^۱ تعیین کرده‌اند. نتایج نشان می‌دهد که میانگین مربعات خطای تخمین تراوایی الگوهای آموزش با استفاده از میانگین‌گیری وزنی ۶٪ نسبت به بهترین فرمول تجربی کاهش پیدا کرده است. همچنین میانگین مربعات خطای تخمین تراوایی 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 [۸].

متأسفانه در کشورمان، تحقیق در مورد کاربرد سیستم‌های چندشبکه‌ای در تخمین و تشخیص ویژگی‌ها و خصوصیات مخزنی در مواردی بسیار معدود و انگشت شمار انجام شده است. بنابراین، مطالعات بیشتر با هدف آشنایی با قابلیت‌ها و کارکردهای آن ضروری به نظر می‌رسد. در این راستا، در مطالعه حاضر، ابتدا تخلخل در مخزن گازی کنگان میدان پارس جنوبی به وسیله شبکه‌های عصبی مصنوعی پس انتشار خطا که به روش اعتبارسنجی همزمان^۲ آموزش دیده‌اند، تخمین زده می‌شود و سپس نتایج آنها با استفاده از روش ترکیب آنسامبلی خطی ترکیب می‌شود تا در نهایت با مقایسه نتایج، ببینیم که آیا ترکیب آنسامبلی به عنوان نوعی از سیستم چندشبکه‌ای می‌تواند نتایج شبکه‌های عصبی مصنوعی منفرد را بهبود بخشد یا خیر و در صورت مثبت بودن پاسخ، چه عواملی بر میزان بهبود و ارتقای نتایج تأثیر می‌گذارند.

مبانی شبکه‌های عصبی مصنوعی

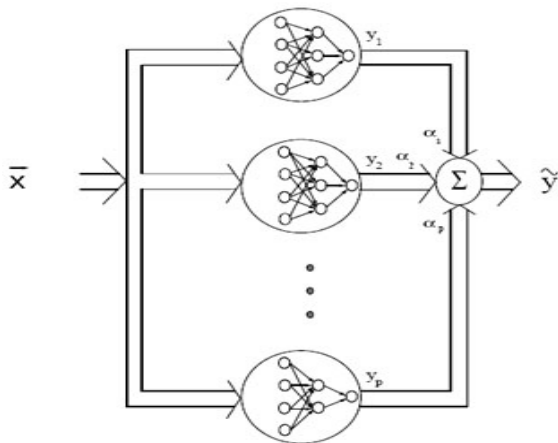
یک شبکه عصبی مصنوعی از تعدادی واحد پردازش به نام نرون^۳ مصنوعی و اتصالات بین آنها تشکیل می‌شود. به طور کلی، نرون‌ها در سه نوع لایه ورودی، میانی و خروجی تعبیه شده‌اند [۱ و ۲].

برای حل هر مسأله، شبکه‌های عصبی سه مرحله را طی می‌کنند: آموزش^۴، تعمیم^۵ و اجرا^۶ [۱]. آموزش فرآیندی

1. Genetic Algorithm (GA)
2. Cross Validation Method
3. Neuron
4. Training
5. Generalization
6. Operation
7. Overfitting
8. Validation Data

انجام شود [۲ و ۴].

در شکل ۲ نمایی از ترکیب آنسامبلی خطی شبکه‌ها نشان داده شده است. چنان‌که در شکل دیده می‌شود، ترکیب خطی خروجی شبکه‌های مختلف، مشابه ایجاد یک شبکه بزرگ است که شبکه‌های سازنده آن به طور موازی با هم عمل می‌کنند و ضرایب ترکیب، همان وزن‌های سیناپسی لایه خروجی هستند. برای ورودی \bar{x} ، خروجی ترکیب \bar{y} ، جمع وزنی خروجی‌های متناظر شبکه‌های جزء (y_i) است و α_i ها ضرایب مربوط به آنها است. اختلاف اصلی در این تشبیه آن است که در ترکیب، ضرایب ارتباطی (α_i ها) ثابت هستند، ولی در شبکه عصبی بزرگ معادل، این ضرایب که نقش وزن‌های سیناپسی لایه آخر را دارند، در طول آموزش تغییر می‌کنند. مزیت این ترکیب نسبت به شبکه هم ارز خود، آن است که در شبکه بزرگ معادل، پارامترهای زیادی وجود دارد که لازم است در فرآیند آموزش تعیین شوند. بنابراین زمان آموزش زیادتر است و از طرفی به خاطر افزایش پارامترها، احتمال بیش‌برازش بیشتر می‌شود [۴ و ۵].



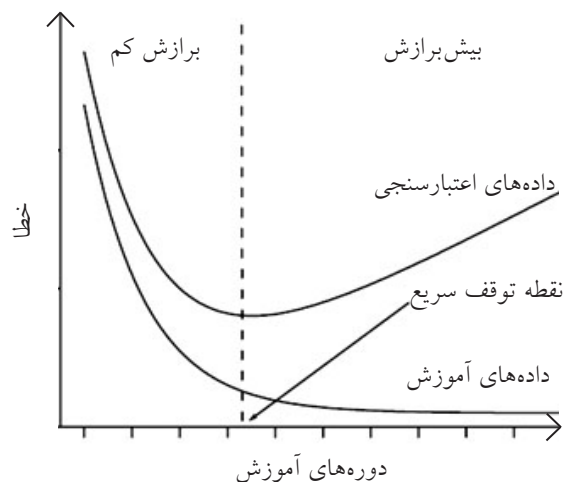
شکل ۲- ترکیب آنسامبلی خطی شبکه‌ها [۴]

و پارامترهای شبکه مقادیری را که در کمینه خطای سری اعتبارسنجی داشتند، می‌گیرند [۲ و ۹]. در شکل ۱ آموزش به روش اعتبارسنجی همزمان به صورت شماتیک نشان داده شده است.

ترکیب آنسامبلی

در یک ترکیب آنسامبلی، هر یک از شبکه‌های جزء، به تنهایی یک راه حل را برای مسأله مورد نظر ارائه می‌دهند و نتایج به روشی مناسب با هم ترکیب می‌شوند. به منظور ایجاد شبکه‌های سازنده ترکیب می‌توانیم یکی از موارد زیر را تغییر دهیم: پارامترهای اولیه شبکه، الگوهای آموزشی، ساختار شبکه و یا الگوریتم آموزش [۲ و ۶].

وقتی که یک سری از شبکه‌ها برای ساخت ترکیب‌های آنسامبلی ایجاد شدند، باید نتایج این شبکه‌ها با روشی مناسب و مؤثر با هم ترکیب شود. چندین روش مختلف ترکیب وجود دارد که انتخاب روش به نوع مسأله‌ای که می‌خواهیم آن را حل کنیم، بستگی دارد. در مسائل تخمین تابع، ترکیب آنسامبلی می‌تواند به طور خطی یا غیرخطی



شکل ۱- تغییرات خطا در سری آموزش و اعتبارسنجی در آموزش به روش اعتبارسنجی همزمان [۲ با اندکی تغییر]

ضرایب بهینه با انجام یک سلسله عملیات ماتریسی و با استفاده از مجموعه‌ای از الگوهای تحقق یافته که در آن به ازای ورودی \bar{x}_n ، مقدار تابع هدف و خروجی هر یک از شبکه‌های جزء آنسامبل، مشخص است و با $\kappa = \{\kappa_n : \kappa_n = (\bar{x}_n, r(\bar{x}_n), \bar{y}(\bar{x}_n)), n=1, \dots, k\}$ نشان داده می‌شود، تعیین می‌گردد. به مجموعه الگوهای k سری الگوهای ترکیب گفته می‌شود. به طور معمول k همان مجموعه الگوهای آموزشی شبکه‌ها انتخاب می‌شود [۴ و ۵].

روش الگوریتم ژنتیک

علاوه بر روش تحلیلی هاشم، الگوریتم ژنتیک نیز می‌تواند به منظور تعیین ضرایب بهینه MSE-OLC با استفاده از سری الگوهای ترکیب به کار رود. الگوریتم ژنتیک با ایجاد یک جمعیت اولیه از جواب‌ها که اصطلاحاً به آنها کروموزم^۱ گفته می‌شود، شروع می‌شود. با هدف دستیابی به یک راه حل بهینه، جمعیت جاری متحول می‌شود. به منظور ایجاد جمعیت جدید از جمعیت جاری، با استفاده از تابع خاصی، مقدار تابع هدف به ازای کروموزم‌های جمعیت جاری مقیاس‌بندی می‌شود تا مقدار برازش هر یک از کروموزم‌ها تعیین شود. آن‌گاه با استفاده از روش‌هایی خاص و بر اساس مقدار برازش، تعدادی از کروموزم‌های جمعیت جاری به عنوان والد انتخاب می‌شوند. بنا بر اصل نخبه‌گرایی^۲، تعداد خاصی از کروموزم‌های جمعیت جاری که بیش‌ترین برازش را داشته‌اند، بدون تغییر در نسل بعد حفظ می‌شوند و بقیه کروموزم‌های جمعیت جدید با استفاده از والدین و بر اساس دو قانون تقاطع^۳ و جهش^۴ تولید شده و برازش آنها تعیین می‌شود. در تقاطع، کروموزم‌های جدید که اصطلاحاً به آنها فرزندان گفته می‌شود، از ترکیب یک جفت از والدها تولید می‌شوند. ولی در جهش، یک فرزند از اعمال تغییرات در یک والد ایجاد می‌شود. بعد از چندین بار تولید نسل، الگوریتم ژنتیک

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

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

که خطای آن:

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

است. بردار ضرایب و مقدار مطلوب برای ورودی \bar{x} است [۴ و ۵].

بنابراین لازم است که باروشی کارآمدی مناسبی برای α_j ‌ها پیدا کنیم که در ادامه به شرح این روش‌ها می‌پردازیم.

روش‌های تحلیلی

یک روش، میانگین‌گیری ساده است که به طور گسترده مورد استفاده قرار می‌گیرد. در این روش فرض می‌شود که همه شبکه‌ها به یک میزان مفید هستند و دقت و توانایی یکسانی در حل مسأله دارند.

روش دیگر یافتن ضرایب بهینه^۱ (OLC) است، به طوری که تابع عملکرد که معمولاً MSE در نظر گرفته می‌شود، کمینه گردد. برای این منظور هاشم روشی تحلیلی با عنوان MSE-OLC را پیشنهاد کرد. در این روش، به منظور تصحیح بایاس خطای α_0 ها، جمله دیگری به رابطه ۱ افزوده می‌شود. این جمله $\alpha_0 y_0(\bar{x})$ است که در آن رابطه $y_0(\bar{x})=1$ برقرار است. بنابراین خروجی ترکیب به صورت زیر در خواهد آمد:

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

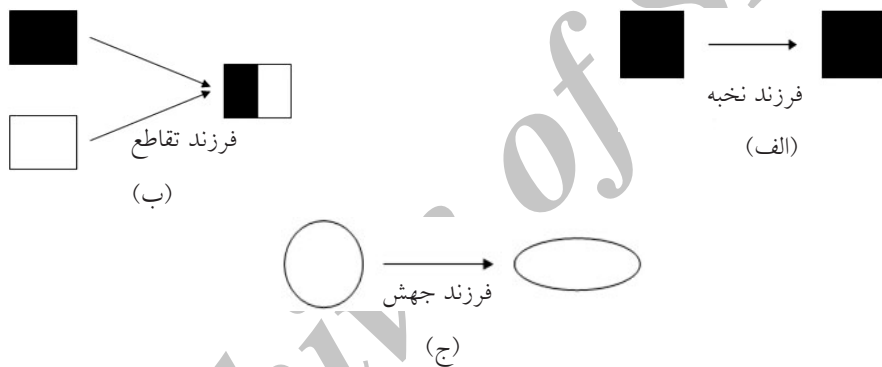
که $\bar{y}(\bar{x})$ و $\bar{\alpha}$ بردارهایی به ابعاد $(p+1) \times 1$ هستند. برای رابطه ۳ که حالت کلی را نشان می‌دهد، سه حالت خاص نیز می‌توان در نظر گرفت. بنابراین در مجموع ۴ حالت خواهیم داشت. تفاوت این ۴ حالت به وجود یا عدم وجود جمله ثابت α_0 و شرط $\sum_{j=1}^p \alpha_j = 1$ مربوط می‌شود. اگر مجموع ضرایب شبکه‌ها ۱ باشد، ترکیب را مقید^۲ و در غیر این صورت ترکیب را نامقید^۳ می‌نامند. این ۴ حالت عبارتند از ۱- نامقید با جمله ثابت (حالت کلی)، ۲- مقید با جمله ثابت، ۳- نامقید بدون جمله ثابت و ۴- مقید بدون جمله ثابت [۴ و ۵].

1. Optimal Linear Combination (OLC)
2. Constrained
3. Unconstrained
4. Chromosomes
5. Elitism
6. Crossover
7. Mutation

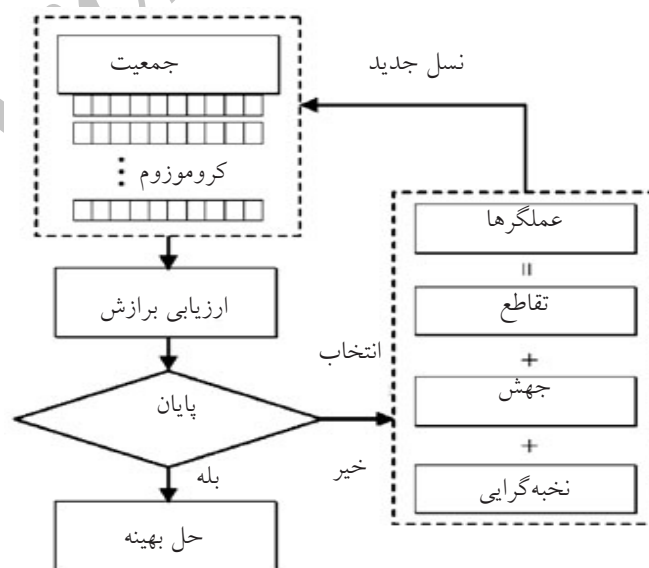
به طور مشابه با شبکه‌های عصبی مصنوعی منفرد، توان تعمیم MSE-OLC عبارت است از کارایی آن برای الگوهایی که در سری ترکیب نبوده‌اند و در برآورد ضرایب بهینه نقشی نداشته‌اند. ترکیب آنسامبلی حاصل زمانی مفید است که علاوه بر الگوهای سری ترکیب، خطای تخمین را برای الگوهای سری آزمون نیز کاهش دهد [۴ و ۶].

در این مقاله به منظور تخمین تخلخل، از روش‌های میانگین‌گیری ساده، روش‌های چهارگانه ترکیب خطی بهینه هاشم و روش غیرتحلیلی ترکیب خطی بهینه با استفاده از الگوریتم ژنتیک برای ساخت ترکیب‌های آنسامبلی خطی استفاده شده است.

همگرا شده و به یک راه حل بهینه تقریباً مطلق با برآزش خوب برای تابع هدف می‌رسد. در الگوریتم ژنتیک، معیارهای مختلفی برای توقف روند تولید نسل وجود دارد که با تحقق یکی از آنها، الگوریتم متوقف می‌شود. مثلاً بیشینه دفعات تولید نسل را می‌توان تعیین کرد و یا زمانی که مقدار برآزش یکی از کروموزم‌های جمعیت جاری از یک حد تعیین شده بیشتر باشد، تولید نسل توقف می‌یابد [۸ و ۱۰]. شکل ۳ نحوه تولید فرزندان را با استفاده از سه عملگر نخبه‌گرایی، تقاطع و جهش نشان می‌دهد. در شکل ۴ فرآیند بهینه‌یابی با استفاده از روش الگوریتم ژنتیک نمایش داده شده است.



شکل ۳- نحوه تولید فرزندان در روش الگوریتم ژنتیک (الف) عملگر نخبه‌گرایی، (ب) عملگر تقاطع و (ج) عملگر جهش [۱۰]



شکل ۴- فرآیند بهینه‌یابی با استفاده از روش الگوریتم ژنتیک [۸ با اندکی تغییر]

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 [۱۳].

الگوهای آموزش و آزمون

در بین چاه‌های حفر شده در میدان پارس جنوبی که داده‌های چاه‌نگاری آنها در دسترس بود و موقعیت آنها نسبت به هم در شکل ۵ نشان داده شده است، از داده‌های چاه‌نگاری چاه‌های SP1، SP3، SP6 و SP13 در بازه عمقی سازند کنگان (حداقل عمق، تقریباً ۲۶۰۰ متر در SP1 و حداکثر عمق، تقریباً ۳۰۰۰ متر در SP13) به منظور طراحی، آموزش و آزمون شبکه‌ها استفاده شد. با توجه به محرمانه بودن موقعیت دقیق جغرافیایی چاه‌ها به لحاظ مشترک بودن این میدان بین ایران و قطر، مقیاس شکل و مختصات چاه‌ها در اختیارمان قرار نگرفت.

از ۲۱۵ الگوی منتخب از داده‌های چاه‌نگاری چاه‌های SP1، SP3 و SP13 به منظور آموزش و از ۸۹ الگوی انتخاب شده از داده‌های چاه‌نگاری چاه SP6 برای آزمون توان تعمیم شبکه‌ها استفاده شد. داده‌های نگارهای صوتی، چگالی، اشعه گاما و تخلخل نوترونی به عنوان ورودی و تخلخل مؤثر به عنوان خروجی شبکه‌ها انتخاب شده است. 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.

شکل ۶ نمودار توزیع فراوانی الگوهای آموزش بر حسب درصد تخلخل مؤثر را نشان می‌دهد. با توجه به این شکل مشخص می‌شود که فراوانی الگوهای با تخلخل زیاد نسبت به الگوهای با تخلخل کم، کمتر است.

تخمین تخلخل با استفاده از شبکه‌های عصبی پس‌انتشار خطای آموزش دیده به روش اعتبارسنجی همزمان و ترکیب آنسامبلی خطی نتایج آن‌ها

کلیه برنامه‌های مورد استفاده در این تحقیق و همچنین کارهای محاسباتی در محیط نرم‌افزار MATLAB تهیه و انجام شده است.

مطالعه موردی: تخمین تخلخل مخزن گازی کنگان در میدان پارس جنوبی

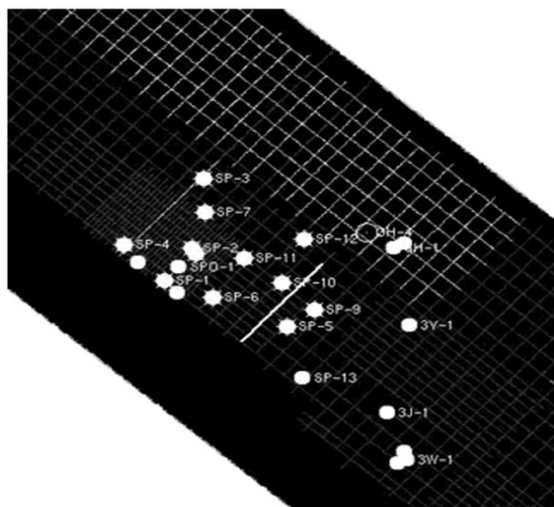
به منظور بررسی قابلیت و توانایی سیستم‌های چند شبکه‌ای در بهبود نتایج حاصل از شبکه‌های عصبی مصنوعی منفرد آموزش دیده به روش اعتبارسنجی همزمان، مطالعه موردی در میدان عظیم هیدروکربوری پارس جنوبی، در بخش‌های گازدار k-1 و k-2 سازند کنگان انجام شد و تخلخل مؤثر ابتدا با استفاده از شبکه‌های عصبی مصنوعی منفرد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

زمین‌شناسی سازند کنگان

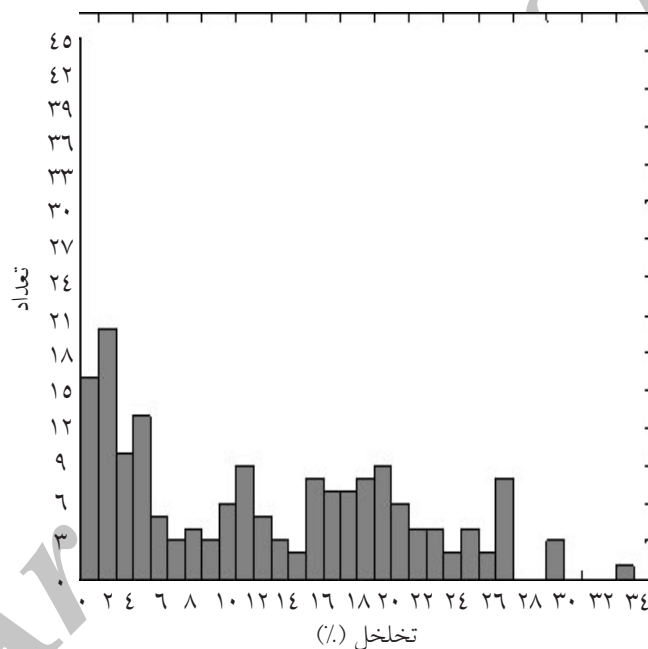
رسوبات این سازند متعلق به دوران تریاس زیرین است که با یک ناپوستگی موازی بر روی سازند دالان به سن پرمین قرار گرفته است. سنگ پوشش آن را شیل آغار و لایه‌های تبخیری سازند دشتک می‌سازند. از نظر زمین‌شناسی ساختمانی، کانسار گازی کنگان تاقدیس نامتقارنی با امتداد شمال غربی - جنوب شرقی است که دامنه جنوب شرقی آن برگشته است [۱۱].

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 تخلخل مناسبی داشته و یکی از بخش‌های مهم گازدار میدان پارس جنوبی می‌باشد. ضخامت این بخش در چاه‌های مختلف به طور متوسط ۱۵۵ متر است [۱۲].

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



شکل ۵- موقعیت چاه‌های میدان پارس جنوبی [۱۲]



شکل ۶- نمودار توزیع فراوانی مقادیر تخلخل مؤثر در ۲۱۵ الگوی آموزشی

گرفت. برای آموزش به روش اعتبارسنجی همزمان تابع آموزش لونیبرگ - مارکوارت^۱ به کار برده شد و پارامترهای آن به گونه‌ای مناسب انتخاب گردید. کمیت max_fail در آموزش به روش اعتبارسنجی همزمان، تعداد دوره‌هایی را مشخص می‌کند که اگر در هنگام آموزش، در این تعداد دوره، خطا به‌طور مکرر برای الگوهای اعتبارسنجی افزایش پیدا کند، آموزش متوقف می‌شود.

ابتدا، تخلخل توسط شبکه‌های عصبی مصنوعی منفرد پس انتشار خطای آموزش دیده به روش اعتبارسنجی همزمان تخمین زده شده و سپس نتایج شبکه‌های منتخب با هم ترکیب شده‌اند.

پیش پردازش داده‌ها به وسیله تابع Prestd در نرم‌افزار MATLAB که مقادیر داده‌ها را به توزیع گوسی استاندارد تبدیل می‌کند، انجام شد. ۷۰٪ از الگوهای آموزش (۱۵۰ الگو) برای آموزش شبکه و ۳۰٪ باقی‌مانده (۶۵ الگو) به منظور اعتبارسنجی هنگام آموزش، مورد استفاده قرار

1. Levenberg-Marquardt Training Function (Trainlm)

پارامترهای شبکه و تعداد دوره‌های آموزش. با توجه به کم‌تر بودن MSE تخمین تخلخل شبکه ۱ برای الگوهای آزمون، این شبکه به عنوان بهترین شبکه عصبی مصنوعی منفرد حاصل در نظر گرفته می‌شود. در شکل ۷ پیشرفت روند آموزش و نمودار همبستگی بین مقدار تخلخل و مقدار تخمینی آن توسط این شبکه برای الگوهای آموزش و آزمون نمایش داده شده است.

شکل (۷ الف) نشان می‌دهد که از دوره ۱۰ به بعد بیش‌برازش شروع شده است و خطای سری اعتبارسنجی افزایش یافته است. به همین دلیل پارامترهای شبکه مقادیری را که در کمینه خطای سری اعتبارسنجی یعنی دوره نهم آموزش داشته‌اند، به خود می‌گیرند.

برای ۱۲۰ ترکیب ممکن از بین ۷ شبکه جدول ۱، شامل ۲۱ ترکیب دو شبکه‌ای، ۳۵ ترکیب سه شبکه‌ای، ۳۵ ترکیب چهارشبکه‌ای، ۱۲ ترکیب پنج شبکه‌ای، ۷ ترکیب شش شبکه‌ای و ۱ ترکیب هفت شبکه‌ای، با هر یک از ۵ روش تحلیلی میانگین‌گیری ساده، ترکیب خطی بهینه نامقید با جمله ثابت، ترکیب خطی بهینه مقید با جمله ثابت، ترکیب خطی بهینه نامقید بدون جمله ثابت و ترکیب خطی بهینه مقید بدون جمله ثابت، ترکیب‌های آنسامبلی خطی ساخته شد.

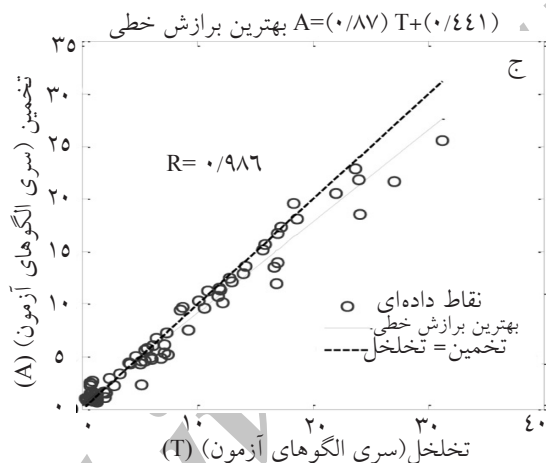
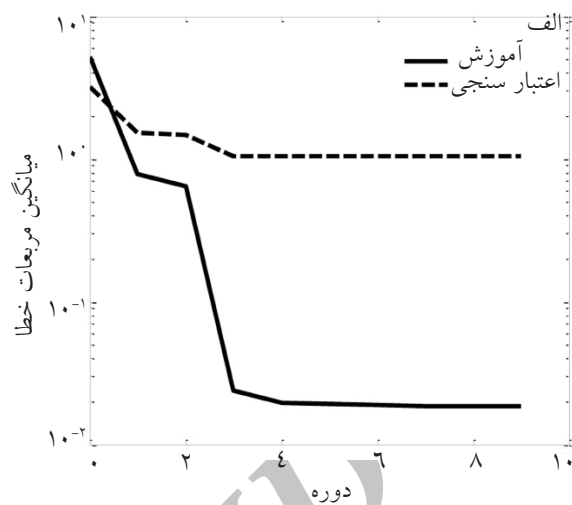
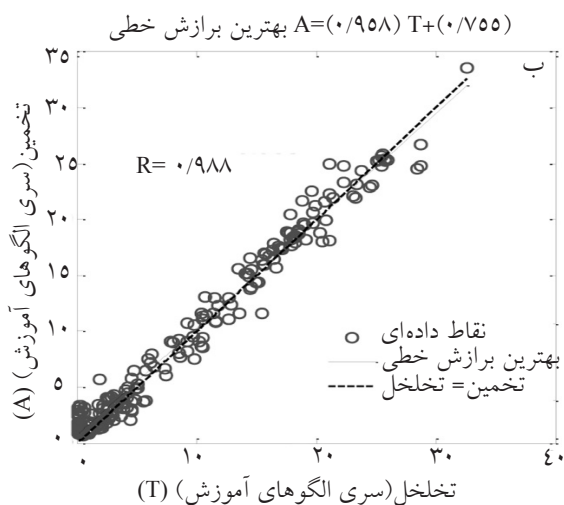
مقدار این کمیت همان مقدار پیش فرض آن در نرم‌افزار، معادل ۵ دوره انتخاب شد.

از شبکه‌هایی با یک یا دو لایه میانی با تعداد نرون‌های مختلف در این لایه‌ها استفاده شد. تابع انتقال نرون‌های لایه‌های میانی تابع تانژانت هیپربولیک^۱ و تابع انتقال نرون‌های لایه خروجی تابع خطی انتخاب گردید. در مورد هر ساختار خاص شبکه، یک مرحله آزمون و خطای طولانی با شرایط اولیه مختلف انجام شده است و بهترین نتیجه که کمترین میانگین مربعات خطای تخمین تخلخل الگوهای آزمون را داشت، ثبت گردید. سپس نتایج ساختارها بر اساس این کمیت مرتب شد و در نهایت از بین ۹۲ شبکه با ساختارهای مختلف، ۷ شبکه که کم‌ترین میانگین مربعات خطای تخمین تخلخل الگوهای آزمون را داشتند، برای ساخت ترکیب‌های آنسامبلی انتخاب شدند.

جدول ۱ نتایج این ۷ شبکه را نشان می‌دهد. ستون‌های این جدول از سمت راست عبارتند از شماره ردیف، ساختار شبکه، میانگین مربعات خطای تخمین تخلخل الگوهای آموزش، ضریب همبستگی بین مقادیر تخلخل و تخمین آن برای الگوهای آموزش، میانگین مربعات خطای تخمین تخلخل الگوهای آزمون، ضریب همبستگی بین مقادیر تخلخل و تخمین آن برای الگوهای آزمون، تعداد کل

جدول ۱- ۷ شبکه منفرد آموزش دیده به روش اعتبارسنجی همزمان 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 دوره‌های آموزش
۱	۴-۲-۱۷-۱	۱/۸۱۳	۰/۹۸۸	۲/۲۵۷	۰/۹۸۶	۷۹	۹
۲	۴-۴-۱-۱	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



شکل ۷- (الف) نمایش پیشرفت روند آموزش و کاهش MSE برای الگوهای آموزش و اعتبارسنجی برای شبکه ۱، (ب) نمودار همبستگی بین تخلخل و تخمین شبکه ۱ برای الگوهای آموزش و (ج) نمودار همبستگی بین تخلخل و تخمین شبکه ۱ برای الگوهای آزمون

نسبی فرزندان تولید شده با استفاده از قانون تقاطع را تعیین می‌کند از ۰/۳ تا ۰/۸ تغییر داده شد. به دیگر گزینه‌های قابل انتخاب مربوط به الگوریتم ژنتیک (یعنی تابع مقیاس‌بندی برازش، تابع انتخاب، تابع نسل، تابع تقاطع و تابع جهش) نیز مقادیر و موارد مختلفی نسبت داده شد و فرآیند سعی و خطا با جمعیت‌های اولیه تصادفی مختلف انجام شد. در نهایت در هر یک از ۶ حالت منتخب، بهترین نتیجه که کم‌ترین میانگین مربعات خطای تخمین تخلخل الگوهای آزمون را داشت، ثبت گردید.

سپس نتایج ۶۰۰ ترکیب آنسامبلی خطی حاصل، برای الگوهای آموزش و آزمون مورد بررسی قرار گرفت. از بین هر یک از دسته ترکیب‌های دو، سه، چهار، پنج، شش و هفت‌شبکه‌ای، بهترین ترکیب آنسامبلی حاصل از ۵ روش تحلیلی که کمترین میانگین مربعات خطای تخمین تخلخل الگوهای آزمون را در دسته داشت، انتخاب شد. برای این ترکیب منتخب، علاوه بر روش‌های تحلیلی، ضرایب ترکیب خطی بهینه با استفاده از روش الگوریتم ژنتیک تعیین گردید و ترکیب آنسامبلی خطی ساخته شد.

نتایج حاصل در جداول ۲ تا ۷ ارائه شده است. در این جدول‌ها، کاهش MSE تخمین تخلخل الگوهای آموزش

در بهینه‌یابی با الگوریتم ژنتیک، جمعیت کروموزم‌ها ۵۰ کروموزوم انتخاب شد که از این تعداد، ۵ کروموزوم نخبه در جمعیت بعدی حفظ و مابقی با استفاده از تقاطع و جهش تولید شد. تعداد دفعات تولید نسل ۱۵۰ بار انتخاب شده است. مقادیر کمیت نسبت تقاطع که درصد یا تعداد

1. Scaling Function
2. Selection Function
3. Stall Generation

از تقسیم MSE تخمین ترکیب بر MSE تخمین شبکه ۱ برای الگوهای آزمون و تفریق مقدار حاصل از ۱ محاسبه شده است.

از تقسیم MSE تخمین ترکیب بر MSE تخمین بهترین شبکه منفرد (شبکه ۱ در جدول ۱) برای الگوهای آموزش و تفریق مقدار حاصل از ۱ به دست آمده است. به طور مشابه، کاهش MSE تخمین تخلخل الگوهای آزمون نیز

جدول ۲- نتایج ترکیب آنسامبلی خطی شبکه‌های ۳ و ۶ جدول ۱

GA - OLC	Hashem - OLC	میانگین گیری ساده	روش ترکیب خطی نتایج
۶ و ۳	۶ و ۳	۶ و ۳	ترکیب آنسامبلی
۱/۵۶۴	۱/۵۱۸	۱/۶۴۹	MSE (الگوهای آموزش)
% ۱۳/۷	% ۱۶/۳	% ۹	کاهش در MSE الگوهای آموزش
۲/۰۶۹	۲/۳۴۶	۲/۲۱۵	MSE (الگوهای آزمون)
% ۸/۳	---	% ۱/۹	کاهش در MSE الگوهای آزمون

جدول ۳- نتایج ترکیب آنسامبلی خطی شبکه‌های ۱، ۳ و ۶ جدول ۱

GA - OLC	Hashem - OLC	میانگین گیری ساده	روش ترکیب خطی نتایج
۶ و ۳، ۱	۶ و ۳، ۱	۶ و ۳، ۱	ترکیب آنسامبلی
۱/۵۴۶	۱/۴۶۶	۱/۵۷۱	MSE (الگوهای آموزش)
% ۱۴/۷	% ۱۹/۱	% ۱۳/۳	کاهش در MSE الگوهای آموزش
۱/۹۷۴	۲/۳۲۸	۲/۱۹۹	MSE (الگوهای آزمون)
% ۱۲/۵	---	% ۲/۶	کاهش در MSE الگوهای آزمون

جدول ۴- نتایج ترکیب آنسامبلی خطی شبکه‌های ۲، ۳، ۵ و ۶ جدول ۱

GA - OLC	Hashem - OLC	میانگین گیری ساده	روش ترکیب خطی نتایج
۶ و ۵، ۳، ۲	۶ و ۵، ۳، ۲	۶ و ۵، ۳، ۲	ترکیب آنسامبلی
۱/۵۵۶	۱/۴۷۳	۱/۶۲۶	MSE (الگوهای آموزش)
% ۱۴/۲	% ۱۸/۸	% ۱۰/۳	کاهش در MSE الگوهای آموزش
۲/۰۴۰	۲/۳۰۱	۲/۱۹۳	MSE (الگوهای آزمون)
% ۹/۶	---	% ۲/۸	کاهش در MSE الگوهای آزمون

جدول ۵- نتایج ترکیب آنسامبلی خطی شبکه‌های ۱، ۲، ۳، ۵ و ۶ جدول (۱)

GA - OLC	Hashem - OLC	میانگین گیری ساده	روش ترکیب خطی نتایج
۶ و ۵، ۳، ۲، ۱	۶ و ۵، ۳، ۲، ۱	۶ و ۵، ۳، ۲، ۱	ترکیب آنسامبلی
۱/۵۶۰	۱/۴۵۰	۱/۵۸۱	MSE (الگوهای آموزش)
% ۱۴	% ۲۰	% ۱۲/۸	کاهش در MSE الگوهای آموزش
۲/۰۰۱	۲/۳۰۹	۲/۱۹۳	MSE (الگوهای آزمون)
% ۱۱/۳	---	% ۲/۸	کاهش در MSE الگوهای آزمون

جدول ۶- نتایج ترکیب آنسامبلی خطی شبکه‌های ۱، ۲، ۳، ۴، ۵ و ۶ جدول (۱)

GA - OLC	Hashem - OLC	میانگین گیری ساده	روش ترکیب خطی نتایج
۶ و ۵، ۴، ۳، ۲، ۱	۶ و ۵، ۴، ۳، ۲، ۱	۶ و ۵، ۴، ۳، ۲، ۱	ترکیب آنسامبلی
۱/۵۸۰	۱/۳۷۰	۱/۶۵۴	MSE (الگوهای آموزش)
% ۱۲/۹	% ۲۴/۴	% ۸/۸	کاهش در MSE الگوهای آموزش
۲/۰۲۷	۲/۳۵۲	۲/۱۹۳	MSE (الگوهای آزمون)
% ۱۰/۲	---	% ۲/۸	کاهش در MSE الگوهای آزمون

جدول ۷- نتایج ترکیب آنسامبلی خطی شبکه ۷ جدول (۱)

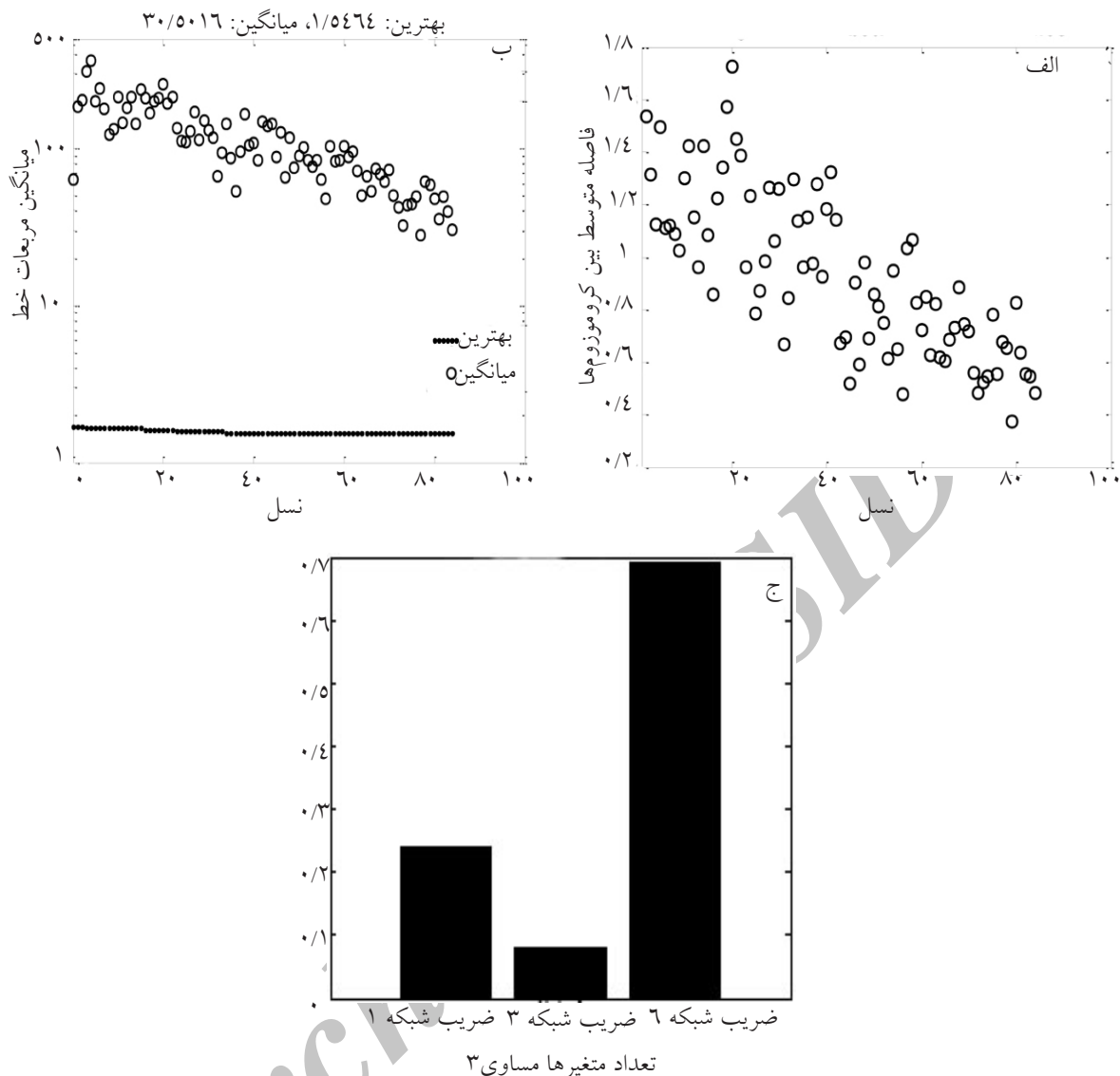
GA - OLC	Hashem - OLC	میانگین گیری ساده	روش ترکیب خطی نتایج
۷ و ۶، ۵، ۴، ۳، ۲، ۱	۷ و ۶، ۵، ۴، ۳، ۲، ۱	۷ و ۶، ۵، ۴، ۳، ۲، ۱	ترکیب آنسامبلی
۱/۵۷۷	۱/۳۶۱	۱/۶۴۱	MSE (الگوهای آموزش)
% ۱۳	% ۲۴/۹	% ۹/۵	کاهش در MSE الگوهای آموزش
۲/۰۰۴	۲/۳۷۶	۲/۲۰۳	MSE (الگوهای آزمون)
% ۱۱/۲	---	% ۲/۴	کاهش در MSE الگوهای آزمون

جدول ۱ بیشتر است و کاهش را ایجاد نکرده و در جدول با نماد --- نمایش داده شده است. این ترکیب MSE تخمین تخلخل الگوهای آموزش را ۱۹/۱٪ کاهش داده است. بنابراین می‌توان گفت که توان تعمیم این ترکیب ضعیف است. ترکیب خطی بهینه این سه شبکه با استفاده از الگوریتم ژنتیک بهترین ترکیبی است که در این تحقیق به دست آمده است. زیرا MSE تخمین تخلخل الگوهای آزمون را نسبت به بهترین شبکه منفرد، بیشترین مقدار یعنی ۱۲/۵٪ کاهش داده است.

در شکل ۸ فرآیند الگوریتم ژنتیک در تعیین ضرایب ترکیب خطی بهینه سه شبکه‌ای شبکه‌های ۱، ۳ و ۶ نشان داده شده است. با توجه به این شکل، فرآیند تولید نسل در تکرار ۸۴ متوقف شده است، زیرا معیارهای توقف تولید نسل به گونه‌ای تعیین شده که اگر تغییر میانگین وزنی MSE طی ۵۰ نسل به صفر برسد، الگوریتم متوقف می‌شود و در اینجا نیز در تکرار ۸۴ چنین اتفاقی رخ داده است. شکل (۸ الف) فاصله متوسط بین کروموزوم‌ها (بردارهای ضرایب) در نسل‌ها را نشان می‌دهد که معیاری برای کمیت تنوع بین کروموزوم‌های جمعیت در یک نسل است. شکل (۸ ب) در هر نسل دو مقدار را نشان می‌دهد. یک مقدار مربوط به بهترین (کم‌ترین) MSE تخمین تخلخل الگوهای آموزش به ازای بهترین کروموزوم در آن نسل است. با توجه به این شکل مشخص است که طی ۸۴ تکرار این مقدار در دفعاتی انگشت شمار به میزان بسیار کمی کاهش و بهبود یافته است و مقدار نهایی آن ۱/۵۴۶۴ می‌باشد. مقدار دیگری که این شکل نشان می‌دهد میانگین MSE تخمین تخلخل الگوهای آموزش به ازای ۵۰ کروموزوم موجود در هر نسل است. شکل (۸ ج) ضرایب بهینه در ترکیب آنسامبلی خطی حاصل از فرآیند الگوریتم ژنتیک را نشان می‌دهد. این ضرایب را می‌توان به صورت [۰/۶۹۲۶، ۰/۰۸۱۹، ۰/۲۴۳۵] نشان داد. بدین معنی که در ترکیب خطی بهینه حاصل از الگوریتم ژنتیک، ضریب شبکه شماره ۱، ۰/۲۴۳۵، ضریب شبکه شماره ۳، ۰/۰۸۱۹ و ضریب شبکه شماره ۶، ۰/۶۹۲۶ است.

جدول ۲ نتایج مربوط به ترکیب‌های دوشبکه‌ای را نشان می‌دهد. ۲۱ انتخاب دوشبکه‌ای ممکن از بین ۷ شبکه جدول ۱ صورت گرفت و با استفاده از ۵ روش تحلیلی مجموعاً ۱۰۵ ترکیب آنسامبلی خطی دوشبکه‌ای به دست آمد که بهترین نتیجه که کمترین میانگین مربعات خطای تخمین تخلخل الگوهای آزمون را داشت، مربوط به میانگین‌گیری ساده خروجی شبکه‌های ۳ و ۶ جدول ۱ می‌باشد که فقط ۱/۹٪ از MSE تخمین تخلخل شبکه ۱ برای الگوهای آزمون کاسته است. ترکیب آنسامبلی خطی بهینه این دو شبکه با استفاده از الگوریتم ژنتیک نیز به دست آمد. از بین ۴ ترکیب خطی بهینه شبکه‌های ۳ و ۶ با استفاده از روش‌های هاشم، ترکیب خطی بهینه نامقید با جمله ثابت، کمترین MSE تخمین تخلخل الگوهای آزمون را دارد که اگر چه MSE تخمین تخلخل الگوهای آموزش را ۱۶/۳٪ کاهش داده، ولی توان تعمیم ضعیفی دارد و MSE تخمین آن برای الگوهای آزمون (۲/۳۴۶) از بهترین شبکه منفرد حاصل (۲/۲۵۷) بیشتر است و کاهش را ایجاد نکرده است که در جدول با نماد --- نمایش داده شده است. ترکیب آنسامبلی خطی بهینه شبکه‌های ۳ و ۶ با استفاده از الگوریتم ژنتیک بهترین ترکیب دوشبکه‌ای است که در این تحقیق به دست آمده است. زیرا MSE تخمین تخلخل الگوهای آزمون را نسبت به بهترین شبکه منفرد، بیشترین مقدار یعنی ۸/۳٪ کاهش داده است.

جدول ۳ نتایج مربوط به ترکیب‌های سه شبکه‌ای را نشان می‌دهد. ۳۵ انتخاب سه شبکه‌ای ممکن از بین ۷ شبکه جدول ۱ صورت گرفت و با استفاده از ۵ روش تحلیلی مجموعاً ۱۷۵ ترکیب آنسامبلی خطی سه شبکه‌ای به دست آمد که بهترین نتیجه که کمترین میانگین مربعات خطای تخمین تخلخل الگوهای آزمون را داشت، مربوط به میانگین‌گیری ساده خروجی شبکه‌های ۱، ۳ و ۶ جدول ۱ می‌باشد. MSE تخمین تخلخل الگوهای آزمون این ترکیب ۲/۱۹۹ می‌باشد که ۲/۶٪ از MSE تخمین تخلخل شبکه ۱ کمتر است. از بین ۴ ترکیب خطی بهینه شبکه‌های ۱، ۳ و ۶ با استفاده از روش‌های هاشم، ترکیب خطی بهینه نامقید بدون جمله ثابت، کمترین MSE تخمین تخلخل الگوهای آزمون با مقدار ۲/۳۲۸ را دارد که از نتیجه شبکه شماره ۱



شکل ۸- الف) فاصله متوسط بین کروموزوم‌ها (بردارهای ضرایب) در نسل‌ها، (ب) بهترین (کم‌ترین) MSE تخمین تخلخل الگوهای آموزش به ازای بهترین کروموزوم در هر نسل و میانگین MSE تخمین تخلخل الگوهای آموزش به ازای ۵۰ کروموزوم موجود در هر نسل و (ج) ضرایب شبکه‌ها در بهترین ترکیب آنسامبلی خطی بهینه

اگر چه میانگین مربعات خطای تخمین تخلخل الگوهای آموزش را نسبت به شبکه ۱، $18/8\%$ کاهش داده، ولی توان تعمیم ضعیفی دارد. به طوری که میانگین مربعات خطای تخمین تخلخل الگوهای آموزش را کاهش نداده است. ترکیب این چهار شبکه با استفاده از الگوریتم ژنتیک نسبت به روش‌های تحلیلی نتیجه بهتری داشته و از MSE تخمین تخلخل الگوهای آموزش و آزمون به ترتیب $14/2\%$ و $9/6\%$ کاسته است.

۵ روش تحلیلی در مورد ۲۱ انتخاب پنج شبکه‌ای ممکن از بین ۷ شبکه به کار برده شد که از بین ۱۰۵ ترکیب تحلیلی

در جدول ۴ نتایج مربوط به ترکیب‌های چهار شبکه‌ای نشان داده شده است. با استفاده از ۵ روش تحلیلی، ۱۷۵ ترکیب آنسامبلی خطی مربوط به ۳۵ انتخاب چهار شبکه‌ای از شبکه‌های جدول ۱ ساخته شد که از بین آنها ترکیب آنسامبلی شبکه‌های ۲، ۳، ۵ و ۶ با استفاده از میانگین‌گیری ساده بهترین نتیجه، یعنی کمترین میانگین مربعات خطای تخمین تخلخل الگوهای آموزش به مقدار $2/193$ را داشته است که $2/8\%$ از شبکه ۱ جدول ۱ کمتر است. بهترین ترکیب آنسامبلی این چهار شبکه با استفاده از روش‌های هاشم که ترکیب خطی نامقید بدون جمله ثابت است،

را دارد، ترکیب خطی بهینه حاصل از الگوریتم ژنتیک است که از MSE تخمین تداخل الگوهای آموزش و آزمون به ترتیب ۱۳٪ و ۱۱/۲٪ نسبت به بهترین شبکه عصبی مصنوعی منفرد حاصل کاسته است.

نتیجه گیری

به طور خلاصه می توان موارد زیر را از این تحقیق استنتاج نمود:

۱- ترکیب آنسامبلی با استفاده از روش های مناسب می تواند نتایج شبکه های عصبی مصنوعی منفرد پس انتشار خطای آموزش دیده به روش اعتبارسنجی همزمان را بهبود دهد.

۲- از روی نتایج ارائه شده در جداول این مقاله می توان گفت که روش ترکیب خطی بهینه هاشم بیشترین کاهش را در MSE تخمین تداخل الگوهای آموزش ایجاد کرده است، اما ترکیب های حاصل از این روش، توان تعمیم خوبی نداشته و در غالب موارد کاهشی در MSE تخمین تداخل الگوهای آزمون ایجاد نکرده است.

۳- از بین ۵ روش تحلیلی به کار رفته در این تحقیق، روش میانگین گیری ساده در مقایسه با چهار روش هاشم، کاهش بیشتری را در MSE تخمین تداخل الگوهای آزمون ایجاد کرده و نتیجه بهتری داشته است.

۴- در همه حالاتی که مورد بررسی قرار گرفته است، ترکیب های آنسامبلی خطی بهینه حاصل از روش الگوریتم ژنتیک نسبت به روش های تحلیلی، MSE تخمین تداخل الگوهای آزمون کم تر و نتایج بهتری داشته اند. این ترکیب ها توان تعمیم خوبی دارند، زیرا علاوه بر کاستن از MSE تخمین تداخل الگوهای آموزش، MSE تخمین تداخل الگوهای آزمون را نیز کاهش داده اند.

۵- بهترین ترکیب آنسامبلی حاصل در این تحقیق، ترکیب آنسامبلی خطی بهینه شبکه های ۱، ۳ و ۶ با استفاده از الگوریتم ژنتیک است که کمترین MSE تخمین تداخل الگوهای آزمون و بیشترین کاهش در MSE تخمین تداخل الگوهای آموزش نسبت به بهترین شبکه عصبی مصنوعی منفرد حاصل، به میزان ۱۲/۵٪ داشته است. این

حاصل، میانگین گیری ساده خروجی شبکه های ۱، ۲، ۳، ۵ و ۶ کمترین میانگین مربعات خطای تخمین تداخل الگوهای آزمون را داشت. ترکیب خطی بهینه حاصل از این ۵ شبکه با استفاده از الگوریتم ژنتیک نیز به دست آمد. جدول ۵ نتایج این مورد را نشان می دهد. اگر چه ترکیب خطی بهینه نامقید بدون جمله ثابت از MSE تخمین تداخل الگوهای آموزش به میزان ۲۰٪ کاسته است، اما در مورد الگوهای آزمون توان تعمیم خوبی ندارد و MSE تخمین این ترکیب برای الگوهای آزمون از بهترین شبکه عصبی منفرد بیشتر است. ترکیب خطی بهینه این ۵ شبکه با استفاده از الگوریتم ژنتیک به خوبی توانسته MSE تخمین تداخل الگوهای آموزش و آزمون را به ترتیب ۱۴٪ و ۱۱/۳٪ کاهش دهد و در بین ترکیب های پنج شبکه ای حاصل بهترین نتایج را داشته است.

جدول ۶ نتایج مربوط به ترکیب های شش شبکه ای را نشان می دهد. برای ۷ حالت شش شبکه ای ممکن از شبکه های جدول ۱، ترکیب های آنسامبلی خطی با استفاده از ۵ روش تحلیلی (مجموعاً ۳۵ ترکیب) ساخته شد که از بین آنها میانگین گیری ساده خروجی شبکه های ۱، ۲، ۳، ۴، ۵ و ۶ کمترین MSE تخمین تداخل الگوهای آزمون و در نتیجه بهترین نتیجه را دارد که ۲/۸٪ از بهترین شبکه عصبی منفرد حاصل کمتر است. از بین ۴ روش ترکیب خطی بهینه هاشم، ترکیب خطی بهینه مقید بدون جمله ثابت، کمترین MSE تخمین تداخل الگوهای آزمون را دارد که در عین حال از MSE تخمین بهترین شبکه عصبی منفرد بزرگ تر است و کاهشی نسبت به آن نداشته است، اگر چه این ترکیب MSE تخمین تداخل الگوهای آموزش را ۲۴/۴٪ کاهش داده است. ترکیب خطی بهینه این شش شبکه با استفاده از الگوریتم ژنتیک، بهترین ترکیب شش شبکه ای حاصل در این تحقیق است که از MSE تخمین تداخل الگوهای آموزش و آزمون نسبت به بهترین شبکه عصبی منفرد به ترتیب ۱۲/۹٪ و ۱۰/۲٪ کاسته است.

نتایج مربوط به تنها ترکیب هفت شبکه ای در جدول ۷ آورده شده است. در این مورد نیز نتایجی مشابه حالات قبلی به دست آمده است. بهترین ترکیب هفت شبکه ای حاصل که کمترین MSE تخمین تداخل الگوهای آزمون

ترکیب MSE تخمین تخلخل الگوهای آموزش را ۱۴/۷٪ نسبت به بهترین شبکه عصبی منفرد کاهش داده است. ۶- به طور کلی توانایی (یا عدم توانایی) ترکیب آنسامبلی در بهبود نتایج شبکه‌های عصبی مصنوعی منفرد و همچنین مقدار بهبود حاصل، از یک مسأله به مسأله دیگر متفاوت است و بنابراین در این مورد نمی‌توان یک اظهار نظر یا نتیجه‌گیری کلی بیان کرد. مجموعه عوامل مؤثر در این زمینه عبارتند از نوع و ماهیت مسأله، شبکه‌های منفرد به کار رفته در ترکیب و روش ترکیب.

تحقیقات نظری و کاربردی بیش‌تر با هدف رفع ابهامات در مورد نقاط قوت و ضعف رهیافت سیستم‌های چندشبکه‌ای ضروری است. از این رو تخمین تخلخل و سایر خصوصیات مخزنی از جمله تراوایی و اشباع سیالات مخزنی و همچنین تخمین نگارهای چاه در میادین هیدروکربوری گوناگون کشور با استفاده مناسب از ترکیب آنسامبلی یا ترکیب مدوله‌ای و یا تلفیق آنها پیشنهاد می‌شود.

تشکر و قدردانی

کارشناسان ارشد مدیریت اکتشاف شرکت ملی نفت، به خاطر راهنمایی‌های با ارزش به نگارندگان، تشکر و قدردانی می‌گردد.

علائم و نشانه‌ها

مراجع

- [۱]. حسنی پاک ع.ا.، شرف‌الدین م.، تحلیل داده‌های اکتشافی، انتشارات دانشگاه تهران، ۱۳۸۰.
- [2]. Bhatt A., *Reservoir properties from well logs using neural networks*, PhD thesis, Department of Petroleum Engineering and Applied Geophysics, Norwegian University of Science and Technology, Norway, 2002.
- [3]. Bhatt A., Helle H. B., "Committee neural networks for porosity and permeability prediction from well logs", *Geophysical Prospecting*, Vol. 50, pp. 645-660, 2002.
- [4]. Hashem Sh., *Optimal linear combinations of neural networks*, PhD thesis, School of Industrial Engineering, Purdue University, America, 1993.
- [5]. Hashem Sh., Schmeiser B. and Yih Y., *Optimal linear combinations of neural networks: an overview*, School of Industrial Engineering, Purdue University, America, 1994.
- [6]. Sharkey A., Sharkey N., Gerecke U. and Chandroth G. O., "The 'test and select' approach to ensemble combination", in: Kittler J., Roli F. (Eds.), *Multiple Classifier Systems, Lecture Notes in Computer Science*, Springer-Verlag, Inc., pp. 30-44, 2000.

MSE : میانگین مربعات خطا
 $MSE- OLC$: روش ترکیب خطی بهینه با تابع هدف MSE
 OLC : روش ترکیب خطی بهینه
 $r(\bar{x})$: مقدار مطلوب (واقعی) برای ورودی \bar{x}
 \bar{x} : بردار ورودی به شبکه عصبی مصنوعی منفرد و یا سیستم چند شبکه‌ای
 \bar{x}_n : n امین بردار ورودی به شبکه عصبی مصنوعی منفرد و یا سیستم چند شبکه‌ای
 $y_r(\bar{x})$: خروجی شبکه r ام جزء ترکیب آنسامبلی برای ورودی \bar{x}
 $\bar{y}(\bar{x})$: بردار خروجی‌های شبکه‌های جزء ترکیب آنسامبلی برای ورودی \bar{x}
 $\bar{y}(\bar{x})$: خروجی ترکیب آنسامبلی برای ورودی \bar{x}
 α_r : ضریب خروجی شبکه r ام جزء ترکیب آنسامبلی
 $\bar{\alpha}$: بردار ضرایب شبکه‌های جزء ترکیب آنسامبلی
 $\delta(\bar{x}; \bar{\alpha})$: خطای تخمین ترکیب آنسامبلی برای ورودی \bar{x} و بردار ضرایب $\bar{\alpha}$
 K_n : n امین الگوی تحقق یافته که در آن مقادیر \bar{x}_{rj} ، $r(\bar{x}_n)$ و $\bar{y}(\bar{x}_n)$ معین است
 K : مجموعه الگوهای تحقق یافته که معمولاً همان مجموعه الگوهای آموزشی انتخاب می‌شود.

- [7]. Fung Ch. Ch., Wong K. W. and Eren H., "Modular artificial neural network for prediction of petrophysical properties from well log data", IEEE transaction on instrumentation and measurement, Vol. 46, pp. 1295-1299, 1997.
- [8]. Chen Ch. H., Lin Z. Sh., "A committee machine with empirical formulas for permeability prediction", Computer & Geosciences., Vol. 32, pp. 485-496, 2006.
- [9]. Demuth H., Beale M. and Hagan M., *Neural network toolbox™ 6 user's guide*, MATLAB Software, www.mathworks.com, 2009.
- [10]. The Math Works™, *Genetic algorithm and direct search toolbox 2 user's guide*, MATLAB Software, www.mathworks.com, 2007.

[۱۱]. افشارحرب ع.، زمین‌شناسی نفت، انتشارات دانشگاه پیام نور، ۱۳۸۱.

[۱۲]. رحیمی م.، 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، ۱۳۸۱.

[۱۳]. زارعی س.، موحد ب.، باقری ع.م.، مردانی ع.، ارزیابی تراوایی سازند کنگان با استفاده از داده‌های نگار CMR و مغزه در میدان پارس جنوبی، بیست و ششمین گردهمایی علوم زمین‌شناسی، تهران، ایران، ۱۳۸۶.

Archive of SID