

## بررسی خواص مخزنی سازند سروک بر اساس نمودارهای پتروفیزیکی و ماشین‌های بردار پشتیبان بهینه‌سازی شده با الگوریتم ژنتیک

حسین سلیمی تیموری، دانشجوی کارشناسی ارشد مهندسی نفت دانشگاه آزاد اسلامی واحد طبس ■ مهدی حسین آبادی\* و کیوان شیعیان، استادیار و عضو هیات علمی دانشگاه آزاد اسلامی واحد طبس |

### چکیده

بررسی و پیش‌بینی خواص مخزنی از فعالیت‌های مهم و کلیدی برای ارزیابی ظرفیت و توانایی تولید از مخزن است. به دست آوردن پارامترهای مهم مخزنی نظیر تخلخل، می‌تواند در تعیین محل بهینه برای حفاری چاه‌های بعدی، طراحی و مدیریت فرآیندهای موثر در توسعه‌ی میداین نفت و گاز و شبیه‌سازی فرآیندهای ازدیاد برداشت مفید باشد. تعیین دقیق این پارامترها همیشه توأم با دشواری‌هایی است. معمولاً برای محاسبه‌ی این پارامترها از آنالیز مغزه‌های سنگ و یا داده‌های چاه‌آزمایی و چاه‌پیمایی استفاده می‌شود که البته مطمئن‌ترین روش برای محاسبه‌ی این پارامترها روش آنالیز مغزه است، اما به دلیل مشکلات اجرایی و هزینه‌های بالا، تعداد کمی از چاه‌های یک میدان، مغزه‌گیری می‌شوند. امروزه استفاده از سیستم‌های هوشمندی نظیر شبکه‌های عصبی جهت تخمین پارامترهای مخزنی، پیشرفت‌های بسیار چشمگیری داشته است.

در این پژوهش با استفاده از داده‌های پتروفیزیکی و الگوریتم ماشین‌های بردار پشتیبان بهینه‌سازی شده با الگوریتم ژنتیک به مدل‌سازی تراوایی و تخلخل سازند سروک در حوضه‌ی زاگرس پرداخته شده است. در این مطالعه رویکرد جدیدی جهت تخمین تراوایی با استفاده از حالات مختلف الگوریتم ماشین بردار پشتیبان معرفی شده است. برای این منظور، الگوریتم‌های SVRP و SSVR اجرا شده‌اند. بر اساس بهینه‌سازی انجام شده، روش‌های SVRL و SVRP به ترتیب با ضرایب همبستگی ۰/۷۹ و ۰/۸۱۶ به پیش‌بینی تراوایی سازند برای این مجموعه از داده‌ها بودند که کارایی بهتر روش SVRP را نشان می‌دهد.

### اطلاعات مقاله

تاریخ ارسال نویسنده: ۹۹/۰۲/۱۷

تاریخ ارسال به داو: ۹۹/۰۲/۲۰

تاریخ پذیرش داو: ۹۹/۱۱/۱۹

### واژگان کلیدی:

مدل‌سازی، ماشین‌های بردار پشتیبان، الگوریتم ژنتیک، تراوایی، تخلخل.

### مقدمه

هدف مشکل است، زیرا هیچ نگاری که بتواند مستقیماً تراوایی و تخلخل را در یک چاه اندازه بگیرد، توسعه پیدا نکرده است. [۳ و ۲] از آنجا که عمده‌ی مخازن نفتی تولیدی ایران از نوع کربناته بوده و این لایه‌ها از ناهمگنی بالایی برخوردار هستند، تخمین خواص پتروفیزیکی آنها در مناطق در حال توسعه، ضروری به نظر می‌رسد.

علاوه بر گران بودن مغزه‌گیری و انجام آزمایش‌های تزریق هلیوم و جیوه در برخی از چاه‌ها مانند چاه‌های افقی، مغزه‌گیری غیرممکن است ولی تقریباً از تمامی چاه‌ها نگارگیری انجام می‌شود. بنابراین بهترین راه‌حل برای رفع این مشکلات استفاده از این نگارها است و این امر صرفه‌جویی قابل توجهی در میزان هزینه‌ها و نیز زمان به همراه خواهد داشت.

برای پیش‌بینی و تخمین پارامترهای مخزنی به کمک داده‌های پتروفیزیکی، روش‌های متعددی وجود دارد که اغلب آنها مبتنی بر روش‌های کلاسیک آماری هستند. بعضی از این روش‌ها با توسعه‌ی تئوری یادگیری آماری سعی در بهبود ویژگی‌های خود داشته‌اند که از میان آنها می‌توان به روش رگرسیون ماشین بردار پشتیبان اشاره کرد. روش رگرسیون ماشین بردار

آگاهی دقیق از تخلخل و تراوایی به متخصصین این امکان را می‌دهد که برای بررسی دقیق مراحل تولید در یک میدان ابزارهای کافی در اختیار داشته باشند و با آگاهی از آنها توانایی بیشتری در توسعه‌ی میداین نفتی خواهند داشت، زیرا با داشتن اطلاعات بهتری از خصوصیات پتروفیزیکی مخزن می‌توان تصمیمات بهتری در مورد محل حفر چاه‌ها گرفت. [۱] تا به امروز سعی بر این بوده که مقادیر صحیح تراوایی و تخلخل، از طریق اندازه‌گیری‌های مستقیم آزمایشگاهی بر روی مغزه‌ها و یا تفسیر نتایج آزمایش چاه به دست آید.

البته هر دوی این روش‌ها صحیح بوده، اما برای توصیف کامل مخزن کافی نیستند، زیرا در هر میدان نفتی به دلیل وقت‌گیر بودن و نیز هزینه‌ی زیاد، فقط تعداد محدودی از چاه‌ها مغزه‌گیری شده و یا آزمایش چاه فقط به تعداد محدود در هر میدان نفتی توصیه می‌شود.

یکی از اهداف مطالعات پتروفیزیکی، تخمین دقیق و صحیح تراوایی و تخلخل در چاه‌هایی است که اندازه‌گیری تراوایی و تخلخل در آنها به هر دلیلی (مثل نبود مغزه یا وجود شکستگی در مغزه‌ها) ممکن نیست. رسیدن به این

\* نویسنده‌ی عهده‌دار مکاتبات (mhsedi1355@iau-tabas.ac.ir)

شده است. [۴] جنس این سازند، آهکی است و دارای تخلخل عمده‌ای از نوع شکستگی است. میدان مهم گازی این مخزن شامل میدان‌های بی‌بی حکیمه، دال‌پری و کوه مُند و میدان‌های مهم نفتی این مخزن شامل میدان‌های اهواز، بی‌بی حکیمه، کلیور کریم، سروستان و سیاه‌مکان است. [۵] سازند سروک در ایران دارای بیش از ۹۰ درصد، نفت خام در جا می‌باشد که دومین مخزن بعد از سازند آسماری در حوضه‌ی زاگرس را تشکیل می‌دهد.

افزون بر این، سازند سروک سفره‌های آب زیرزمینی و چشمه‌های زیادی را در حوضه‌ی زاگرس در خود جای داده است. [۶] این سازند به صورت هم‌شیب بر روی سازند کژدمی قرار گرفته است در حالی که مرز بالایی آن با سازند ایلام به طور ناپیوستگی فرسایشی است. از این رو مخزن سروک، افق اصلی در تولید نفت در نواحی نفت‌خیز ایران محسوب می‌شود که نرخ تولید در چاه‌های مختلف متغیر است که دلیل این تفاوت در خصوصیات پتروفیزیکی موجود در این سازند می‌باشد. با توجه به مخزنی بودن بخش‌های مختلف سازند سروک، بررسی در ناحیه‌ی مورد مطالعه ضرورت دارد. [۱]

میدان نفتی مورد مطالعه، بزرگ‌ترین میدان نفتی ایران و یکی از بزرگ‌ترین میدانی نفتی جهان است که کشف آن به سال ۱۳۷۶ باز می‌گردد. بر روی نقشه‌ی هم‌تراز سطحی، طول میدان در بخش ایرانی حدود ۴۵ تا ۶۰ کیلومتر، عرض آن ۱۷ کیلومتر و مساحت آن به طور تقریبی ۹۰ کیلومتر مربع است. مدیریت راهبردی این میدان به شرکت نفت و گاز اروندان واگذار شد که در حال حاضر ۵۰ هزار بشکه در روز از ۲۱ حلقه چاه، تولید می‌شود. [۱] تاق‌دیس میدان مورد مطالعه، در ناحیه‌ی دزفول شمالی در مجاورت خط مرزی ایران و عراق به فاصله‌ی ۵۰ کیلومتری جنوب غربی ساختمان جفیر در مجاورت ساختمان‌های مجنون و نهر عراق، ۸۰ کیلومتری غرب اهواز، ۶۰ کیلومتری شمال غرب میدان دارخوین و ۷۰ کیلومتری غرب سوسنگرد در منطقه‌ی هورالعظیم قرار گرفته است. این تاق‌دیس دارای روند شمالی-جنوبی است.

امتداد تاق‌دیس شمالی-جنوبی با ساختمان‌های دارخوین، خرمشهر هم‌جهت می‌باشد. [۷] ظرفیت این میدان ۳۰ تا ۳۳ میلیارد بشکه تخمین زده شده است که حدود ۶ درصد آن به صورت تخلیه‌ی طبیعی قابل برداشت خواهد بود. با اجرای توسعه‌ی کامل این میدان، روزانه ۲۶۰ هزار بشکه قابل تولید خواهد بود. در این میدان چهار لایه‌ی سروک، کژدمی، گدوان و فهلیان به عنوان سازندهای قابل بهره‌برداری شناخته شده است و اخیراً با حفر چاه شماره‌ی ۷ این میدان، سازندهای نفتی ایلام و لایه‌های ماسه‌سنگی کژدمی نیز حاوی نفت قابل تولید می‌باشند. سازند سروک دارای نفت سنگین است و سازندهای فهلیان، کژدمی و گدوان دارای نفت سبک هستند.

اهمیت میدان مورد مطالعه را بزرگی میدان از نظر وسعت، وجود مخازن نفتی مختلف، میزان نفت در جا و مجاورت آن با کشور عراق، می‌توان عنوان کرد و احتمال این که میدان مورد مطالعه با میدان مجنون عراق در بخش‌هایی مشترک باشد، وجود دارد. [۸]

پشتیبان از روش‌های موفق به کار گرفته شده در زمینه‌ی پیش‌بینی حل مسائل عددی است که در دهه‌ی اخیر مورد توجه اغلب مهندسين از جمله مهندسين نفت قرار گرفته است.

این روش که جزو روش‌های آموزش نظارت شده (Supervised) است، بین داده‌های ورودی و مقدار پارامتر وابسته، بر اساس کمینه‌سازی ریسک ساختاری، ارتباط برقرار می‌کند. [۲] دقت بالای این روش در یافتن تابع رگرسیون داده‌های با پیچیدگی بالا، یادگیری ساده نسبت به سایر الگوریتم‌های یادگیری ماشین، ارجحیت کمینه‌سازی ریسک ساختاری نسبت به روش متداول کمینه‌سازی ریسک تجربی الگوریتم‌هایی مانند شبکه‌ی عصبی و روش‌های کلاسیک آماری به سبب قابلیت تعمیم بهتر از برتری‌های روش رگرسیون ماشین بردار پشتیبان نسبت به سایر روش‌ها است. [۱۰] همچنین این روش بر خلاف روش‌هایی چون شبکه‌ی عصبی به پاسخ‌های موضعی همگرا نمی‌شود و در برخورد با حجم کم داده‌های آموزشی نسبت به سایر روش‌های کلاسیک، عملکرد بهتری دارد. [۱۱] روش رگرسیون بردار پشتیبان در مواجهه با مسائلی با ابعاد بالا (تعداد متغیرهای وابسته‌ی بالا) کارکرد مناسبی دارد، زیرا ابعاد داده‌های آموزشی مستقیماً در آموزش آن دخالت ندارند، بنابراین حجم محاسبات برای تعداد داده‌های بالا در حد معقولی است. [۸] در این مطالعه برای تخمین تراوایی و تخلخل از روی نگاره‌ی چاه از الگوریتم ماشین‌های بردار پشتیبان بهینه‌سازی شده با الگوریتم ژنتیک در سازند سروک استفاده شده است.

ویژگی‌های منحصربه‌فرد الگوریتم ژنتیک به یافتن مقادیر بهینه‌ی پارامترهای الگوریتم رگرسیون ماشین بردار پشتیبان کمک می‌کند و در نتیجه تابع رگرسیون به سرعت و با دقت مناسبی حاصل می‌شود. از مقایسه‌ی نتایج پیش‌بینی شده توسط تابع رگرسیون و مقادیر آزمایشگاهی پارامتر تراوایی و بررسی ضریب همبستگی میان تراوایی تخمینی و تراوایی آزمایشگاهی می‌توان دریافت که اگر اطلاعات کمی و کیفی نسبتاً خوبی از داده‌های نگاره‌ی چاه در اختیار باشد، الگوریتم رگرسیون ماشین بردار پشتیبان به کمک این داده‌ها در پیش‌بینی تراوایی موفق عمل می‌کند و می‌توان به کمک آن تراوایی را برای سایر نقاط مخزن که بعضاً ممکن است به سبب مشکلاتی نظیر شکستگی و... امکان مغزه‌گیری نداشته باشند، پیش‌بینی کرد.

## ۲- زمین‌شناسی منطقه‌ی مورد مطالعه

سازند سروک یکی از سازندهای زمین‌شناسی گروه بنگستان در زاگرس با سن کرتاسه‌ی میانه (آلبین-تورونین) است. این سازند یکی از مخازن مهم هیدروکربنی در حوضه‌ی زاگرس محسوب می‌شود. سازند سروک به صورت هم‌شیب بر روی سازند کژدمی قرار گرفته است، در حالی که مرز بالایی آن با سازند ایلام به صورت ناپیوستگی فرسایشی است. [۴] مقطع نمونه‌ی سازند سروک در تنگ سروک، در بخش مرکزی دامنه‌های جنوبی کوه بنگستان واقع در استان کهگیلویه و بویراحمد شهرستان بهمئی انتخاب

## ۳- رگرسیون بردار پشتیبان

ماشین‌های بردار پشتیبان به دو گروه عمده‌ی: الف) مدل طبقه‌بندی ماشین بردار پشتیبان و ب) مدل رگرسیون ماشین بردار پشتیبان تقسیم‌بندی می‌شوند. از مدل طبقه‌بندی ماشین بردار پشتیبان جهت حل مسائل طبقه‌بندی داده‌هایی که در کلاس‌های مختلف قرار می‌گیرند، استفاده می‌شود. مدل رگرسیون ماشین بردار پشتیبان در حل مسائل عددی پیش‌بینی کاربرد دارد.

همان‌گونه که بیان شد، ماشین بردار پشتیبان مبتنی بر مینیمم کردن ریسک ساختاری است. همچنین برای کاربرد ماشین‌های بردار پشتیبان در مسائل رگرسیون می‌توان از تابع خطایی استفاده کرد که خطاهایی که در یک فاصله‌ی معین به نام *Insensitive* از مقادیر واقعی می‌باشند، را نادیده می‌گیرد. [۹] این تابع به صورت رابطه‌ی ۱ تعریف می‌شود.

رابطه‌ی (۱):

$$L(y, f(x, \alpha)) = |y, f(x, \alpha)|_{\varepsilon} = \begin{cases} 0 & \text{if } |y, f(x, \alpha)| \leq \varepsilon \\ |y, f(x, \alpha)| - \varepsilon & \text{otherwise} \end{cases}$$

این تابع خطا، مقادیر خطای کمتر از  $\varepsilon$  را در نظر نمی‌گیرد. تابع رگرسیون به وسیله‌ی تابع رابطه‌ی ۲ تخمین زده می‌شود:

$$f(x) = \langle w, x \rangle + b \quad \text{رابطه‌ی (۲)}$$

که در معادله‌ی ۲،  $\langle \cdot \rangle$  ضرب داخلی است. تابع بهینه‌ی رگرسیون به وسیله‌ی مینیمم تابع زیر بیان می‌شود.

$$\varphi(w, \xi) = \frac{1}{2} \|w\|^2 + c \sum_{i=1}^N (\xi_i^+ + \xi_i^-) \quad \text{رابطه‌ی (۳)}$$

$$\text{Subject to } \begin{cases} y_i - (\langle w_i, x_i \rangle + b) \leq \varepsilon + \xi_i^+ \\ (\langle w_i, x_i \rangle + b) - y_i \leq \varepsilon + \xi_i^- \\ \xi_i^+, \xi_i^- \geq 0 \end{cases} \quad \text{رابطه‌ی (۴)}$$

که  $C$  یک مقدار از پیش تعیین شده، بوده و متغیرهایی هستند که قیود بالا و پایین خروجی سیستم را تعیین می‌کنند. اگر داده‌ها به صورت خطی مجزا از هم باشند، یک سطح بهینه که داده‌ها را بدون خطا و با حداکثر فاصله میان صفحه و نزدیک‌ترین نقاط آموزشی (بردارهای پشتیبان) تفکیک می‌کند، آموزش می‌دهد. اگر نقاط آموزشی را به صورت  $[x_i, y_i]$  و بردار ورودی را به صورت  $x_i \in \mathbb{R}^n$  تعریف کنیم، در حالتی که داده‌ها به صورت خطی قابل تفکیک باشند، معادله به صورت رابطه‌ی ۵ است:

$$y = f(x) = \text{sign}[\sum_{i=1}^N y_i \alpha_i \langle x_i, x \rangle + b] \quad \text{رابطه‌ی (۵)}$$

که در آن  $y$ ، خروجی معادله و  $y_i$  ارزش طبقه‌ی نمونه‌ی آزمایشی  $x_i$  است. بردار  $x = x_1, x_2, \dots, x_n$  نشان‌دهنده‌ی یک داده‌ی ورودی و بردارهای  $x_i, i=1, 2, \dots, N$  بردارهای پشتیبان هستند.

اگر داده‌ها به صورت خطی قابل تفکیک نباشند، می‌توان با اعمال پیش پردازش، نمونه‌ها را به فضایی بالاتر برد. در این حالت، رابطه‌ی ۵ به رابطه‌ی ۶ تغییر می‌یابد. [۱۰]

$$y = f(x) = \text{sign}[\sum_{i=1}^N y_i \alpha_i \cdot k(x_i, x_j) + b] \quad \text{رابطه‌ی (۶)}$$

تابع  $k(x_i, x_j)$  تابع کرنلی است که برای ایجاد ماشین‌هایی با انواع مختلفی از سطوح غیر خطی در فضای داده‌ها، تولید می‌کند. برای مدل رگرسیون بردار پشتیبان از کرنل‌های مختلفی استفاده می‌شود که عبارتند از: خطی، درجه‌ی دوم، تابع پایه‌ی شعاعی و چند جمله‌ای. معمولاً، تابع پایه‌ی شعاعی برای پیش‌بینی عملکرد بهتری دارد. معادله‌ی تابع پایه‌ی شعاعی که در این مقاله مورد استفاده قرار گرفته است به صورت رابطه‌ی ۷ است:

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right), \quad \sigma > 0 \quad \text{رابطه‌ی (۷)}$$

در ساخت یک مدل کارآمد رگرسیون بردار پشتیبان، پارامترهای مدل باید با استفاده از یک روش بهینه‌سازی به طور دقیق محاسبه شود. این پارامترها عبارتند از: نوع تابع کرنل، پارامتر تابع کرنل  $\sigma$ ، پارامتر تنظیم‌کننده‌ی  $C$  و پارامتر دقت  $\varepsilon$  مربوط به حداکثر خطا.

به طور کلی قابلیت روش رگرسیون بردار پشتیبان شدیداً به آموزش پارامترهای مورد اشاره بستگی دارد. مقدار پارامتر تنظیم‌کننده  $C = [2^{-5}, 2^{15}]$ ، پارامتر تابع کرنل  $\sigma = [2^{-5}, 2^3]$  و پارامتر دقت  $\varepsilon \in [0.01, 0.6]$  باید در این دامنه تنظیم شوند. [۲] انتخاب خوب پارامترها نقش برجسته و مهمی را در عملکرد مطمئن رگرسیون بردار پشتیبان ایفا می‌کند که می‌توان با استفاده از الگوریتم‌های بهینه‌سازی، مقدار بهینه‌ی این پارامترها را به دست آورد.  $\sigma$  ضرب کرنل برای توابع هسته‌ی چندجمله‌ای، سیگنوییدی و پایه‌ی شعاعی است.

هر چه مقدار  $\sigma$  بیشتر باشد، الگوریتم تلاش می‌کند برازش را دقیقاً بر اساس داده‌های آموزشی انجام دهد و این امر موجب تعمیم یافتن خطا و وقوع مشکل بیش برازش می‌شود.

$C$  پارامتر تنظیم‌کننده یا جریمه برای جمله‌ی خطا است. این پارامتر همچنین برقراری تعادل بین مرزهای تصمیم‌گیری هموار و پیچیدگی تابع رگرسیون نقاط داده‌ی آموزشی را کنترل می‌کند. پارامتر  $\varepsilon$  نیز حداکثر خطای قابل قبول را نشان می‌دهد. در صورتی که اختلاف مقادیر حقیقی و مقادیر پیش‌بینی شده توسط تابع رگرسیون از مقدار  $\varepsilon$  کمتر باشد، مدل رگرسیون بردار پشتیبان از این خطا صرف‌نظر می‌کند. در این مطالعه از الگوریتم ژنتیک برای پیدا کردن مقادیر بهینه‌ی پارامترهای رگرسیون بردار پشتیبان استفاده شده است.

معیار بهینه بودن این متغیرها دستیابی به بیشترین ضریب همبستگی میان داده‌های اندازه‌گیری شده و داده‌های تخمین زده شده توسط مدل و همچنین کمترین خطای میانگین مربعات است، زیرا در این صورت مدل ساخته شده در حین مواجهه با داده‌هایی که مدل بر اساس آنها ساخته نشده (داده‌های آزمایشی)، بهترین عملکرد را خواهد داشت. به همین دلیل برای تشخیص عملکرد مدل از شاخص‌های بیشترین ضریب

تصادفی درون داده‌های آزمایشی، بازه‌ی متغیر C، متغیر ε و متغیر σ تعیین می‌شود.

۲- اندازه‌گیری برازندگی هریک از داده‌های آموزشی.

۳- ایجاد چرخ رولت استاندارد برای انتخاب بهترین کروموزوم‌ها به منظور تبادل ژن‌ها و تشکیل کروموزوم نسل جدید با ژن‌های بهتر.

۴- در صورتی که نسل جدید داده‌های تولید شده به نحوی باشند که کمترین میزان میانگین مربعات و ضریب همبستگی مدنظر را داشته باشند، عملیات خاتمه می‌یابد، در غیر این صورت با جمعیت جدید به وجود آمده مراحل از آغاز مجدداً تکرار می‌شوند.

### ۱-۳- آماده‌سازی داده

آماده‌سازی داده‌ها اغلب پیچیده‌ترین بخش کاربرد الگوریتم‌های پیشین است. [۳ و ۴] قبل از ورود داده‌ها به ماشین بردار پشتیبان، یک سری تصحیح و عملیات مقدماتی بر روی داده‌ها جهت حصول جواب قابل اطمینان انجام گرفته است. این عملیات شامل حذف داده‌های نامناسب، هماهنگ‌سازی بین نتایج اندازه‌گیری‌ها در مغزه و نگارها و همچنین، تطابق عمق بین نگارهای چاه و مغزه است.

با توجه به این که آنالیز مغزه در اعماق خاصی (امکان مغزه‌گیری سالم) انجام شده، در ابتدا مقادیر تراوایی و تخلخل با نگارهای چاه با توجه به عمق مشترک (اعماق جدول ۱)، در یک فایل جداگانه مرتب شد که نمونه‌ای از آن در جدول ۲ دیده می‌شود.

در این پژوهش سعی شده با استفاده از حداکثر متغیرهای مرتبط و در دسترس (شامل نمودارهای پتروفیزیکی، طیف پرتو گاما، تخلخل، نوترون، چگالی سازند، صوتی، عامل فتوالکتریک، داده‌های تخلخل و...) به عنوان ورودی و همچنین استفاده از الگوریتم‌های یادگیرنده‌ی رگرسیون بردار پشتیبان رویکرد جدیدی برای تخمین تراوایی ارائه شود.

نوع تابع	عبارت ریاضی
خطی	$(u \times v) / n$
چندجمله‌ای	$(\gamma \times u \times C \epsilon_0)^{degree} / n$
سیگموئیدی	$\tanh(\gamma \times u \times v + C \epsilon_0) / n$
پایه‌ی شعاعی	$\exp(\gamma \times  u - v ^2) / n$

همبستگی و کمترین مقدار خطای میانگین مربعات در هر مرحله استفاده شده است. برای ترکیب الگوریتم ماشین بردار پشتیبان و الگوریتم ژنتیک در گام اول پارامترهای مدل SVR را برای ساختن کروموزوم‌ها به صورت تصادفی درون داده‌های آزمایشی کدگذاری می‌کنیم.

در ابتدا به منظور انجام کدگذاری در الگوریتم ژنتیک، باید یک بازه‌ی اولیه برای پارامترهای تنظیم‌کننده‌ی C، پارامتر تابع کرنل σ و پارامتر دقت ε تعیین شود. در این پژوهش بازه‌ی اولیه‌ی مقدار پارامتر تنظیم‌کننده  $C = [1, 100]$ ، پارامتر تابع کرنل  $\sigma \in [0, 1]$  و پارامتر دقت  $\epsilon \in [0, 0.001, 0.01]$  انتخاب می‌شوند. در گام بعدی باید برازندگی هر یک از داده‌های آموزشی اندازه‌گیری شود.

به منظور جلوگیری از مسئله‌ی بیش برازش، داده‌های آموزشی به صورت تصادفی به پنج دسته تقسیم می‌شوند. از چهار زیرمجموعه به عنوان داده‌ی آموزشی به همراه  $C$  و  $\sigma$  از قبل تعیین شده استفاده و تابع رگرسیون محاسبه می‌شود. عملکرد این مجموعه پارامترها به وسیله‌ی مجموعه‌ی پنجم باقی‌مانده با استفاده از معیار خطای میانگین مربعات سنجیده می‌شود. این روند برای هر یک از چهار مجموعه‌ی دیگر نیز تکرار می‌شود. داده‌های مربوط به رگرسیونی که کمترین خطای میانگین مربعات را دارند به عنوان داده‌های اصلح برگزیده می‌شوند و شانس بیشتری برای انتخاب خواهند داشت.

در گام سوم یک چرخ رولت استاندارد برای انتخاب بهترین کروموزوم‌ها به منظور تبادل ژن‌ها و تشکیل کروموزوم نسل جدید با ژن‌های بهتر ایجاد می‌شود. در صورتی که نسل جدید داده‌های تولید شده به نحوی باشند که کمترین میزان میانگین مربعات و ضریب همبستگی مدنظر را داشته باشند، عملیات خاتمه می‌یابد، در غیر این صورت با جمعیت جدید به وجود آمده، مراحل از آغاز مجدداً تکرار می‌شوند تا خروجی دلخواه مسئله که ساخت مدل رگرسیونی با کمترین خطای میانگین مربعات و بیشترین ضریب همبستگی است، به دست آید. پارامترهای C، ε و σ مربوط به مدل رگرسیون با خطای کمترین مربعات و بیشترین ضریب همبستگی به عنوان پارامترهای بهینه انتخاب می‌شود.

به صورت خلاصه سازوکار اصلی SVM در حل مسئله‌ی رگرسیون به این صورت بیان می‌شود:

**الف)** ماشین بردار پشتیبان، تابع رگرسیون را با به‌کارگیری یک دسته تابع خطی تخمین می‌زند.

**ب)** ماشین بردار پشتیبان، عملیات رگرسیون را با تابعی که انحراف از مقدار واقعی در آن، کمتر از تابع تلفات تعیین شده است، انجام می‌دهد.

**ج)** ماشین بردار پشتیبان با کمیته کردن ریسک ساختاری، بهترین جواب را ارائه می‌دهد. [۲ و ۳]

همچنین بهینه‌سازی الگوریتم رگرسیون ماشین بردار پشتیبان با استفاده از الگوریتم ژنتیک در ۴ مرحله‌ی زیر انجام می‌شود:

۱- کدگذاری پارامترهای مدل SVR برای ساختن کروموزوم‌ها به صورت

نمونه‌ای از فایل مرتب شده‌ی داده‌ها



Core depth	Log depth	GR	DT(US/T)	RHOB(G/CC)	NPHI(%)	PEF(U)	MSLF	LLD(OHMM)	LLS(OHMM)	KH(mD)	KV(mD)	LKH	LKV
۲۷۷۸/۰۸	۲۷۷۷/۰۲۲	۱۳/۱۲۲	۴۶/۲۳۸	۲/۹۱۷	۰/۷	۴/۲۲۶	۷۰۲/۱۵۷	۳۰۱/۷۴	۲۶۵/۸۷۹	۰/۰۱۶	۰/۰۰۹	- ۱/۷۹۵۹	- ۲/۰۴۵۸
۲۷۷۸/۰۵	۲۷۷۸/۸۹	۱۳/۳۹۳	۴۷/۴۹۶	۲/۹۲	۰/۹	۴/۶۴	۱۲۴/۷۸۸	۷۲/۲۱۶	۶۸/۲۸۴	۰/۰۱۵	۰/۰۱۱	- ۱/۸۲۳۹	- ۱/۹۵۸۶
۲۷۷۹/۱	۲۷۷۹/۱۵۶	۱۴/۹۷۷	۴۸/۴۲۸	۲/۸۶۵	۲/۳	۳/۷۹۷	۴۰/۳۲۵	۶۱/۰۸۴	۵۶/۶۸۲	۰/۰۳۶	۰/۰۰۸	- ۱/۴۴۳۷	- ۲/۰۹۶۹
۲۷۸۱/۶۳	۲۷۸۱/۵۹۴	۱۵/۵۴۸	۴۶/۷۷۲	۲/۶۳۵	۰/۶	۱۲/۶۷۹	۱۰/۱۸	۲۶۹/۴۱	۱۶۲/۱۹۳	۰/۰۰۸	۰/۰۰۶	- ۲/۰۹۶۹	- ۲/۲۲۱۸
۲۷۸۵/۷۸	۲۷۸۵/۷۰۸	۱۶/۹۸	۴۹/۳۰۲	۲/۸۸۸	۱	۳/۹۴۳	۵۵۵/۲۸۷	۵۶/۳۳۷	۵۱/۱۸۳	۰/۰۰۸	۰/۰۰۶	- ۲/۹۶۹	- ۲/۲۲۱۸
۲۷۸۸/۶۲	۲۷۸۸/۶۰۴	۱۰/۸۹۵	۴۷/۰۲۵	۲/۸۹۳	۰/۷	۳/۷۹۹	۴۵/۷۳۶	۲۳۱/۳۵۳	۲۱۴/۶۵۵	۰/۰۰۶	۰/۰۰۶	- ۲/۲۲۱۸	- ۲/۲۲۱۸
۲۷۸۹/۹۶	۲۷۸۹/۹۷۸	۱۲/۴۲۲	۴۶/۵۷۹	۲/۸۷۳	۰/۸	۴/۱۵۳	۱۰۵/۶۰۲	۳۱۵/۸۷۹	۲۹۲/۷۳۸	۰/۰۰۴	۰/۰۰۶	- ۲/۳۹۷۹	- ۲/۲۲۱۸
۲۷۹۰/۷۹	۲۷۹۰/۷۳۷	۱۲/۲۶۸	۴۷/۱۹۱	۲/۶۲۶	۰/۶	۸/۶۱۱	۰/۵۹۲	۲۴۶/۷۵	۲۰۹/۰۵۹	۰/۰۱	۰/۰۰۸	- ۲	- ۲/۰۹۶۹
۲۷۹۲/۶۴	۲۷۹۲/۵۶۶	۱۱/۲۴	۴۷/۴۱۹	۲/۶۸۷	۱/۸	۷/۷۶۱	۰/۶۲۳	۶۵۸/۲۲۶	۲۶۳/۲۴۲	۰/۰۰۷	۰/۱۵۲	- ۲/۱۵۹۴	- ۰/۸۱۸۲
۲۷۹۴/۶۷	۲۷۹۴/۷	۱۱/۹۷	۴۶/۳۳۶	۲/۹۲۴	۰/۷	۳/۷۳	۷۱۳/۵۴۲	۲۹۷/۴۷۱	۲۶۱/۲۷	۰/۰۱	۰/۰۴۲	- ۲	- ۱/۳۷۶۸
۲۷۹۶/۶۸	۲۷۹۶/۶۸	۱۴/۶۹۱	۴۹/۵۸۳	۲/۹۱۳	۱/۴	۴/۱۲۴	۱۷۷۲/۸۰۸	۴۲/۴۳۸	۴۸/۵۳۲	۰/۰۲۲	۰/۰۰۵	- ۱/۶۵۷۶	- ۲/۳۰۱
۲۸۰۰/۳۹	۲۸۰۰/۳۳۸	۲۵/۰۷۹	۴۹/۱۷۵	۲/۷۹۳	۳/۸	۴/۵۴۴	۱/۱۳۶	۴۳/۴۴	۵۲/۴۲۸	۰/۱۳۳	۰/۰۰۸	- ۰/۸۷۶۱	- ۲/۰۹۶۹
۲۸۰۳/۸۷	۲۸۰۳/۸۷	۱۱/۰۸	۴۸/۱۶۳	۲/۸۹	۱/۳	۴/۲۱۲	۱۰۴/۱۳۷	۲۳۹/۴۸۵	۲۰۷/۶۱۵	۰/۰۵۳	۰/۰۰۶	- ۱/۲۷۵۷	- ۲/۲۲۱۸

ماتریس ضرایب همبستگی بین متغیرهای مستقل و لگاریتم تراوایی



	LKH	LKV	X	Y	Z	GR	DT	RHOB	NPHI	PEF	MSFL	LLD	LLS
LKH	1												
LKV	۰/۹۵۷	1											
X	- ۰/۱۱۲	۰/۱۶۱	1										
Y	۰/۱۱	۰/۱۶۲	- ۰/۹۹۹	1									
Z	۰/۲۸۵	۰/۲۵۸	۰/۶۵۵	- ۰/۶۵۵	1								
GR	- ۰/۰۱۵	۰/۰۲۵	۰/۲۱۹	- ۰/۲۱۶	۰/۴۹	1							
DT	۰/۷۴۲	۰/۷۵۱	- ۰/۰۲۵	- ۰/۰۲۶	۰/۳۴۷	۰/۰۸۸	1						
RHOB	- ۰/۶۶۴	- ۶۵۹	- ۰/۰۱۱	- ۰/۰۱	- ۰/۳۷۹	- ۰/۱۸۲	- ۰/۹۱۵	1					
NPHI	۰/۸۲	۰/۸۴۸	- ۰/۱۱	۰/۱۰۷	۰/۲۴	۰/۱۶۷	۰/۷۴۶	- ۶۱۵	1				
PEF	۰/۱۲۵	۰/۰۹۸	۰/۱۲۵	- ۰/۱۱۹	۰/۳۶۱	۰/۲۱	۰/۳۵	- ۰/۵۷۶	۰/۰۰۹	1			
MSFL	- ۰/۳۱۸	- ۰/۳۱۱	۰/۰۷۲	۰/۰۷	- ۰/۱۶۴	- ۰/۲۵۹	- ۰/۳۲۷	۰/۴۰۱	- ۰/۳۱۸	- ۲۰۶	1		
LLD	- ۰/۱۰۶	- ۰/۱	۰/۱۷۷	- ۰/۱۱	۰/۱۰۶	- ۰/۰۴۱	- ۰/۱۵۳	۰/۱۶۹	- ۰/۱۰۱	- ۰/۰۷۸	۰/۲۷۴	1	
LLS	- ۰/۲۰۶	- ۰/۱۸۷	۰/۰۷۷	- ۰/۰۸	- ۰/۰۲۳	- ۰/۱۰۲	- ۰/۲۳۹	۰/۲۴۸	- ۰/۲۰۳	- ۰/۱۱۵	۰/۲۳۵	۰/۷۶۷	1

جدول ۴ نمونه‌ای از فایل نرمال شده‌ی داده‌های ورودی که به صورت تصادفی از داده‌های موجود انتخاب شده

R (ضریب رگرسیون)	۰,۳۷۹	۰,۲۱۸	۰,۸۳۸	۰,۷۴۵	۰,۸۹۱	۰,۲۶۵	۰,۴۲۶	۰,۱۸۶	۰,۲۸۹
Z	GR	DT	RHOB	NPHI	PEF	MSFL	LLD	LS	
	۰/۴۳۴۷۵	- ۰/۵۹۱۲۸	۰/۹۵۷۲۷	۰/۹۴۴۲۴	۰/۹۴۵۵۳	- ۰/۸۵۳۵۹	- ۰/۲۹۸۰۶	۰/۹۳۲۶۵	- ۰/۹۲۷۸۷
	۰/۴۲۷۱۵	۰/۴۹۶۲	- ۰/۸۴۷۱۸	۰/۸۱۸۱۸	- ۰/۸۲۱۰۱	- ۰/۹۴۷۶۶	- ۰/۹۶۰۰۱	- ۰/۹۶۶۷۴	۰/۹۶۵۱۹
	۰/۹۱	- ۰/۴۶۶۹۳	۰/۹۳۰۴۳	۰/۲۶۰۶۱	- ۰/۹۵۳۳۱	۱	- ۰/۹۹۰۱۶	- ۰/۹۳۹۹۲	- ۰/۹۵۶۲۸
	۰/۴۰۳۸	۰/۳۹۳۵۳	۰/۸۰۳۲۵	۰/۸۷۳۹۴	۰/۹۲۲۱۸	- ۰/۹۱۵۶۴	- ۰/۴۴۴۹۶	۰/۹۸۷۸	۰/۹۸۶۴
	- ۰/۳۶۴۷۱	- ۰/۵۱۰۶۶	۰/۷۸۹۱۲	۰/۹۳۴۵۵	۰/۸۹۱۰۵	- ۰/۸۷۵۹۵	۰/۷۷۲۷۷	۰/۹۹۰۹۳	- ۰/۹۸۷۴۳
	۰/۳۳۱۹۱	- ۰/۶۵۵۹۵	- ۰/۸۶۰۵۵	۰/۸۷۸۷۹	۰/۸۹۸۸۳	- ۰/۸۵۶۶۶	۰/۸۶۹۱۸	- ۰/۹۴۶۶۴	۰/۹۴۳۸۴
	- ۰/۳۳۴۸۴	- ۰/۰۷۶۵۶	۰/۷۳۹۷۶	۰/۸۰۱۲۱	- ۰/۷۸۲۱	- ۰/۹۲۸۸	- ۰/۹۵۵۱	۰/۹۹۰۳۷	- ۰/۹۸۷۵
	- ۰/۳۳۱۵۹	- ۰/۶۳۳۰۵	۰/۸۵۲۸۶	۰/۸۷۱۵۲	- ۰/۹۲۲۱۸	- ۰/۹۴۵۹	- ۰/۹۰۰۳۴	۰/۹۶۱۱۲	۰/۹۵۸۵۱
	۰/۲۶۴۸	- ۰/۸۱۵۱۷	۰/۰۶۳۳۸۹	- ۰/۱۷۵۷۶	۰/۱۲۰۶۲	- ۰/۷۲۰۲۶	- ۰/۹۹۷۵۳	- ۰/۹۹۴۲	- ۰/۹۹۸۸۹
۱	- ۰/۶۴۳۴۶	- ۰/۶۵۷۶۲	۰/۳۳۰۹۱	- ۰/۸۶۸۸۹	- ۰/۳۱۷۲۲	- ۰/۹۹۷۲۲	۰/۹۴۶۳۷	۰/۹۵۵۶۲	

کروی میکرو (MSFL)، نگار مقاومت ویژه‌ی القایی عمیق (LLD)، نگار مقاومت ویژه‌ی القایی کم‌عمق (LLS)، نگار شاخص فتوالکتریک (PEF) و نگار مقاومت ویژه‌ی واقعی سازند (KH)، تراوایی افقی، (KV) تراوایی عمودی، (LKH) لگاریتم تراوایی افقی و (LKV) لگاریتم تراوایی عمودی باشد.

خروجی شبکه هم ماتریس تک سطری یا تک ستونی از تراوایی (قائم یا افقی) است که می‌توان آنها را در محیط برنامه‌ای به طور دلخواه مرتب کرد. با توجه به جدول ۳ می‌توان به همبستگی مثبت بالا بین لگاریتم تراوایی و نگارهای نوترون اصلاح شده و همچنین همبستگی منفی بالا بین لگاریتم تراوایی با نگار چگالی پی برد. نگارهای دیگر دارای همبستگی متوسطی با تراوایی هستند.

همچنین برای تخمین تراوایی از نرم‌افزار متلب و جعبه ابزار مربوط به SVM استفاده شده که تمامی حالات و مقادیر ممکن و قابل آزمایش توسط برنامه‌نویسی و ایجاد حلقه‌های متناسب مورد بررسی و تحلیل قرار گرفته است، به نحوی که درصدهای مختلفی از اطلاعات را انتخاب کرده و آنها را توسط توابع هسته‌ی خطی و سیگموئیدی آموزش داده است. در نهایت از مجموع ۳۱۶ داده‌ی موجود به عنوان الگو برای سازند، تعداد ۱۲۲ داده برای آموزش انتخاب شد. هر بار درصد متفاوتی از این تعداد داده به الگوریتم ماشین بردار پشتیبان آموزش داده می‌شد.

با افزایش تعداد داده‌های آموزشی، ضریب همبستگی تابع رگرسیون نیز افزایش می‌یافت. برای جلوگیری از مشکل بیش برآزش، حد بهینه‌ای از تعداد داده‌ها برای این منظور انتخاب شدند که تعداد آنها ۱۲۲ داده بود.

در این مطالعه ابتدا اطلاعات نمودارگیری و مغزه در ۴ چاه A, B, C و D جمع‌آوری شد. سپس این داده‌ها و اطلاعات بر اساس متغیرهای ورودی و متغیر خروجی پالایش و داده‌های اضافه حذف شد. سپس با استفاده از فرمول‌های مربوط به نرمالیزه کردن، داده‌های خارج از محدوده نیز حذف و در نهایت اطلاعات هر ۳ چاه در یک پایگاه داده (شامل ۱۲۲ نقطه) جمع‌آوری شد.

تغییرات بعضی از نگارها در مقابل عمق برای ۴ چاه با توجه به عمق شروع و خاتمه‌ی ۲ سازند، در جدول ۲ آمده است. در نهایت تعداد ۳۱۶ داده (الگو) برای این سازند، برای طراحی الگوریتم، مورد استفاده قرار گرفت.

سپس توسط نرم‌افزار SPSS مقادیر همبستگی بین متغیرهای مستقل (نگارهای چاه) و متغیرهای وابسته (لگاریتم تراوایی و تخلخل) به دست آمد. [۶۵] چون بعد از بررسی مشخص شد نگارها با تراوایی رابطه‌ی لگاریتمی، اما با تخلخل رابطه‌ی خطی دارند، بنابراین مقادیر تراوایی به صورت لگاریتمی در نظر گرفته شده و نتایج آن در جدول ۳ آورده شده است.

با توجه به ماتریس همبستگی بین متغیرها، ورودی‌های لازم می‌تواند انتخاب شود که می‌توان آنها را در محیط برنامه‌نویسی برای طراحی شبکه به طور دلخواه مرتب کرد.

ورودی شبکه می‌تواند ماتریسی از بین متغیرهای طول جغرافیایی (X)، عرض جغرافیایی (Y)، عمق مغزه‌گیری شده (Core Depth)، عمق نمودارگیری (Log Depth)، نگار صوتی (DT)، نگار گاما (GR)، نگار نوترون اصلاح شده (NPHI)، نگار چگالی (RHOB)، نگار مقاومت ویژه‌ی



حذف می‌شوند. نتیجه‌ی این عملیات ۱۲۲ داده می‌شود که در آموزش مدل رگرسیون بردار پشتیبان استفاده می‌شوند. نمونه‌ای از این داده‌ها در جدول ۴ آورده شده است.

برای نرمال کردن داده‌ها در بازه‌ی ۱ تا -۱ از دستور  $premnmx$  در طراحی شبکه پس انتشار (BPNN) استفاده شد که از فرمول: 
$$p_n = 2 \left( \frac{p - P_{MIN}}{P_{MAX} - P_{MIN}} \right) - 1$$
 برای نرمال کردن داده‌ها استفاده می‌کند که  $p$  داده‌ی اولیه و  $P_n$  داده‌ی نرمال شده است.

البته در این مطالعه فقط داده‌های ورودی، نرمال شده‌اند. تخلخل به صورت درصد و تراوایی با همان حالت اولیه یعنی لگاریتم تراوایی به کار رفته است. نمونه‌ای از فایل داده‌های ورودی در جدول ۴ آمده است.

### بحث و نتایج

ضریب همبستگی بین تراوایی پیش‌بینی شده و تراوایی افقی برای سه مجموعه داده‌ی آموزشی، اعتبارسنجی و تست، با استفاده از تابع هسته‌ی خطی محاسبه و در شکل ۱ و برای تراوایی قائم شکل ۲ نشان داده شده است.

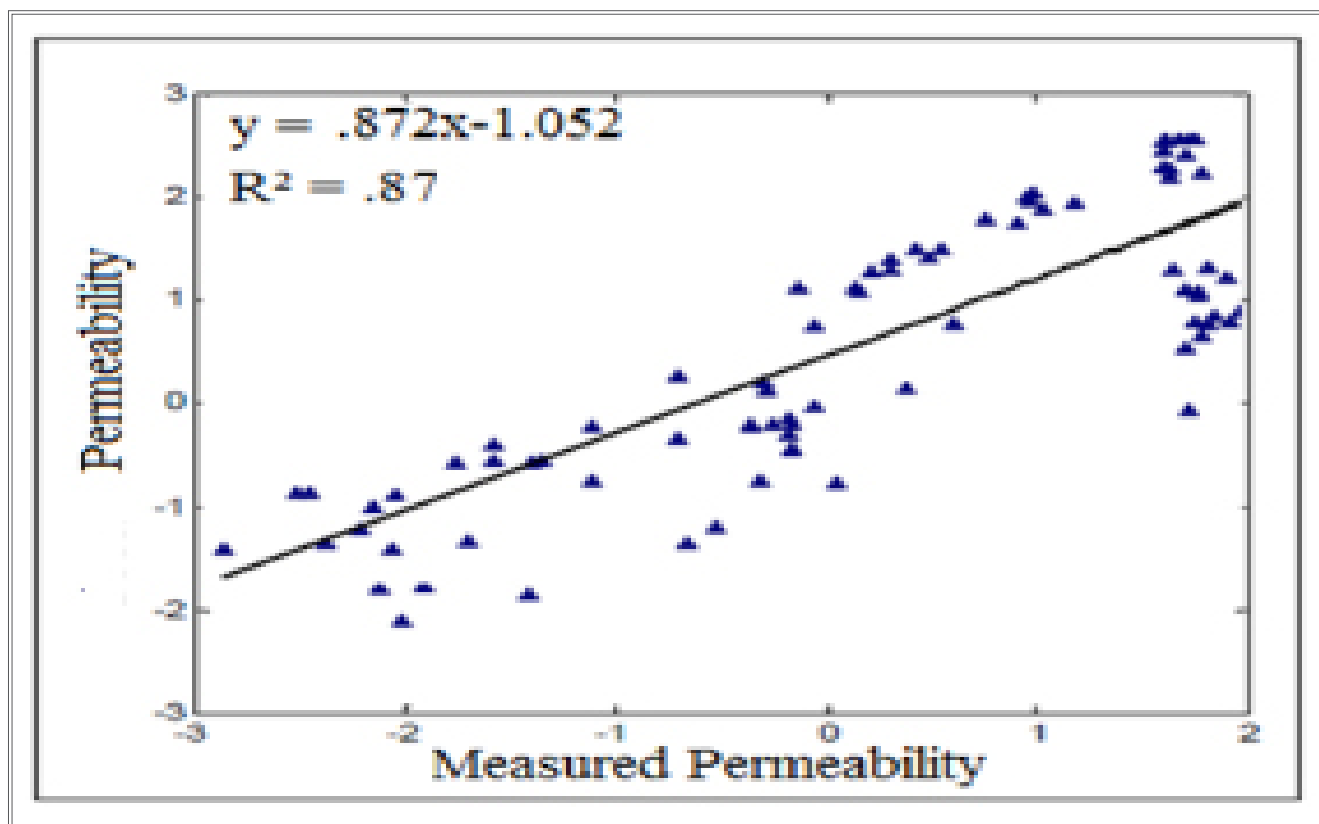
برای در نظر گرفتن تمامی پارامترهای تاثیرگذار بر پارامتر تراوایی از اطلاعات مربوط به تمامی متغیرهای پتروفیزیکی ذکر شده در جدول ۲ استفاده شد.

به این ترتیب مدل ساخته شده از جامعیت بیشتری بر خوردار شد و پارامتر تراوایی را با استفاده از پارامترهای مستقل تاثیرگذار توصیف می‌کرد. جدول ۳ اهمیت هر یک از این پارامترها در پیش‌بینی مقدار تراوایی را نشان می‌دهد.

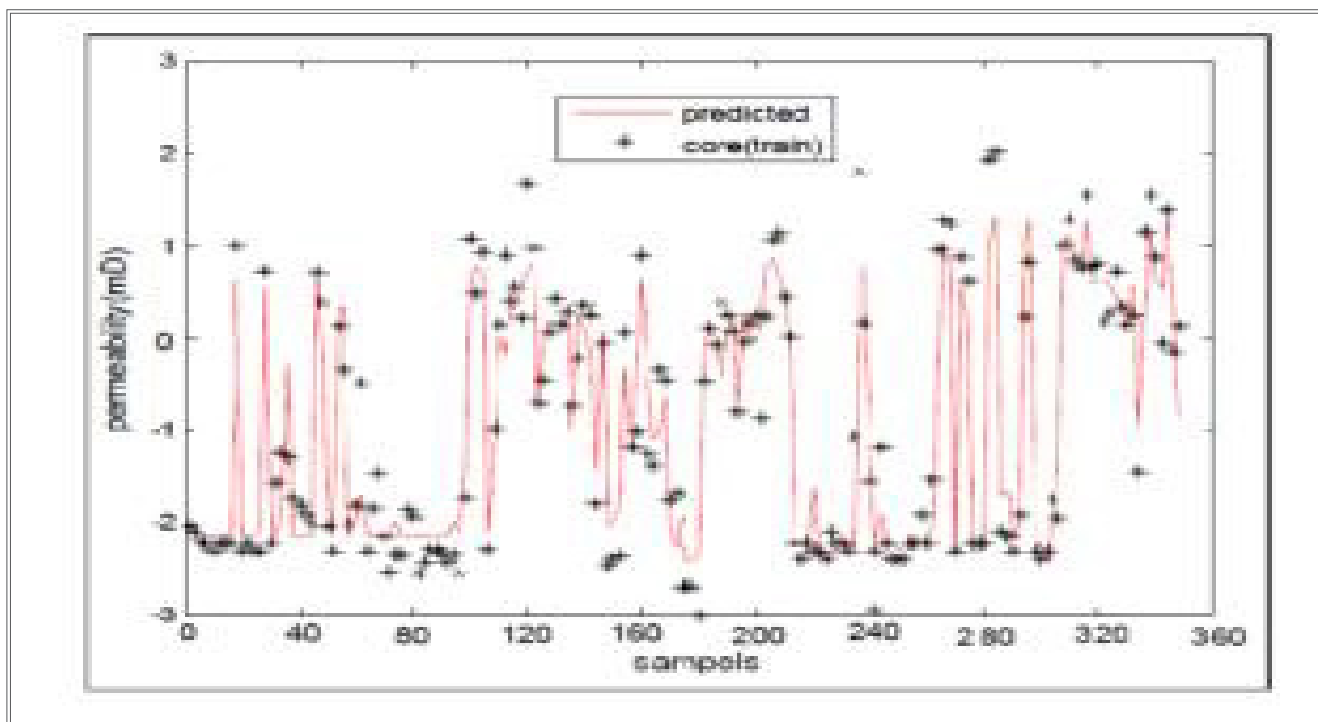
از آنجا که در این پژوهش هدف ساخت مدلی جامع برای توصیف تراوایی است، از تمامی متغیرها، حتی متغیرهای وابسته‌ای که ضریب همبستگی آنها کم است، نیز استفاده می‌شود.

البته متغیرهای مستقلی نظیر تخلخل نوترونی، نمودار صوتی و چگالی سازند به ترتیب بیشترین ضریب همبستگی را دارند که نشان‌دهنده‌ی وزن بیشتر این متغیرها در تعیین تابع تراوایی است. انتخاب ۱۲۲ داده به عنوان داده‌ی آموزشی با هدف بیشتر شدن ضریب همبستگی متغیرهای مستقل نسبت به پارامتر تراوایی انجام می‌شود.

به بیانی دیگر تنها دسته‌ای از هر یک از متغیرهای مستقل انتخاب می‌شود که ضریب همبستگی بالاتری با پارامتر تراوایی دارند و سایر آنها



شکل ۱ | مقایسه‌ی همبستگی تراوایی افقی شبکه با تراوایی مغزه در داده‌های اعتبارسنجی رگرسیون بردار پشتیبان با تابع هسته‌ی خطی



شکل ۲ | مقایسه‌ی بین لگاریتم تراوایی افقی شبکه با تراوایی مغزه در داده‌های تست رگرسیون بردار پشتیبان با تابع هسته‌ی خطی

اجرای رگرسیون بردار پشتیبان با تابع هسته‌ی پایه‌ی شعاعی است که داده‌های مغزه با علامت ستاره و داده‌های پیش‌بینی به صورت منحنی نشان داده شده است.

ضریب همبستگی بین داده‌های پیش‌بینی و داده‌های مغزه برای سه سری داده‌های آموزش، اعتبارسنجی و تست برای تعیین تراوایی به ترتیب برابر ۰/۸۷، ۰/۹۰ و ۰/۹۱ است. با توجه به ضریب همبستگی بالای سه دست آمده توسط شبکه در سه مرحله، می‌توان به قابلیت خوب شبکه‌ی طراحی شده اشاره کرد.

در شکل ۴ نیز که مقایسه‌ی بین لگاریتم تراوایی افقی شبکه با لگاریتم تراوایی مغزه در سه سری داده در مقابل تعداد داده‌های هر سری آورده شده است و داده‌های مغزه با علامت ستاره و داده‌های پیش‌بینی به صورت منحنی نشان داده شده‌اند، مشاهده می‌شود که علامت ستاره به منحنی داده‌های پیش‌بینی خیلی نزدیک و در بیشتر موارد مماس شده است.

این نتایج نیز به قابلیت خوب شبکه‌ی طراحی شده اشاره می‌کنند. بنابراین شبکه‌ی عصبی طراحی شده به روش توقف سریع برای پیش‌بینی تراوایی موفق عمل کرده است و نتایج آن با داده‌های مغزه بسیار نزدیک هستند و دارای اعتبار بالایی می‌باشند. بنابراین از این شبکه می‌توان برای تعمیم به چاه‌هایی که مغزه‌گیری انجام نشده و تراوایی نامشخص است، استفاده کرد.

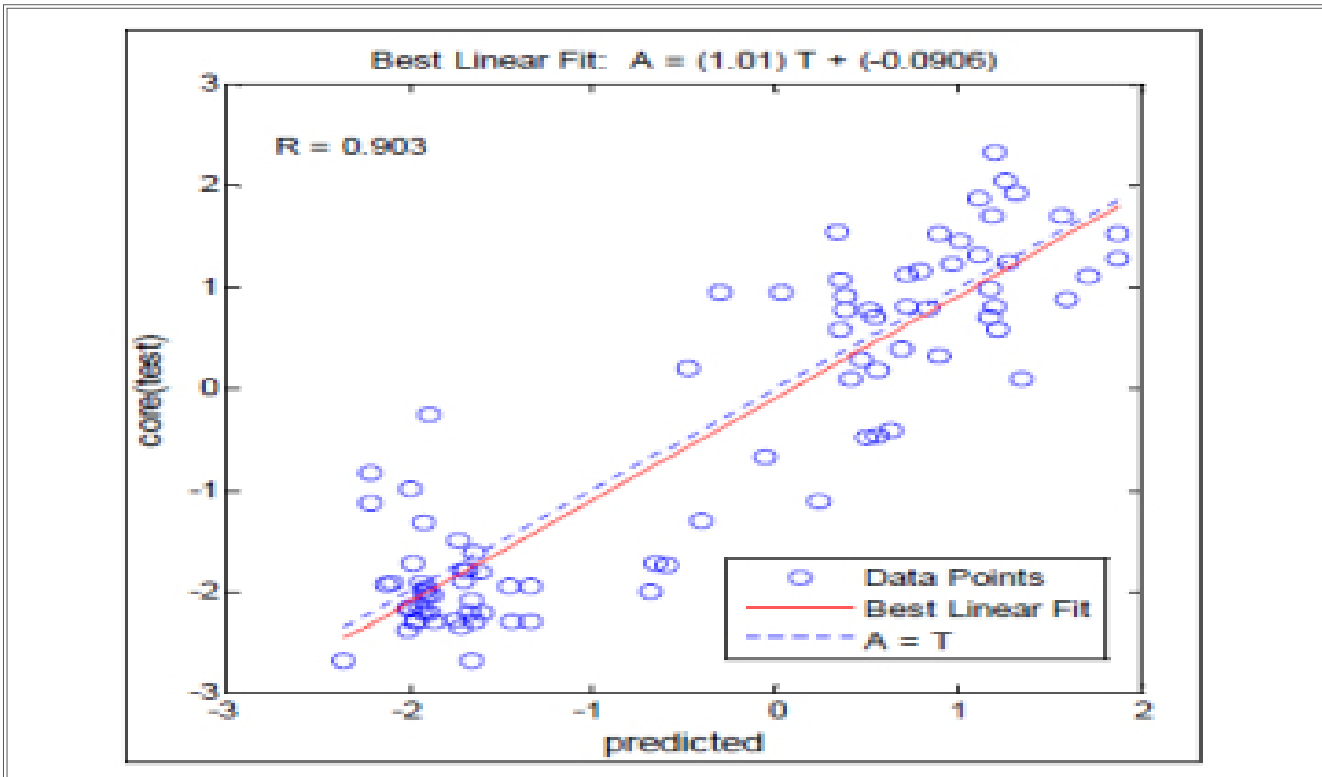
همان‌طور که در شکل ۱ دیده می‌شود، هر کدام از این نگارها دارای خط برازش شده‌ای بین پاسخ شبکه و داده‌های مغزه هستند. این خط دارای یک شیب و ضریب همبستگی است که پراکندگی داده‌ها را حول خط برازش شده، نشان می‌دهند.

هر چه مقدار ضریب همبستگی و شیب خط به عدد ۱ نزدیک‌تر باشند (A=T)، پراکندگی داده‌ها حول خط برازش کمتر و نشان از پیش‌بینی خوب است. ضریب همبستگی بین داده‌های پیش‌بینی و داده‌های مغزه برای سه سری داده‌های آموزش، اعتبارسنجی و تست به ترتیب برابر ۰/۹۵، ۰/۹ و ۰/۹۲ است.

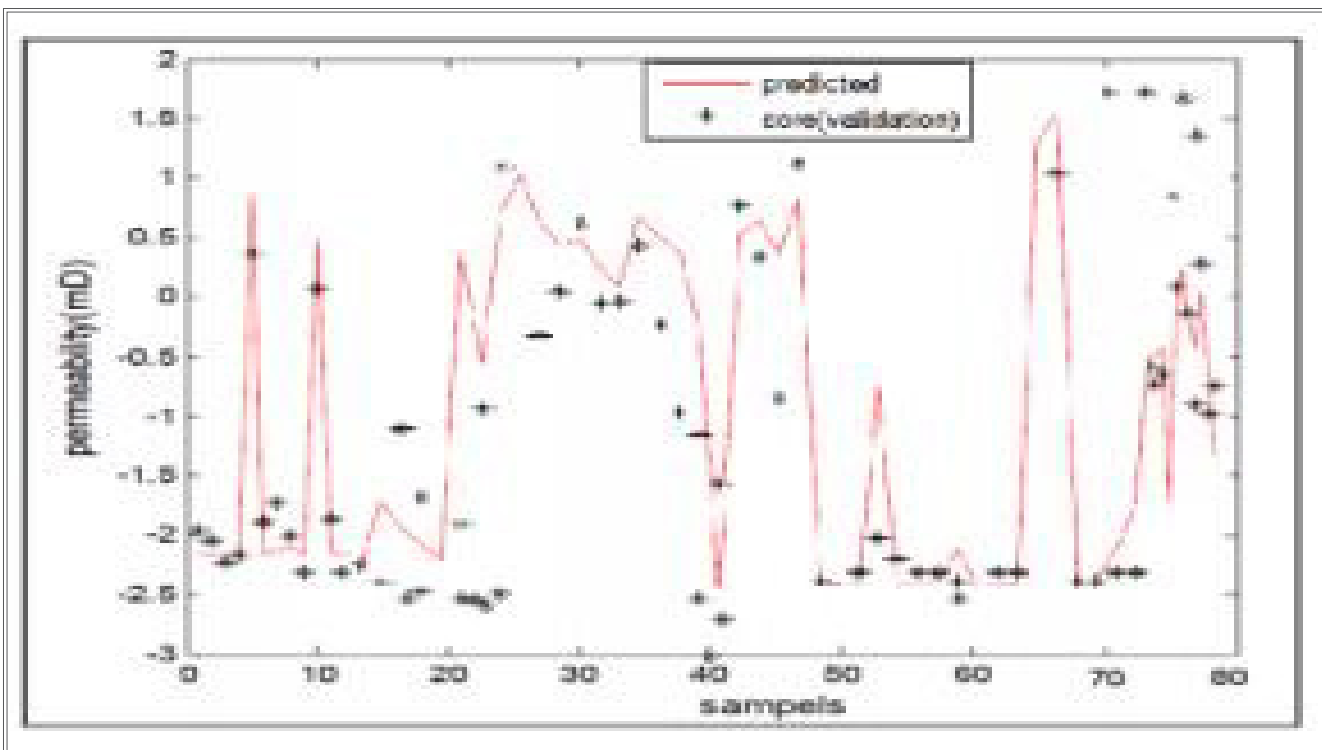
با توجه به ضریب همبستگی بالای سه دست آمده توسط شبکه در سه مرحله، می‌توان به قابلیت خوب شبکه‌ی طراحی شده اشاره کرد. در شکل ۲ نیز که مقایسه بین لگاریتم تراوایی افقی شبکه با لگاریتم تراوایی مغزه در سه سری داده در مقابل تعداد داده‌های هر سری آورده شده است و داده‌های مغزه با علامت ستاره و داده‌های پیش‌بینی شده به صورت منحنی نشان داده شده‌اند، مشاهده می‌شود که علامت ستاره به منحنی داده‌های پیش‌بینی در هر سه مورد داده‌های مرحله‌ی آموزش، تست و اعتبارسنجی خیلی نزدیک و در بیشتر موارد مماس شده است.

این نتایج نیز به قابلیت خوب شبکه‌ی طراحی شده اشاره می‌کنند. اشکال ۳ و ۴ هم نشانگر مقایسه‌ی بین لگاریتم تراوایی شبکه با لگاریتم تراوایی مغزه در سه سری داده در مقابل تعداد داده‌های هر سری برای





شکل ۳ | همبستگی بین لگاریتم تراوایی شبکه با تراوایی مغزه در داده‌های تست اجرای رگرسیون بردار پشتیبان با تابع هسته‌ای (SVRP)



شکل ۴ | مقایسه‌ی بین لگاریتم تراوایی شبکه با تراوایی مغزه در داده‌های تست اجرای رگرسیون بردار پشتیبان با تابع هسته‌ای (SVRP)

### نتیجه گیری

علاوه بر این در بهترین حالت با آموزش ۸۰ درصد کل داده‌های موجود می‌توان با استفاده از رگرسیون ماشین بردار پشتیبان یکنواخت، تراوایی را با ضریب همبستگی ۹۴/۲ درصد و خطای میانگین مربعات ۰/۰۰۱۱ تخمین زد. مقادیر واقعی تراوایی حاصل از اندازه‌گیری‌های آزمایشگاهی روی نمونه‌های مغزه‌ی سنگ هستند. بر اساس بهینه‌سازی انجام شده، روش‌های SVRL و SVRP به ترتیب با ضرایب همبستگی ۰/۷۹ و ۰/۸۱۶ به پیش‌بینی تراوایی سازند برای این مجموعه از داده‌ها بودند که کارایی بهتر روش SVRP را نشان می‌دهد. ■

در این روش از مجموعه‌ی داده‌ای یکسان شامل سری کامل داده‌های پتروفیزیکی (نمودارهای مقاومتی، صوت، هسته‌ای، تخلخل و اشباع) مخزن آماری در سازند سروک استفاده شد. نتایج این تحقیق نشان داد که اگر اطلاعات کمی و کیفی نسبتاً خوبی از داده‌های نگارهای چاه در اختیار باشد، می‌توان خواص مخزن مانند تراوایی را به خوبی و با ضریب خطای نسبتاً کمتری برای داده‌های جدید با داده‌های دچار اشکال شده محاسبه کرد.

### منابع

- [1]. فریدونی، ب، مختاری، م، ۱۳۹۶، شبیه‌سازی پارامترهای مخزنی تخلخل و تراوایی با استفاده از روش شبیه‌سازی گوسی متوالی توام در یکی از میادین نفتی خلیج فارس، نشریه علمی- پژوهشی زمین شناسی نفت ایران، شماره ۱۳، ص ۹۴-۷۸.
- [2]. پاک‌پرور، س، ادهمیان، آ، ۱۳۹۶، بررسی کیفیت مخزنی و عوامل کنترل‌کننده‌ی آن در بخش مخزنی سازند سروک در یکی از میادین جنوب‌غربی ایران، ماهنامه‌ی علمی-ترویجی اکتشاف و تولید نفت و گاز، شماره‌ی ۱۴۴، ص ۶۰-۵۴.
- [3]. همتی‌نیک، ح، فضلی، ل، نظری فرد، م، طباطبائی رئیس‌ی، م، ۱۳۹۶، مدل‌سازی سه بعدی زمین‌شناسی سازندهای ایلام و سروک در یکی از میادین نفتی، ماهنامه‌ی علمی-ترویجی اکتشاف و تولید نفت و گاز، شماره‌ی ۱۴۸، ص ۶۵-۵۸.
- [4]. Siamak Nazari, Heidi A. Kuzma, 2011, Predicting Permeability from Well Log Data and Core Measurements Using Support Vector Machines, SEG San Antonio 2011 Annual Meeting, pp. 2004-2008.
- [5]. Mohseni, H., & Javanmard, R. Z, 2020, New data on sequence stratigraphy of the Sarvak Formation in Malekshahi city, (Ilam province) Zagros basin, Iran. Marine and Petroleum Geology, 112, 104035.
- [6]. Cristianini Nello., Shawe Taylor John., 2003, An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods, Cambridge University Press, Chapter 6.
- [7]. Te-Ming Hung, Vojislav Kecman, Ivica Kopriva, 2006, Kernel Base Algorithms for Mining Huge Data Set, Springer, Chapter 2.
- [8]. Ethem. Alpaydin, 2010, Introduction to Machine Learning Second Edition, The MIT Press Cambridge, pp. 21361.
- [9]. Bishop.M, Octobr 2007, Pattern Recognition and Machine Learning, Springer, Vol 3, pp. 137154, Vol 6, pp. 291311-, Vol 7, pp. 326356.
- [10]. Moradi, M., Tokhmechi, B., & Masoudi, P, 2019. Inversion of well logs into rock types, lithofacies and environmental facies, using pattern recognition, a case study of carbonate Sarvak Formation. Carbonates and Evaporites, 34(2), pp. 335-347.
- [11]. Vanderlooy.S, 2008, Matlab Toolbox for Machine Learning, Maastricht University of the Netherlands Technical Report 0803-, Vol 6.2.
- [12]. Chih-Chung Chang, Chih-Jen Lin, March 4, 2013, LIBSVM: A Library for Support Vector Machines, Department of Computer Science National Taiwan University, pp. 135-138.
- [13]. Arbogast, J.S., Franklin, M.H, 1999. Artificial neural networks and high speed resistivity modeling software speed reservoir characterization, Petroleum engineering international, Vol. 72, No. 5, pp. 57-61.
- [14]. Wang, P., Tamhane, D., Aminzadeh, F, 2000. A soft computing Approach to Integrate Well logs and geological clusters for Petrophysical Prediction, Proceedings 3rd International Conference on Petroleum Geophysics, New Delhi, India.
- [15]. Helle, H.B., A. Bhatt, and B. Ursin, 2001. Porosity and permeability prediction from wireline logs using artificial neural networks: a North Sea case study. Geophysical Prospecting”, 49(4): pp. 431-444.
- [16]. Huang, Y., Wong, P.M., Gedeon, T.D, 1998. Prediction of reservoir permeability using genetic algorithms, I Applications, Vol. 12, No. 13, pp. 67- 75.
- [17]. Assadi, A., Honarmand, J., Moallemi, S. A., Abdollahie-Fard, I., 2018. An integrated approach for identification and characterization of paleo-exposure surfaces in the upper Sarvak Formation of Abadan Plain, SW Iran. Journal of African Earth Sciences, Vol. 145, pp. 32-48.