نشریه علمی–پژوهشی زمین شناسی نفت ایران
Iranian Joural of Petrolum Geology

سال دهم، شماره ۲۰، پائیز و زمستان ۱۳۹۹ص۲۹-۴۰
No. 20, Autumn & Winter 2020, pp 29-40

طراحی مدل گروهی تخمین تراوایی مخزن هیدروکربوری با استفاده از نگاره-های پتروفیزیکی بر اساس تفکیک لیتولوژیکی

عباس سلحشور*[۱]، احمد گائینی[۲]، علیرضا شاهین[۳]، مصیب کمری[۴]

۱.دانشجوی دکتری؛ دانشکده‌ی مهندسی صنایع، مهندسی صنایع، دانشگاه ایوانکی

۲. استادیار؛ دانشکده‌ی مهندسی صنایع، مهندسی صنایع، دانشگاه ایوانکی

۳. استادیار؛ گروه زمین‌شناسی، زمین‌شناسی، دانشگاه اصفهان

۴. مهندس ارشد؛ شرکت ملی مناطق نفتخیز جنوب

* salahshor.a@gmial.com

دریافت مرداد ۱۴۰۰ ،پذیرش آبان ۱۴۰۰

چکیده

تراوایی یا نفوذپذیری، یکی از خصوصیات مهم مخازن نفت و گاز است که پیش‌بینی آن دشوار می‌باشد. در حال حاضر از مدل‌های تجربی و رگرسیونی برای پیش‌بینی تراوایی استفاده می‌شود که شامل صرف زمان و هزینه‌های زیاد مرتبط با اندازه‌گیری‌های آزمایشگاهی است. در چند وقت اخیر، به دلیل قابلیت پیش‌بینی بهتر، از الگوریتم‌های یادگیری ماشین برای پیش‌بینی تراوایی استفاده شده است. در این مطالعه، مدل گروهی یادگیری ماشین جدیدی برای پیش‌بینی تراوایی در مخازن نفت و گاز معرفی شده است. در این روش، داده‌های ورودی با استفاده از اطلاعات لیتولوژی لاگ‌ها برچسب‌گذاری شده و به تعدادی از دسته‌ها تفکیک شده است و هر دسته توسط الگوریتم یادگیری ماشین مدل‌سازی شد. برخلاف مطالعات قبلی که به صورت مستقل روی مدل‌ها کار می‌کردند در اینجا ضمن طراحی یک مدل گروهی با استفاده از الگوریتم‌های ETR وDTR و GBR و داده‌های پتروفیزیکی، توانستیم صحت و دقت پیش‌بینی همچنین خطای میانگین مربعات را به طرز چشم‌گیری بهبود ببخشیم و تراوایی را با دقت ۹۹/۸۲ درصد پیش‌بینی کنیم. نتایج نشان داد که مدل‌های گروهی در بهبود دقت پیش‌بینی تراوایی در مقایسه با مدل‌های انفرادی تاثیر فراوانی دارند و همچنین تفکیک نمونه‌ها بر اساس اطلاعات لیتوژی دلیلی بر بهینه نمودن تخمین ترواییِ نسبت به تحقیقات گذشته بود.

کلمات کلیدی: تراوایی، مدل گروهی، لیتولوژی، یادگیری ماشین، نگاره‌های پتروفیزیکی.

۱- مقدمه

تراوایی توانایی سنگ‌های متخلخل در انتقال مایع است. کاربرد این متغیر مهم در مطالعات مخازن زیرزمینی نفت و گاز جهت مدیریت مخزن و فرآیندهای توسعه‌ی میادین نفت و گاز، اجتناب‌ناپذیر است[۴،۳]. از لحاظ ماهیت و مشکلات عملیاتی، تخمین این متغیر نسبت به برخی متغیرهای دیگر سنگ و سیال، دشوارتر است[۵]. شیوه‌های که در حال حاضر معمول است تعیین تراوایی و تخلخل مغزه در آزمایشگاه، به دست آوردن یک رگرسیون خطی بین این دو متغیر و تعمیم آن به کل چاه است که چندان مناسب نیست. تلاش‌های مختلفی جهت تخمین تراوایی صورت گرفته که می‌توان آن‌ها را به دو دسته‌ی روش‌های تجربی و آماری تقسیم کرد. روش‌های آماری، روش‌های نوینی مبتنی بر هوش مصنوعی، شبکه‌های عصبی، منطق فازی و غیره هستند که بسیار کارآمد ظاهر شده‌اند. در این مقایسه باید تعداد متغیرهای مخزن مدنظر، کیفیت وحجم و تنوع اطلاعات آزمایشی انتخاب شده، همگن یا ناهمگن بودن مخزن که همگی بر دقت روش تاثیر گذارند را درنظر گرفت.

در این مقاله مخزن آسماری میدان نفتی مارون که از لحاظ پیچیدگی و ناهمگنی یکی از پیچیده‌ترین مخازن نفتی دنیا می‌باشد مورد مطالعه قرار گرفته است، این ناهمگونی که بین بخش‌های کربناته و ماسه سنگی همراه با درصدهایی از رس پراکنده نیز می‌باشد، در داده‌های لیتولوژی به وضوح مشهود است، بنابراین هدف اصلی مطالعه پیش رو حذف این ناهمگونی توسط دسته‌بندی داده‌ها به گروه‌های همگن از طریق اطلاعات لیتولوژی و مدل سازی و تخمین تراوایی توسط ارائه مدل جدید با افزایش دقت تخمین تراوایی از ترکیب گروهی الگوریتم‌های یادگیری ماشین [۴،۵] (درخت تصمیم، گرادیان تقویت شده و درخت تصمیم افزوده)[۷،۶]، می‌باشد. در اغلب مطالعات گذشته، که در خصوص استفاده از الگوریتم‌های ماشین جهت تخمین تراوایی بوده است کمتر به روش‌های مبتنی بر الگوریتم درخت تصمیم پرداخته شده است. در صورتی که درخت‌های تصمیم از رایج‌ترین روش‌های دسته‌بندی محسوب می‌شوند، چرا که قابل فهم بودن آن‌ها توسط کاربران غیرحرفه‌ای از مزیت اصلی آن به شمار می‌رود. از درخت تصمیم می‌توان به منظور ایجاد مدل‌های پیش‌بینی خودکار استفاده کرد که در حوزه یادگیری ماشین، استخراج داده و آمار کاربردی هستند. این روش که تحت عنوان Decision Tree Learning شناخته می‌شود، به بررسی مشاهدات در مورد یک آیتم به جهت پیش‌بینی مقدار آن می‌پردازد. روش یادگیری درخت تصمیم توسعه‌های فراوانی از جمله جنگل درختان تصادفی[1] و تقویت گرادیان برای مسائل رگرسیون و کلاسه بندی دارد که یک مدل پیش‌بینی کننده را در قالب مجموعه‌ای از مدل‌های پیش‌بینی کننده ضعیف ایجاد می‌کند. دقت یک مدل پیش‌بینی کننده را می‌توان به دو روش، استفاده از مهندسی ویژگی و اعمال مستقیم الگوریتم‌های تقویت‌کننده همچون Gradient Boosting، XGBoost و AdaBoost تقویت کرد[۶].

الگوریتم‌های تقویت‌کننده یکی از قدرتمندترین ایده‌های یادگیری هستند که در ۲۰ سال گذشته ارائه شده‌اند. این الگوریتم‌ها برای مسائل کلاسه‌بندی طراحی شده‌اند، اما می‌توان از آن‌ها برای رگرسیون نیز استفاده کرد. هدف از تقویت گرادیان در این الگوریتم‌ها ایجاد روشی است که خروجی‌های چند کلاسه‌بند ضعیف را ترکیب کند تا به یک کمیسیون قدرتمند دست یابند. کلاسه‌بند ضعیف (به عنوان مثال درخت تصمیم گیری)، کلاسه‌بندی است که میزان خطای آن، تنها از یک حدس تصادفی بهتر باشد. هدف از الگوریتم تقویت کننده، اعمال مکرر الگوریتم کلاسه‌بندی ضعیف بر روی

[1] Random Forest Regressor

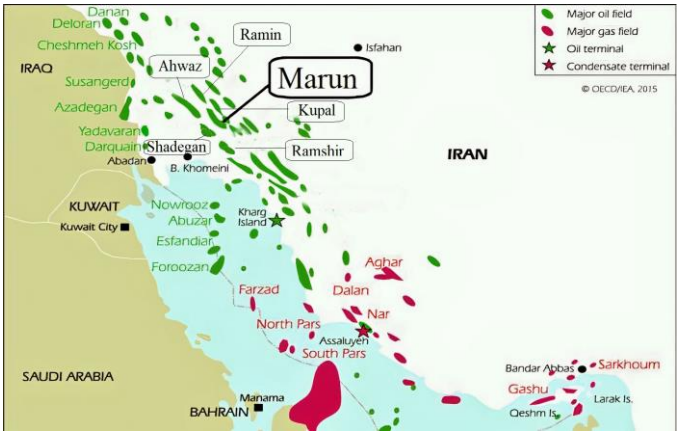
نسخه‌های تغییریافته داده‌ها و در نتیجه تولید دنباله‌ای از کلاسه‌بندها است که درگرادیان تقویت شده، هر درخت تصمیم خطای درخت تصمیم قبلی را پیش‌بینی می‌کند، در نتیجه میزان خطا (گرادیان) را (تقویت) بهبود می‌دهد[۶].

در سال‌های گذشته مطالعات و پژوهش‌هایی در خصوص تخمین تراوایی مخزن با استفاده از روش‌های یادگیری ماشین انجام شده است. که از جمله این پژوهش‌ها در سال ۲۰۱۸، احمدی و همکاران، مقایسه روش‌های یادگیری ماشین برای تخمین تراوایی و تخلخل مخازن نفتی با استفاده از نگاره‌های پتروفیزیکی، روش ترکیبی HGAPSO-LSSVM را ارائه دادند[۱۰]، در سال ۲۰۱۹، ادنیران یک مدل گروهی رقابتی برای پیش‌بینی تراوایی در مخازن ناهمگن نفت و گاز ارائه کرد نتایج نشان داد که مدل‌های گروهی رقابتی، دقت پیش‌بینی بهتری در تخمین تراوایی دارد، آن‌ها از یک مدل گروهی برای تخمین تراوایی شامل ماشین‌های ANN و SVM و ANFIS بهره بردند که موفق شدند با ضریب همبستگی ۹۵/۳۴ درصد تراوایی را تخمین بزنند [۱۱]، سباستین و همکاران در مقاله تخمین تراوایی مطلق با استفاده از شبکه‌های عصبی بر اساس لاگ‌های چاه و داده‌های آزمایشگاهی در سال ۲۰۱۹، توانستند با ضریب همبستگی ۷۳ در صدی بین تراوایی حاصل از داده‌های آزمایشگاه و تراوایی حاصل از شبکه عصبی کار خود را ارائه دهند[۱۲]. در سال ۲۰۱۶، عبدالعزیز و همکاران تحقیقی با عنوان پیش‌بینی تراوایی بهبود یافته به وسیله داده‌های لرزه نگاری و داده‌های لاگ با استفاده از تکنیک‌های هوش مصنوعی انجام دادند. در این تحقیق، یک مطالعه مقایسه‌ای در پیش‌بینی تراوایی مخزن در اثر ترکیب مجموعه داده‌های لاگ چاه و لرزه‌نگاری انجام شد. این مطالعه با به کارگیری سه تکنیک پیشرفته هوش مصنوعی شامل FN، SVM و T2FLS، روی مجموعه داده‌های لرزه‌نگاری و لاگ به صورت ترکیبی، انجام شده است. مدل‌ها برای عملکرد مطلوب بهینه‌سازی شدند و مجموعه داده‌ها به ترتیب به زیرمجموعه‌های آموزش و آزمایش به دنبال الگوی یادگیری ماشین استاندارد قرار گرفتند. نتایج حاصل از این مطالعه نشان داد که ترکیب داده‌های لرزه‌نگاری و لاگ در مقایسه با استفاده سنتی از داده‌های لاگ و لرزه‌نگاری به صورت مجزا عملکرد بهتری داشتند. آن‌ها مدل‌های ترکیبی از سیستم فازی نوع دوم (T2FLS)، ماشین بردار پشتیبان (SVM) و شبکه‌های تابعی (FN) را تحت عنوان الگوریتم انتخاب ویژگی غیرخطی ارائه دادند، مقایسه نتایج مدل‌های ترکیبی از این سه الگوریتم نشان داد که روش ترکیبی FN-SVM با ضریب تعیین ۸۲ درصد و خطای میانگین مربعات ۰/۴۶ بهترین تخمین را به دست آورند [۱۳]. در سال ۲۰۱۴ نیز سلحشور و همکارانش مقاله‌ای با عنوان بهینه سازی الگوریتم‌های یادگیری ماشین جهت تخمین تراوایی مخازن زیرزمینی نفت و گاز ارائه کردند، آن‌ها روش ماشین بردار پشتیبان با توابع هسته مختلف را جهت تخمین تراوایی استفاده کردند که در نهایت توانستند با استفاده از روش رگرسیون بردار پشتیبان با تابع هسته پایه شعاعی بهترین تخمین را با ضریب تعیین ۹۶/۱ داشته باشند [۱۴]. طیبی در مقاله‌ای با عنوان مطالعه موردی تعیین تراوایی مخازن کربناته با استفاده از لاگ NMR در یکی از میادین جنوب غربی ایران از روش‌های SDR و Timur برای محاسبه تراوایی استفاده کردند دقت و صحت تخمین آن‌ها ۴۴,۱ درصد بود [۱۵]. حکیمی‌نژاد و همکاران در سال ۲۰۱۷، پژوهشی بر روی داده‌های چهار چاه کنگان و دالان واقع در میدان پارس جنوبی ایران انجام دادند. آن‌ها با انتخاب ۸ ویژگی از هر چاه، با استفاده از روش انتخاب ویژگی مبتنی بر همبستگی، چهار ویژگی مؤثر در هر چاه را انتخاب کردند. سپس از روش‌های رگرسیون، شبکه عصبی چندلایه، مدل درخت خطی، سیستم فازی نوع یک و سیستم فازی نوع دو برای پیش‌بینی تراوایی استفاده کردند. نتایج نشان داد که با توجه به وجود عدم قطعیت در

پارامترهای پتروفیزیکی و تراوایی، سیستم فازی نوع دو عدم قطعیت‌ها را بهتر پوشش می‌دهد. این روش در حالت پایه، تراوایی را با دقت ۰/۹۴۸۱ و ریشه دوم میانگین مربعات خطا ۰/۳۰۶۰ پیش‌بینی کرد که با استفاده از روش ترکیبی الگوریتم ژنتیک و الگوریتم جستجوی گرانشی، توابع عضویت فازی بهبود یافت و این بهبودها منجر به افزایش دقت پیش‌بینی تراوایی با ضریب تعیین ۰/۹۷۶۸ و کاهش ریشه دوم میانگین مربعات خطا به مقدار ۰/۱۶۰۲ شد [۱۶].

۲- موقعیت جغرافیایی و ویژگی‌های ساختمانی میدان نفتی مارون

میدان نفتی مارون در سال ۱۹۶۳ به روش لرزه‌نگاری دو بعدی کشف شد. این میدان بزرگ نفتی در فرو افتادگی دزفول حدود ۴۵ کیلومتری جنوب شرقی اهواز قرار دارد [۱]. روند محوری این میدان شمال غربی-جنوب شرقی ا ست. میدان مارون در مخزن آسماری دارای طول ۶۷ کیلومتر و عرض متوسط ۵/۵ کیلومتر می باشد که بیشترین و کمترین عرض آن به ترتیب ۷ و ۳/۵ کیلومتر می باشد. موقعیت جغرافیایی این میدان نسبت به میدان‌های مجاور همان‌طور که در شکل (۱) نشان داده شده است، از شمال توسط میدان رامین، از شرق توسط میدان کوپال، از غرب و شمال غرب توسط میدان‌های شادگان و اهواز و از جنوب توسط میدان رامشیر محدود می شود [۲].

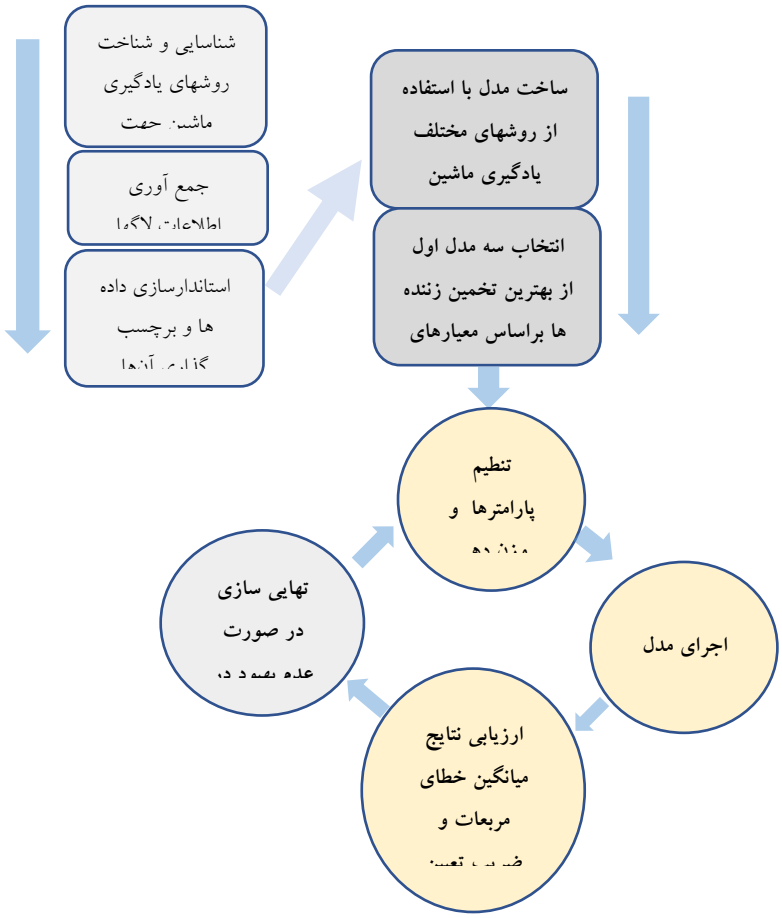


شکل ۱. موقعیت جغرافیایی میدان نفتی مارون [۱۷].

۳- روش کار

دراین مطالعه روش تحلیل مطابق شکل (۲) ارائه شده است. به این ترتیب که ابتدا روش‌های یادگیری ماشین غیرخطی که تا کنون در تخمین تراوایی کمتر استفاده شده و همچنین جزو الگوریتمهای قوی در یادگیری می‌باشند انتخاب شده، سپس اطلاعات مربوط به لاگهای پتروفیزیکی در اختیار قرار داده می‌شود، اطلاعات مربوطه پس از استانداردسازی در بازه ۰ تا ۱، بر اساس حجم مقادیر لیتولوژی در هر رکورد، یکی از برچسب‌های Sandstone، Calcite، Dolomite، Shale و Anhydrite را به خود اختصاص می‌دهد و داده‌ها بر این اساس تفکیک می‌شوند. آنگاه به ازاء هر دسته، مدل‌های یادگیری ماشین KNR, RFR, DTR, GBR, ADR, ETR با پارامترهای مختلف در هر مدل و در تکرارهای مشابه بررسی خواهد شد. سپس سه مدل از الگوریتم‌هایی که بهترین عملکرد را دارند برای تشکیل یک ماشین گروهی انتخاب می‌شوند که معیار ارزیابی عملکرد کارایی مدل‌ها، ضریب همبستگی بین مقادیر تخمین‌زده شده و مقادیر واقعی تراوایی و همچنین خطای

میانگین مربعات می‌باشد. از اینجا به بعد مدل ترکیبی گروهی ساخته شده به ازای مقادیر مختلف پارامترهای ورودی و همچنین وزن دهی متفاوت به مدل‌های درون ماشین که سه مدل می‌باشد ارزیابی خواهد شد و بهترین عملکرد ماشین بعنوان مدل نهایی طراحی شده برگزیده خواهد شد.



شکل(۲) گام‌های روش تحلیل

۳-۱ شناخت و آماده‌سازی داده‌ها در این مطالعه ابتدا تعداد ۷۷۸۵ رکورد مربوط به ویژگی‌های پرتو گاما[1]، چگالی[2]، مقاومت الکتریکی[3]، نوترون[4]، اشباع آب[5]، فتوالکتریک[6]، تخلخل موثر[7]و اطلاعات لیتولوژی شامل حجم کلسیت[8]، حجم دولومیت[9]، حجم ماسه سنگ[10]، حجم شیل[11]، حجم انیدریت[12]، از لاگ‌های چاه‌های به همراه تراوایی[13] مربوطه، مشتمل بر تعداد ۱۴ چاه از مخزن نفتی مارون ایران انتخاب شد.

[1] Gamma Ray
[2] Density
[3] Resistivity Environmentally
[4] Neutron
[5] Water Saturation
[6] Photoelectric Factor
[7] Effective Porosity
[8] Calcite Volume
[9] Dolomite Volume
[10] Sand Stone Volume
[11] Shale Volume
[12] Anhydrite Volume
[13] Permeability

مجموعه‌ای آموزشی به صورت تصادفی از روی داده های لاگ‌های مربوط به ۱۴ چاه انتخاب شده و مجموعه آزمایشی نیز به صورت تصادفی از بین داده‌های کل ۱۴ چاه انتخاب شده است. در هر نمونه آزمایشی و آموزشی دو مجموعه داده وجود دارد، که درنهایت به صورت زیر دسته‌بندی خواهند شد:

$$\{ (x_1 ,y_2) , \ldots . \quad (x_1 ,y_2) \in R^N , y\in \} \qquad (1)$$

که xi به صورت زیر می‌باشند:

$$Xi = (DT, CGR, PEF, RHOB, NPHI, RT, SW, POR)$$

و

$$Yi = Perm$$

به این ترتیب که تراوایی به عنوان پارامتر خروجی و به مقادیر لاگ پرتو گاما، تخلخل، پرتو الکتریکی، چگالی، نوترون، سونیک، مقدار اشباع، فتو الکتریک به عنوان بردار ورودی می‌باشد.

۳-۲ برچسب‌گذاری داده‌ها بر حسب اطلاعات لیتولوژی

در این قسمت داده‌های ورودی را به ازای هر رکورد مطابق جدول (۱) برچسب‌گذاری می‌کنیم. بدین‌گونه که اطلاعات لیتولوژی شامل ستون‌های حجم کربنات، حجم دولومیت، حجم ما سه سنگ، حجم شل در هر نمونه را برر سی کرده و چنانچه مقدار هر کدام از آن‌ها نسبت به سایر ستونهای دیگر بیشتر باشد برچسب آن رکورد را مطابق ستون گونه سنگ در جدول (۱) به خود اختصاص می‌دهد. در نهایت داده‌ها در چهار د سته تقسیم‌بندی می شود که هر د سته به صورت مجزا برای مدل تخمین‌زننده به عنوان ورودی انتخاب خواهد شد.

جدول (۱) نحوه برچسب‌گذاری اطلاعات ورودی

عمق لاگ	عمق مغزه	حجم کلسیت	حجم دولومیت	حجم ماسه	حجم شیل	حجم انیدریت	گونه سنگ
۲۹۵۸،۵٤۱	۲۹۵۸،۲۵	۰،۸۵۳	۰،۱٤۷	۰	۰	۰	Calcite
۲۹٦۰،٦۷۵	۲۹٦۰،۷۱	۰،٤۷٤٦	۰،۵۲۵٤	۰	۰	۰	Dolomite
۳۰۲۳،۹۲۱	۳۰۲۳،۵۳	۰،۲۳۳٦	۰	۰،۷٦٦٤	۰	۰	Sand
۳۱۱۳،۰۷۵	۳۱۱۳،۱۱	۰	۰،۰۱٦۳	۰	۰،۹۸۳۷	۰	Shale

۳-۳ انتخاب روش استانداردسازی داده‌ها

روش‌های مختلفی برای استاندارد سازی وجود دارد که در تمام این روشها داده‌ها طوری تبدیل می‌شوند که شرایط مورد نظر تأمین گردد. یکی از مهمترین روشهای استاندارد سازی، تبدیل داده ها به مجموعه جدیدی است که در آن همه‌ی مقادیر بین صفر و یک می‌باشند:

$$Z_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \qquad (2)$$

که در آن Z_i مقدار استاندارد شده، x_i مقدار هریک از داده ها، x_{min} حداقل کل داده‌ها x_{max} حداکثر مقدار داده‌ها می‌باشد.

۳-۴ روش ارزیابی استفاده شده

نتایج به دست آمده از الگوریتم‌های مورد استفاده از لحاظ دقت ، صحت، مقدار و کیفیت مورد بررسی قرار گیرند تا بتوان بهترین حالت برای مدل‌سازی را انتخاب نمود لذا روش به دست آوردن این کمیت‌ها در زیر توضیح داده شده است.

خطای میانگین مربعات [1] :

$$MSE= (\sum \text{Actual-Forecast})\ 2 / n =\sum(\text{Error})2 / n \qquad (۳)$$

رابطه‌ی-۳ ،خطای میانگین مربعات است که در آن n ، تعداد نمونه‌های تخمینی، Actual، مقدار واقعی هر برچسب و Forecast، مقدار تخمینی برای هر یک است. هر چه مقدار کمیت MSE کمتر باشد، ضریب خطا نیز کمتر خواهد بود. ضریب همبستگی، شدت رابطه و همچنین نوع رابطه (مستقیم یا معکوس) را نشان می‌دهد. این ضریب بین (۱) تا (۱ -) است و در حالت عدم وجود رابطه بین دو متغیر، برابر صفر می‌باشد. همبستگی بین دو متغیر تصادفی X و Y به صورت رابطه‌ی-۴ تعریف می‌شود:

$$\text{Corr (X, Y)} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)}{\sigma_X \sigma_Y} \qquad (٤)$$

که در آن E عملگر امید ریاضی، cov به معنای کوواریانس، corr نماد معمول برای همبستگی پیرسون و σ نماد انحراف معیار است.

۳-۵ آموزش و ساخت مدل

در این مرحله ابتدا داده‌ها را بر اساس اطلاعات برچسب گذاری شده در بخش ۲-۲ به چهار کلاس A, B, C , D تقسیم‌بندی می‌کنیم و هر کلاس را به دو مجموعه ۸۰ در صدی آموزش و ۲۰ در صدی آزمایش بصورت تصادفی تقسیم کرده، سپس با استفاده فریم‌ورک و زبان برنامه نویسی پایتون مدل‌های یادگیری ماشین رگرسیونی درخت تصمیم، درخت تصمیم افزوده، رگرسیون جنگل تصادفی، کی-نزدیکترین همسایه، رگرسیون انطباقی تقویت شده، رگرسیون گرادیانی تقویت شده، مدل‌های داده را بصورت انفرادی آموزش دادیم که نتایج عملکردی مطابق جدول (۲) به ترتیب برای هر کلاس به دست آمد.

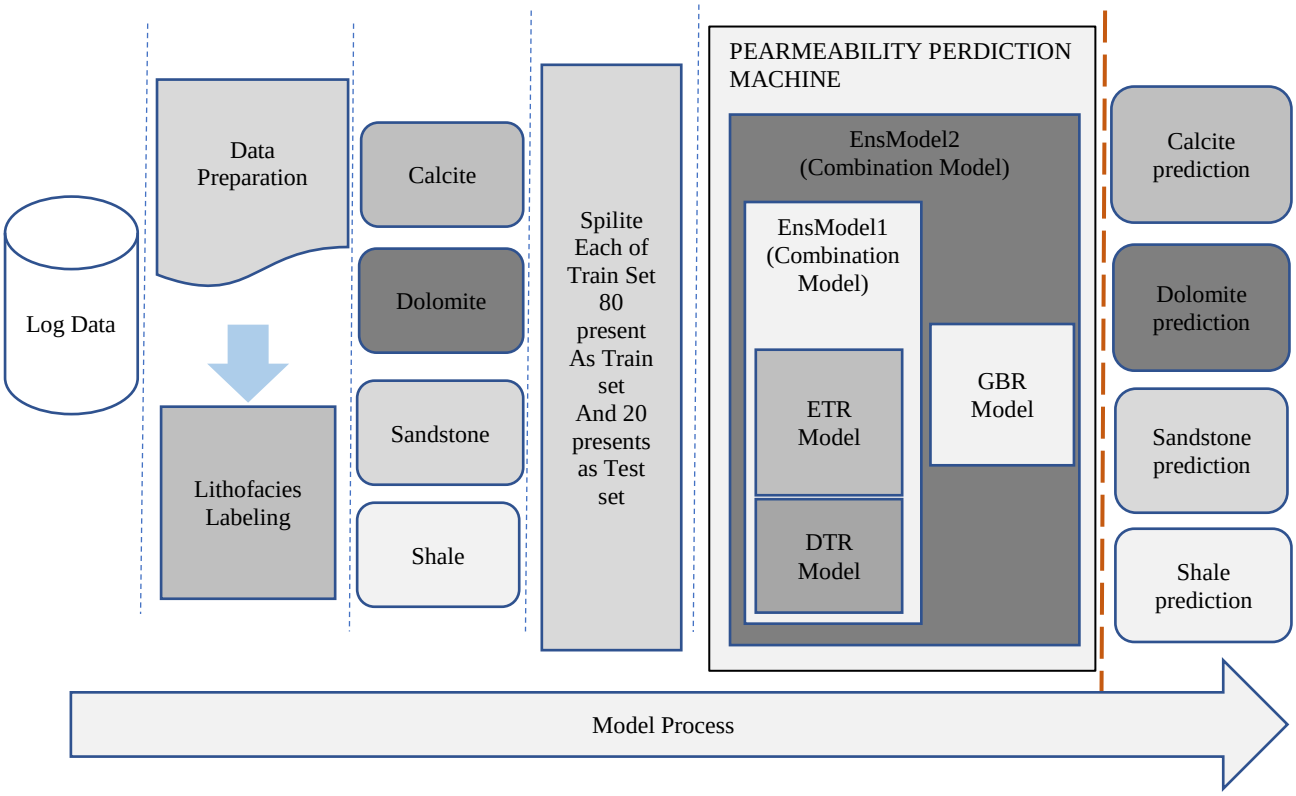
جدول۲ نتایج اجرای روشهای یادگیری شش گانه بر روی ۴ دسته اطلاعات ورودی

ML Algorithm	Evaluate	Class A	Class B	Class C	Class D
		Count: ۴۱۲	Count: ۳۵۶۷	Count: ۹۵۶	Count: ۲۸۵۰
ETR	R^2	۸۳/۶	۸۱/۷	۷۴/۵	۸۸/۰۵
	MSE	۳۴(e-۶)	۳۶(e-۶)	۴۲(e-۶)	۳۱(e-۶)
RFR	R^2	۷۹/۵	۶۰/۱۳	۵۳/۳۷	۷۱/۸۸
	MSE	۶۳(e-۶)	۵۴(e-۵)	۸۶(e-۵)	۶۸(e-۶)
KNR	R^2	۷۵/۴	۵۵/۷	۶۴/۹	۷۳/۶
	MSE	۶۹(e-۶)	۸۱ (e-۵)	۵۱(e-۶)	۶۹(e-۶)
ABR	R^2	۳۸/۶۶	۴۹/۴۴	۲۳/۸۹	۴۷/۰۲

[1] MEAN SQUARE ERROR

	MSE	۸۹(e-۵)	۸۱(e-۵)	۹۴(e-۴)	۷۹(e-۵)
GBR	R^2	۶۴/۰	۶۸/۳۳	۷۸/۲۲	۷۸/۱۱
	MSE	۵۱(e-۶)	۵۵(e-۵)	۳۸(e-۶)	۳۹(e-۶)
DTR	R^2	۹۶/۲	۷۷/۲۲	۷۱/۳	۸۲/۹
	MSE	۱/۸(e-۶)	۴۱(e-۶)	۶۸(e-۶)	۳۸(e-۶)

با توجه به نتایج جدول(۲) مدل‌های ETR و GBR و DTR برای طراحی مدل گروهی تخمین زننده تراوایی انتخاب شد. سپس مدل ترکیبی با استفاده از ترکیبات مختلف و وزن‌های متفاوت از مدل‌های ETR و GBR و DTR در تکرارهای مختلف ایجاد گردید، نتایج حاصل از اجرا های مختلف توسط شاخص‌های ارزیابی R^2 و MSE در هر تکرار بررسی شد و با توجه به نتایج بهترین تکرار، مدلی ترکیبی به صورت ترکیب وزنی ۹۰ درصد از روش ETR و ۱۰ درصدی روش DTR به عنوان مدل گروهی اول تشکیل شده که با نام Ensemble1 نام‌گذاری شد و در مرحله بعد نیز با روش GBR با درصد وزنی ۱۰ به ۹۰ ترکیب شده و با نام Ensemble2 که در شکل (۳) به نمای کلی مدل طراحی شده، اشاره می‌شود. درصد وزنی ترکیب مدل‌ها توسط حلقه تکرار و بررسی همه حالات مطابق با فرایند شکل(۲) به دست آمده است که ارزیابی خروجی حالات مختلف توسط محاسبه ضریب همبستگی و خطای میانگین مربعات در هر تکرار محاسبه شده و در نهایت بهینه‌ترین برازش به عنوان مدل نهایی مطابق شکل (۳) برگزیده شد.

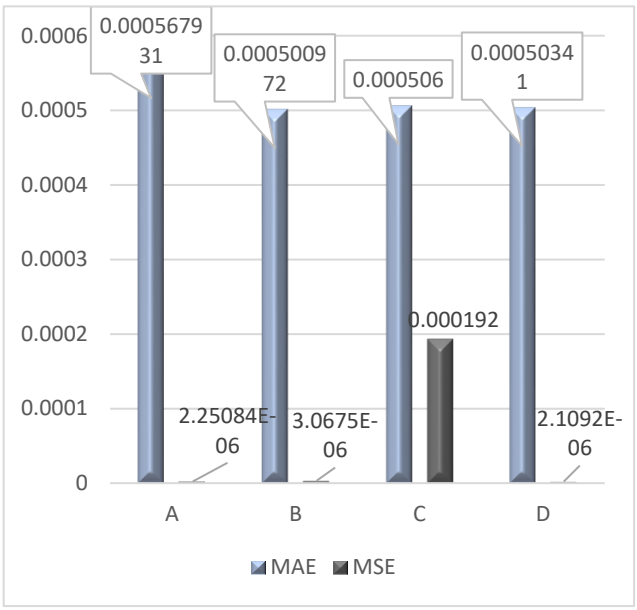


شکل(۳) شمای اجرا شده در آموزش مدل و تخمین تراوایی

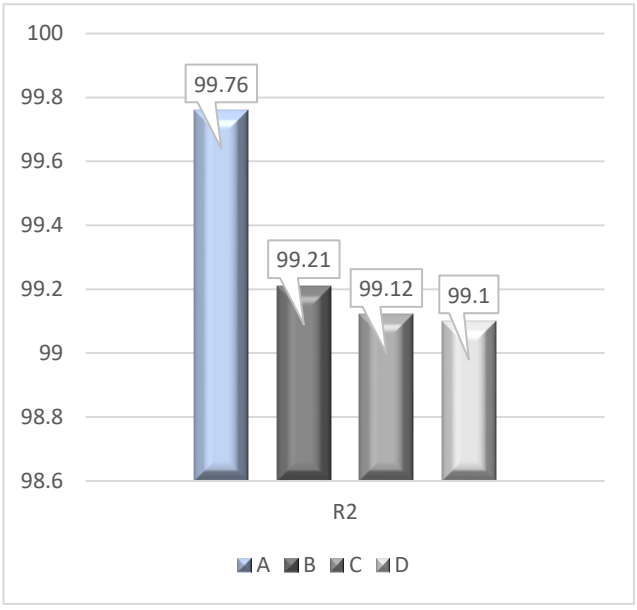
۴- بحث و نتایج

در این مقاله با ا ستفاده از روش‌های یادگیری شش‌گانه یک مدل ترکیبی گروهی شامل سه روش ETR، DTR و GBR مطابق شکل(۳)، طراحی شد که در جدول (۳) نتایج اجرای مدل به تفکیک برچسب‌گذاری داده‌ها و میزان کمیت هر دسته، همچنین ضریب تعیین و خطای میانگین مربعات و خطای مطلق میانگین آورده شده است.

جدول(۳) نتایج اجرای مدل گروهی تخمین تراوایی

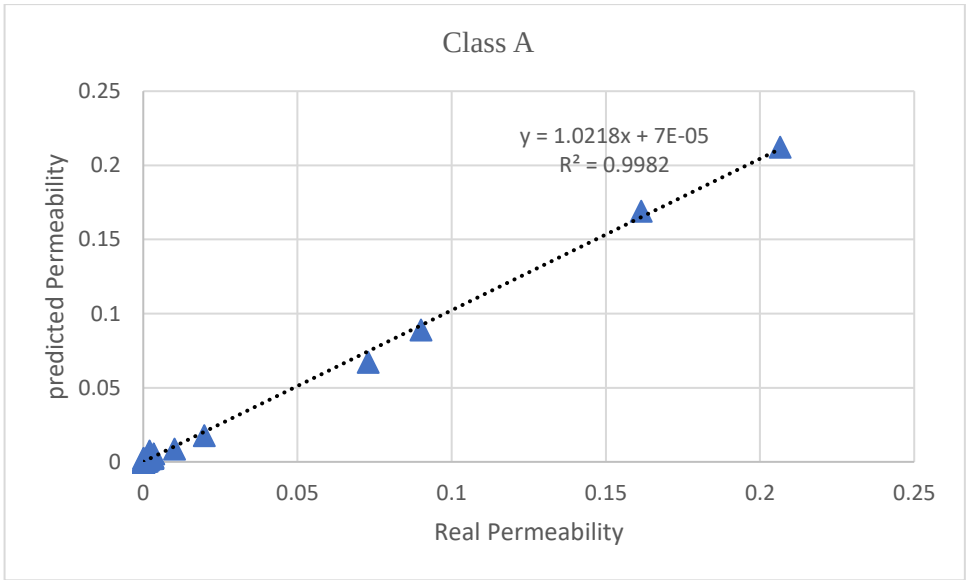


شکل (۴) نمودار مقایسه خطا برای هر دسته

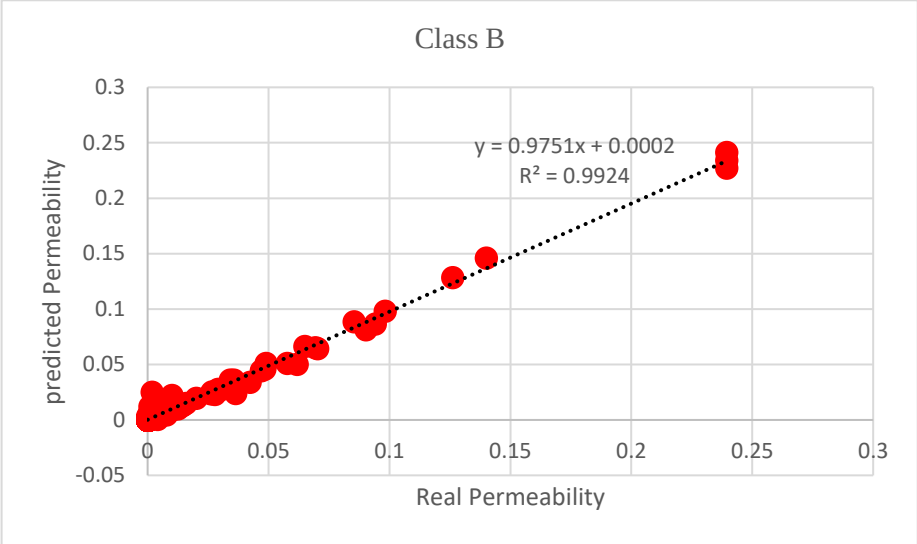


شکل (۵) نمودار مقایسه R^2 برای هر دسته

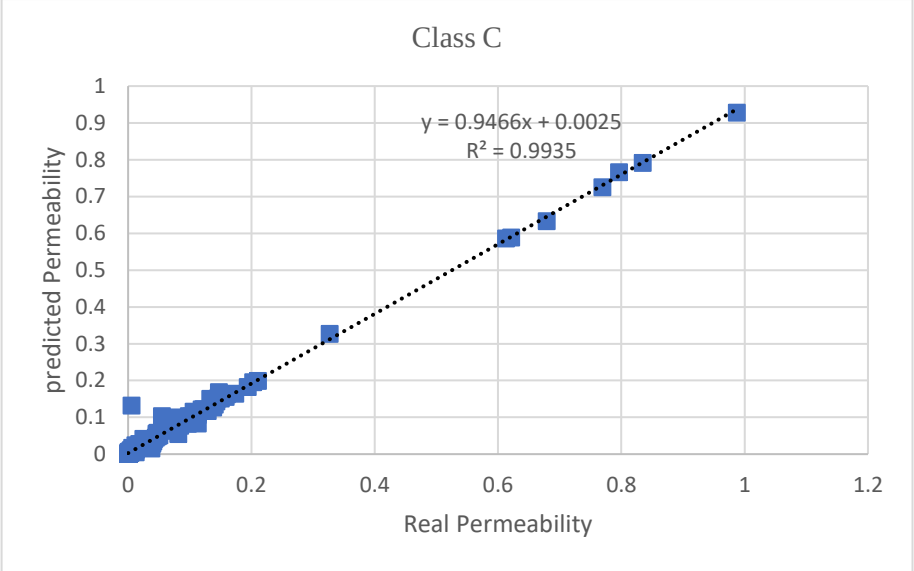
در شکل (۴) خطای میانگین مربعات و خطای مطلق میانگین توسط مدل گروهی ایجاد شده بر اساس ۴ گونه سنگ A، B، C و D نمایش داده شده ا ست. همچنین در شکل (۵) مقدار ضریب تعیین که نشان‌دهنده دقت تخمین بر ا ساس ۴ گونه سنگ می‌باشد قابل مشاهده است.



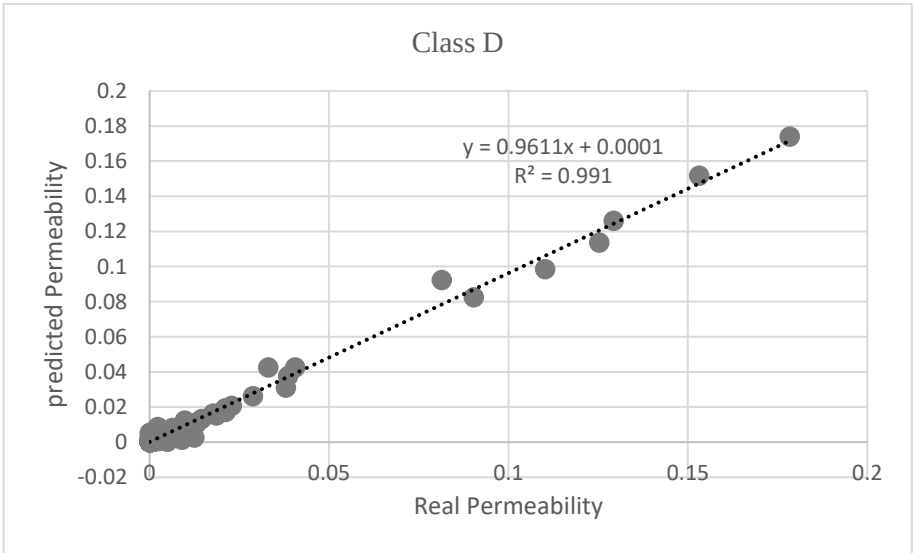
شکل (۶) نمودار تخمین تراوایی توسط مدل گروهی برای دسته A (شیل)



شکل (۷) نمودار تخمین تراوایی توسط مدل گروهی برای دسته B (کلسیت)



شکل (۸) نمودار تخمین تراوایی توسط مدل گروهی برای دسته C (ماسه سنگ)



شکل (۹) نمودار تخمین تراوایی توسط مدل گروهی برای دسته D (دولومیت)

در شکل‌های ۶ و ۷ و ۸ و ۹ به ترتیب مقایسه میزان تراوایی تخمین زده شده توسط مدل گروهی و میزان تراوایی واقعی بر اساس گونه سنگ‌های A، B، C و D نمایش داده شده است. همچنین بررسی نتایج جدول شماره (۳) نشان می‌دهد که مدل ترکیبی گروهی نتایج بسیار خوبی برای همه دسته‌ها که شامل تخمین تراوایی با ضریب تعیین میانگین ۹۹/۳۷۷ و میانگین خطای میانگین مربعات ۰/۰۰۰۰۴۹۸ بوده و وبهترین تخمین را برای دسته A با شاخص ضریب تعیین ۹۹/۸۲ و خطای میانگین مربعات (e-06) ۲/۲۵۰۸۴حاصل شده است. همچنین با نگاهی برنتایج دسته‌های دیگر یک عامل جهت اختلاف بین دقت تخمین، تعداد نمونه‌های هر دسته می‌باشد، البته کیفیت مقادیر نمونه‌ها نیز در تخمین بی‌تاثیر نبوده است. همچنین جهت برازش مدل با استفاده از شاخص ریشه میانگین مربعات باقیمانده(RMSR[1])،چنانچه این شاخص کمتر از ۰/۰۵ باشد نشان‌دهنده نیکویی برازش مدل بر داده‌ها است و مطابق جدول (۴)، این موضوع مشخص است. از سوی دیگر همان‌طور که در شکل‌های ۶تا ۹ نمایش داده شده، داده‌های هر دسته متنوع بوده و بیشتر تمرکز داده از لحاظ کمیت بین ۰ تا ۰/۰۵ می‌با شد. نتایج دقت تخمین این پژوهش در مقایسه با مطالعات بخش (۱) بیانگر توانایی بالای مدل طراحی شده نسبت به سایر روش‌ها می‌باشد.

جدول(۴) برازش مدل بر دسته‌ها

Class Name	Class Count	RMSR
A	۴۱۲	۰/۰۰۱۵
B	۳۵۶۷	۰/۰۰۱۷
C	۹۵۶	۰/۰۱۳۸
D	۲۸۵۰	۰/۰۰۱۴

۵- نتیجه‌گیری

در این مطالعه با توجه به هدف مشخص شده که افزایش دقت در تخمین تراوایی مخزن نفتی مارون از طریق دسته‌بندی داده‌ها به گروه‌های همگن با استفاده از اطلاعات لیتولوژی بود، یک مدل جدید از ترکیب گروهی الگوریتم‌های یادگیری ماشین شامل رگرسیون درخت تصمیم، رگرسیون گرادیان تقویت شده و رگرسیون درخت تصمیم افزوده ارائه گردید. نتایج تخمین نشان داد که این مدل می‌تواند به عنوان روشی جایگزین برای روش‌های تجربی و آزمایشگاهی باشد، همچنین مقادیر تخمین تروایی با ضریب تعیین ۹۹/۷۶ درصد و خطای میانگین مربعات (e-۰۶)۲/۲۵۰۸۴، نشان‌دهنده عملکرد بسیار خوب آن نسبت به دیگر روش‌های رگرسیونی یادگیری ماشین استفاده شده تا کنون می‌باشد. از سوی دیگر تفکیک سنگها بر اساس اطلاعات لیتولوژی لاگ‌ها جهت حذف ناهمگونی نمونه‌ها علاوه بر پیچیدگی مخزن مارون در دستیابی به نتایج تاثیر فراوانی داشت.

[1] Root Mean Square Residual

سپاس و قدردانی

نویسندگان بر خود لازم می‌دانند از مدیران و مسولین شرکت ملی مناطق نفتخیز جنوب به خصوص اداره مهندسی پتروفیزیک به سبب همکاری و در اختیار گذاشتن داده‌های مورد استفاده تشکر نمایند. از داوران مقاله سرکار خانم دکتر بیتا ارباب و جناب آقای دکتر ایمان زحمتکش تشکر و قدردانی می گردد.

منابع

[۱] میرزا قلی پور، علی؛ حقی، عبدالحمید (۱۳۶۹). مطالعه زمین شناسی میدان نفتی مارون، گزارش شماره پ-۴۲۱۰ ،اداره کل زمین شناسی گسترشی، شرکت ملی مناطق نفتخیز جنوب، ص ۵۵.

[۲] سراج، م.، (۱۳۸۴). تحلیل ساختاری مقدماتی میادین نفتی مناطق نفتخیز جنوب (محدوده فروافتادگی دزفول شمالی)، گزارش شماره پ-۵۶۱۳، شرکت ملي مناطق نفتخیز جنوب.

[3] North F.K (1985) Petroleum Geology, Allen & Unwin.

[4] Ramzi H.R (1998) Well Logging, SANAM Publishing.

[5] Oberto S. (1984) Fundamentals of Well-Log Interpretation - The Acquisition of Logging Data, Elsevier.

[6] Ayyadevara V. (2018) Pro Machine Learning Algorithms, Apress.

[7] Lakhmi C. Jain (2016) Foundations and Methods in Combinatorial and Statistical Data Analysis and Clustering, Springer Press.

[8] Shalev-Schwartz Sh. (2014) understanding-machine-learning-theory-algorithms, Cambridge University Press.

[9] Mitchell M. (1997) Machine-Learning-Tom-Mitchell, McGraw-Hill.

[10] Ahmadi Mohammad ali Zhangxing Chen (2018), Comparison of machine learning methods for estimating permeability and porosity of oil reservoirs via Petro-physical logs, KeAi Elsevier,**6 (2)**:831-844.

[11] Adeniran A. (2019) A competitive ensemble model for permeability prediction in heterogeneous oil and gas reservoirs, *Elsevier Paper*, **4 (80)**:353-367.

[12] Waszkiewicz S., rakowska-Madejska P., Puskarczyk E. (2019), Estimation of absolute permeability using artificial neural networks (Multilayer perceptron) based on well logs and laboratory data from Silurian and Ordovician deposits in SE Poland, *Springer*, **6 (67)**:1885–1894.

[13] Fatai A., Abdulazeez A., Abdullatif A. (2016), Improved Permeability Prediction from Seismic and Log Data using Artificial Intelligence Techniques, *Journal of Petroleum Science and Engineering*, **145**: 230-237.

[14] Salahshoor A., Soheyli F., Kamari m. (2014), Optimization of machine learning algorithms for estimating the permeability of underground oil and gas reservoirs ", *Journal of Iranian Oil Exploration and Production*, **3(120)**: 56-62.

[15] Tayebi H., Habibniyan B. (2018), A case study of carbonate reservoir permeability determination using NMR log in one of the southwestern fields of Iran, Iranian Journal of Petroleum Geology, 7, 13:43-62.

[16] Hakiminezhad H., Mirzarezaei M., Najar B. (2019), Estimation of permeability using uncertainty methods: type -2 fuzzy system, *Iranian Journal of Geophysics*, **12 (1)**: 82-91.

[17] Amiri Bakhtiyar, H., A. Telmadarreie, M. Shayesteh, M. H. Heidari Fard, H. Talebi, and Z. Shirband. (2018). Estimating Total Organic Carbon Content and Source Rock Evaluation, Applying ΔlogR and Neural Network Methods: Ahwaz and Marun Oilfields SW of Iran. *Petroleum Science and Technology*, **29 (16)**: 1691–1704.

نشریه علمی–پژوهشی زمین شناسی نفت ایران
سال دهم، شماره ۲۰، پائیز و زمستان ۱۳۹۹ص۲۹-۴۰
Iranian Joural of Petrolum Geology
No. 20, Autumn & Winter 2020, pp 29-40

Designing an ensemble model for estimating the permeability of a hydrocarbon reservoir by petrophysical lithology labeling

Abbas Salahshoor[1], Ahmad Gaeini[2], Alireza Shahin[3], Mossayyeb Kamari[4]

1- Ph.D. Candidate; Faculty of Industrial Eng., Industrial Engineering, Eyvanaki University
2- Assistant Professor; Faculty of Industrial Eng., Industrial Engineering, Eyvanaki University
3- Assistant Professor; Faculty of Geology, Geology, Isfahan University
4- Senior Engineer in National Iranian Oil Company; M.Sc. in Petrophysics Engineering

*Salahshor.a@gmail.com

Received: July 2021, Accepted: October 2021

Abstract

Permeability is one of the important characteristics of oil and gas reservoirs that is difficult to predict. In the present solution, experimental and regression models are used to predict permeability, which includes time and high costs associated with laboratory measurements. Recently, machine learning algorithms have been used to predict permeability due to better predictability. In this study, a new ensemble machine learning model for permeability prediction in oil and gas reservoirs is introduced. In this method, the input data are labeled using the lithology information of the logs and divided into a number of categories and each category was modeled by machine learning algorithm. Unlike previous studies that worked independently on models, here we were able to predict the accuracy of such a square mean error by designing a group model using ETR, DTR, GBR algorithms and petrophysical data. Improve dramatically and predict permeability with 99/82% accuracy.

The results showed that group models have a great effect on improving the accuracy of permeability prediction compared to individual models and also the separation of samples based on lithology information was a reason to optimize the Trojan estimate compared to previous studies

.

Keywords: Permeability, Ensemble model, Lithology, Machine learning, Petrophysical logs.