

مقاله پژوهشی

خوشه‌بندی خصوصیات مخزن جهت تشخیص نوع سنگ با استفاده از روش کی-میانگین در یکی از میادین نفتی جنوب غرب ایران

عباس سلحشور^{۱*}؛ احمد گابینی^۲؛ علیرضا شاهین^۳؛ مصیب کمری^۴

۱. دانشجوی دکترای تخصصی؛ مهندسی صنایع، دانشکده مهندسی صنایع، دانشگاه ایوانکی

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

۳. دکترای تخصصی؛ زمین‌شناسی، دانشکده علوم، دانشگاه اصفهان

۴. کارشناسی ارشد؛ مهندس ارشد پتروفیزیک، شرکت ملی مناطق نفت خیز جنوب

دریافت مقاله: ۱۴۰۰/۰۵/۰۷ پذیرش مقاله: ۱۴۰۰/۱۰/۲۲

شناسه دیجیتال (DOI): 10.22107/jpg.2022.297325.1147

چکیده

واژگان کلیدی

تعیین گونه‌های سنگ در ساخت مدل استاتیک و پویای مخازن هیدروکربنی از اهمیت ویژه‌ای برخوردار می‌باشد. تخمین دقیق خواص سنگ‌های مخزنی، باعث افزایش دقت و صحت در پیش‌بینی میزان ذخیره مخزن و عملکرد آن می‌شود. تاکنون مدل‌های متعددی برای تعیین گونه‌های سنگ‌های مخزن توسط متخصصین پیشنهاد شده است؛ اما اکثر مدل‌های پیشنهادی بر اساس روش‌های متداول مدل محور مهندسی و زمین‌شناسی سنگ‌های مخزنی کرناشته استوار بوده است. بنابراین استفاده از یک روش یادگیری ماشین برای تعیین گونه‌های سنگی در مقایسه با روش‌های پیشین و مقایسه کارایی و عملکرد آن با سایر روش‌ها ضروری به نظر می‌رسد. در این مطالعه داده‌های مغزه و لاگ در مخزن نفتی مارون پس از آماده‌سازی، با استفاده از تکنیک سری‌های زمانی پویا (DTW) هم عمق سازی شده‌اند. سپس داده‌های مغزه توسط روش یادگیری ماشین غیر نظارتی کی-میانگین خوشه‌بندی شدند. همچنین فرآیند خوشه‌بندی داده‌های مغزه توسط روش‌های متداول مدل محور از قبیل روش شاخص منطقه جریان (FZI) و وینلند نیز انجام گردید. در ادامه نتایج خوشه‌بندی با استفاده از روش‌های کی-میانگین، شاخص منطقه جریان و وینلند با در اختیار داشتن اطلاعات لیتولوژی لاگ‌ها صحت سنجی شده و با یکدیگر مقایسه شده است. روش کی-میانگین با معیار صحت‌سنجی ۹۳.۵ درصد، موفق به انجام بیشترین تفکیک‌پذیری خوشه‌ها شد که نشان داد روش یادگیری ماشین مبتنی بر داده کی-میانگین، جایگزین مناسبی برای روش‌های متداول مدل محور برای خوشه‌بندی گونه‌های سنگی می‌باشند.

شیوه‌نامه، نگارش،
نشریه‌ی علمی-پژوهشی
ژئومکانیک نفت،
انجمن ژئومکانیک نفت

۱. پیش‌گفتار

آب و تراوایی بسیار ضروری بوده و منجر به جامع‌تر شدن مدل مخزن و تخمین بهتر مقدار ذخیره‌ی درجا می‌گردد. روش‌های مختلفی برای تقسیم‌بندی گونه‌های سنگی ارائه شده‌اند که مبنای بسیاری از آن‌ها داده‌های تخلخل و تراوایی است (Potter, 2010). از جمله روش‌های تجربی (مدل محور) برای تعیین گونه‌های سنگ مخزن می‌توان به روش‌های وینلند و FZI^۱ اشاره کرد (Merandy & Muhammad, 2016). با ظهور تکنیک‌های تحلیل داده

تعیین گونه‌های سنگی روشی برای تفکیک سنگ‌های مخزنی به گونه‌های مجزا بر اساس خواص مشابه مانند تراوایی و تخلخل می‌باشد. به‌طور معمول تعیین اولیه‌ی گونه‌های سنگی بر اساس مطالعات اولیه‌ی میکروفاسیس یا پتروفاسیس است (Lucia et al., 1999). تفکیک گونه‌های سنگی به علت تأثیرگذاری بر خصوصیات مخزنی مانند اشباع

¹ Flow Zone Indicator



شکل ۱. موقعیت جغرافیایی میدان نفتی مارون (Amiri Bakhtiyar et al., 2018)

۲.۱ استفاده از یادگیری ماشین در تعیین گونه‌های سنگ مخزن

در سال‌های گذشته مطالعات و پژوهش‌هایی در حوزه تعیین گونه‌های سنگ مخزن با استفاده از روش‌های یادگیری ماشین انجام شده است. از جمله این پژوهش‌ها، آل کتان و همکاران در خصوص رویکرد تحلیل خوشه‌ای برای تحلیل گونه‌های سنگی در میدان نفتی خباز (Al Kattan et al., 2018)، همچنین توصیف لاگ‌های چاه نفت با استفاده از تحلیل خوشه‌ای کی-میانگین ۲ توسط امجد علی (Ali & Sheng, 2020)، بود. مراندی روش‌های تعیین گونه‌های سنگ برای بهبود توصیف مخازن کربناته را تحلیل و بررسی کرده است (Merandy & Muhammad, 2016). همچنین ریل نیز مروری انتقادی بر روش‌های خوشه‌بندی جهت تعیین گونه‌های سنگ در مخازن کربناته انجام داده است (Rebelle & Lalanne, 2014). میرزایی در مطالعه‌ای تحت عنوان ارائه چارچوبی جدید برای انتخاب روش جایگزین در تحلیل ویژه مغزه (Mirzaei et al., 2020) به تعیین گونه‌های سنگی پرداخته است. مریم ذیفن با استفاده از داده‌کاوی، تعیین واحدهای جریان‌ی هیدرولیکی و دسته‌بندی نوع سنگ در مخازن نفتی را انجام داد (ذیفن، ۱۳۹۸). علاوه بر این سپهری و همکاران نیز پژوهشی تحت عنوان تعیین الکتروفاسیس‌های مخزن A در یکی از میادین نفتی ایران با استفاده از روش خوشه‌بندی چند کیفیتی بر پایه نمودار (MRGC) انجام دادند (سپهری و زینالی، ۱۳۹۶).

تحت عنوان داده‌کاوی، شیوه‌های جدیدی برای تحلیل و پیش‌بینی مقادیر داده‌ها ارائه شده‌اند. این شیوه‌ها بر پایه اشتراک روش‌های محاسباتی هوش مصنوعی، یادگیری ماشین، آمار و بازیابی اطلاعات هستند (Lakhami, 2016). هدف اصلی فرآیند داده‌کاوی، استخراج اطلاعات توصیفی و قابل پیش‌بینی از مجموعه داده با کاوش الگوها، دسته‌بندی بر اساس روابط، خوشه‌بندی بر اساس شباهت‌ها و تخمین بر پایه روندها و سری‌های زمانی است. طبق پژوهش‌های صورت گرفته از به‌کارگیری فرآیند داده‌کاوی در حوزه‌های کاربردی مختلف، کسب نتایج با دقت مناسب، پردازش در زمان کمینه و قابلیت پیش‌بینی مطلوب و عدم محدودیت‌های روش‌های سنتی، از جمله مزیت‌های استفاده از فرآیند داده‌کاوی است (Wang, 2003).

در این پژوهش، به مقایسه عملکرد روش‌های وینلند و FZI به‌عنوان روش‌های تجربی تخمینی و الگوریتم یادگیری ماشین خوشه‌بندی کننده کی-میانگین در حوزه شناسایی و تفکیک نوع سنگ در مخازن نفتی پرداخته شده است. ابتدا به مفاهیم و پژوهش‌های مرتبط پرداخته خواهد شد، سپس تبیین مسئله و مقایسه نتایج کسب‌شده از هر دو روش، مورد بحث قرار خواهد گرفت.

۱.۱ موقعیت جغرافیایی و ویژگی‌های ساختمانی میدان نفتی مارون

میدان نفتی مارون در سال ۱۹۶۳ به روش لرزه‌نگاری دوبعدی کشف شد. این میدان بزرگ نفتی در فروافتادگی دزفول حدود ۴۵ کیلومتری جنوب شرقی اهواز قرار دارد (میرزاقلی-پور و حقی، ۱۳۶۹). روند محوری این میدان شمال غربی-جنوب شرقی است. میدان مارون در مخزن آسماری دارای طول ۶۷ کیلومتر و عرض متوسط ۵/۵ کیلومتر می‌باشد که بیشترین و کمترین عرض آن به ترتیب ۷ و ۳/۵ کیلومتر می‌باشد. 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 می‌شود (سراج، ۱۳۸۴).

2 K means Clustering

3 Multi Resolution Graph-based Clustering

تعیین گونه‌های سنگی جمع‌آوری شده‌اند. سپس ضمن آماده‌سازی و استانداردسازی داده‌ها به دلیل اینکه اطلاعات از دو منبع مختلف تهیه شده‌اند، عملیات هم عمق‌سازی روی آن‌ها اجرا می‌شود. ساخت مدل بر اساس روش یادگیری ماشین غیر نظارتی جهت خوشه‌بندی با استفاده از تکنیک کی-میانگین انجام خواهد شد. به صورت هم‌زمان روش‌های تجربی خوشه-بندی گونه‌های سنگ مخزن همچون روش *FZI* و وینلند نیز بررسی خواهد شد و در نهایت بهترین مدل را جهت خوشه-بندی بر اساس شاخص‌های ارزیابی خوشه‌بندی که شامل بیشترین شباهت داده‌ها در هر خوشه و همچنین تعداد خوشه‌ها در مقایسه با اطلاعات لیتولوژی مخزن می‌باشد، توسعه داده خواهد شد.

در این مطالعه ابتدا تعداد ۷۸۳۵ رکورد از ۱۴ چاه میدان نفتی مارون ایران انتخاب شد. اطلاعات تراوایی^۴ و تخلخل^۵ نمونه-های مغزه از اعماق مختلف چاه و همچنین مقادیر مربوط به لاگ‌های پرتو گاما^۶، چگالی^۷، مقاومت الکتریکی^۸، نوترون^۹، اشباع آب^{۱۰}، فتوالکتریک^{۱۱}، تخلخل مؤثر^{۱۲} و اطلاعات لیتولوژی شامل حجم کلسیت^{۱۳}، حجم دولومیت^{۱۴}، حجم ماسه‌سنگ^{۱۵}، حجم شیل^{۱۶}، حجم انیدریت^{۱۷}، انتخاب گردید. پس از استانداردسازی، داده‌های مغزه و لاگ با استفاده از روش *DTW*^{۱۸}، هم عمق‌سازی شدند.

۱.۲ روش *FZI* جهت تعیین گونه‌های سنگ

یکی از روش‌هایی که جهت تفکیک انواع گونه‌های سنگ بکار گرفته می‌شود روش تعیین واحدهای جریان‌ی هیدرولیکی است. هر واحد جریان‌ی توسط پارامتری به نام شاخص منطقه جریان ۱۹ (*FZI*) معرفی می‌شود. شاخص منطقه جریان‌ی از



شکل ۲. گام‌های تحلیل مسئله

۲. روش بررسی داده‌ها

در این بخش گام‌های تحلیل مسئله مطابق شکل ۲ ارائه می‌شود. به این ترتیب که ابتدا اطلاعات مربوط به مغزه گیری و نمودار گیری (چاه‌نگاری) از یکی از مخازن نفتی ایران با هدف

⁴ Permeability

⁵ Porosity

⁶ Gamma Ray

⁷ Density

⁸ Resistivity Environmentally

⁹ Neutron

¹⁰ Water Saturation

¹¹ Photoelectric Factor

¹² Effective Porosity

¹³ Calcite Volume

¹⁴ Dolomite Volume

¹⁵ Sand Stone Volume

¹⁶ Shale Volume

¹⁷ Anhydrite Volume

¹⁸ Dynamic Time Wrapping

¹⁹ Flow Zone Indicator

۲.۲ روش وینلند جهت تعیین گونه‌های سنگ

در ۱۹۷۲ وینلند با ترکیب داده‌های تخلخل، تراوایی و اشباع آب با فشار موئینگی تزریق جیوه، معادله‌ای تجربی را بین تخلخل، تراوایی هوا و فضاهای خالی مرتبط با اشباع ۳۵ درصد جیوه ($R35$) به دست آورد که از آن پس مبنای بسیاری از مطالعات تعیین گونه‌های سنگی قرار گرفته است. روش وینلند بین داده‌های تخلخل، تراوایی و فضاهای خالی مرتبط با حالت اشباع ۳۵ درصد جیوه ($R35$) ارتباط برقرار می‌کند. این روش برای درصد‌های دیگر (۳۰، ۴۰ و ۵۰) نیز قابل استفاده است؛ اما ۳۵ درصد بالاترین دقت را دارد (Potter, 2010). شاخص $R35$ شعاع گلوگاه فضای خالی محاسبه‌شده در اشباع ۳۵ درصد جیوه در یک آزمون فشار موئینگی تزریق جیوه، تراوایی و تخلخل است که با استفاده از رابطه‌ی ۶ (موسوم به رابطه‌ی وینلند) محاسبه می‌شود.

$$\text{LOG}(R35) = 732 + 0.588 \times \text{LOG}(K) - 0.864 \times \text{LOG}(\phi) \quad (۶)$$

۳.۲ خوشه‌بندی

تحلیل خوشه‌ای^{۲۲} یا به‌طور خلاصه خوشه‌بندی، فرآیندی است که به کمک آن می‌توان مجموعه‌ای از اشیاء را به گروه‌های مجزا افراز کرد. هر افراز یک خوشه نامیده می‌شود. اعضاء هر خوشه با توجه به ویژگی‌هایی که دارند به یکدیگر بسیار شبیه هستند و در عوض میزان شباهت بین خوشه‌ها کمترین مقدار است. در چنین حالتی هدف از خوشه‌بندی، نسبت دادن برچسب‌هایی به اشیاء است که نشان‌دهنده عضویت هر شیء به خوشه است. در تحلیل خوشه‌ای برخلاف دسته‌بندی^{۲۳}، به دلیل عدم وجود برچسب‌های اولیه به‌عنوان یک روش غیر نظارتی می‌باشد. در این روش، بر اساس ویژگی‌های مشترک و روش‌های اندازه‌گیری فاصله یا شباهت بین اشیاء، باید برچسب‌هایی به‌طور خودکار نسبت داده شوند. درحالی‌که در دسته‌بندی برچسب‌های اولیه موجود است و باید با استفاده از الگوهای پیش‌بینی قادر به برچسب‌گذاری برای مشاهدات جدید باشیم (Aggarwal & Reddy, 2014).

طریق رابطه بین نسبت تخلخل به حجم دانه‌های سنگ (ε) و توزیع فضای منافذ که به‌عنوان شاخص کیفیت مخزن ۲۰ (RQI) شناخته‌شده، محاسبه می‌گردد (ذیفن، ۱۳۹۸). این روابط به‌صورت زیر تعریف می‌شود (ذیفن، ۱۳۹۸).

$$RQI = 0.0314 \times \sqrt{\frac{k}{\phi_e}} \quad (۱)$$

$$\varepsilon = \frac{\phi_e}{1 - \phi_e} \quad (۲)$$

$$FZI = \frac{RQI}{\varepsilon} \quad (۳)$$

که در رابطه‌ی (۱)، پارامتر k مقدار تراوایی سنگ مخزن با واحد میلی داری و پارامتر ϕ_e نشان‌دهنده میزان تخلخل مؤثر (واحد کسری) است. در رابطه‌ی (۲)، پارامتر ε بیانگر تخلخل نرمال شده، است که در محاسبات شاخص منطقه جریانی استفاده می‌شود و در نهایت با استفاده از فرمول رابطه‌ی (۴) می‌توان بر روی نمودار $\log\text{-}\log$ مناطق مختلف مخزنی را نشان داد، به‌طوری‌که مناطقی با FZI یکسان بر روی شیب خط ثابت قرار می‌گیرند.

$$\text{Log } RQI = \text{Log } FZI + \text{Log } \varepsilon \quad (۴)$$

برای تعیین مرزهای FZI ، از رسم نمودار تجمعی فراوانی استفاده می‌شود. در این نمودار، در بازه‌هایی که شیب به‌صورت ناگهانی تغییر کرده است، به‌عنوان دسته‌بندی نوع سنگ در نظر گرفته می‌شوند. بعد از محاسبه پارامتر FZI ، میزان فراوانی تجمعی آن نیز محاسبه می‌شود و با رسم مقادیر این دو پارامتر نسبت به هم می‌توان تعداد بهینه نواحی جریانی را به دست آورد (ذیفن، ۱۳۹۸).

همچنین پس از محاسبه مقادیر FZI ، به دلیل اینکه این مقادیر پیوسته و اعشار هستند، برای تفکیک آن‌ها به‌صورت گسسته لازم است مقدار گسسته نوع سنگ (DRT^{*1}) را با استفاده از رابطه‌ی (۵) به دست آوریم (Guo et al., 2007).

$$DRT = \text{Round}(2 \times \ln(FZI) + 10.6) \quad (۵)$$

²⁰ Reservoir Quality Indicator

²¹ Discrete Rock Type

²² Cluster Analysis

²³ Classification

موضوع مطرح گردید. در این تحقیق ما از روش محاسباتی DTW برای یافتن نقاط هم عمق با استفاده از دو ویژگی عمق و تداخل که در هر دو دسته اطلاعات به صورت مشترک وجود داشت استفاده نموده و نزدیکترین نقاط را به یکدیگر جهت هم عمق سازی انتخاب کرده‌ایم.

سری‌های زمانی پویا (DTW)، الگوریتمی برای اندازه‌گیری شباهت بین دو دنباله زمانی هستند که ممکن است در سرعت یا زمان متفاوت باشند. برای نمونه، DTW می‌تواند شباهت بین دو الگوی راه رفتن را بیابد حتی اگر سرعت یا شتاب راه رفتن آن‌ها در بازه‌های زمانی یکسان نباشد. همچنین DTW توانایی تحلیل دنباله‌های زمانی مربوطه به داده‌های صوتی، ویدئویی و تصویری را دارد. به‌طور کلی، هر داده‌ای که در قالب یک سری یا دنباله زمانی قرار گیرد، پتانسیل تحلیل توسط تکنیک DTW را خواهد داشت. از کاربردهای DTW می‌توان به تشخیص صدا، تشخیص دستخط و تشخیص تطبیق اشکال هندسی اشاره کرد. در مجموع، DTW روشی است که بهینه‌ترین تطبیق بین دو دنباله زمانی با محدودیت‌های معین را پیدا می‌کند. دنباله‌ها به صورت غیرخطی در محور زمان کش‌وقوس پیدا می‌کنند تا معیاری برای شباهت آن‌ها مستقل از برخی تغییرات غیرخطی در محور زمان به دست آورد. این روش تنظیم دنباله بسیاری از اوقات در دسته‌بندی سری زمانی مورد استفاده قرار می‌گیرد. اگرچه DTW معیاری همانند فاصله بین دو دنباله داده شده را می‌یابد، لیکن بیشترین تلاش‌ها برای هم عمق سازی خودکار بر اساس برنامه‌نویسی پویا بوده است. به همین دلیل لازم است که این الگوریتم پرکاربرد برای مقایسه سری‌های زمانی مختلف بیشتر بررسی کنیم. فرض کنید که دو سری زمانی، شامل جستار S، با تعداد M عنصر و جستار مرجع به نام R، شامل N عنصر باید با یکدیگر مقایسه شوند. DTW، به منظور تطبیق هر چه بهتر جستار S با مرجع R که هیچ تغییری نمی‌کند با حفظ یکنواختی بین نقاط مختلف جستار، از یک تابع هزینه بین دو نقطه از جستار و مرجع که دارای کمترین مقدار برای C(Si, Rj) باشد را به عنوان نقطه بهینه انتخاب می‌کند. در واقع وقتی هر دو نقطه بیشترین شباهت را به همدیگر دارند، مقدار این تابع که فاصله اقلیدسی بین دو نقطه است، کمترین باشد.

اگرچه بیشتر الگوریتم‌ها یا روش‌های خوشه‌بندی مبنای یکسانی دارند ولی تفاوت‌هایی در شیوه اندازه‌گیری شباهت یا فاصله و همچنین انتخاب برچسب برای اشیاء هر خوشه در این روش‌ها وجود دارد. به‌طور کلی چهار گروه اصلی برای الگوریتم‌های خوشه‌بندی وجود دارد. الگوریتم‌های خوشه‌بندی تفکیکی، الگوریتم‌های خوشه‌بندی سلسله مراتبی، الگوریتم‌های خوشه‌بندی بر مبنای چگالی و الگوریتم‌های خوشه‌بندی بر مبنای مدل (Lakmi, 2016). در این مطالعه روش خوشه‌بندی کی-میانگین (خوشه‌بندی تفکیکی)، مورد بررسی قرار گرفته است (Wang, 2003).

۴.۲ روش آرنج ۲۴ برای تعیین پارامتر k در روش

کی-میانگین

تعیین تعداد خوشه‌ها در روش کی-میانگین، یکی از مهم‌ترین مراحل خوشه‌بندی می‌باشد. به این ترتیب که برای بازه‌ای از مقادیر K الگوریتم کی-میانگین را اجرا می‌کنیم، سپس مجموع خطای مربعات (SSE) را به ازای هر مقدار K محاسبه می‌شود. SSE به عنوان مجموع فاصله مربعات بین اعضای هر خوشه و مرکز آن تعریف شده است و از طریق رابطه‌ی (۷) محاسبه می‌شود (Ayyadevara, 2018).

$$SSE = \sum_{i=1}^k \sum_{x \in C_i} d(C_i - x)^2 \quad (7)$$

با استفاده از رابطه‌ی (۷) نمودار مقادیر SSE برحسب K را رسم می‌نماییم، آنگاه روی نمودار آنجا که SSE حداکثر کاهش را دارد به عنوان شکست آرنج بوده و تعداد بهینه برای پارامتر K می‌باشد (Ayyadevara, 2018).

۵.۲ هم عمق سازی

به‌طور کلی مغزه گیری و نمودار گیری دو فرآیند جدا از هم هستند که در دو زمان متفاوت و توسط ابزارها و روش‌های متفاوت انجام می‌شوند و در نتیجه عدم تطابق عمق مغزه‌ها با نمودارها می‌تواند امری طبیعی باشد (قربان پور، ۱۳۸۶). با توجه به اینکه اطلاعات موجود برای مخزن از دو منبع مختلف شامل اطلاعات چاه پیمایی (لاگ) و اطلاعات نمونه‌گیری مغزه به دست آمده است، لزوم هم عمق سازی نقاط به عنوان یک

²⁴ Elbow

²⁵ Sum of Square Error

در این بخش با استفاده از چارچوب و زبان برنامه‌نویسی ویژوال بیسیک^{۲۶} یک برنامه کاربردی تولید کردیم. اطلاعات مربوط به ۱۷ هزار نقطه از نمونه‌های مغزه مرتبط با ۱۸ چاه نفت در مخزن مارون به کار گرفته شد. پس از بررسی ویژگی‌های مشترک و مورد نیاز بین مغزه و لاگ‌ها، تعداد ۱۴ چاه منطبق بر ۷۷۸۵ نمونه مغزه نهایی شد. این تعداد نمونه مغزه بر اساس رابطه‌ی (۸) با نمونه لاگ‌ها موجود در هر چاه به صورت مستقل هم عمق شده‌اند:

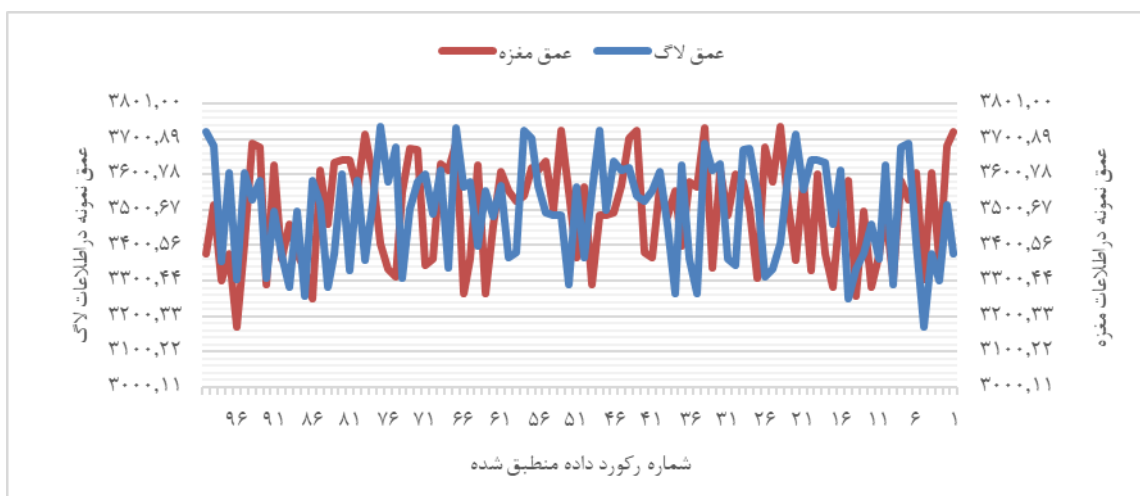
$$D_{Euclidean} = \sqrt{(depth_{core} - depth_{log})^2 + (porosity_{core} - porosity_{log})^2} \quad (8)$$

تا دوباره در مقایسات بعدی دخالت داده نشوند و از بروز داده‌های تکراری جلوگیری شود و در نهایت داده‌های هم عمق شده در یک جا به عنوان خروجی ذخیره می‌شوند. شکل (۳)، هم عمق سازی داده‌های لاگ و مغزه را بر اساس عمق و فاصله‌ها برای ۱۰۰ نقطه نمونه پس از انجام عملیات را نشان می‌دهد.

اگرچه می‌توان از معیارهایی مثل فاصله منهتن و سایر معیارهای مشابه نیز برای محاسبه تابع هزینه استفاده نمود. در خاتمه‌ی انجام محاسبات یک ماتریس $M \times N$ حاصل می‌شود که در آن هر عنصر $C(i, j) = c(s_i, t_j)$ نشان داده خواهد شد (Garcia, 2020).

۶.۲ هم عمق سازی داده‌های مغزه و لاگ با استفاده از روش DTW

در این روش پس از محاسبه فاصله‌های اقلیدسی بین هر نمونه از مغزه با کلیه نقاط لاگ در همان چاه، نتیجه محاسبات به صورت صعودی مرتب‌سازی شده و نزدیک‌ترین مقادیر به عنوان دو نقطه هم عمق انتخاب خواهد شد و این دو نقطه که یکی از داده‌های مغزه و دیگری از داده‌های لاگ می‌باشد از نمونه‌ها خارج شده و در یک فایل جداگانه ذخیره می‌شوند



شکل ۳. نمودار برشی از هم عمق سازی مغزه و لاگ

۱.۳ بررسی روش وینلند

جهت اجرای روش وینلند با توجه به رابطه‌ی (۶) میزان $R35$ محاسبه شد. سپس نتایج محاسبه شده به صورت صعودی مرتب گردید. نواحی مختلف آنالیز وینلند، بر اساس اندازه‌ی

۳. نتایج و بررسی

در این بخش نتایج اجرای روش‌های خوشه‌بندی مورد بررسی و ارزیابی قرار می‌دهیم و آن‌ها را مقایسه نموده‌ایم.

²⁶ Microsoft Visual Basic.Net

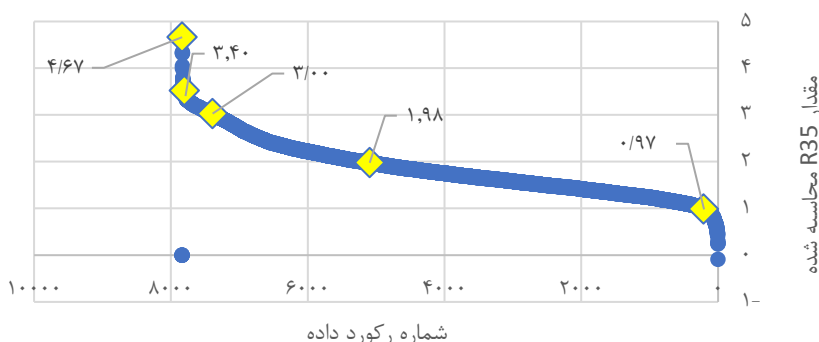
یا اینکه واحد محاسبه $R35$ تغییر مقدار داشته (به عنوان مثال تغییر $R35$ از مقدار ۱ به ۲) است که تعیین کننده خوشه جدید می باشد.

در نهایت بر اساس ۵ بازه فوق، داده ها در خوشه ها، توزیع شد. جدول (۱) پراکندگی داده ها را در هر خوشه بر اساس فراوانی هر برجسب لیتولوژی، مطابق روش وینلند نمایش می دهد. بر اساس جدول (۱)، در خوشه های A و B و C ، اغلب گونه ها از نوع کلسیت و در خوشه های D و E بیشتر از گونه ماسه سنگ می باشد.

شعاع گلوگاه تعیین شد که با توجه به شکل (۴)، خوشه ها به شرح زیر انتخاب شدند:

واحدهایی با $R35$ کمتر از ۰.۹۷ میکرون
واحدهایی با $R35$ بین ۰.۹۷ تا ۱.۹۸ میکرون
واحدهایی با $R35$ بین ۱.۹۸ تا ۳ میکرون
واحدهایی با $R35$ بین ۳ تا ۳.۴ میکرون
واحدهایی با $R35$ بین ۳.۴ تا ۴.۶۷ میکرون

همان طور که در شکل (۴) مشاهده می شود، مقادیر $R35$ به صورت تجمعی جهت تعیین تعداد خوشه ها قرار گرفت، در نقاطی که با علامت لوزی مشخص شده، نمودار تغییر کرده و



شکل ۴. نمودار تجمعی برای تعیین مرزها و محدوده خوشه ها

گونه های سنگ

در این بخش با توجه به بررسی های اولیه روی داده ها از طریق روش های خوشه بندی، روش کی-میانگین به صورت جدی بررسی شد. در پیاده سازی مدل از زبان برنامه نویسی پایتون نسخه ۳، اکوسیستم آناکوندا و چارچوب اسپایدر به صورت ویژه استفاده گردید.

نکته مهم در الگوریتم کی-میانگین تعیین تعداد خوشه ها (پارامتر K) می باشد. روش های آماری مختلفی برای تعیین K وجود دارد که در مدل ما از روش آرنج، استفاده نمودیم. به طوری که مقدار K را بین ۱ تا ۲۰ خوشه در نظر گرفتیم و به کمک رابطه ی (۷)، ۲۰ مرحله الگوریتم کی-میانگین را با استفاده از داده های تراوایی و تخلخل مغزه اجرا کردیم و مقدار SSE را برای هر مقدار K ، در هر مرحله محاسبه نمودیم، سپس نمودار تعداد K بر حسب SSE را رسم کردیم، در نقطه $K=4$ بیشترین شکست نمودار مطابق شکل (۷) مشاهده می شود. در نتیجه تعداد ۴ خوشه بر اساس آنالیز

مطابق شکل (۵) داده های مغزه در خوشه های A, B, C, D, E تقسیم شده اند. بر این اساس محور عمودی شکل (۵)، نشان دهنده لگاریتم تراوایی و محور افقی نشان دهنده تخلخل می باشد و تفکیک رنگ های شکل (۵) بر اساس خوشه های جدول (۱) است.

۲.۲ بررسی روش FZI

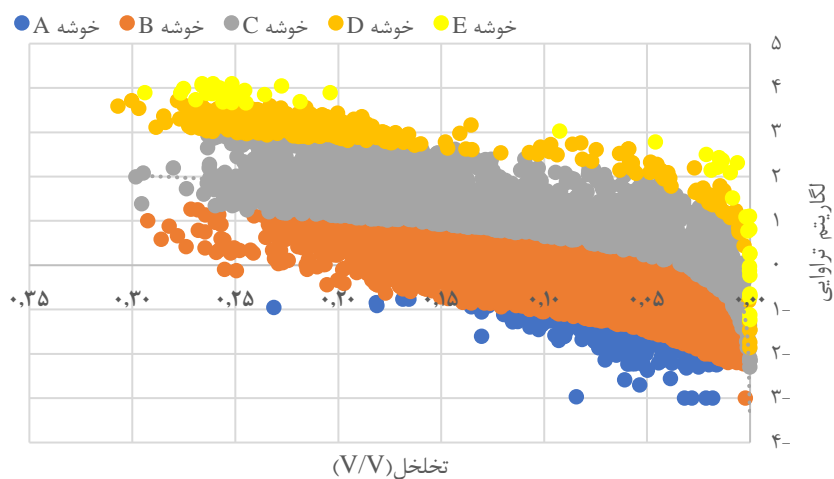
جهت اجرای روش FZI با توجه به رابطه ی (۳)، ابتدا مقادیر FZI محاسبه شده، سپس اعداد به دست آمده به صورت صعودی مرتب گردید، در ادامه برای خوشه بندی مقادیر FZI و تفکیک گسسته گونه های سنگ، مقادیر DRT ، بر اساس رابطه ی (۵) محاسبه شد؛ و پس از آن با رسم نمودار تجمعی شکل (۶)، نقاطی که در خوشه های مشابه قرار داشتند مطابق جدول (۲) قرار گرفت.

۳.۳ بررسی روش کی-میانگین جهت تعیین

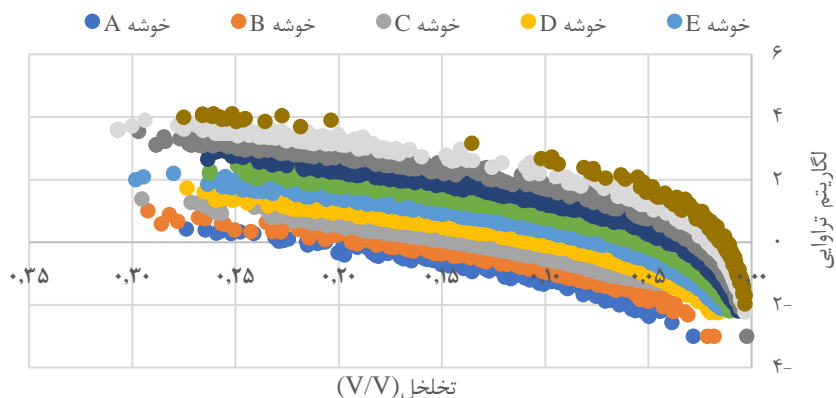
آرنج، برای داده‌ها تعیین شد. پس از اجرای مدل کی- میانگین ضمن بهینه کردن پارامترهای ورودی مدل، نتایج جدول (۳) حاصل گردید:

جدول ۱. فراوانی تعداد نقاط در خوشه‌ها در روش وینلند

خوشه	کلسیت	دولومیت	ماسه	شیل	انیدریت
A	۹۷	۷۰	۵	۷	۱
B	۳۰۳۶	۱۴۰۹	۲۵۱	۲۰۶	۱۱
C	۱۱۲۶	۵۸۱	۴۳۰	۱۱۳	۲
D	۱۱۴	۴۱	۲۵۱	۲۵	۰
E	۱۶	۵	۲۵	۴	۰



شکل ۵. نمودار تخلخل-تراوایی در روش وینلند



شکل ۶. نمودار تخلخل-تراوایی در روش FZ

جدول ۲. فراوانی تعداد نقاط در خوشه‌ها در روش FZI

نام خوشه	کلسیت	دولومیت	ماسه	شیل	انیدریت
A	۱	۸۳	۹۴	۱	۳
B	۱	۳۲۹	۲۳۸	۱۰	۶
C	۴	۷۷۰	۴۰۸	۳۴	۵۲
D	۳	۸۴۹	۴۲۸	۱۱۷	۷۳
E	۱	۷۳۴	۳۳۸	۱۲۰	۶۴
F	۴	۵۳۲	۲۳۲	۱۲۱	۴۸
G	۰	۳۹۱	۱۲۸	۱۵۵	۳۸
H	۰	۳۰۱	۹۵	۲۳۸	۲۸
I	۰	۱۸۱	۶۵	۱۳۳	۱۲
J	۰	۸۷	۳۲	۲۴	۷

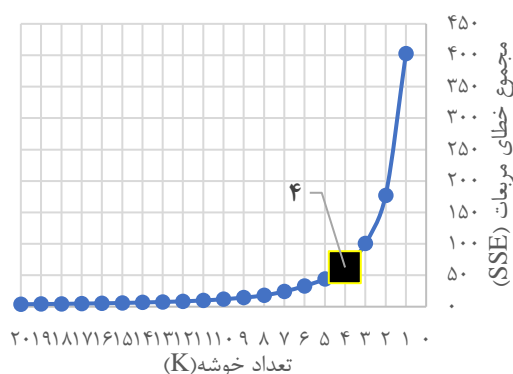
جدول ۳. فراوانی تعداد نقاط در خوشه‌ها در روش کی-میانگین

نام خوشه	کلسیت	دولومیت	ماسه	شیل	انیدریت
A	۵	۰	۱	۳۵۳	۰
B	۶۴۱	۲۱۱۵	۲	۱	۸
C	۴	۰	۹۵۷	۱	۰
D	۳۷۳۹	۰	۲	۰	۶

همچنین در شکل (۸) نمودار تخلخل بر مبنای لگاریتم تراوایی برای خوشه‌های مشخص شده در جدول (۳)، نمایش داده شده است.

۴.۳ ارزیابی و مقایسه

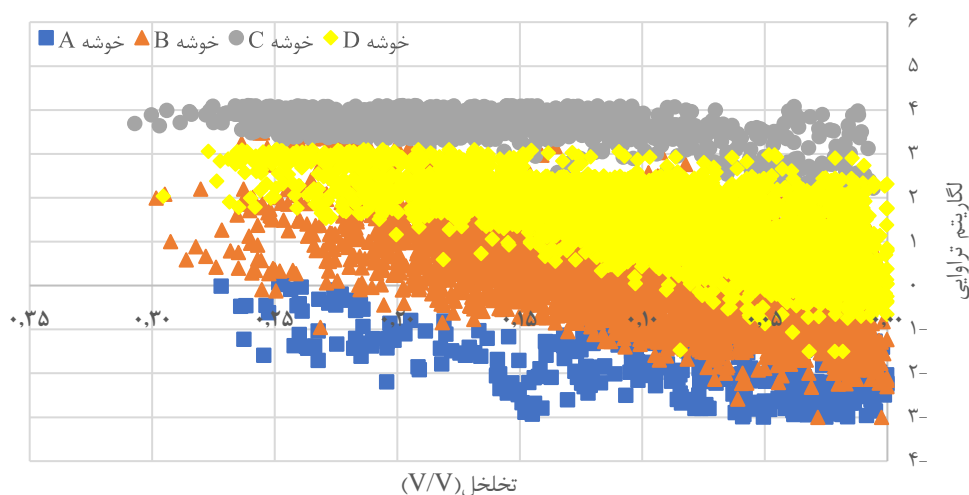
جدول (۴) نتایج تفسیر گونه‌های سنگی مختلف بر اساس داده‌های چاه‌نگاری را نمایش می‌دهد. مطابق این اطلاعات، برچسب‌گذاری نمونه‌ها در اعماق مختلف بر اساس برش^{۲۷} پیشنهادی انجام شد. انتظار می‌رود که چنانچه خوشه‌بندی با کیفیت عالی انجام شود، فاصله بین خوشه‌ها بیشترین و فاصله بین اعضای هر خوشه کمترین باشد



شکل ۷. نمودار تعیین مقدار K با استفاده از روش آرنج

در جدول (۳) همان‌طور که مشاهده می‌کنید در خوشه A بیشترین داده‌ها از گونه سنگی با برچسب شیل هستند و در خوشه B از نوع دولومیت و در خوشه‌های C و D به ترتیب از گونه‌های ماسه و کلسیت می‌باشند که به خوبی تفکیک شده‌اند.

²⁷ Cut Off



شکل ۸. نمودار تخلخل-تراوایی در روش کی-میانگین

جهت اعتبارسنجی نتایج خوشه‌بندی در این پژوهش از دو معیار زیر استفاده شده است: معیار اول: مطابق رابطه‌ی (۹)، تعداد خوشه‌های منطبق بر تعداد برچسب‌های دسته سنگ‌ها در لاگ‌ها، با توجه به اطلاعات لیتولوژی که شامل پنج نوع کلسیت، دولومیت، شیل، ماسه و انیدریت می‌باشد.

جدول ۴. برچسب‌های تعیین شده برای گونه‌های سنگ بر اساس حداکثر حجم ماده

شماره برچسب گونه سنگ	۵	۴	۳	۲	۱
نوع سنگ	شیل	ماسه	کلسیت	انیدریت	دولومیت

$$(9) \quad E_{cc}^{28} = \text{تعداد مورد انتظار خوشه‌ها} \quad \text{می‌نمایم.}$$

(۱۱) معیار دوم: میانگین درصد تشخیص موارد درست منطبق بر

$$Accuracy\ Total = \sum_{i=1}^n TPP_i / \text{count of clusters}$$

اطلاعات لیتولوژی به ازای همه نمونه‌های هر خوشه بر مبنای بیشترین تعداد برچسب‌ها در هر خوشه

$$(10) \quad TPP^{29} = \frac{TP}{TP+FP}$$

در رابطه‌ی (۱۱)، حاصل جمع میانگین‌های درصد تشخیص صحیح نقاط (TPP) را به ازای هر روش خوشه‌بندی بر تعداد خوشه‌ها در همان روش تقسیم شده است که مقدار حاصل شده، شاخص صحت‌سنجی کلی برای آن روش خوشه‌بندی خواهد بود.

TP: تعداد نقاط درست خوشه‌بندی شده بر مبنای تعداد حداکثر نمونه‌ها در هر خوشه

FP: تعداد نقاط نادرست خوشه‌بندی شده بر مبنای تعداد حداکثر نمونه‌ها در هر خوشه

در جدول (۵) نتایج ارزیابی اجرای روش‌های خوشه‌بندی وینلند، FZI و کی-میانگین آورده شده است؛ که نشان می‌دهد تعداد خوشه‌های مورد انتظار در ستون Ecc، بر اساس

پس از محاسبه TPP برای همه خوشه‌ها، به کمک رابطه‌ی (۱۱)، میانگین صحت‌سنجی کل^{۳۰} در هر روش را محاسبه

²⁸ Expected Count of Cluster

²⁹ TRUE POINT PRECISION

³⁰ Accuracy Total

دارای برچسب انیدریت بود و این مقدار نسبت به حجم فراوانی از سایر نمونه سنگ‌ها بسیار ناچیز است، پس می‌توان نتیجه گرفت به خوبی تعداد خوشه‌ها منطبق بر تعداد برچسب‌ها تشخیص داده شده است.

جدول (۴) می‌بایست ۵ خوشه باشد. در روش کی-میانگین همان‌طور که در بخش (۳.۳) روش آرنج، برای تعیین تعداد خوشه‌های مدل کی-میانگین اشاره شد، داده‌ها در ۴ خوشه قرار گرفتند و همان‌گونه که فقط تعداد ۱۴ نمونه از سنگ‌ها

جدول ۵. مقایسه نتایج ارزیابی روش‌های خوشه‌بندی

روش تعیین گونه سنگ	تعداد خوشه‌های مورد انتظار	تعداد خوشه‌های به‌دست آمده	دقت کل
کی-میانگین	۵	۴	۰.۹۳۵
<i>FZI</i>	۵	۵	۰.۵۴۶
وینلند	۵	۱۰	۰.۵۴۳

بررسی شده بهترین عملکرد را با دقت ۹۳.۵ درصد داشت و در این مورد صحت خوشه‌ها با تطبیق بر اطلاعات لیتولوژی رؤیت شد.

موارد زیر در نتایج این پژوهش قابل توجه هستند:

۱- روش یادگیری ماشین بدون نظارت می‌تواند در خوشه-بندی ویژگی‌های مخزن استفاده شود و جایگزین خوبی برای روش‌های تجربی باشد.

۲- روش کی-میانگین در مقایسه با سایر روش‌های یادگیری نیاز به برچسب‌گذاری اطلاعات از قبل ندارد و مستقل عمل می‌کند.

۳- کاربرد هم عمق سازی اطلاعات مغزه و لاگ قبل از خوشه-بندی بسیار با اهمیت است.

برای کارهای آتی پیشنهاد می‌شود ویژگی‌های بیشتری از پارامترهای مغزه که به‌صورت مشترک در مغزه و لاگ وجود دارد را جهت خوشه‌بندی استفاده نمود که در این مطالعه محدودیت دسترسی به آن اطلاعات وجود داشت. همچنین نتایج این تحقیق در آینده می‌تواند جهت تخمین تراوایی مخازن نفتی مورد استفاده قرار گیرد.

۵. سپاس‌گزاری

نویسندگان بر خود لازم می‌دانند از مدیران و مسئولین شرکت ملی مناطق نفت‌خیز جنوب به‌خصوص اداره مهندسی پتروفیزیک به سبب همکاری و در اختیار گذاشتن داده‌های مورد استفاده تشکر نمایند

در نتایج ارزیابی جدول (۵) مشخص شد که روش یادگیری ماشین کی-میانگین عملکرد خوبی در تشخیص نمونه‌ها و خوشه‌بندی اطلاعات مخزن داشت و در بین روش‌های

۵.۳ مقایسه و بررسی نمودارهای تخلخل-تراوایی

مطابق شکل‌های ۵ و ۶ و ۸ همان‌گونه که در شکل (۵) نمایش داده شده، تعداد ۵ خوشه با ۵ رنگ مختلف نشان داده شده است و در شکل (۶) تعداد ۱۰ خوشه با ۱۰ رنگ متفاوت نمایش داده شده و به همین ترتیب در شکل (۸) نیز تعداد ۴ خوشه نمایش داده شده است. تفاوت نمودارها مربوط به تعداد خوشه‌ها و مرزهای هر خوشه و همچنین مرتب‌سازی داده‌ها بر اساس برچسب خوشه‌ها می‌باشد به‌عنوان مثال در روش *FZI* مقادیر داده‌ها بر اساس مقدار محاسبه شده برای *FZI* مرتب‌سازی شده است. البته پراکندگی در نمودار شکل (۸) به دلیل انجام خوشه‌بندی دقیق‌تر به‌صورت خطی نمایش داده شده است.

۴. جمع‌بندی و پیشنهادات

در این مطالعه داده‌های مغزه و لاگ در میدان نفتی مارون در اختیار قرار گرفت و پس از هم عمق سازی و استانداردسازی داده‌ها جهت تعیین گونه‌های سنگی با استفاده از روش‌های محاسباتی و روش یادگیری ماشین کی-میانگین بررسی شد. مدل یادگیری غیر نظارتی کی-میانگین توانست ضمن خوشه-بندی داده‌ها در چهار دسته، بهترین تطبیق را در مقایسه با اطلاعات لیتولوژی مخزن داشته باشد و با شاخص صحت سنجی ۹۳.۵ درصد داده‌ها را تفکیک نمود. خطای باقی مانده به دلیل ناهمگن بودن و پیچیدگی مخزن می‌باشد. همچنین

۶. مراجع

Lucia, F., Martin, A.J., Solomon, S.T., Hartmann, D.J. (1999). Characterization of petrophysical flow units in carbonate reservoirs: Discussion, AAPG bulletin, Vol. 83, No. 7, pp. 1161-1163.

Potter, G. (2010). Core analysis uncertainly and rock typing, in SPWLA Carbonate Workshop, Abu Dhabi.

Merandy, p., Muhammad, N. (2016). An Analysis of Rock Typing Methods in Carbonate Rocks for Better Carbonate Reservoir Characterization: A Case Study of Minahaki Carbonate Formation, Banggai Sula Basin, Central Sulawesi, 41th Scientific Annual Meeting of Indonesian Association of Geophysicists, Lampung, Indonesia, pp. 68-82 .

Garcia, A. (2020). Automatic Depth Matching for Petrophysical Borehole Logs. Delft University of Technology, Delf, MA.

Aggarwal, C. C. & Reddy, C. K. (eds.) (2014). Data Clustering: Algorithms and Applications. CRC Press.

Ayyadevara, V. K. (2018). Pro Machine Learning Algorithms: A Hands-On Approach to Implementing Algorithms in Python and R. In Apress.

Lakhmi C. Jain. (2016). Foundations and Methods in Combinatorial and Statistical Data Analysis and Clustering, Fourth Edition, Springer Press.

Wang, J. (2003). Data Mining: Opportunities and Challenges, USA Montclair, IRM Press.

Al Kattan, W., AL Jawad, S. N., & Jomaah, H. (2018). Cluster Analysis Approach to Identify Rock Type in Tertiary Reservoir of Khabaz Oil Field Case Study, Iraqi Journal of Chemical and Petroleum Engineering, Vol. 19, No. 2, pp. 9-13.

Ali, A., Sheng-Chang, C. (2020). Characterization of well logs using K-mean cluster analysis. Journal of Petroleum Exploration and Production Technology, Springer, Vol. 10, No. 6, pp. 2245-2256.

Rebelle, M., Lalanne, B. (2014). Rock-typing in carbonates: A critical review of clustering methods, Society of Petroleum Engineers - 30th Abu Dhabi International Petroleum Exhibition and Conference, Abu Dhabi, UAE.

Mirzaei, A., Asadolahpour, S. R., Saboorian-Jooybari, H. (2020). A new framework for selection of representative samples for special core analysis, Petroleum Research Publishing Services by Elsevier, Vol. 5, No. 3, pp 210-226.

Guo, G., Diaz, M. A., Paz, F., Smalley, J., & Waninger, E. A. (2007). Rock typing as an effective tool for permeability and water-saturation modeling: A case study in a clastic reservoir in the Oriente basin, SPE Reservoir Evaluation and Engineering, Vol. 10, No. 6. <https://doi.org/10.2118/97033-pa>.

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 $\Delta\log R$ and Neural Network Methods: Ahwaz and Marun Oilfields SW of Iran. Petroleum Science and Technology, Vol. 29, No. 16, pp 1691–1704. <https://doi.org/10.1080/10916461003620495>.

قربان پور، ح؛ (۱۳۸۶)، افزایش دقت ارزیابی-های پتروفیزیکی با استفاده از کالیبراسیون داده-های نمودارهای پتروفیزیکی با مغزه در یکی از میادین شرکت نفت مناطق مرکزی ایران، پایان‌نامه کارشناسی ارشد.

ذیفن، م؛ (۱۳۹۸)، تعیین واحدهای جریانیهیدرولیکی و دسته-بندی نوع سنگ در مخازن نفتی با استفاده از داده-کاوی (مطالعه موردی: یکی از مخازن نفتی جنوب غربی ایران). ماهنامه علمی اکتشاف و تولید نفت و گاز، شماره ۱۷۳، صفحات ۴۴-۵۰.

سپهری، ق؛ زینالی، م؛ (۱۳۹۶). تعیین الکتروفاسیس-های مخزن A در یکی از میادین نفتی ایران با استفاده از روش MRGC، ماهنامه علمی اکتشاف و تولید نفت و گاز، شماره ۱۴۹، صفحات ۴۴-۴۸.

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.

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