

شناسایی منطقه هیدروکربنی با استفاده از روش یادگیری ماشین بردار پشتیبان

محمدرضا قاسمی نژاد^۱، مجید باقری^{۲*} و مجید نبی بیدهندی^۳

^۱ دانش آموخته کارشناسی ارشد ژئوفیزیک، مؤسسه ژئوفیزیک دانشگاه تهران، تهران، ایران

^۲ استادیار گروه فیزیک زمین، مؤسسه ژئوفیزیک دانشگاه تهران، تهران، ایران

^۳ استاد گروه فیزیک زمین، مؤسسه ژئوفیزیک دانشگاه تهران، تهران، ایران

(تاریخ دریافت: ۱۴۰۰/۰۱/۱۳، تاریخ پذیرش: ۱۴۰۰/۰۳/۱۰)

چکیده

تعیین منطقه هیدروکربنی یکی از اساسی‌ترین مراحل اکتشاف است. روش‌های گوناگونی برای تعیین منطقه هیدروکربنی وجود دارد که پیدا کردن مناسب‌ترین روش یک مجموعه داده خاص، علاوه بر افزایش سرعت محاسبات، دقت پیش‌بینی‌ها و تخمین‌های حاصل از تحلیل را بهبود می‌بخشد. ماشین بردار پشتیبان از جمله روش‌های تحلیل داده است که با استفاده از تابع کرنل می‌تواند ارتباط بهتری بین منطقه هیدروکربنی و پارامترهای داده‌ها برقرار کند؛ بنابراین نتیجه بهتری برای تخمین و دسته‌بندی داده‌ها به دست می‌آید. این روش به دلیل حاشیه تفکیک‌کننده در دسته‌بندی داده‌ها در مرحله آموزش، توانایی بیشتری در تعمیم‌پذیری و تفکیک داده‌های واقعی نسبت به بقیه تفکیک‌کننده‌ها از جمله شبکه عصبی دارد. در این مقاله تعیین منطقه هیدروکربنی با استفاده از داده‌های چاه و لرزه‌ای انجام شده است. در ابتدا با استفاده از نگاره‌های چاه، گزارش تفسیری چاه منطقه هیدروکربنی در بخش‌های مختلف سازند آسماری مشخص شد و با استفاده از روش ماشین بردار پشتیبان سعی شد که منطقه هیدروکربنی با استفاده از نگاره‌ها به دست آید. نتایج تفکیک‌کننده بردار پشتیبان با گزارش تفسیری چاه مقایسه شد. نتایج دسته‌بندی این روش با منطقه هیدروکربنی چاه سازگاری خوبی داشت. در مرحله بعد، تعیین منطقه هیدروکربنی با استفاده از داده‌های لرزه‌ای به روش ماشین بردار پشتیبان انجام شد که در ابتدا نشانگرهای لرزه‌ای از رد لرزه‌ای در محل چاه استخراج شد. سپس با استفاده از ماتریس کوواریانس و رسم سرتاسری مربوط به نشانگرهای لرزه‌ای، نشانگرهای لرزه‌ای مناسب برای مشخص کردن منطقه هیدروکربنی تعیین شدند. در ادامه، به منظور اعتبارسنجی نتایج این روش از رد کناری چاه برای داده اعتبارسنجی استفاده شد. نتایج این روش با بخش‌های هیدروکربنی با درصد خطای کمی همخوانی داشت.

واژه‌های کلیدی: تعیین منطقه هیدروکربنی، ماشین بردار پشتیبان، نگاره‌های چاه، نشانگرهای لرزه‌ای، اعتبارسنجی

۱ مقدمه

هدف از تحلیل رخساره، به دست آوردن پارامترهای پتروفیزیکی مهم مخزن و شناسایی مرزهای ناهمگنی در زیر سطح زمین است. نتایج تفسیر پارامترهای پتروفیزیکی، ورودی فرایند مدل سازی سه بعدی مخزن است و توزیع پارامترهای مخزنی در فضای سه بعدی با استفاده از این پارامترها صورت می گیرد. این مدل کاربرد زیادی در بخش های مختلف از جمله اکتشاف و حفاری چاه های جدید، ازدیاد برداشت از یک مخزن، تعیین مناطق مناسب تر برای ازدیاد برداشت، کاهش خطر حفاری، تعیین سنگ شناسی مخزن، تعیین چاه های کلیدی و تعمیر آن به سایر چاه های موجود در منطقه دارد. مهم ترین پارامترهای پتروفیزیکی عبارت اند از: حجم شیل؛ تخلخل؛ تراوایی؛ اشباع سیالات مخزن؛ سنگ شناسی مخزن.

بررسی رخساره های لرزه ای از دهه ۹۰ میلادی آغاز شده است و با توجه به اهمیت و کاربرد آن در توصیف مخزن، همواره پژوهشگران متعددی به آن توجه کرده اند. برای مثال ماتئو و رایس (۱۹۹۹) از تحلیل عامل تفکیک کننده برای به دست آوردن نسبت ماسه به شیل در منطقه های مخزنی گوناگون استفاده کردند. هاگان (۱۹۸۲) از تحلیل مؤلفه های اصلی برای بررسی تغییرات جانبی در تخلخل استفاده کرد. ماتلاگ و همکاران (۱۹۸۵) برای تعیین مرز تغییر رخساره های ماسه ای از یک تابع تصمیم گیری خطی استفاده کردند. دومای و فورنیر (۱۹۸۸) هر دو روش تحلیل مؤلفه های اصلی و تحلیل عامل تفکیک کننده را برای تعیین رخساره های لرزه ای به کار گرفتند. سیمان (۱۹۹۱) سامانه ای کارا بر پایه اطلاعات طراحی کرد تا مقاطع لرزه ای را بر اساس ساختارشان قطعه بندی کند.

فورنیر و همکاران (۲۰۰۲) از رخساره های لرزه ای برای منطقه بندی افقی مخزن استفاده کردند. آلیانا و هنس (۲۰۰۲)، وست و همکاران (۲۰۰۲) و سقاف و همکاران

(۲۰۰۳) با استفاده از شبکه های عصبی و داده های چاه، تحلیل رخساره را انجام دادند و از آنها در به نقشه در آوردن ناهمگنی های مخزن استفاده کردند. لیناری و همکاران (۲۰۰۳) و فرزادی (۲۰۰۶) از نشانگرهای سه بعدی داده های لرزه ای و خوشه بندی فضای ویژگی برای مدل سازی رخساره های لرزه ای در محدوده مخزن استفاده کردند. الکسیس و تانوی (۲۰۰۸) برای تحلیل رخساره مخزن از اطلاعات زمین شناسی و طبقه بندی سلسله مراتبی استفاده کردند. ماروکوئین و همکاران (۲۰۰۹) یک داده - کاوی بصری برای تحلیل رخساره انجام دادند. همچنین باردینی و همکاران (۲۰۱۰) از نشانگرهای لرزه ای چند - گانه برای مدل سازی سه بعدی رخساره های لرزه ای استفاده کردند. حسین و موکرچی (۲۰۱۱) و پاپاروزی و همکاران (۲۰۱۱) از روش های آماری و احتمالاتی برای شبیه سازی رخساره ها در محدوده مخزن استفاده کردند. باقری و همکاران (۲۰۱۴) تحلیل رخساره لرزه ای را با استفاده از داده های چاه و روش های بانظارت در روش های مختلف یادگیری ماشین انجام دادند. استفاده از داده های چاه برای تعیین منطقه هیدروکربنی به صورت محلی و در ادامه، استفاده از داده های لرزه ای به منظور بررسی پراکندگی بخش های هیدروکربنی، گام مؤثری در گسترش میدان های نفتی است. این مهم در پژوهش حاضر با استفاده از ماشین بردار پشتیبان حاصل شده است که روشی نوین با قدرت تعمیم پذیری بیشتر نسبت به دیگر روش های دسته بندی است. نتایج این تحقیق نشان می دهد ماشین بردار پشتیبان، روشی قدرتمند در دسته بندی داده های چاه و داده های لرزه ای برای تعیین منطقه هیدروکربنی است.

۲ روش تحقیق

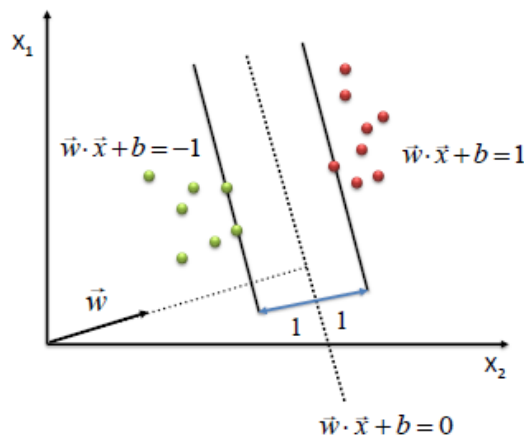
وینیک (۱۹۹۵) الگوریتم تعمیم یافته ای را برای ساخت ابرصفحه های تفکیک کننده با حاشیه بهینه معرفی کردند.

مقایسه کرد. علت این قضیه را نیز می‌توان توانایی این روش در حل مسائل گوناگون و تعمیم زیاد آن دانست. اگر رابطه (۱) معادله یک خط در صفحه فرض شود:

$$\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \mathbf{W} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}, \quad (1)$$

$$\mathbf{W}^T \mathbf{X} + b = 0$$

در رابطه (۱)، \mathbf{X} بردار مختصات نقاط و \mathbf{W} بردار نرمال بر خط تفکیک کننده می‌باشد. خط تفکیک کننده، صفحه را به دو قسمت تقسیم می‌کند که در یک طرف آن مقدار $\mathbf{W}^T \mathbf{X} + b$ مثبت و در طرف دیگر منفی است. برای ایجاد حاشیه در دو طرف خط تفکیک کننده می‌توان دو خط موازی آن را با دادن مقادیر دلخواه مثبت و منفی ایجاد کرد (شکل ۱).



شکل ۱. خط تفکیک کننده صفحه را به دو بخش تقسیم کرده است. با ایجاد حاشیه مثبت و منفی دلخواه که در این حالت +۱ و -۱ است، روش ماشین بردار پشتیبان قدرت تعمیم‌پذیری بیشتری پیدا خواهد کرد.

در نهایت، با داشتن \mathbf{W} و b و معادله تفکیک کننده می‌توان کلاس داده‌های جدید \mathbf{X} را بررسی کرد. با استفاده از تابع علامت و معادله شروط داده‌ها کلاس داده‌های جدید مشخص می‌شود:

$$y_i = +1 \rightarrow \mathbf{W}^T \mathbf{X}_i + b \geq +1 \rightarrow \text{sign}(\mathbf{W}^T \mathbf{X}_i + b) = 1 \quad (4)$$

$$y_i = -1 \rightarrow \mathbf{W}^T \mathbf{X}_i + b \leq -1 \rightarrow \text{sign}(\mathbf{W}^T \mathbf{X}_i + b) = -1 \quad (5)$$

الگوریتم ماشین بردار پشتیبان (Support Vector Machines, SVM) تعمیم غیرخطی آن الگوریتم است. وپنیک (1998) ایده بیشینه حاشیه اصلاح شده را مطرح کردند. در این روش، خطای طبقه‌بندی نادرست در مدل SVM در نظر گرفته شده است و الگوریتم بیشینه حاشیه با تعریف متغیر جریمه در مدل بهینه‌سازی، به مسائلی تعمیم داده شد که به صورت خطی جداپذیر نیستند.

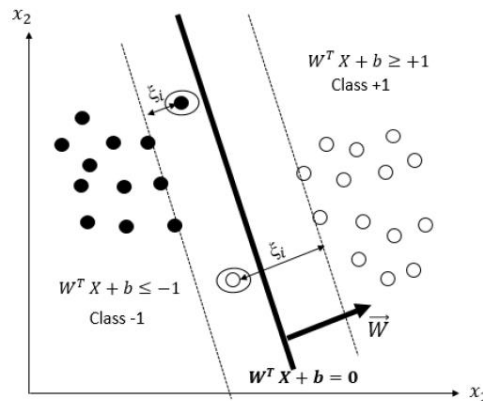
یکی از روش‌هایی که در حال حاضر به صورت گسترده برای مسئله دسته‌بندی استفاده می‌شود، روش SVM است. شاید به گونه‌ای بتوان محبوبیت کنونی روش ماشین بردار پشتیبان را با محبوبیت شبکه‌های عصبی در دهه گذشته

برای به دست آوردن معادله ابرصفحه جداکننده با بیشترین حاشیه باید \mathbf{W} و b را به نحوی تعیین کرد که مقدار حاشیه بیشینه شود. پس با یک مسئله بهینه‌سازی روبه‌رو هستیم که هدف آن بیشینه‌سازی مقدار $\frac{1}{\|\mathbf{W}\|}$ است. برای این کار کافی است مقدار کمینه $\|\mathbf{W}\|$ را به دست آورد.

برای داده‌ها دو معادله زیر برقرار است:

$$\mathbf{W}^T \mathbf{X}_i + b \geq +1 \quad \forall \mathbf{X}_i \text{ with } y_i = +1 \quad (2)$$

$$\mathbf{W}^T \mathbf{X}_i + b \leq -1 \quad \forall \mathbf{X}_i \text{ with } y_i = -1 \quad (3)$$



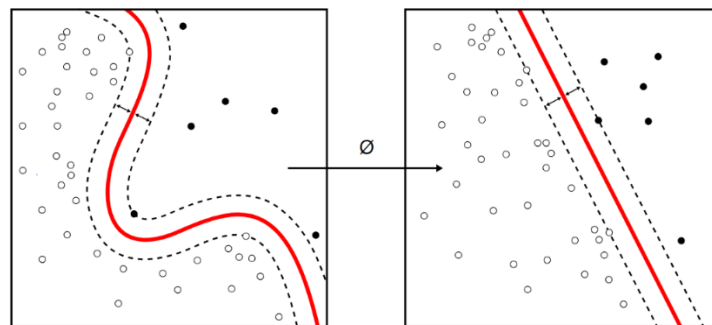
شکل ۲. هنگامی که داده‌ها به صورت خطی تفکیک‌پذیر هستند اما تعداد کمی از آنها از حاشیه ایجادشده تخطی می‌کنند، جداسازی دو کلاس با ضریب جریمه ξ_i صورت می‌گیرد.

انتقال این داده‌ها از فضای ورودی به فضای ویژگی (feature space) با ابعاد بیشتر است به طوری که در فضای ویژگی، این داده‌ها از یکدیگر به صورت خطی جدایی‌پذیر باشند. فرض کنید تابع φ داده‌های آموزشی را از فضای ورودی (X) به فضای ویژگی (Z) می‌نگارد. برای هر نقطه از داده‌های ورودی رابطه زیر برقرار است:

$$\varphi = X \rightarrow \varphi(X) \quad (۶)$$

شکل (۳) نشان می‌دهد هر نقطه در فضای ورودی که دو

حال موردی را در نظر بگیرید که داده‌های آموزشی نمی‌توانند بدون خطا به صورت خطی از یکدیگر تفکیک شوند و تعدادی داده نوفه‌دار وجود دارد که در کلاس مورد نظر خود دسته‌بندی نشده‌اند و خارج از محدوده کلاس خود قرار می‌گیرند. در مواردی که تعداد این داده‌های نوفه‌دار کم باشد، می‌توان باز هم داده‌ها را به صورت خطی از یکدیگر تفکیک کرد. برای این کار متغیر کمکی جدید ξ_i یا متغیر جریمه (slack variable) معرفی می‌شود: اگر تفکیک داده‌ها به صورت خطی از یکدیگر امکان‌پذیر نباشد، ایده روش ماشین بردار پشتیبان،



شکل ۳. داده‌هایی که به صورت خطی تفکیک‌پذیر نیستند (شکل سمت چپ)، با انتقال به فضای ویژگی با ابعاد بیشتر به صورت خطی تفکیک‌پذیر می‌شوند (شکل سمت راست) (فخاری و همکاران، ۱۳۹۶).

تبعیت می‌کند. میدان مورد مطالعه دو برآمدگی در افق آسماری دارد. امتداد و ساختار این تاق‌دیس در جهت شمال غرب - جنوب شرق قرار دارد. در این میدان، توالی رسوبی به صورت طبیعی وجود دارد و هیچ گونه اثری از گسل و روران‌دگی یا بهم‌ریختگی ساختمانی مشاهده نشده است. سازند مخزنی در این میدان، سازند آسماری است. سازند آسماری در چاه مورد مطالعه این تاق‌دیس به لحاظ ویژگی‌های سنگ‌شناسی دربرگیرنده سنگ‌های کربناته از نوع سنگ آهک دولومیتی، سنگ آهک، دولومیت و همچنین شیل و مقداری ماسه‌سنگ است. مرز زیرین این سازند با سازند پابده به صورت تدریجی و هم‌شیب در عمق ۴۲۵۲ متری و مرز بالایی آن با بسته شدن حوضه و ته‌نشست تبخیری‌ها، با سازند گچساران در عمق ۳۸۱۸/۵۰ متری به صورت ناپیوسته و هم‌شیب است. مرز زیرین با سازند پابده در عملیات سرچاهی در عمق ۴۱۲۶ متری تعیین شده است.

۲-۳ نتایج روش

در این پژوهش، از ماشین بردار پشتیبان به عنوان یک روش نظارتی در دسته‌بندی داده‌ها استفاده شده است. در اولین مرحله سعی شد ماتریس داده‌های ورودی و ماتریس نتیجه یا خروجی داده‌ها برای اعمال روش ماشین بردار پشتیبان روی داده‌ها و بررسی عملکرد این روش آماده شود. با توجه به اینکه هدف این تحقیق، تعیین منطقه هیدروکربنی با استفاده از نگاره‌های چاه و داده‌های لرزه‌ای است، با بررسی گزارش تفسیری چاه مورد نظر و نگاره‌های مؤثر در تعیین منطقه هیدروکربنی، بخش‌های هیدروکربنی در چاه مشخص شد. در شکل ۴ نگاره‌ها در مقابل منطقه هیدروکربنی نمایش داده شده است. از آنجاکه سعی شده است بخش‌های هیدروکربنی بر اساس ویژگی‌های هرکدام از نگاره‌ها مشخص شود، این بخش‌ها دقیق‌تر مشخص شده‌اند.

بعد X_1 و X_2 دارد، با نگاشت به فضای ویژگی با ابعاد Z_1 ، Z_2 و Z_3 انتقال داده می‌شود. این انتقال سبب می‌شود داده‌های آموزشی که در فضای ورودی به صورت خطی تفکیک‌پذیر نبودند، در فضای ویژگی به صورت خطی با صفحه نمایش داده‌شده تفکیک‌پذیر شوند. روابط جدید با جاگذاری به صورت زیر نوشته می‌شوند:

$$(7) \quad W^T \varphi(X_i) + b = 0 \quad \text{فضای ویژگی}$$

$$(8) \quad \frac{1}{2} W^T W + \text{تابع اولیه جهت کمیته‌سازی}$$

$$(9) \quad (W^T \varphi(X_i) + b) \geq 1 - \xi_i \quad \text{و} \quad \xi_i \geq 0, \quad i = 1, \dots, n$$

قیدهای مسئله در فضای جدید هستند. چون تفکیک‌کننده در فضای ویژگی جدید خطی است، با معرفی ضرایب لاگرانژ و ایجاد تابع اولیه لاگرانژین L_p و سپس ایجاد فرم دوگان L_D ، مسئله بهینه‌سازی به وجود آمده به صورت معادله (۹) است:

$$(9) \quad W = \sum_{i=1}^n \alpha_i y_i \varphi(X_i) \quad \text{تابع وزنی}$$

$$(10) \quad \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \varphi(X_i^T) \varphi(X_j) - \sum_{i=1}^n \alpha_i \quad \text{تابع هدف جهت کمیته‌سازی}$$

که دقیقاً مانند آنچه در جداکننده خطی بود، مسئله بهینه‌سازی یک مسئله بهینه‌سازی درجه دو است که با حل آن، ضرایب α_i به دست می‌آیند. خروجی تابع جداکننده به صورت رابطه زیر است:

$$(11) \quad \text{sign}(\sum_{i=1}^n \alpha_i y_i \varphi(X_i^T) \varphi(X_k) + b) = \text{class}(X_k)$$

۳ اعمال روش

۱-۳ زمین‌شناسی میدان مورد مطالعه

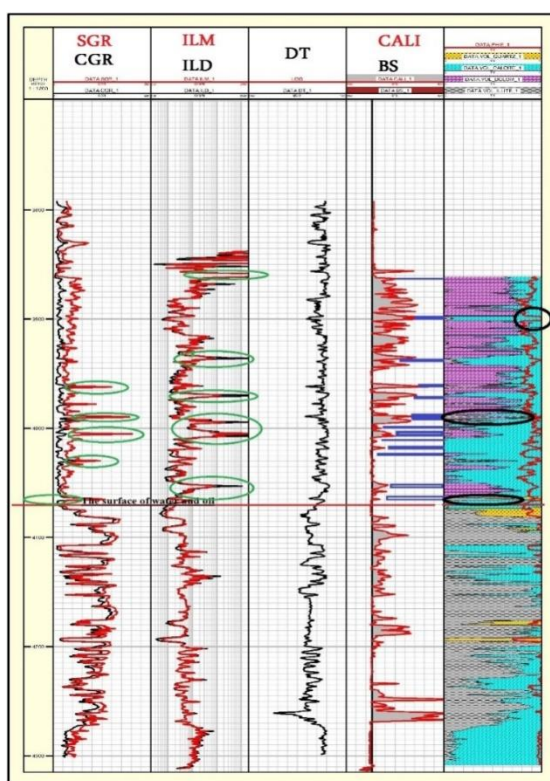
میدان نفتی مورد مطالعه یک تاق‌دیس زیرسطحی است که در سطح زمین با رسوبات آبرفتی عهد حاضر پوشیده شده است و از نظر روند ساختمانی، از روند کلی زاگرس

بگیرد. هرچه حجم داده آزمون کمتر شده است، خطا افزایش یافته است؛ چون طبقه‌بندی نکردن هر داده به دلیل کاهش حجم داده‌های آزمون تأثیر بیشتری بر خطا دارد (شکل ۵).

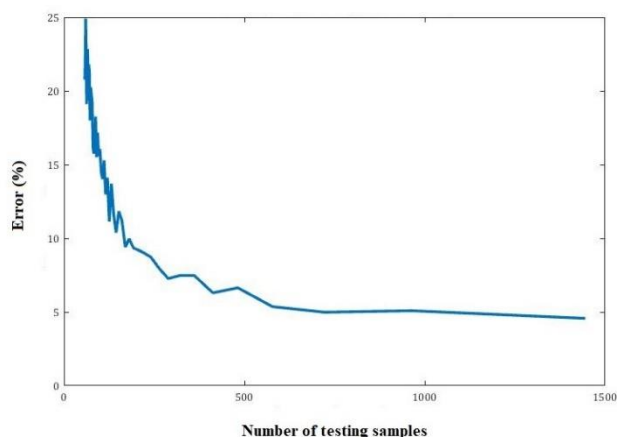
در شکل (۶) کلاس واقعی داده‌ها و نتیجه از روش ماشین بردار پشتیبان به روش اعتبارسنجی متقاطع در کنار هم نمایش داده شده است. داده‌ها در چهار مرحله بررسی شد و بر اساس اصول روش‌های یادگیری ماشین، در هر مرحله ۷۵ درصد داده‌ها برای آموزش و ۲۵ درصد برای اعتبارسنجی استفاده شد. میزان همبستگی بین خروجی ماشین بردار پشتیبان و نتایج گزارش تفسیری چاه و نگاره‌ها ۹۴ درصد است.

برای تشخیص منطقه تولید با استفاده از نگاره‌های گاما و مقاومت و نتایج تخمین سنگ‌شناسی، بخش‌های حاوی هیدروکربن از دیگر قسمت‌های چاه تشخیص داده شده‌اند. با کمک گرفتن از نگاره‌های چاه می‌توان بخش‌های هیدروکربنی را بهتر مشخص کرد؛ برای مثال بالا بودن نگاره‌های گاما و مقاومت در چاه نشان‌دهنده وجود هیدروکربن است. در بخش‌هایی که شیل موجود است نیز به دلیل کاهش تخلخل، احتمال وجود هیدروکربن کاهش می‌یابد.

در این مورد نیز تعداد داده‌های آموزش و آزمون در هر مرحله با هم چه از نظر مقدار و چه از نظر نقطه داده‌برداری متفاوت هستند و سعی شده است که با استفاده از تغییر حجم داده‌ها در هر مرحله اعتبارسنجی نیز صورت



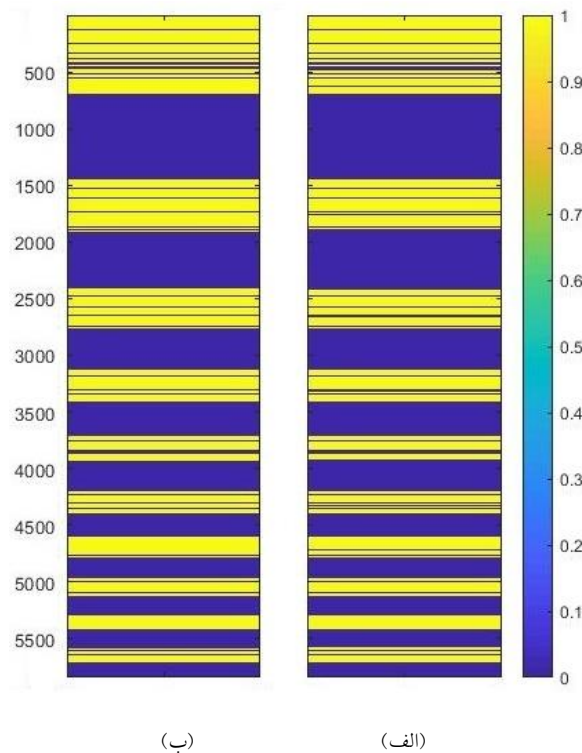
شکل ۴. ترسیم نگاره‌های مؤثر در تعیین منطقه هیدروکربنی. از راست به چپ: نگاره کلیپر (قطریاب)، نگاره قطر مته حفاری، نگاره صوتی، نگاره مقاومت ویژه عمیق و متوسط و نگاره‌های گامای طبیعی. در مقابل، ستون سنگ‌شناسی برای نمایش بخش‌های هیدروکربنی با کمک گزارش تفسیری چاه نمایش داده شده است.



شکل ۵. تغییرات میزان خطا بر اساس حجم داده‌های آموزش. حجم داده‌های آموزشی در مرحله آموزش ماشین بردار پشتیبان مؤثر است و باعث افزایش قدرت تعمیم‌پذیری تفکیک‌کننده می‌شود. هر اندازه حجم داده‌های آموزشی افزایش پیدا کند، میزان خطا کاهش می‌یابد.

می‌کنند. نشانگر مؤثر، مقاومت صوتی است که اصلی‌ترین ویژگی در تعیین منطقه هیدروکربنی فرض می‌شود. ویژگی‌های بعدی، نشانگرهایی هستند که همبستگی کمتری با مقاومت صوتی دارند.

برای تعیین منطقه هیدروکربنی با استفاده از داده‌های لرزه‌ای، در ابتدا تعدادی از نشانگرهای رایج استخراج می‌شود. سپس از بین این نشانگرها، نشانگرهایی انتخاب می‌شوند که کمک بیشتری به تفکیک منطقه هیدروکربنی



شکل ۶. (الف) دسته‌بندی واقعی داده‌ها در محل چاه (ب) نتایج دسته‌بندی با روش SVM. بر اساس اعتبارسنجی متقاطع که از ۷۵ درصد داده‌ها برای آموزش و ۲۵ درصد داده‌ها برای آزمون استفاده شده است، میزان همبستگی ۹۶ درصد به دست آمده است.

جدول ۱. همبستگی میان چهار نشانگر مورد استفاده در جهت تحلیل رخساره هیدروکربنی شامل قطبش ظاهری، پوش، مقاومت صوتی نسبی و جذب. میزان همبستگی این نشانگرها با یکدیگر بسیار کم است و به همین دلیل، قدرت زیادی در تفکیک‌پذیری و شناسایی منطقه هیدروکربنی دارند.

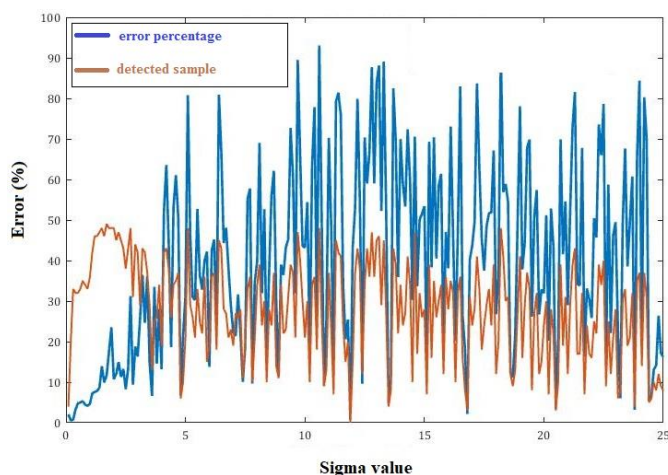
نشانگر	مقاومت صوتی	پوش	قطبش ظاهری	جذب
قطبش ظاهری	۰/۱۰۹۱	۰/۲۳۷۳	۱	۰/۰۶۲۰
پوش	۰/۰۶۳۳	۱	۰/۲۳۷۳	۰/۰۲۳۱
مقاومت صوتی	۱	۰/۰۶۳۳	۰/۱۰۹۱	۰/۲۳۶۶
جذب	۰/۲۳۷۳	۰/۱۷۱۲	۰/۰۶۲۰	۱

می‌شود تا مرزهای توالی‌های رسوبی را نمایان سازد. تباین زیاد مقاومت صوتی، سطوح دگرشیبی و ناپیوستگی‌ها را نشان می‌دهد. تباین زیاد همچنین می‌تواند با تخلخل زیاد ساختار و وجود محتوای سیالی در یک مخزن هیدروکربنی ارتباط داشته باشد. نشانگر جذب از جمله نشانگرهای فیزیکی است و ارتباط زیادی با تخلخل، نفوذپذیری و شکستگی دارد (تانر، ۲۰۰۱). با توجه به اجتماع هیدروکربن‌ها در نواحی متخلخل و تضعیف زیاد دامنه در این نواحی، به‌ویژه در مناطق گازدار، این نشانگر برای پیدا کردن مخازن هیدروکربنی از توانایی زیادی برخوردار است و می‌تواند مفید واقع شود. برای حصول اطمینان از صحت نتیجه تعیین منطقه هیدروکربنی با استفاده از داده‌های لرزه‌ای، این نتیجه با نتایج مطمئن به‌دست آمده از چاه مقایسه می‌شود.

۳-۳ بررسی تأثیر پارامترهای مختلف روش

پارامترهای مختلفی در نتیجه ماشین بردار پشتیبان مؤثر هستند و انتخاب آنها مهم است. از جمله این پارامترها می‌توان به نوع کرنل، مقدار سیگما و مقدار ضریب جریمه اشاره کرد. در ابتدا با سعی و خطا مشخص شد که کرنل بر پایه شعاعی (*RBF*)، بهترین نوع کرنل است. در مرحله بعد، پارامتر مربوط به کرنل و ضریب جریمه آن قدر تغییر داده شدند تا بهترین حالت پیدا شود. با تغییر پارامتر کرنل *RBF* و بررسی مقدار خطا متناسب با پارامتر کرنل، مشخص شد پارامترهای کوچک‌تر، خطای کمتری دارند.

نشانگرهایی که با نشانگرهای اصلی همبستگی کمی دارند، باید با یکدیگر نیز همبستگی کمتری داشته باشند؛ در غیر این صورت، از بین چندین نشانگر که همبستگی کمی با امپدانس صوتی دارند و همبستگی زیادی با همدیگر دارند، یکی به‌عنوان ورودی انتخاب می‌شود. از آنجاکه حجم محاسبات زیاد است، فقط از چهار نشانگر استفاده شده است که سه نشانگر دیگر شامل نشانگر قطبیت ظاهری (*apparent polarity*) ردلرزه، نشانگر پوش (*envelope*) و نشانگر جذب هستند. نشانگر قطبیت ظاهری ردلرزه از علامت ضرایب بازتاب متأثر است. از این نشانگر می‌توان برای افزایش پیوستگی پدیده‌ها و ارتقاء تفسیر بازتاب‌های لرزه‌ای استفاده کرد. تغییرات جانبی قطبیت ظاهری، تغییرات جانبی سنگ‌شناسی را نشان می‌دهد. همچنین این نشانگر می‌تواند آشکارسازی لایه‌های نازک را ساده‌تر کند و برای جداسازی نقاط روشن مفید باشد. نشانگر پوش با محتوای انرژی ارتباط دارد. هنگامی که آشکارسازی تغییرات جزئی و اساسی سنگ‌شناسی از روی ردلرزه اصلی مشکل است، نشانگر پوش مفید است. مقادیر زیاد پوش، نشان‌دهنده تغییرات اساسی سنگ‌شناسی و مرزهای توالی‌های رسوبی است که با لکه‌های روشن ارتباط دارد و در اثر وجود گاز حاصل می‌شود. تغییرات جانبی پوش، تغییرات جانبی سنگ‌شناسی را نشان می‌دهد. نشانگر مقاومت صوتی یک خاصیت فیزیکی ذاتی سنگ‌ها است. در تفسیر داده‌های لرزه‌ای، مقاومت صوتی با داده لرزه‌ای اصلی ترکیب

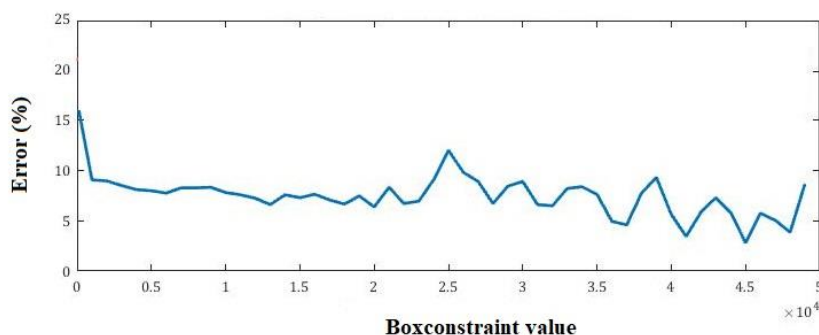


شکل ۷. تغییرات خطا متناسب با تغییرات پارامتر سیگما برای مشخص کردن منطقه هیدروکربنی. نمودار آبی رنگ، تغییرات درصد خطا و نمودار قهوه‌ای رنگ، تعداد نمونه‌های منطقه هیدروکربنی را نشان می‌دهد که درست تشخیص داده شده‌اند. بهینه‌ترین مقدار سیگما $2/7$ است و علاوه بر اینکه درصد خطای کمی دارد، تعداد نمونه‌های منطقه هیدروکربنی درست تشخیص داده شده، بیشترین مقدار ممکن است.

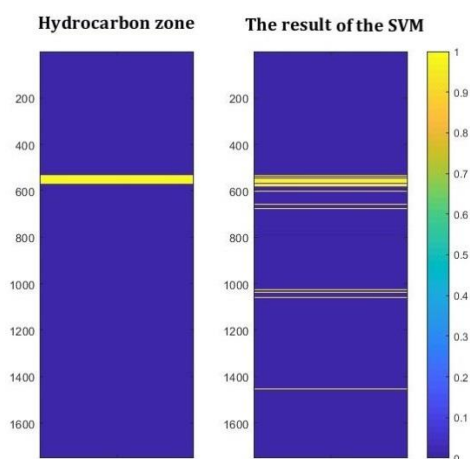
مقدار این ضریب برای یک داده زیادتر باشد، آن داده نقش کمتری در آموزش دارد. برای یافتن بهترین مقدار ضریب جریمه، تغییرات خطا متناسب با تغییرات ضریب جریمه رسم شده است.

همان‌طور که در شکل ۸ مشاهده می‌شود، با افزایش مقدار ضریب جریمه، درصد خطا روندی نزولی طی کرده است ولی باز هم نوسان دارد و با هر میزان تغییرات ضریب جریمه، هنوز مقدار کمی خطا وجود دارد.

با افزایش مقدار پارامتر سیگما در الگوریتم، آموزش با مشکل مواجه می‌شود و مقدار خطا افزایش پیدا می‌کند. همان‌طور که در شکل ۷ دیده می‌شود، وقتی سیگما بیشتر از ۳ است، نوسان خطا زیاد می‌شود، درحالی‌که برای مقادیر کمتر از ۳، مقدار خطا ثبات بیشتری دارد. به ازای سیگما برابر با $0/3$ ، مقدار خطا کمینه است و به نظر می‌رسد این مقدار برای سیگما بهینه است. علاوه بر پارامتر RBF باید مقدار ضریب جریمه را نیز مشخص کرد. هر چه



شکل ۸. تغییرات خطا بر اساس تغییرات ضریب جریمه برای تشخیص منطقه هیدروکربنی. ضریب جریمه بهینه برای این گروه از داده‌ها 45000 است. در این مقدار بهینه، عملکرد تفکیک‌کننده بردار پشتیبان با کمترین خطای ممکن همراه است.



شکل ۹. (راست) بهترین خروجی روش ماشین بردار پشتیبان در جهت تخمین منطقه هیدروکربنی (چپ) محل دقیق قرارگیری منطقه هیدروکربنی در چاه مورد نظر. توانایی زیاد ماشین بردار پشتیبان برای شناسایی محدوده مخزنی بارز است.

استفاده شد. تعیین منطقه هیدروکربنی در دو بخش درون چاهی و لرزه‌ای صورت گرفت که در هر بخش، نتایج تفکیک‌کننده بردار پشتیبان نسبت به گزارش چاه و محدوده مخزنی مقایسه شد. با توجه به تغییر پارامترها، میزان عملکرد این روش در بخش‌های مختلف متفاوت بود. در بخش درون‌چاهی با انتخاب نگاره‌های متناسب با منطقه هیدروکربنی، با توجه به گزارش چاه، عملکرد تفکیک‌کننده بردار پشتیبان در تعیین منطقه هیدروکربنی مناسب بود و بخش‌های هیدروکربنی را به‌خوبی از دیگر قسمت‌های چاه تفکیک کرده است. روش استفاده‌شده، ۹۴ درصد داده‌ها را درست دسته‌بندی کرد.

تعیین منطقه هیدروکربنی با استفاده از داده‌های لرزه‌ای به روش ماشین بردار پشتیبان به عوامل مختلفی از جمله تعداد نشانگرهای لرزه‌ای بستگی دارد. در صورتی که تعداد نشانگرهای استفاده‌شده در مرحله آموزش زیاد یا خیلی کم باشد، فرایند آموزش کامل نمی‌شود و تفکیک به‌خوبی صورت نمی‌گیرد. علاوه بر این، پارامترهای الگوریتم باید متناسب با حجم داده‌ها تنظیم و تابع کرنل مناسب انتخاب شود تا تفکیک‌کننده بهترین عملکرد را داشته باشد. در این تحقیق با توجه به حجم کم داده‌ها در مرحله آموزش، الگوریتم توانسته است تعیین محدوده

در ادامه، پس از اینکه پارامترهای مناسب برای ماشین بردار پشتیبان به‌دست‌آمد، با استفاده از نشانگرهای استخراج‌شده از ردلرزه نزدیک چاه مورد نظر، منطقه هیدروکربنی از دیگر قسمت‌های داده‌های لرزه‌ای جدا شد. با توجه به کمبود داده‌های آموزشی و پیچیدگی مسئله، روش ماشین بردار پشتیبان تا حد قابل قبولی توانست منطقه هیدروکربنی را از بقیه قسمت‌ها جدا کند. سپس برای افزایش دقت خروجی و تشخیص دادن منطقه هیدروکربنی با احتمال بهتر، چندین خروجی مناسب با تغییر پارامترهای روش ماشین بردار پشتیبان با هم ترکیب شدند تا خروجی قابل درک‌تری به‌دست‌بیاید. در شکل ۹ در ستون سمت چپ محل دقیق قرارگیری منطقه هیدروکربنی در داده لرزه‌ای مشخص شده است و در ستون سمت راست، بهترین خروجی ماشین بردار پشتیبان نمایش داده شده است که با دقت قابل قبولی توانسته است محل منطقه هیدروکربنی را تخمین بزند و توانایی زیاد خود را نشان دهد.

۴ نتیجه‌گیری

در این مقاله، از روش ماشین بردار پشتیبان برای تعیین منطقه هیدروکربنی در یکی از میدان‌های نفتی جنوب

- Linari, V., Santiago, M., Pastore, C., Azbel, K., and Poupon, M., 2003, Seismic facies analysis based on 3D multi-attribute volume classification, La Palma field, Maracaibo, Venezuela: *The Leading Edge*, **22**, 32-36.
- Marroquin, I. D., Brault, J. J., and Hart, B. S., 2009, A visual data-mining methodology for seismic facies analysis: *Geophysics*, **74**, 13-23.
- Mathieu, P. G., Rice, G. W., 1999, Multivariate analysis used in the detection of stratigraphic anomalies from seismic data: *Geophysics*, **31**, 401-444.
- Matlock, R. J., McGowen, R. S., and Asimakopoulos, G., 1985, Can seismic stratigraphy problems be solved using automated pattern analysis and recognition: 55th Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, session S17, 7.
- Paparozi E., Grana, D., Mancini, S., and Tarchiani, C., 2011, Seismic driven probabilistic classification of reservoir facies and static reservoir modeling, 13rd EAGE Conference & Exhibition Incorporating SPE EUROPEC Vienna, Austria, May, 80440.
- Saggaf, M. M., Toksoz, M. N., and Marhoon M. I., 2003, Seismic facies classification and identification by competitive neural networks: *Geophysics*, **92**, 4321-4333, 8003.
- Simaan, M. A., 1991, A knowledge-based computer system for segmentation of seismic sections based on texture: 61st Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 289-292.
- Taner, M. T., 2001, Seismic attributes: Recorder, **26**, 48-56.
- Vapnik, V., 1995, *The nature of statistical learning theory*, Springer-Verlag, New York, 314 pp.
- Vapnik, V., 1998, *Statistical Learning Theory*, Wiley, New York, NY, USA.
- West, B., May, S., Eastwood, J. E., and Rossen, C., 2002, Interactive seismic facies classification using textural and neural networks: *The Leading Edge*, **21**, 1042-1049.
- هیدروکربنی را در تمام داده با دقت قابل قبول انجام دهد، اما در محدوده هیدروکربنی واقع در نزدیک چاه، آموزش به خوبی صورت گرفته و نتیجه پذیرفتنی بوده است.
- ### منابع
- فخاری، م.، هاشمی، ح.، ۱۳۹۶، تفکیک دودکش های گازی از رسوبات سطحی حاوی گاز با استفاده از کلاس بندی ماشین بردار پشتیبان با تابع هسته شعاعی، پایان نامه کارشناسی ارشد، دانشگاه تهران.
- Alexis, C., and Tanwi, B., 2008, Integrated geological and geophysical analysis by hierarchical classification combining seismic stratigraphic and AVO attributes: *Petroleum Geoscience*, **14**, 339-354.
- Alpana, B., and Hans B., 2002, Determination of facies from well logs using modular neural networks: *Petroleum Geoscience*, **8**, 217-228.
- Bagheri, M., and Riahi, M. A., 2014, Seismic facies analysis from well logs based on supervised classification scheme with different machine learning techniques: *Arabian Journal of Geosciences*, **8**(9), DOI 4004001/s48441-041-4934-4.
- Bardini, S., Grana, D., and Maffioletti, F., 2010, 3D Geological and Seismic Modelling for Reservoir Characterization: 72nd EAGE Conference & Exhibition incorporating, Barcelona, Spain.
- Dumay, J., and Fournier, F., 1988, Multivariate statistical analyses applied to seismic facies recognition: *Geophysics*, **53**, 1151-1159.
- Farzadi, P., 2006, Seismic facies analysis based on 3D multi-attribute volume classification, Dariyan Formation, SE Persian Gulf: *Journal of Petroleum Geology*, **8398**, 443-411.
- Fournier, F., Dequierez, P. Y., Macrides, G. C., and Rademakers, M., 2002, Quantitative lithostratigraphic interpretation of seismic data for characterization of the Unayzah Formation in central Saudi Arabia: *Geophysics*, **67**, 1372-1381.
- Hagan, D. C., 1982, The applications of principal component analysis to seismic data sets: *Geoexploration*, **20**, 93-111.
- Hossain, Z., and Mukerji, T., 2011, Statistical Rock Physics and Monte Carlo Simulation of Seismic Attributes for Greensand: 73rd EAGE Conference & Exhibition incorporating, Vienna, Austria.

Hydrocarbon zone identification using support vector machine learning method

Mohammadreza Ghaseminezhad¹, Majid Bagheri^{2*} and Majid Nabi-Bidhendi³

¹*M.Sc, Institute of Geophysics, University of Tehran, Tehran, Iran*

²*Assistant Professor Institute of Geophysics, University of Tehran, Tehran, Iran*

³*Professor, Institute of Geophysics, University of Tehran, Tehran, Iran*

(Received: 02 April 2021, Accepted: 31 May 2021)

Summary

One of the basic steps of oil exploration is to define the hydrocarbon zone. Different methods have been used so far for defining such zones. For a specific dataset, finding the most appropriate method leads to more accurate estimates and predictions of analysis besides improving the speed of calculations. Support Vector Machine (SVM), which is one of the methods for analyzing the data, uses kernel functions. It finds a better relationship between data factors and hydrocarbon zone leading to better estimates and classifications.

In this article, hydrocarbon zone detection has been done using seismic and well data.

The purpose of facies analysis is to obtain important petrophysical parameters of the reservoir and to identify heterogeneous boundaries below the ground. The results of the interpretation of petrophysical parameters are the input of the three-dimensional reservoir modeling process and through these parameters, the reservoir parameters are distributed in three-dimensional space. This model is widely used in various sections such as exploration and drilling of new wells, overdraft from a reservoir, determination of suitable areas for overdraft, reduction of drilling risk and risk, determination of reservoir lithology and identification of key well and its extension to other wells in the region. The most important petrophysical parameters are shale volume, porosity, permeability, reservoir fluid saturation and reservoir lithology.

The study of seismic facies has been started since the 90's and due to its importance and application in reservoir description, it has always been considered by many researchers.

To perform the analysis above, first, the hydrocarbon zones were spotted across the Asmari Formation using well logs and well geology reports. Next, the SVM method was used to detect each hydrocarbon zone using well logs. There was an acceptable agreement between the results of SVM method and well geology reports. Second, hydrocarbon zones detection was done using seismic data by SVM. At this stage, seismic attributes were extracted from the seismic trace in the well location. Then, covariance matrix and cross plots of seismic attributes used to identify the most effective attributes to hydrocarbon zones detection. In order to validate the results, the seismic attributes of another trace near the well location were used for hydrocarbon zone detection. SVM results matched hydrocarbon zones with low error.

Keywords: Hydrocarbon zones detection, Support Vector Machine (SVM), well logs, seismic attributes, validation

*Corresponding author:

majidbagheri@ut.ac.ir