



DOI: 10.30479/jmre.2019.10524.1255

تخمین عیار کانسار فسفات اسفوردی با روش رگرسیون بردار پشتیبان

ملیحه عباس زاده^{*}

۱- استادیار، گروه مهندسی معدن، دانشگاه کاشان، کاشان

(دریافت ۱۳۹۸/۰۲/۰۴، پذیرش ۱۳۹۸/۰۸/۲۹)

چکیده

امروزه تخمین متغیر با استفاده از روش‌های مبتنی بر هوش مصنوعی از جمله رویکردهای جدیدی است که فرآیند تصمیم‌گیری موثر را در بسیاری از علوم میسر ساخته است. تخمین عیار نیز از مسایل مهم در ارزیابی ذخایر معدنی در علوم زمین به شمار می‌رود. روش‌های زمین‌آماری از جمله روش‌های متداول تخمین متغیر در علوم زمین محسوب می‌شوند. از آنجایی که این روش‌ها در رابطه با داده‌هایی که تعداد آن‌ها محدود است و ماهیت پراکندگی و غیرخطی بودن دارند تا حدودی دچار مشکل می‌شوند، در این مطالعه از روش رگرسیون بردار پشتیبان به عنوان یکی از روش‌های هوشمند در حیطه الگوریتم‌های یادگیری ماشین برای تخمین عیار در کانسار فسفات اسفوردی استفاده شده است. دقت مدل‌سازی انجام گرفته با این روش بر اساس داده‌های آزمایش در حدود ۸۴ درصد شد که نشان‌دهنده کارایی مدل‌سازی انجام شده است. بر اساس نتایج به دست آمده از مدل‌سازی انجام گرفته به روش رگرسیون بردار پشتیبان، اقدام به تخمین عیار در محدوده مدل بلوکی کانسار فسفات اسفوردی شد. نواحی معرفی شده به عنوان مناطق پریپتانسیل در مدل بلوکی تخمین زده شده به روش رگرسیون بردار پشتیبان می‌تواند در ادامه فرآیند اکتشاف به عنوان محل گمانه‌های تکمیلی مورد برداشت قرار گیرد. همچنین بر اساس نتایج به دست آمده از روش رگرسیون بردار پشتیبان در کانسار فسفات اسفوردی، مدل تناژ-عیار متوسط تهیه شد. به عنوان نمونه بر اساس این مدل به ازای عیار حد ۶ درصد، تناژ ذخیره حدود ۱۵/۳۶ میلیون تن با عیار متوسط ۱۳/۵۹ درصد به دست آمد.

کلمات کلیدی

الگوریتم یادگیری ماشین، رگرسیون بردار پشتیبان، مدل‌سازی، تخمین، فسفات اسفوردی.

۱- مقدمه

گرفتن ماهیت رابطه فضایی غیرخطی بین داده‌ها ندارند [۱]. تحقیقات گسترده در طول سالیان متمادی منجر به توسعه روش‌های نوین قابل استفاده در تخمین مقدار متغیر شده است. از جمله این روش‌ها الگوریتم‌های یادگیری ماشین^۱ است که روش‌های جدیدی برای حل مساله تخمین متغیر پیش روی محققان قرار داده است [۳]. این الگوریتم‌های یادگیری، رابطه‌ای را که به طور ذاتی در بین داده‌ها وجود دارد از روی نمونه‌هایی که در دسترس الگوریتم قرار داده می‌شود، یاد می‌گیرند. جذابیت این تخمین‌گرهای غیرخطی در توانایی کاری آن‌ها به صورت یک جعبه سیاه است. با دادن داده‌های کافی به این الگوریتم‌ها و سپس آموزش آن‌ها، این الگوریتم‌ها می‌توانند رابطه‌ای را که بین نمونه‌های ورودی (به عنوان مثال مختصات نمونه‌ها) و نمونه‌های خروجی (به عنوان مثال عیار کانسنگ در نقاط اندازه‌گیری شده) وجود دارد یاد بگیرند. با این روش، لازم نیست فرضیاتی مانند خطی بودن برای مولفه‌ها، ضرایب و یا روابط در نظر گرفته شود [۱]. الگوریتم‌های یادگیری ماشین به دو دسته اصلی تقسیم می‌شوند که با عنوان روش‌های یادگیری نظارت شده و روش‌های یادگیری نظارت نشده شناخته می‌شوند [۵، ۴]. هر کدام از این دو دسته اصلی روش‌ها و الگوریتم‌های متعددی را شامل می‌شوند که هر یک با توجه به ماهیت مساله مورد بررسی و نیز تعداد و ماهیت داده‌ها، مورد استفاده قرار می‌گیرند. با این وجود از بین روش‌های بسیار متنوع و کارا در حوزه هوش مصنوعی در اغلب موارد از روش‌های شبکه‌های عصبی مصنوعی [۱۵-۶]، استنتاج فازی عصبی [۱۱]، جن‌گل تصادفی [۱۶] و ماشین بردار پشتیبان [۱۷-۱۹، ۳] برای تخمین عیار استفاده شده است.

در مهندسی معدن از شبکه‌های عصبی مصنوعی علاوه بر کاربرد در تخمین عیار کاربردهای متنوع دیگری نیز گزارش شده است که برای نمونه می‌توان به تخمین ابعاد شبکه آتشفشانی [۲۰]، طراحی کارخانه‌های فرآوری [۲۱، ۲۲]، طبقه بندی زمین‌شناسی [۲۳]، طبقه بندی داده‌های دورسنجی [۲۴، ۲۵]، شناسایی مدل‌های شکست در بازکردن معادن زیرزمینی [۲۶، ۲۷] و نظایر آن اشاره کرد. پس از شبکه‌های عصبی مصنوعی، بیشترین کاربرد از روش‌های تخمین هوشمند در تخمین عیار در مورد ماشین بردار پشتیبان ارایه شده است. این روش با وجود اینکه در مقایسه با روش شبکه‌های عصبی مصنوعی، روش نسبتاً جدیدی محسوب می‌شود، به دلیل پشتوانه ریاضی قوی‌تری که دارد مقبولیت گسترده‌ای را در زمان کوتاهی به دست آورده است [۳۱-۲۸].

مساله تخمین متغیر، مساله‌ای آماری است که می‌توان آن را به سادگی به صورت پیش‌بینی مقدار متغیر در نقاط نمونه‌برداری نشده $(Z^*(x))$ بر اساس مجموعه داده‌های معلوم $Z_1, Z_2, Z_3, \dots, Z_n$ حاصل از نمونه‌برداری (معمولاً نمونه‌های برداشت شده از مغزه‌های حفاری) در نقاط $X_1, X_2, X_3, \dots, X_n$ تعریف کرد. این فرآیند بر این فرض استوار است که نمونه‌های مورد استفاده تصادفی و مستقل از یکدیگرند [۱]. روش‌های کلاسیک و خطی متعددی برای تخمین متغیر ارایه شده است که از منظر ریاضی می‌توان تمام آن‌ها را به صورت رابطه ۱ بیان کرد:

$$Z^*(v) = \sum_{i=1}^n \lambda_i Z(x_i) \quad (1)$$

که در آن:

λ_i : وزن نمونه i ام برای بلوک مورد تخمین v است.
تفاوت تخمین‌گرهای خطی در شیوه تعیین اوزان نمونه‌ها است. تا پیش از توسعه روش‌های محاسباتی و ارایه نرم‌افزارهای کامپیوتری، روش‌های بلوکی (مثلی، چندضلعی و نامنظم) و مقاطع موازی (عمودی، افقی و شیبدار) برای تخمین متداول بودند که بیشتر ماهیت تجربی داشتند ولی پس از توسعه نرم‌افزارهای محاسباتی، روش‌هایی همچون روش‌های زمین آماری به روش‌های معمول برای تخمین متغیر تبدیل شدند. به طور کلی تخمین زمین آماری روشی است که با آن می‌توان مقدار یک متغیر را در نقاطی با مختصات معلوم و با استفاده از مقدار همان متغیر در نقطه دیگری با مختصات معلوم به دست آورد [۲]. اما در شرایطی که ویژگی غیرخطی بودن در بین داده‌ها وجود داشته باشد باید از تخمین‌گرهای غیرخطی برای بهبود تخمین و اطمینان از دقیق بودن تخمین انجام گرفته، استفاده شود [۱]. همچنین روش‌های زمین آماری به شدت به نوع توزیع و تعداد داده‌ها حساس است و در شرایطی که تعداد داده‌ها کم باشد و یا وابستگی فضایی بین آن‌ها ضعیف باشد، فرآیند تخمین مشکل و تقریباً غیرممکن است. بنابراین بدیهی است که مقدار خطای ایجاد شده نیز بسیار زیاد خواهد بود [۳]. اگرچه نسخه‌های مختلفی از کریجینگ همچون لاگ کریجینگ و کریجینگ چند شاخصه ارایه شده‌اند که با اعمال یکسری تبدیلات ویژه بر روی داده‌ها، تلاش دارند تا روابط غیرخطی را در تخمین لحاظ کنند اما در برخی از موارد این روش‌ها کارایی بالایی در خصوص در نظر

ساختاری ایران مرکزی و درون سنگ‌هایی با سن اینفراکامبرین که تناوبی از شیل، ماسه‌سنگ، آهک و دولومیت همراه با سنگ‌های آتشفشانی که از بازالت تا ریولیت تغییر می‌کند، قرار دارد. به لحاظ کانی‌شناسی، کانه‌زایی در سه افق متفاوت با نام‌های افق آهنی، افق فسفات‌دار و افق دگرگونی رخ داده است. کانی‌زایی اصلی در ارتباط با کانی‌های مگنتیت و آپاتیت بوده و شواهد صحرایی سه گونه کانی‌زایی آپاتیت به صورت رگه‌ای، انتشاری و توده‌ای را نشان داده است [۴۱]. توده‌های نفوذی در منطقه، با کانی‌زایی گسترده مگنتیت، آپاتیت-مگنتیت و گاه غنی از عناصر نادر خاکی همراه‌اند. خاستگاه بیشتر ذخایر آهن ناحیه بافق را می‌توان به توده‌های نفوذی یاد شده نسبت داد. این توده‌ها هم خود دارای مقدار قابل توجهی از آهن‌اند و هم این که سنگ‌های آتشفشانی سری ریزو (اینفراکامبرین) که سنگ‌های فراگیر توده‌ها را تشکیل می‌دهند دارای زمینه بالایی از آهن‌اند [۴۲]. نقشه زمین‌شناسی منطقه اسفوردی در شکل ۱ نمایش داده شده است.

۳- مواد و روش‌ها

۳-۱- مجموعه داده‌ها

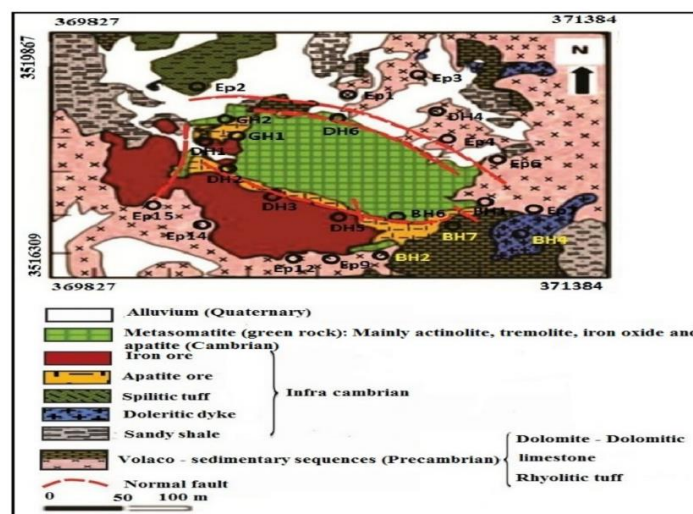
داده‌های مورد استفاده در این مطالعه، اطلاعات به دست آمده از ۵۹ گمانه در محدوده کانسار اسفوردی است که در مجموع شامل ۳۷۸۱ داده و شامل ویژگی‌هایی همچون مختصات جغرافیایی (X,Y,Z) و درصد P_2O_5 است. ویژگی‌های آماری داده‌های مورد استفاده در این مطالعه در جدول ۱ آورده

این روش بر مبنای تئوری یادگیری آماری^۲ و حداقل سازی ریسک^۳ ساختاری استوار است و نخستین بار توسط وپنیک^۴ در دهه ۹۰ میلادی معرفی شد [۳۳،۳۲،۱۷].

تاکنون مطالعات اندکی در رابطه با کاربرد روش ماشین بردار پشتیبان در علوم زمین نسبت به کاربرد این روش در سایر علوم انجام گرفته است که از آن جمله می‌توان به مواردی همچون استفاده از روش ماشین بردار پشتیبان در رابطه با تحلیل داده‌های مکانی [۱]، تهیه نقشه رادیواکتیویته طبیعی [۳۴]، استفاده از روش ماشین بردار پشتیبان برای تخمین غلظت آرسنیک در سنگ بستر مشتق شده از رسوبات رودخانه‌ای [۳۵]، طبقه‌بندی و جداسازی زون‌های دگرسانی مرتبط با کانی‌سازی [۳۸،۳۶] و طبقه‌بندی تصاویر ماهواره‌ای [۳۸-۴۰] اشاره کرد. در زمینه تخمین عیار نیز از این روش در تخمین عیار کانسارهای اسلیت [۱۷]، پلاتین نوع یخچالی [۱۸]، طلا [۳] و آهن [۱۹] استفاده شده است. در این مقاله سعی شده است تا پس از مقایسه کارایی روش رگرسیون بردار پشتیبان با روش زمین آماری کریجینگ معمولی، به کارگیری روش رگرسیون بردار پشتیبان برای تخمین عیار فسفات در محدوده کانسار اسفوردی مورد بررسی قرار گیرد.

۲- منطقه مورد مطالعه

کانسار مگنتیت-آپاتیت اسفوردی در ۳۵ کیلومتری شمال شرقی بافق در مختصات ۳۸°، ۵۵° طول جغرافیایی و ۴۷'، ۳۱° عرض جغرافیایی واقع شده است. محدوده مورد مطالعه در زون



شکل ۱: نقشه زمین‌شناسی ساده شده منطقه اسفوردی در مقیاس ۱:۵۰۰۰۰ [۴۱]

شده است. برای استفاده از روش رگرسیون بردار پشتیبان، مجموعه داده‌ها به گونه‌ای به دو زیرمجموعه آموزش (حدود ۸۰ درصد داده‌ها) و آزمایش (حدود ۲۰ درصد داده‌ها) تقسیم شد که تا حد امکان ویژگی‌های آماری این دو زیرمجموعه مشابه باشند.

بر این اساس تعداد ۳۰۲۵ داده در زیر مجموعه آموزش و تعداد ۷۵۶ داده در زیر مجموعه آزمایش قرار گرفتند. مشخصات آماری داده‌های آموزش و آزمایش استفاده شده در مدل رگرسیون بردار پشتیبان به صورت تفکیک شده در جدول ۲ ارائه شده است.

۲-۳- روش رگرسیون بردار پشتیبان

ماشین‌های بردار پشتیبان به مجموعه‌ای از مدل‌های خطی تعمیم‌یافته تعلق دارند که در آن‌ها بر اساس مقدار مربوط به ترکیب خطی ویژگی‌ها، تصمیماتی در مورد طبقه بندی و رگرسیون اتخاذ می‌شود. ماشین‌های بردار پشتیبان مشابه با شبکه‌های عصبی مصنوعی قادر به تقریب با درجه دقت دلخواه برای توابع چند متغیره‌اند و می‌توان از آن‌ها برای مدل کردن انواع فرآیندهای غیرخطی و پیچیده استفاده کرد [۴۳].

رگرسیون بردار پشتیبان نوعی از ماشین‌های بردار پشتیبان است و برای تخمین در مسایل مختلف کاربرد دارد. این روش شامل الگوریتم‌های آموزشی بوده و بر مبنای روش ماشین‌های

ماشین‌های بردار پشتیبان به مجموعه‌ای از مدل‌های خطی تعمیم‌یافته تعلق دارند که در آن‌ها بر اساس مقدار مربوط به ترکیب خطی ویژگی‌ها، تصمیماتی در مورد طبقه بندی و رگرسیون اتخاذ می‌شود. ماشین‌های بردار پشتیبان مشابه با شبکه‌های عصبی مصنوعی قادر به تقریب با درجه دقت دلخواه برای توابع چند متغیره‌اند و می‌توان از آن‌ها برای مدل کردن انواع فرآیندهای غیرخطی و پیچیده استفاده کرد [۴۳].

رگرسیون بردار پشتیبان نوعی از ماشین‌های بردار پشتیبان است و برای تخمین در مسایل مختلف کاربرد دارد. این روش شامل الگوریتم‌های آموزشی بوده و بر مبنای روش ماشین‌های

$$f(x) = w^T \cdot \varphi(x) + b \quad (2)$$

که در آن:

w : بردار وزن

b : مقدار بایاس است.

این تابع باید نسبت به تغییرات کوچک در مدل حساس نباشد و مدل رگرسیونی را به خوبی توسعه دهد. تابع مورد نظر با عنوان تابع زیان E - غیر حساس^۵ شناخته می‌شود. بنابراین حل تابع f می‌تواند به شکل مساله درجه دوم بیان شود (رابطه ۳):

جدول ۱: ویژگی‌های آماری داده‌های مورد استفاده در مطالعه

پارامترهای آماری	X	Y	Z	P ₂ O ₅
میانگین	۱۰۸۴۷٫۳۳	۲۲۰۳۳٫۷۹	۱۶۸۴٫۶۳	۶٫۵۵
مینیمم	۱۰۱۵۰٫۸۵	۲۱۸۱۵٫۵	۱۵۸۲٫۸	۰
ماکزیمم	۱۱۰۷۵٫۶	۲۲۲۵۳٫۷	۱۷۵۷٫۱۵	۴۰٫۰۴
انحراف معیار	۱۶۳٫۴۳	۱۱۰٫۲۸	۲۷٫۹۸	۸٫۶۱۵

جدول ۲: ویژگی‌های آماری داده‌های آموزش و آزمایش استفاده شده در مدلسازی

پارامترهای آماری	X		Y		Z		P ₂ O ₅	
	آموزش	آزمایش	آموزش	آزمایش	آموزش	آزمایش	آموزش	آزمایش
میانگین	۱۰۸۸۶٫۱۲	۱۰۸۹۵٫۸۷	۲۲۰۳۱٫۰۲	۲۲۰۳۶٫۷۶	۱۶۸۵٫۴۳	۱۶۸۳٫۵۹۴	۶٫۷۹۰۵۳۴	۷٫۰۸۱۹۴۵
مینیمم	۱۰۱۵۰٫۵۸	۱۰۱۵۳٫۱۷	۲۱۸۱۵٫۵	۲۱۸۱۵٫۵	۱۵۸۸٫۱۵	۱۵۸۲٫۸	۰	۰
ماکزیمم	۱۱۰۷۵٫۶	۱۱۰۷۵٫۶	۲۲۲۵۳٫۷	۲۲۲۵۳٫۷	۱۷۵۷٫۱۵	۱۷۵۴٫۵۷۵	۴۰٫۰۴	۳۶٫۷۹
انحراف معیار	۱۶۲٫۰۲۲	۱۶۸٫۸۳	۱۱۰٫۲۳	۱۱۰٫۲۸۵	۲۷٫۸۴	۲۸٫۵۰	۸٫۶۸	۸٫۳۲

مدلسازی و تخمین ارایه شده است [۴۶-۴۸]. اگر عیار به صورت یک تابع تصادفی پایا مرتبه دوم در نظر گرفته شود $\{Z(x), x \in D \subset R^d; d=3\}$ که x یک موقعیت در فضای D است. آنگاه متوسط عیار یک بلوک v را می‌توان با استفاده از روش کریجینگ معمولی بر اساس عیار نمونه‌های واقع در پیرامون آن $Z(x_i), i=1, \dots, n$ بر اساس رابطه ۱ که پیشتر در بخش مقدمه به آن اشاره شد، تخمین زد [۴۹، ۵۰]. در روش کریجینگ اوزان نمونه‌ها λ_i از حل دستگاه معادلات خطی زیر قابل محاسبه است (رابطه ۶):

$$\sum_{j=1}^n \lambda_j \gamma(x_i, x_j) + \mu = \bar{\gamma}(v, x_i) \quad (۶)$$

$$\sum_{j=1}^n \lambda_j = 1, \quad i = 1, \dots, n$$

که در آن:

$\gamma(x_i, x_j)$: واریوگرام بین نمونه واقع در موقعیت x_j و x_i
 $\bar{\gamma}(v, x_i)$: میانگین واریوگرام نقاط واقع درون بلوک v موقعیت x_i و μ ضریب لاگرانژ است.

همچنین واریانس کریجینگ حاصل را نیز می‌توان براساس رابطه ۷ محاسبه کرد:

$$\sigma_k^2(v) = \sum_{j=1}^n \lambda_j \bar{\gamma}(x_j, v) + \mu - \bar{\gamma}(v, v) \quad (۷)$$

لازمه اجرای روش کریجینگ محاسبه واریوگرام تجربی و برازش یک مدل مناسب به آن است. در جدول ۴ متداول‌ترین مدل‌های واریوگرام ارایه شده‌اند. یکی از پیش شرط‌های استفاده از روش کریجینگ معمولی، تبعیت متغیر هدف از تابع توزیع نرمال است. در صورتی که این شرط صادق نباشد، می‌توان از ابزارهای تبدیل به تابع توزیع نرمال [۵۱] و یا روش مدل زمین آماری خطی تعمیم یافته [۵۲] استفاده کرد.

جدول ۴: متداول‌ترین مدل‌های واریوگرام [۵۰]

مدل واریوگرام	تابع واریوگرام
خطی	$\gamma(h) = \alpha h^\lambda \quad 0 < \lambda < 2$
کروی	$\gamma(h) = c_0 + c_1 \left(\frac{3h}{2a} - \frac{1}{2} \left(\frac{h}{a} \right)^3 \right)$
نمایی	$\gamma(h) = c_0 + c_1 \left(1 - e^{-h/\alpha} \right)$
گوسی	$\gamma(h) = c_0 + c_1 \left(1 - e^{-(h/\alpha)^2} \right)$
h : فاصله، α : دامنه، c : سقف، c_0 اثر قطعه ای و α دامنه موثر	

$$\begin{cases} \text{Min } \frac{1}{2} w^T w + C \sum_{i=1}^l (\xi_i - \xi_i^*) \\ \text{s. t. } y_i - w \cdot \varphi(x_i) - b \leq \varepsilon + \xi_i \\ w \cdot \varphi(x_i) + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i \geq 0 \\ \xi_i^* \geq 0 \end{cases} \quad (۳)$$

که در آن:

C : پارامتر جریمه

ξ_i : متغیر کمکی است.

با معرفی ضرایب لاگرانژ و استفاده از تابع کرنل به جای ضرب داخلی عملگرها در فضای ویژگی، مساله بهینه‌سازی فوق را می‌توان به شکل مساله دوگان زیر نوشت (رابطه ۴):

$$\begin{cases} \text{Max } \left[-\frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l (\alpha_i - \alpha_j)(\alpha_j - \alpha_i) K(x_i, x_j) - \varepsilon \sum_{i=1}^l (\alpha_i - \alpha_i^*) + \sum_{i=1}^l y_i (\alpha_i - \alpha_i^*) \right] \\ \text{s. t. } \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0, \quad i = 1, \dots, l \\ 0 \leq \alpha_i \leq C \\ 0 \leq \alpha_i^* \leq C \end{cases} \quad (۴)$$

با حل این مساله دوگان حل مساله رابطه ۲ به شکل رابطه ۵ خواهد بود:

$$f(x) = \sum_{x_i \in SV} (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad (۵)$$

که در آن:

α_i و α_i^* : ضرایب لاگرانژ

$K(x_i, x_j)$: تابع کرنل است.

متداول‌ترین انواع تابع کرنل در جدول ۳ ارایه شده است.

جدول ۳: متداول‌ترین توابع کرنل به کار گرفته شده در مسایل تفکیک‌ناپذیر خطی [۴۵]

نوع خطی ^۶	تابع کرنل
چند جمله ای ^۷	$K(x_i, x_j) = (\gamma x_i x_j + r)^d, \gamma > 0$
پایه شعاعی ^۸	$K(x_i, x_j) = \exp\{-\gamma \ x_i - x_j\ ^2\}, \gamma > 0$
سیگموئید ^۹	$K(x_i, x_j) = \tanh(\gamma x_i x_j + r), \gamma > 0$

۳-۳- روش کریجینگ

امروزه کریجینگ به عنوان متداول‌ترین روش تخمین در مدلسازی و ارزیابی ذخایر معدنی شناخته می‌شود. به طور کلی این روش با هدف در نظر گرفتن همبستگی فضایی در فرآیند

۴- نتایج و بحث

۴-۱- تخمین عیار فسفات به روش رگرسیون بردار پشتیبان

در این مقاله برای تخمین عیار در کانسار فسفات اسفوردی به روش رگرسیون بردار پشتیبان، تابع پایه شعاعی یا RBF به عنوان تابع کرنل در نظر گرفته شده است. کرنل RBF نمونه‌ها را به صورت غیرخطی به فضایی با ابعاد بیشتر می‌نگارد. این کرنل در مقایسه با سایر انواع کرنل دارای مشکلات محاسباتی کمتری است و قادر است مجموعه داده‌هایی با ابعاد بالا را به نحو مطلوبی مورد تجزیه و تحلیل قرار دهد. همچنین این تابع کرنل در مقایسه با توابعی همچون توابع چند جمله‌ای تعداد پارامتر کمتری دارد که باعث کاهش پیچیدگی مدل می‌شود [۵۴، ۵۳]. کارایی یک مدل مبتنی بر روش ماشین بردار پشتیبان تا حد زیادی وابسته به پارامترهای انتخاب شده برای مدل است. برای دستیابی به مدلی که قدرت تعمیم بالایی داشته باشد، لازم است پارامترهای مدل با دقت تعیین شوند. نحوه انتخاب پارامترهای بهینه مدل می‌تواند کیفیت عملکرد مدل را تحت تاثیر قرار دهد و در نتیجه مدلی با پارامترهای نامناسب ممکن است نتایج نامطلوبی را حاصل کند [۵۳، ۳۰]. برای تعیین پارامترهای بهینه مدل (شامل C، γ و ϵ) در دو مرحله از روش جستجوی شبکه‌ای^{۱۱} بر مبنای اعتبارسنجی متقابل^{۱۱} ۱۰ تایی استفاده شد. روش جستجوی شبکه‌ای یک روش جستجوی فراگیر^{۱۲} است. در این روش شبکه‌ای یکنواخت در فضای پارامتر مورد جستجو تعریف می‌شود، سپس تمام نقاط در شبکه برای پیدا کردن یک نقطه بهینه کلی مورد ارزیابی قرار می‌گیرد. در نهایت، جستجوی شبکه‌ای نقطه بهینه کلی تمام نقاط در شبکه پارامتر مورد نظر را پیدا خواهد کرد. اساس روش جستجوی شبکه‌ای به این صورت است که ابتدا شبکه‌ای درشت^{۱۳} در فضای پارامتر تشکیل داده می‌شود، با نزدیک شدن به نقطه بهینه در مرحله بعد شبکه ریزتری^{۱۴} تعریف شده و در نهایت به نقطه بهینه کلی در فضای پارامتر مورد جستجو نزدیک می‌شود [۵۵]. در جدول ۵ بازه جستجو برای هر یک از پارامترهای مدل آورده شده است. ایده اصلی این روش یافتن پارامترهای بهینه مدل به نحوی است که خطای مدل تا حد امکان کاهش یابد. برای این منظور همان‌طور که اشاره شد روش اعتبارسنجی متقابل به همراه روش جستجوی شبکه‌ای به کار گرفته می‌شود.

در اعتبارسنجی k -تایی ابتدا مجموعه آموزش به k زیرمجموعه با اندازه یکسان تقسیم می‌شود. به طور متوالی هر

بار یکی از زیرمجموعه‌ها به وسیله مدل که با توجه به $k-1$ زیرمجموعه دیگر آموزش دیده است، آزمایش می‌شود. بنابراین هر یک از نمونه‌های موجود در مجموعه آموزش یک مرتبه مورد تخمین قرار می‌گیرد. در نتیجه دقت اعتبارسنجی متقابل درصدی از داده‌ها خواهد بود که به درستی پیش‌بینی شده‌اند. در واقع برای پارامترهای مورد نظر میزان RMSE در هر یک از زیرمجموعه‌های آزمایش با استفاده از روش اعتبارسنجی متقابل به دست می‌آید. میانگین RMSE مربوط به k زیرمجموعه آزمایش به عنوان معیار ارزیابی عملکرد مدل در نهایت ارایه می‌شود. به این ترتیب پارامترهایی که منجر به کمترین میزان RMSE شوند به عنوان پارامترهای بهینه مدل معرفی می‌شوند. یکی دیگر از قابلیت‌های روش اعتبارسنجی متقابل، جلوگیری از ایجاد مشکل بیش‌برازشی در مدل است [۵۵، ۵۴].

در جدول ۶ مقادیر بهینه پارامترهای C، γ و ϵ به دست آمده بر اساس روش جستجوی شبکه‌ای همراه با اعتبارسنجی ۱۰ تایی ارایه شده است. پس از آن که پارامترهای بهینه مدل که به ازای آن‌ها خطای به دست آمده از روش اعتبارسنجی متقابل کمترین مقدار خود را دارد، انتخاب شد، تمام داده‌های آموزش (۸۰ درصد کل داده‌ها) بر اساس مدل بهینه، آموزش داده شد. فرآیند آموزش مدل با استفاده از تابع LIBSVM [۵۶] و در نرم‌افزار Weka انجام گرفت. پس از مرحله آموزش مدل بر اساس پارامترهای بهینه، عملکرد و کارایی مدل رگرسیون بردار پشتیبان بر اساس داده‌های آزمایش (۲۰ درصد از کل داده‌ها) مورد ارزیابی قرار گرفت.

جدول ۵: بازه جستجو برای هر یک از پارامترهای مدل در تخمین عیار در کانسار فسفات اسفوردی [۵۵]

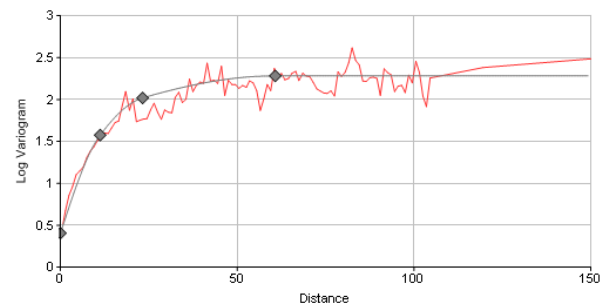
پارامتر مدل رگرسیون بردار پشتیبان	بازه جستجو
C	{ 2^{-15} , , 2^5 }
γ	{ 2^{-15} , , 2^3 }
ϵ	{ 2^{-8} , , 2^{-1} }

جدول ۶: مقادیر بهینه پارامترهای C، γ و ϵ به دست آمده با استفاده از روش جستجوی شبکه‌ای بر مبنای اعتبارسنجی ۱۰ تایی

پارامترهای بهینه مدل	مقدار بهینه پارامتر
C	۱۶
γ	3.629×10^{-5}
ϵ	۰/۱۴۹

۴-۲- تخمین عیار فسفات به روش کریجینگ

باتوجه به اینکه روش کریجینگ به عنوان متداول ترین روش تخمین عیار شناخته می شود، به عنوان روش مبنای مقایسه در این تحقیق انتخاب شده است. یکی از پیش شرطهای تعیین کننده در این روش، تبعیت توزیع متغیر مورد نظر از تابع توزیع نرمال است. بررسی انجام گرفته بر اساس شکل تابع توزیع و آزمون کولموگروف-اسمیرنوف، نشان داد که داده ها از تابع توزیع نرمال تبعیت نمی کنند و از تابع توزیع لاگ نرمال تبعیت می کنند. بر این اساس، لازم است تا از روش لاگ کریجینگ برای تخمین متغیر استفاده شود. در گام بعد، واریوگرام تجربی لگاریتمی متغیر عیار فسفات محاسبه و به آن مدل مناسب برازش یافت. برای انتخاب بهترین مدل ممکن، از روش اعتبارسنجی متقابل استفاده شد. همانطور که در شکل ۲ نمایش داده شده است، یک مدل مرکب سه ساختار برای برازش به واریوگرام تجربی انتخاب شده است. در جدول ۷ پارامترهای مدل ارائه شده است. در گام بعد اقدام به ارزیابی کارایی روش کریجینگ بر روی داده های آموزش و آزمایش شد. نتایج حاصل از به کارگیری مدل رگرسیون بردار پشتیبان و همچنین روش کریجینگ برای داده های آموزش و آزمایش در جدول ۸ آورده شده است. همان طور که در این جدول مشاهده می شود، مقادیر ضریب همبستگی و خطای ریشه میانگین مربعات به دست آمده به عنوان شاخص های ارزیابی عملکرد مدلسازی انجام گرفته برای مجموعه های آموزش و آزمایش در مدل رگرسیون بردار پشتیبان در مقایسه با روش کریجینگ وضعیت مطلوب تری دارد که کارایی مدل رگرسیون بردار پشتیبان استفاده شده در مقایسه با روش کریجینگ در تخمین عیار فسفات در محدوده مورد مطالعه را نشان می دهد.



شکل ۲: واریوگرام تجربی لگاریتمی عیار فسفات و مدل برازش یافته به آن

جدول ۷: پارامترهای مدل واریوگرام برازش یافته

ساختار	نوع مدل	سقف	دامنه
	اثر قطعه ای	۰/۴۰۶	-
اول	کروی	۰/۴۳۱	۱۱
دوم	کروی	۰/۸۶۳	۲۳/۲
سوم	کروی	۰/۵۸۲	۶۰/۴

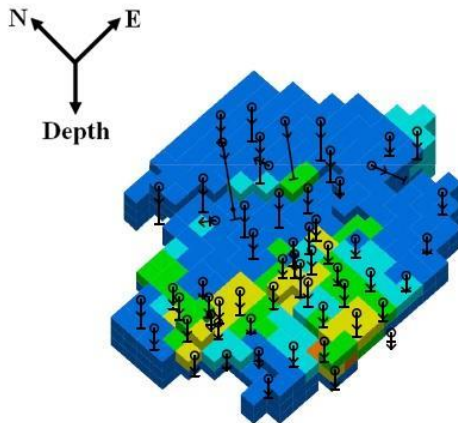
جدول ۸: نتایج حاصل از به کارگیری روش های رگرسیون بردار پشتیبان و کریجینگ برای داده های آموزش و آزمایش

	رگرسیون بردار پشتیبان		کریجینگ	
	آزمایش	آموزش	آزمایش	آموزش
ضریب همبستگی	۰/۸۴۶۶	۰/۸۵۲۶	۰/۸۲۰	۰/۸۳۰
خطای ریشه میانگین مربعات	۴/۴۴۷۵	۴/۵۶۶۶	۵/۳۸۶۸	۵/۳۳۴۶
تعداد داده ها	۷۵۶	۳۰۲۵	۷۵۶	۳۰۲۵

۴-۳- تخمین مدل بلوکی و تهیه مدل تناژ-عیار متوسط بر اساس مدل رگرسیون بردار پشتیبان

مقایسه اعتبارسنجی انجام گرفته برای روش های کریجینگ و مدل رگرسیون بردار پشتیبان در جدول ۸، نشانگر عملکرد بهتر روش رگرسیون بردار پشتیبان است. از این رو در گام بعد، از این روش در تخمین عیار فسفات در مدل بلوکی متناظر با محدوده تخمینی کانسار فسفات اسفوردی استفاده شد. در این مرحله، ابتدا مدل زمین شناسی کانسار بر اساس ابعاد بلوک ۳۰×۳۰×۳۰ متر به مدل بلوکی تبدیل و سپس با استفاده از مدل رگرسیون بردار پشتیبان آموزش یافته، عیار تخمینی هر یک از بلوک ها تخمین زده شد. در شکل ۳ مدل بلوکی عیار حاصله به روش رگرسیون بردار پشتیبان در قالب یک مدل سه بعدی نشان داده شده است. همچنین مقطع افقی تخمین زده شده متناظر با تراز ۱۶۶۰ متری به عنوان یک مقطع نمونه در شکل ۴ ارائه شده است.

از آنجا که در ارزیابی و تخمین یک کانسار معدنی لازم است، ارتباط بین تناژ و عیار متوسط ذخیره برای عیار حد های مختلف مشخص شود، در مرحله بعد مدل تناژ-عیار متوسط-تهیه شد. بر اساس مقادیر عیار حد ارائه شده در جدول ۹ منحنی تغییرات تناژ و عیار متوسط نسبت به عیار حد بر اساس

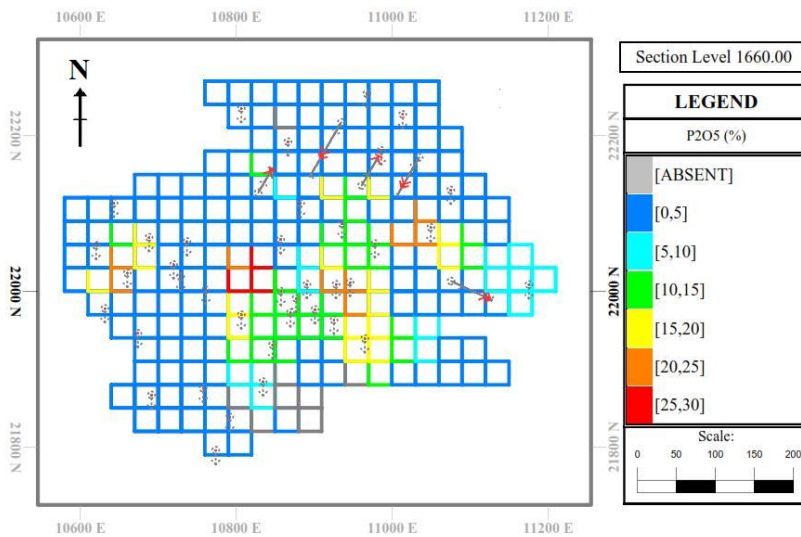


شکل ۳: مدل بلوکی تخمین زده شده به روش رگرسیون بردار پشتیبان در محدوده کانسار فسفات اسفوردی

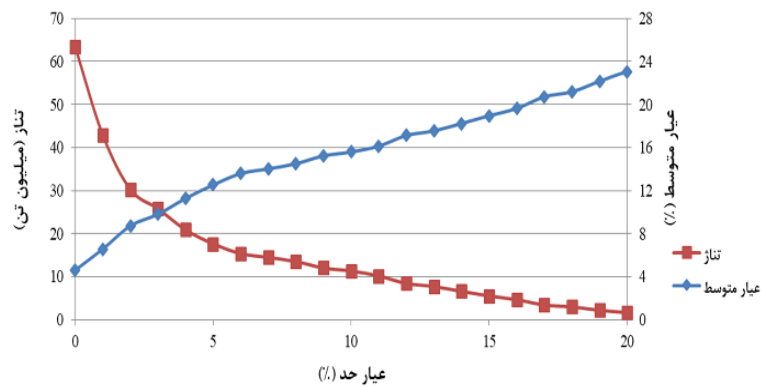
نتایج به دست آمده از روش رگرسیون بردار پشتیبان در کانسار فسفات اسفوردی در شکل ۵ آورده شده است. منحنی عیار-تناژ تهیه شده با در نظر گرفتن وزن مخصوص ۳٫۲۵ تن بر متر مکعب کانسار تهیه شده است.

جدول ۹: مقادیر تناژ و عیار متوسط در عیار حدهای مختلف در کانسار فسفات اسفوردی

عیار حد (درصد)	تناژ (میلیون تن)	عیار متوسط (درصد)
۴	۲۰٫۸۸	۱۱٫۲۷
۶	۱۵٫۳۶	۱۳٫۵۹
۸	۱۳٫۵۱	۱۴٫۵
۱۰	۱۱٫۳۲	۱۵٫۵۸



شکل ۴: مقطع افقی تخمین زده شده به روش رگرسیون بردار پشتیبان، متناظر با تراز ۱۶۶۰ متری



شکل ۵: تغییرات تناژ و عیار میانگین نسبت به عیار حد در کانسار فسفات اسفوردی بر اساس نتایج روش رگرسیون بردار پشتیبان

۵- نتیجه گیری

امروزه تخمین متغیر از جمله رویکردهای جدیدی است که فرآیند تصمیم‌گیری موثر را در بسیاری از علوم میسر ساخته است. تخمین عیار از مسایل مهم در ارزیابی ذخایر معدنی در علوم زمین به شمار می‌رود. برای تخمین این متغیر کلیدی روش‌های متفاوتی وجود دارد. متداول‌ترین این روش‌ها تکنیک‌های زمین آماری است. به طور کلی تخمین زمین آماری روشی است که با آن می‌توان مقدار یک متغیر را در نقاطی با مختصات معلوم و با استفاده از مقدار همان متغیر در نقطه دیگری با مختصات معلوم به دست آورد اما در شرایطی که ویژگی غیرخطی بودن در بین داده‌ها وجود داشته باشد، باید از تخمین‌گرهای غیرخطی برای بهبود تخمین و اطمینان از دقیق بودن تخمین انجام گرفته، استفاده شود. تحقیقات گسترده در طول سالیان متمادی منجر به توسعه روش‌های نوین قابل استفاده در تخمین مقدار متغیر در کانسار شده است. از جمله این روش‌ها الگوریتم‌های یادگیری ماشین است که روش‌های جدیدی را برای حل مشکل تخمین متغیر پیش روی محققان قرار داده است. در این مطالعه از روش رگرسیون بردار پشتیبان که از جمله کارآمدترین روش‌های الگوریتم یادگیری ماشین است برای تخمین عیار در محدوده کانسار فسفات اسفوردی استفاده شده است. بر این اساس ابتدا با استفاده از روش جستجوی شبکه‌ای مقادیر بهینه پارامترهای مدل به دست آمده و سپس با استفاده از تابع LIBSVM مدل‌سازی مورد نظر بر اساس پارامترهای بهینه به دست آمده انجام شد. ارزیابی مدل‌سازی انجام گرفته بر اساس داده‌های آزمایش و مقایسه با نتایج اعتبارسنجی روش کریجینگ به عنوان متداول‌ترین روش زمین آماری در تخمین و ارزیابی ذخایر، نشان می‌دهد که مدل ایجاد شده توانایی مطلوبی در تخمین عیار به عنوان پارامتری کلیدی در ارزیابی ذخیره کانسار خواهد داشت. در نتیجه از این مدل در تخمین عیار بر اساس داده‌های به دست آمده از مدل بلوکی کانسار فسفات اسفوردی استفاده شد. نتایج حاصل از این مدل‌سازی در محدوده مدل بلوکی کانسار نشان می‌دهد که این روش کارایی مطلوبی داشته است و در ادامه، فرآیند تحلیل اقتصادی و طراحی معدن می‌تواند بر اساس مدل عیار تخمینی به دست آمده انجام پذیرد. همچنین بر اساس ارزیابی ذخیره انجام گرفته بر اساس نتایج روش رگرسیون بردار پشتیبان میزان ذخیره کانسار و عیار متوسط کانسار نسبت به عیار حدهای مختلف تخمین زده شد.

۶- مراجع

- [1] Dutta, S. (2006). "Predictive performance of machine learning algorithms for ore reserve estimation in sparse and imprecise data". PhD thesis, University of Alaska Fairbanks, pp. 189.
- [2] Hassani pak, A. A., and Sharafodin, M. (2001). "Exploration Data Analysis". Tehran University Press, pp. 991.
- [3] Dutta, S., Bandopadhyay, S., Ganguli, R., and Misra, D. (2010). "Machine Learning Algorithms and Their Application to Ore Reserve Estimation of Sparse and Imprecise Data". Journal of Intelligent Learning Systems & Applications, 2: 86-96.
- [4] Alpaydin, E. (2010). "Introduction to Machine Learning". Second Edition, Cambridge, Mass.: MIT Press, pp. 584.
- [5] Nilsson, N. J. (1998). "Introduction to Machine Learning". An Early Draft of a Proposed Textbook, Department of Computer Science, Stanford University, Stanford, CA 94305, pp. 188.
- [6] Samanta, B., Ganguli, R., and Bandopadhyay, S. (2005). "Comparing the predictive performance of neural networks with ordinary kriging in a bauxite deposit". Mining Technology, 114: 129-139.
- [7] Chatterjee, S., Bhattacharjee, A., Samanta, B., and Pal, S. K. (2006). "Ore Grade Estimation of a Limestone Deposit in India Using an Artificial Neural Network". Applied GIS, 2: 1-20.
- [8] Chatterjee, S., Bandopadhyay, S., and Machuca, D. (2010). "Ore Grade Prediction Using a Genetic Algorithm and Clustering Based Ensemble Neural Network Model". Math Geosci, 42: 309-326.
- [9] Guo, W. W. (2010). "A novel application of neural networks for instant iron-ore grade estimation". Expert Systems with Applications, 37(12): 8729-8735.
- [10] Li, X-l., Xie, Y-l., Guo, Q-j., and Li, L-h. (2010). "Adaptive ore grade estimation method for the mineral deposit evaluation". Mathematical and Computer Modelling, 52(11-12): 1947-1956.
- [11] Tahmasebi, P., and Hezarkhani, A. (2010). "Application of Adaptive Neuro-Fuzzy Inference System for Grade Estimation; Case Study, Sarcheshmeh Porphyry Copper Deposit, Kerman, Iran". Australian Journal of Basic and Applied Sciences, 4(3): 408-420.
- [12] Tahmasebi, P., and Hezarkhani, A. (2012). "A hybrid neural networks-fuzzy logic-genetic algorithm for grade estimation". Computers & Geosciences, 42: 18-27.

- [24] Miller, D. M., Kaminsky, E. J., and Rana, S. (1995). "Neural Network Classification of Remote Sensing Data". *Computers & Geosciences*, 21: 377-386.
- [25] Wang, Y. G., and Li, H. P. (2010). "Remote sensing image classification based on artificial neural network: A case study of Honghe Wetlands National Nature Reserve". *International Conference on Computer, Mechatronics, Control and Electronic Engineering (CMCE)*, 17-20.
- [26] Lee, C., and Sterling, R. (1992). "Identifying probable failure modes for underground openings using a neural network". *International Journal of Rock Mechanics and Mining Science & Geomechanics Abstracts*, 29(1): 49-67.
- [27] Shahin, M. A., Jaksa, M. B., and Maier, H. R. (2008). "State of the art of artificial neural networks in geotechnical engineering". *Electronic Journal of Geotechnical Engineering*, 8: 1-26.
- [28] Kecman, V. (2000). "Learnig and Soft Computing: Support Vector Machines, Neural Network and Fuzzy Logic Models". MIT Publishers, pp. 576.
- [29] Kecman, V. (2004). "Support Vector Machines Basics". The University of Auckland, School of Engineering, Report 616, pp. 58.
- [30] Smola, A. J., and Scholkopf, B. (1998). "A Tutorial on Support Vector Regression". *NeuroCOLT Technical Report NC-TR-98-030*, Royal Holloway College, University of London, UK.
- [31] Smola, A. J., and Scholkopf, B. (2004). "A Tutorial on Support Vector Regression". *Statistics and Computing*, 14: 199-222.
- [32] Sugumaran, V., and Ramachandran, K. I. (2011). "Effect of number of features on classification of roller bearing faults using SVM and PSVM". *Expert Systems with Applications*, 38: 4088-4096.
- [33] Zhen-yuan, J., Jian-wei, M., Fu-ji, W., and Wei, L. (2011). "Hybrid of simulated annealing and SVM for hydraulic valve characteristics prediction". *Expert Systems with Applications*, 38: 8030-8036.
- [34] Pozdnoukhov, A. (2005). "Support Vector Regression for automated Robust Spatial Mapping of Neural Radioactivity". *Journal of Applied GIS*, 1(2): 1-10.
- [35] Twarakavi, N. C., Misra, D., and Bandopadhyay, S. (2006). "Prediction of Arsenic in Bedrock Derived Stream Sediments at a Gold Mine Site under Conditions of Sparse Data". *Natural Resources Research*, 15: 15-26.
- [36] Abbaszadeh, M., Hezarkhani, A., and Soltani- [13] Mahmoudabadi, H., Izadi, M., and Menhaj, M. B. (2009). "A hybrid method for grade estimation using genetic algorithm and neural networks". *Computational Geosciences*, 13: 91-101.
- [14] Nezamolhosseini, S. A., Mojtahedzadeh, S. H., and Gholamnejad, J. (2017). "The Application of Artificial Neural Networks to Ore Reserve Estimation at Choghart Iron Ore Deposit". *Analytical and Numerical Methods in Mining Engineering*, 6(Special Issue): 73-83.
- [۱۵] صیادی، ا. ر.، منجزی، م.، شهرآبادی، ح.؛ ۱۳۸۸؛ "برآورد ذخیره کانسار فسفات اسفوردی با استفاده از روش های زمین آماری و شبکه عصبی مصنوعی". فصلنامه علوم زمین، دوره هجدهم، شماره ۶۹، ص ۱۰۹-۱۰۲.
- [16] Jafarsteh, B., Fathianpour, N., and Suárez, A. (2018). "Comparison of machine learning methods for copper ore grade estimation". *Computational Geosciences*, 22: 1371-1388.
- [17] Matias, J. M., Vaamonde, A., Taboda, J., and González-Manteiga, W. (2004). "Support vector machines and gradient boosting for graphical estimation of a slate deposit". *Stochastic Environmental Research and Risk Assessment*, 18: 309-323.
- [18] Tenorio, V. O., Bandopadhyay, S., Misra, D., Naidu, S., and Kelley, J. (2015). "Support Vector Machines Applied for Resource Estimation of Underwater Glacier-Type Platinum Deposits". *Society for Mining, Metallurgy, and Exploration*, 889-902.
- [19] Maleki, S., Ramazia, H. R., and Moradi, S. (2014). "Estimation of Iron concentration by using a support vector machine and an artificial neural network - the case study of the Choghart deposit southeast of Yazd, Yazd, Iran". *Geopersia*, 4: 201-212.
- [20] Bakhshandeh Amnieh, H., Siamaki, A., and Soltani, S. (2012). "Design of blasting pattern in proportion to the peak particle velocity (PPV): Artificial neural networks approach". *Safety Science*, 50: 1913-1916.
- [21] Kotake, N., Suzuki, K., Asahi, S., and Kanda, Y. (2002). "Experimental study on the grinding rate constant of solid materials in a ball mill". *Powder Technology*, 122: 101-108.
- [22] Singh, V., Banerjee, P. K., Tripathy, S. K., Saxena, V. K., and Venugopal, R. (2013). "Artificial Neural Network Modeling of Ball Mill Grinding Process". *Powder Metallurgy & Mining*, 2(2): 2-4.
- [23] Braspenning P. J., and Thuijsman F. (1995). "Artificial Neural Networks: An Introduction to ANN Theory and Practice". Springer Science & Business Media, New York, pp. 295.

- time series, and spatial data". Springer Science & Business Media, New York, pp. 318.
- [48] Oliver, M. A., and Webster, R. (2014). "A tutorial guide to geostatistics: Computing and modelling variograms and kriging". CATENA, 113: 56-69.
- [49] Journel, A., and Huijbregts, C. J. (1978). "Mining geostatistics". Academic Press, London, New York, pp. 600.
- [50] Webster, R., and Oliver, M. A. (2007). "Geostatistics for Environmental Scientists". 2nd Edition, John Wiley & Sons, Chichester, pp. 330.
- [51] Diggle, P. J., Tawn, J. A., and Moyeed, R. (2007). "Model-based geostatistics". Journal of the Royal Statistical Society: Series C (Applied Statistics), 47(3): 299-350.
- [52] Brown, P. (2015). "Model-Based Geostatistics the Easy Way". Journal of Statistical Software, 63(12): 1-24.
- [53] Lin, S. W., Lee, Z. J., Chen, S. C., and Tseng, T. Y. (2008). "Parameter determination of support vector machine and feature selection using simulated annealing approach". Applied Soft Computing, 8: 1505-1512.
- [54] Che, X. L., and Hu, L. (2008). "Grid Resource Prediction Approach Based On Nu-Support Vector Regression". Proceedings of the 7th International Conference on Machine Learning and Cybernetics, Kunming, 12-15 July.
- [55] Hsu, C. W., Chang, C. C., and Lin, C. J. (2010). "A Practical Guide to Support Vector Classification". Technical Report, Department of Computer Science and Information Engineering, University of National Taiwan, Taipei, 1-12.
- [56] Chang, C. C., and Lin, C. J. (2011). "LIBSVM: a library for support vector machines". ACM Transactions on Intelligent Systems and Technology, 2: 1-27.
- Mohammadi, S. (2013). "An SVM based machine learning method for the separation of alteration zones in Sungun porphyry copper deposit". Chem Erde-Geochemistry, 73: 545-554.
- [37] Abbaszadeh, M., Hezarkhani, A., and Soltani-Mohammadi, S. (2015). "Classification of Alteration Zones Based on Whole-Rock Geochemical Data using Support Vector Machine". Journal of the Geological Society of India, 85: 500-508.
- [38] Soliman, O. S., and Mahmoud, A. S. (2012). "A classification system for remote sensing satellite images using support vector machine with non-linear kernel functions". 8th International Conference on Informatics and Systems (INFOS), Cairo, 181-187.
- [39] Soliman, O. S., Mahmoud, A. S., and Hassan, S. M. (2012). "Remote Sensing Satellite Images Classification Using Support Vector Machine and Particle Swarm Optimization". Third International Conference on Innovations in Bio-Inspired Computing and Applications, Kaohsiung, 280-285.
- [40] Manthira Moorthi, S., Misra, I., Kaur, R., Darji, N. P., and Ramakrishnan, R. (2011). "Kernel based learning approach for satellite image classification using support vector machine". IEEE Recent Advances in Intelligent Computational Systems, Trivandrum, 107-110.
- [۴۱] رجب زاده، م. ع.، حسینی، ک.، موسوی نسب، ز.؛ ۲۰۱۵؛ "مطالعه کانی‌شناسی و زمین شیمی آپاتیت و سنگ‌های میزبان فسفات در کانسار اسفوردی استان یزد با هدف تعیین منشأ و خاستگاه آپاتیت". زمین شناسی اقتصادی، دوره ششم، شماره ۲، ص ۳۳۱-۳۵۳.
- [۴۲] سهیلی، م.، مهدوی، م.؛ ۱۳۷۰؛ "نقشه زمین‌شناسی اسفوردی، مقیاس ۱:۱۰۰۰۰۰". سازمان زمین‌شناسی کشور.
- [۴۳] شهرابی، ج.، ذوالقدر شجاعی، ع.؛ ۱۳۹۰؛ "داده‌کاوی پیشرفته: مفاهیم و الگوریتم‌ها". انتشارات جهاد دانشگاهی واحد دانشگاه صنعتی امیر کبیر، ۴۷۲ صفحه.
- [44] Zhang, D., Liu, W., Wang, A., and Deng, Q. (2011). "Parameter Optimization for Support Vector Regression Based on Genetic Algorithm with Simplex Crossover Operator". Journal of Information & Computational Science, 8: 911-920.
- [45] Zuo, R., and Carranza, M. E. J. (2011). "Support vector machine: A tool for mapping mineral prospectivity". Computers & Geosciences, 37: 1967-1975.
- [46] Matheron, G. (1969). "Le krigeage universel (Universal kriging)". Cahiers du Centre de Morphologie Mathematique, Ecole des Mines de Paris, Fontainebleau, 1: pp. 83.
- [47] Christensen, R. (1991). "Linear models for multivariate,

¹ Machine learning algorithms

² Statistical Learning Theory (SLT)

³ Structural Risk Minimization (SRM)

⁴ Vapnik

⁵ ϵ -Insensitive loss function

⁶ Linear

⁷ Polynomial

⁸ Radial basis

- ⁹ Sigmoid
- ¹⁰ Grid Search
- ¹¹ Cross validation
- ¹² Exhaustive search
- ¹³ Coarse grid
- ¹⁴ Fine grid



Grade Estimation in Esfordi Phosphate Deposit Using Support Vector Regression Method

Abbaszadeh M.^{1*}

1- Assistant Professor, Dept. of Mining Engineering, University of Kashan, Iran
Abbaszadeh@kashanu.ac.ir

(Received: 24 Apr. 2019, Accepted: 20 Nov. 2019)

Abstract: Nowadays, artificial intelligence methods have been broadly developed and applied for variable estimation to facilitate decision making in many fields. Grade estimation is an important issue in evaluating mineral deposits. Geostatistical methods are among the most commonly used approaches for variable estimation. Since these methods are somewhat defective in relation to limited numbers of dispersed nonlinear data, in this study, the support vector regression, a machine learning method, has been used for grade estimation in Esfordi phosphate deposit. The modeling accuracy was 84% according to the test data. Based on the results obtained from the modeling using the support vector regression method, grade estimation has been made within the block model in Esfordi phosphate deposit. The proposed potential areas in the block model can be taken as the the additional borehole sites in the further exploration stage. The tonnage-grade model was also prepared based on the results obtained by using the support vector regression modeling procedure. For example, based on this model, for a 6% cutoff grade, the reserve is about 15.36 million tons with an average grade of 13.59%.

Keywords: Machine learning algorithms, Support vector regression, Modeling, Estimation, Esfordi phosphate.

INTRODUCTION

There are several classic and linear estimation methods that can be expressed in mathematical form as below:

$$Z^*(v) = \sum_{i=1}^n \lambda_i Z(x_i) \quad (1)$$

Where λ_i is the weight of i th sample for block v . linear estimators are different in how they determine the weights of samples. After development of computational software, some methods such as geostatistics became common for variables estimation. Geostatistics is generally referred to method of estimating one unknown variable using known variables surrounded it [1]. In the case of encountering nonlinear data, nonlinear estimators should be used to improve the accuracy of the estimation [2]. Furthermore, geostatistical methods are highly sensitive to distribution and quantity of data, so that estimations become

very hard or even impossible in case of low data quantity or spatial dependency [3]. As novel variable estimation approaches, machine learning algorithms were used in this study to estimate the unknown variable. These algorithms can reveal the relationship between input (e.g. sample coordinates) and targeted variable (e.g. ore grade at known points) after being well trained with enough data. Assumptions such as linearity of components, coefficients, or relations are not necessary to be considered for these algorithms [2]. Artificial neural networks [4], random forest [5], and support vector machines [6] are of mostly used intelligent methods for grade estimation. Support vector regression (SVR) was used in this study to estimate the grade in Esfordi phosphate deposit. Robust mathematical reasoning of support vector regression has led to a widespread acceptance. Compared to the other scientific branches, few studies have considered the application of SVR in earth sciences. In this research, support vector regression was firstly compared with kriging and then implemented to estimate phosphate grade in Esfordi deposit.

METHODS

A total of 3781 data comprising geographical coordinates (x, y, z) and P_2O_5 percent, were obtained from 59 boreholes in Esfordi Phosphate deposit.

Support Vector Regression

SVR is a type of support vector machines that is commonly used for function approximation. Method includes training algorithms and is based on but more complete than the support vector classifier machines [7]. The SVR, which is based on the statistical learning theory and the structural risk minimization, was firstly introduced by Vapnik in 1990s[8].

Kriging

Kriging is a well-known estimation method in modeling and estimating mineral deposits. It is proposed with the aim of considering spatial correlation in modeling and estimation processes [9].

FINDINGS AND ARGUMENT

Grade Estimation Using Support Vector Regression

After preparing and determination of training and test datasets (80% and 20%, respectively), radial basis kernel function (RBF) was selected in order to estimate and model the phosphate grade using the support vector regression method. Better functionality was the main reason of using this kernel. In the next step, parameters were needed to be carefully determined to obtain a model with high generalization ability. In this regard, the grid search method with cross validation was used to determine optimal values for the model parameters. Model was then trained using the training dataset and finally evaluated on the test dataset.

Grade Estimation Using Kriging Method

As one of the most widely recognized grade estimation methods, kriging was used in this study to compare with the results obtained from SVR. The distribution function curve and Kolmogorov-Smirnov test proved the log-normality of the dataset, and hence, log-kriging method was performed to estimate the grade. Afterwards, the logarithmic experimental variogram of the phosphate grade was calculated and fitted to the model. Cross validation method was then used to select the best possible model. After obtaining the optimal parameters, kriging performance was assessed on training and test datasets. Results of applying support vector regression and kriging methods on both training and test datasets are tabulated in Table 1. For both training and testing sets, as can be seen in the Table 1, support vector regression was outperform to kriging in terms of the evaluation indices of correlation coefficient and root mean square

Table 1. Evaluation results of SVR and Kriging methods (training and test datasets)

	Kriging		Support Vector Regression	
	Train	Test	Train	Test
Correlation Coefficient	0.830	0.820	0.8526	0.8466
RMSE	5.3346	5.3868	4.5666	4.4475
No. of data	3025	756	3025	756

error, demonstrating the higher efficiency of SVR compared to the kriging method in estimating phosphate content in study area.

Block model estimation and preparation of tonnage-average grade based on SVR model

The evaluation results of SVR and kriging methods listed in Table 1, indicating better performance of the SVR method. Therefore, SVR was used to estimate phosphate grade in the block model of Esfordi phosphate deposit. At this stage, the geological model of the deposit was converted to block models of 30x30x30 m dimensions, and then, using the trained SVR model, grade of each block was estimated. As an example, Figure 1 shows the estimated horizontal section relating to level 1660-meter.

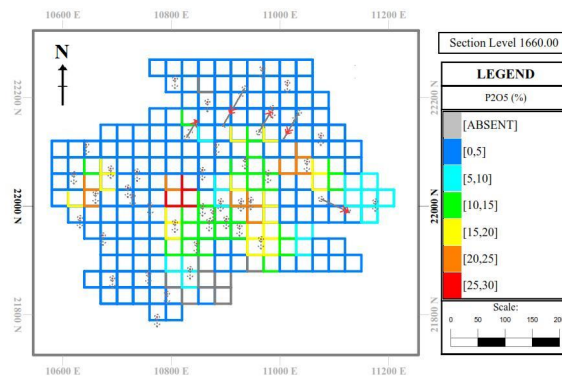


Figure 1. Horizontal section estimated by SVR, corresponding to level of 1660 m

Preparation of the tonnage-grade model was the next step, since there is always a need to determine the relationship between tonnage and average grade of the deposit for different cutoff grades. Figure 2, shows curves of tonnage and average grades of the Esfordi phosphate deposit versus the cutoff grade, according to the cutoff grades presented in Table 2 and the results obtained from the SVR method. The grade-tonnage curves was prepared with the specific gravity of 3.25 ton/m³.

Table 2. Average tonnage and grade values at different cutoff grades in Esfordi phosphate deposit

Average Grade (%)	Tonnage (million ton)	Cutoff Grade (%)
11.27	20.88	4
13.59	15.36	6
14.5	13.51	8
15.58	11.32	10

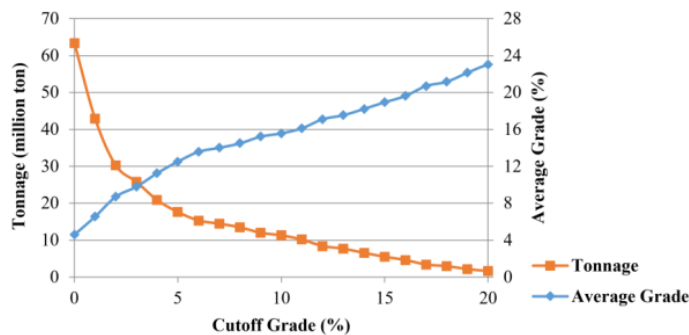


Figure 2. Changes in tonnage and average grade versus the cutoff grade in Esfordi phosphate deposit (based on results of the support vector regression)

CONCLUSIONS

As a new and effective approach, variable estimation has enabled decision making in many branches of science. Grade estimation is one of the most important issues in evaluating mineral reserves in earth sciences. Different methods can be considered to estimate the grade, the most common of which is geostatistics. In case of nonlinearity, however, nonlinear estimators should be used to improve and ensure the accuracy of the estimation. As one of the most efficient machine learning methods, support vector regression was used to estimate the grade in Esfordi phosphate deposit. Accordingly, the optimal values of the model parameters were firstly obtained by means of the grid search method, and then modeling was performed using the LIBSVM function and optimal parameters. The model obtained was then evaluated on test data and compared with results obtained from Kriging as one of the most common geostatistical techniques. SVR showed a high capability to estimate grade as one of the key parameters of reserve evaluation. Obtaining results were then used to estimate the grade based on the other data parameters of Esfordi block model. Results obtained from SVR are acceptable and can be incorporated in future mine economic assessments and design processes. The amount and average grade of the deposit were then determined based on the SVR estimated grades and different cutoff values.

REFERENCES

- [1] Hassani pak, A. A., and Sharafodin, M. (2001). *“Exploration Data Analysis”*. Tehran University Press, pp. 991.
- [2] Dutta, S. (2006). *“Predictive performance of machine learning algorithms for ore reserve estimation in sparse and imprecise data”*. PhD thesis, University of Alaska Fairbanks, pp. 189.
- [3] Dutta, S., Bandopadhyay, S., Ganguli, R., and Misra, D. (2010). *“Machine Learning Algorithms and Their Application to Ore Reserve Estimation of Sparse and Imprecise Data”*. Journal of Intelligent Learning Systems & Applications, 2: 86-96.
- [4] Chatterjee, S., Bandopadhyay, S., and Machuca, D. (2010). *“Ore Grade Prediction Using a Genetic Algorithm and Clustering Based Ensemble Neural Network Model”*. Math Geosci, 42: 309–326.
- [5] Jafrasteh, B., Fathianpour, N., and Suárez, A. (2018). *“Comparison of machine learning methods for copper ore grade estimation”*. Computational Geosciences, 22: 1371-1388.
- [6] Maleki, S., Ramazia, H. R., and Moradi, S. (2014). *“Estimation of Iron concentration by using a support vector machine and an artificial neural network - the case study of the Choghart deposit southeast of Yazd, Yazd, Iran”*. Geopersia, 4: 201-212.
- [7] Smola, A. J., and Scholkopf, B. (2004). *“A Tutorial on Support Vector Regression”*. Statistics and Computing, 14: 199-222.
- [8] Matías, J. M., Vaamonde, A., Taboda, J., and González-Manteiga, W. (2004). *“Support vector machines and gradient boosting for graphical estimation of a slate deposit”*. Stochastic Environmental Research and Risk Assessment, 18: 309-323.
- [9] Oliver, M. A., and Webster, R. (2014). *“A tutorial guide to geostatistics: Computing and modelling variograms and kriging”*. CATENA, 113: 56-69.