

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده مهندسی معدن، نفت و ژئوفیزیک
گروه اکتشاف معدن

رساله دکتری

" ارزیابی نقشه پراکندگی ژئوشیمیایی عمقی غلظت عناصر بر اساس روش‌های

شناسایی الگو به منظور بهبود تخمین عیار و مدل کانسار در عمق "

نگارنده: محرم جهانگیری

استاد راهنما

دکتر سید رضا قوامی ریابی

استاد مشاور

دکتر بهزاد تخم چی

دی ۱۳۹۶



فرم شماره ۱۲: صورت جلسه نهایی دفاع از رساله دکتری (Ph.D)
(ویژه دانشجویان ورودی های ۹۴ و ما قبل)

بدینوسیله گواهی می شود آقای محرم جهانگیری دانشجوی دکتری رشته مهندسی معدن - اکتشاف به شماره دانشجویی ۹۱۲۴۹۹۵ و ورودی مهر ماه سال ۹۱ در تاریخ ۱۳۹۶/۱۰/۱۶ از رساله نظری / عملی خود با عنوان: "ارزیابی نقشه پراکندگی ژئوشیمیایی عمقی غلظت عناصر بر اساس روش های شناسایی الگو به منظور بهبود تخمین عیار و مدل کانسار در عمق"

دفاع و با اخذ نمره ۱۹.۳ به درجه عالی نائل گردید.

<input checked="" type="checkbox"/> الف) درجه عالی: نمره ۱۹-۲۰	<input type="checkbox"/> ب) درجه بسیار خوب: نمره ۱۸/۹۹ - ۱۷
<input type="checkbox"/> ج) درجه خوب: نمره ۱۶/۹۹ - ۱۵	<input type="checkbox"/> د) غیر قابل قبول و نیاز به دفاع مجدد دارد
<input type="checkbox"/> ه) رساله نیاز به اصلاحات دارد	

ردیف	هیئت داوران	سمت	مرتبۀ علمی	امضاء
۱	دکتر سید رضا قوامی ریایی	استاد راهنما	دانشیار	
۲	دکتر بهزاد تخم چی	مشاور	دانشیار	
۳	دکتر احمد رضا مختاری	استاد مدعو خارجی	دانشیار	
۴	دکتر ابولقاسم کامکار روحانی	استاد مدعو داخلی	دانشیار	
۵	دکتر مسعود علیپور اصل	استاد مدعو داخلی	استادیار	
۶	دکتر امین روشندل کاهو	سرپرست (نماینده) تحصیلات تکمیلی دانشکده	دانشیار	

مدیر محترم تحصیلات تکمیلی دانشگاه:

ضمن تأیید مراتب فوق مقرر فرمائید اقدامات لازم در خصوص انجام مراحل دانش آموختگی آقای محرم جهانگیری بعمل آید.

نام و نام خانوادگی رئیس دانشکده: دکتر علیرضا عرب امیری
تاریخ و امضاء و مهر دانشکده:



تقدیم به:

پدر، مادر، همسر و خانواده عزیزم

ماحصل آموخته‌هایم را تقدیم می‌کنم به آنان که مهر آسمانی شان آرام بخش آلام زمینی‌ام است
به استوارترین تکیه گاهم، دستان پرمهر پدرم
به سبزترین نگاه زندگیم، چشمان سبز مادرم
به زلال‌ترین عنصر زندگی‌ام، وجود همسر مهربانم
به بهترین تعریف عشق و محبت، کانون گرم خانواده‌ام
که هرچه آموختم در مکتب عشق شما آموختم و هرچه بکوشم قطره‌ای از دریای بی‌کران مهربانیتان را سپاس
نتوانم بگویم .

امروز هستی‌ام به امید شماست و فردا کلید باغ بهشت‌م رضای شما
ره آوردی گران سنگ‌تر از این ارزان نداشتم تا به خاک پایتان نثار کنم، باشد که حاصل تلاش‌م نسیم گونه غبار
خستگی‌تان را بزدايد .
بوسه بر دستان پرمهرتان.

تشکر و قدردانی

جناب آقایان دکتر سیدرضا قوامی ریابی و دکتر بهزاد تخم چی اساتید محترم راهنما و مشاورم:

شما روشنایی بخش تاریکی جان هستید و ظلمت اندیشه را نور می‌بخشید. چگونه سپاس گویم مهربانی و لطف شما را که سرشار از عشق و یقین است. چگونه سپاس گویم تأثیر علم آموزی شما را که چراغ روشن هدایت را بر کلبه‌ی محقر وجودم فروزان ساخته است. آری در مقابل این همه عظمت و شکوه شما، مرا نه توان سپاس است و نه کلام وصف.

تعهد نامه

اینجانب محرم جهانگیری دانشجوی دوره دکتری رشته مهندسی معدن - اکتشاف دانشکده معدن، نفت و ژئوفیزیک دانشگاه صنعتی شاهرود نویسنده رساله " ارزیابی نقشه پراکندگی ژئوشیمیایی عمقی غلظت عناصر بر اساس روش‌های شناسایی الگو به منظور بهبود تخمین عیار و مدل کانسار در عمق " تحت راهنمایی جناب آقای دکتر سیدرضا قوامی ربایی و مشاوره آقای دکتر بهزاد تخم چی متعهد می شوم:

- تحقیقات در این رساله توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش‌های محققان دیگر به مرجع مورد استفاده استناد شده است.
- مطالب مندرج در رساله تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق معنوی این اثر متعلق به دانشگاه شاهرود می‌باشد و مقالات مستخرج با نام « دانشگاه شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی رساله تأثیرگذار بوده‌اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این رساله، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری، ضوابط و اصول اخلاق انسانی رعایت شده است.



تاریخ: ۱۳۹۶/۱۱/۰۲

امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم‌افزارها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می‌باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در پایان‌نامه بدون ذکر مرجع مجاز نمی‌باشد.

چکیده

با توجه به اینکه امروزه روش‌های هوشمند و در راس آن روش‌های تشخیص‌الگو از روش‌های متداول تحلیل داده‌ها در اکثر رشته‌های علوم و مهندسی می‌باشد، ژئوشیمی اکتشافی نیز از این روش‌ها بهره برده و در تحلیل داده‌های ژئوشیمیایی از این روش‌ها برای جداسازی آنومالی از زمینه داده‌های سطحی استفاده شده است. در تحقیق حاضر سعی شده با استفاده از الگوریتم‌های خوشه‌بندی و طبقه‌بندی و روش‌های تشخیص‌الگو، ضمن تلفیق اطلاعات سطحی و عمقی کانسار مس پورفیری سوناجیل، تخمین غلظت عناصر در نقاط فاقد اندازه‌گیری و تخمین غلظت افق‌های مختلف عمقی انجام گیرد و تغییرات عمقی غلظت ژئوشیمیایی عناصر در سطوح مختلف، مدل‌سازی و بررسی گردد. با استفاده از شاخص‌های اعتبارسنجی، تعداد خوشه بهینه داده‌ها شناسایی و خوشه‌بندی داده‌ها بوسیله الگوریتم‌های خوشه‌بندی انجام شده است و تخمین مقادیر غلظت عناصر بوسیله شبکه عصبی و داده‌های خوشه‌بندی شده صورت گرفته است. نتایج نشان داد که استفاده از ترکیب الگوریتم خوشه‌بندی FCM و شبکه عصبی دقت تخمین را به خوبی افزایش داده و میانگین دقت تخمین ۲۲ عنصر مورد مطالعه را از ۷۵ درصد به ۸۸ درصد افزایش داده است. پس از تخمین، نقشه‌های غلظت عناصر در سطوح مختلف عمقی مورد بررسی قرار گرفته و عناصر مرتبط با کانی‌سازی، فوق و تحت‌کانساری شناسایی شد. بررسی رفتار عناصر نشان داد که عنصر آرسنیک رفتار فوق‌کانساری و عنصر وانادیوم رفتار تحت‌کانساری در کانسار مس پورفیری سوناجیل از خود نشان داده است. بر اساس طبقه‌بندی تغییرات غلظت مس توسط روش فرکتالی عیار - مساحت و تعیین کلاس‌های مس در الگوریتم‌های طبقه‌بندی، مدل کانسار در عمق ترسیم گردید. نتایج بررسی هاله‌های عمقی مختلف نشان داد کانسار در قسمت جنوب‌غربی منطقه فرسایش یافته و کانی‌سازی در قسمت شمال‌مرکزی منطقه در عمق ادامه دارد. از بین روش‌های مختلف طبقه‌بندی استفاده شده، بیشترین دقت مربوط به الگوریتم شبکه عصبی با دقت طبقه‌بندی ۹۰/۶ درصد بود. اعتبارسنجی نتایج بوسیله اطلاعاتی ژئوفیزیکی و گمانه‌ها، نتایج حاصل را تایید کرده است.

کلمات کلیدی: الگوریتم‌های تشخیص الگو، دقت تخمین، مقاطع عمقی ژئوشیمیایی، مدلسازی عمقی کانسار.

مجلات:

1. Jahangiri, M., Ghavami Riabi, S.R. and Tokhmechi, B., 2017. **Estimation of geochemical elements using a hybrid neural network-Gustafson-Kessel algorithm.** Journal of Mining and Environment.

۲. جهانگیری، محرم؛ قوامی ریابی، سید رضا؛ تخم چی، بهزاد؛ ۱۳۹۶؛ ترکیب الگوریتم خوشه‌بندی Fuzzy c-means با شبکه عصبی پرسپترون چند لایه برای افزایش دقت تخمین غلظت عناصر ژئوشیمیایی، مثال موردی - محدوده شرقی کانسار مس پورفیری سوناجیل؛ نشریه علمی-پژوهشی " فصلنامه زمین‌شناسی ایران " .

کنفرانسی:

۱. جهانگیری، محرم؛ قوامی ریابی، سید رضا؛ تخم چی، بهزاد؛ ۱۳۹۵، جداسازی آنومالی از زمینه شرق سوناجیل بوسیله روش فرکتال و مقایسه آن با نتایج خوشه‌بندی Fuzzy c- و k-means و means، هشتمین همایش انجمن زمین‌شناسی اقتصادی ایران - دانشگاه زنجان.
۲. جهانگیری، محرم؛ قوامی ریابی، سید رضا؛ تخم چی، بهزاد؛ ۱۳۹۵، بهبود شناسایی هاله ژئوشیمیایی طلا در شرق سوناجیل هریس بوسیله شبکه‌عصبی چندلایه با استفاده از الگوریتم خوشه‌بندی Gustafson kessel، هشتمین همایش انجمن زمین‌شناسی اقتصادی ایران - دانشگاه زنجان.
۳. جهانگیری، محرم؛ قوامی ریابی، سید رضا؛ تخم چی، بهزاد؛ ۱۳۹۶، خوشه بندی داده های ژئوشیمیایی کانسار مس پورفیری سوناجیل با استفاده از الگوریتم خوشه بندی Gustafson kessel، دهمین همایش ملی زمین‌شناسی دانشگاه پیام نور، تبریز، دانشگاه پیام نور استان آذربایجان شرقی - مرکز تبریز.

فهرست مطالب

۱- کلیات	۱
۱-۱- مقدمه	۲
۲-۱- کانسارهای مس پورفیری	۳
۳-۱- بیان مساله و ضرورت انجام تحقیق	۶
۴-۱- اهداف و سوالات تحقیق	۸
۵-۱- فرض‌های تحقیق	۹
۶-۱- روش اجرای تحقیق	۱۰
۷-۱- زمین‌شناسی منطقه	۱۱
۱-۷-۱- ائوسن	۱۳
۱-۷-۱-۱- واحد آندزیتی ائوسن (Ean)	۱۳
۱-۷-۱-۲- واحد آندزیت مگاپورفیری (Eam)	۱۴
۱-۷-۱-۳- واحد آندزیت پورفیری (Eap)	۱۵
۲-۷-۱- الیگوسن	۱۶
۱-۷-۲-۱- سنگ‌های گرانیتوئیدی الیگوسن (Og)	۱۶
۳-۷-۱- کواترنر	۱۸
۱-۷-۳-۱- واحد بازالتی کواترنر (Qv)	۱۸
۸-۱- داده‌های ژئوشیمیایی استفاده شده	۱۸
۹-۱- ساختار پایان نامه	۲۰
۲- معرفی روش‌ها	۲۳
۱-۲- مقدمه	۲۴
۲-۲- ماشین بردار پشتیبان (SVM)	۲۵
۳-۲- طبقه‌بند نزدیک‌ترین k همسایگی	۳۰

۳۲	۴-۲ طبقه‌بندی داده‌ها بر اساس تئوری بی‌زین
۳۲	۱-۴-۲ مبانی علمی روش
۳۳	۲-۴-۲ روش طبقه‌بندی بی‌زین
۳۴	۵-۲ خوشه بندی Fuzzy C-Means (FCM)
۳۶	۶-۲ الگوریتم گوستافسون کسل
۳۸	۷-۲ الگوریتم گت‌جوا (GG)
۴۰	۱-۷-۲ شاخص‌های اعتبارسنجی الگوریتم‌های خوشه‌بندی فازی
۴۰	۱-۱-۷-۲ شاخص PC
۴۱	۲-۱-۷-۲ شاخص CE
۴۱	۳-۱-۷-۲ شاخص زی - بنی XB
۴۲	۴-۱-۷-۲ شاخص SC
۴۲	۵-۱-۷-۲ شاخص جدایش S
۴۳	۸-۲ روش خوشه بندی K-means
۴۵	۱-۸-۲ شاخص‌های اعتبارسنجی الگوریتم خوشه‌بندی K-means
۴۵	۱-۱-۸-۲ شاخص دان (DI)
۴۶	۲-۱-۸-۲ شاخص دان جایگزین (ADI)
۴۶	۹-۲ شبکه عصبی مصنوعی
۴۹	۱۰-۲ روش فرکتالی عیار - مساحت
۵۱	۳- مروری بر مطالعات پیشین
۵۲	۱-۳ مقدمه
۵۲	۲-۳ تخمین عیار و ذخیره در کانسارهای معدنی
۵۴	۳-۳ بهبود تخمین گر‌ها
۵۶	۴-۳ جدایش آنومالی از زمینه

- ۵۷.....۳-۵- الگوریتم‌های شناسایی الگو.....
- ۵۹.....۳-۶- تاریخچه‌ای از کاربرد روش‌ها در ژئوشیمی.....
- ۶۷.....۳-۷- مثال‌هایی از کاربرد روش‌های شناسایی الگو.....
- ۶۷.....۳-۷-۱- کاربرد ماشین‌های بردار پشتیبان.....
- ۷۳.....۳-۷-۲- کاربرد طبقه‌بندی KNN.....
- ۷۷.....۳-۷-۳- کاربرد روش‌های بیزین.....
- ۸۷.....۴- بهبود دقت تخمین غلظت عناصر ژئوشیمیایی.....
- ۸۸.....۴-۱- مقدمه.....
- ۸۹.....۴-۲- طراحی تخمینگر بر اساس کل داده‌ها برای عناصر ژئوشیمیایی.....
- ۹۱.....۴-۳- طراحی تخمینگر بر اساس داده‌های خوشه‌بندی شده برای عناصر ژئوشیمیایی.....
- ۹۲.....۴-۳-۱- تعیین تعداد خوشه بهینه.....
- ۹۴.....۴-۳-۲- افزایش دقت تخمین با الگوریتم خوشه‌بندی FCM.....
- ۹۶.....۴-۳-۲-۱- مقایسه دقت تخمینگر با حالت کلی.....
- ۱۰۰.....۴-۳-۲-۲- مقایسه خطای تخمینگر با حالت کلی.....
- ۱۰۲.....۴-۳-۳- افزایش دقت تخمین با الگوریتم خوشه‌بندی GK.....
- ۱۰۴.....۴-۳-۳-۱- مقایسه دقت تخمینگر با حالت کلی.....
- ۱۰۸.....۴-۳-۳-۲- مقایسه خطای تخمینگر با حالت کلی.....
- ۱۱۰.....۴-۳-۴- افزایش دقت تخمین با الگوریتم خوشه‌بندی GG.....
- ۱۱۲.....۴-۳-۴-۱- مقایسه دقت تخمینگر با حالت کلی.....
- ۱۱۵.....۴-۳-۴-۲- مقایسه خطای تخمینگر با حالت کلی.....
- ۱۱۶.....۴-۳-۵- افزایش دقت تخمین با الگوریتم خوشه‌بندی K means.....
- ۱۱۸.....۴-۳-۵-۱- مقایسه دقت تخمینگر با حالت کلی.....
- ۱۲۲.....۴-۳-۵-۲- مقایسه خطای تخمینگر با حالت کلی.....

- ۴-۴- انتخاب بهترین تخمینگر ۱۲۳
- ۴-۴-۱- بررسی کیفیت داده‌های تخمینی ۱۲۶
- ۵- نقشه غلظت عناصر در سطوح مختلف عمقی و مدل کانسار ۱۲۹
- ۵-۱- مقدمه ۱۳۰
- ۵-۲- عناصر تحت و فوق کانساری در سطوح مختلف عمقی ۱۳۰
- ۵-۲-۱- پروفیل ژئوشیمیایی AA' ۱۳۴
- ۵-۲-۲- پروفیل ژئوشیمیایی EE' ۱۴۹
- ۵-۲-۳- پروفیل ژئوشیمیایی BB' ۱۵۵
- ۵-۲-۴- پروفیل ژئوشیمیایی CC' ۱۵۷
- ۵-۲-۵- پروفیل ژئوشیمیایی DD' ۱۶۰
- ۵-۲-۶- نسبت عنصر فوق کانساری به تحت کانساری ۱۶۲
- ۵-۳- طبقه‌بندی تغییرات غلظت مس در سطوح مختلف عمقی ۱۶۳
- ۵-۳-۱- نقشه فرکتالی تغییرات غلظت مس ۱۶۳
- ۵-۳-۲- تحلیل مولفه‌های اصلی ۱۶۶
- ۵-۳-۳- طبقه‌بندی داده‌ها با روش بیزین ۱۶۹
- ۵-۳-۳-۱- استفاده از کل داده‌ها ۱۶۹
- ۵-۳-۳-۲- استفاده از داده‌های خوشه‌بندی ۱۷۰
- ۵-۳-۴- طبقه‌بندی داده‌ها با روش SVM ۱۷۲
- ۵-۳-۴-۱- استفاده از کل داده‌ها ۱۷۲
- ۵-۳-۴-۲- استفاده از داده‌های خوشه‌بندی ۱۷۳
- ۵-۳-۵- طبقه‌بندی داده‌ها با روش KNN ۱۷۶
- ۵-۳-۵-۱- استفاده از کل داده‌ها ۱۷۶
- ۵-۳-۵-۲- استفاده از داده‌های خوشه‌بندی ۱۷۷

- ۱۷۹ ANN طبقه‌بندی داده‌ها با روش ANN ۵-۳-۶
- ۱۷۹ استفاده از کل داده‌ها. ۵-۳-۶-۱
- ۱۸۰ استفاده از داده‌های خوشه‌بندی. ۵-۳-۶-۲
- ۱۸۳ انتخاب بهترین طبقه‌بند. ۵-۳-۷
- ۱۸۵ اعتبارسنجی، نتیجه‌گیری و پیشنهادات. ۶-۱-۱۸۵
- ۱۸۶ داده‌های اکتشافی منطقه و اعتبارسنجی. ۶-۱-۱۸۶
- ۱۸۶ روش پلاریزاسیون القایی. ۶-۱-۱-۱۸۶
- ۱۹۰ اعتبارسنجی. ۶-۲-۱۹۰
- ۱۹۰ نتیجه‌گیری. ۶-۳-۱۹۰
- ۱۹۱ پیشنهادات. ۶-۴-۱۹۱
- ۱۹۳ مراجع. ۷-۱۹۳

فهرست شکل‌ها

- شکل ۱-۱: مدل ذخیره مس پرفیری لاول و گیلبرت شامل الف) مقطع عرضی مناطق آلتراسیون‌های هیدروترمال و ب) مواد معدنی همراه این مناطق ۴
- شکل ۱-۲: تبدیل فضای ناهمگن کل داده‌ها به فضای همگن با استفاده از الگوریتم‌های خوشه‌بندی ۷
- شکل ۱-۳: نقشه زمین‌شناسی شرق محدوده اکتشافی سوناچیل هریس ۱۲
- شکل ۱-۴: بافت میکرولیتی پورفیریک در آندزیت مگاپورفیرهای منطقه (XPL, $\times 25$) ۱۴
- شکل ۱-۵: بافت پورفیریک با خمیره درشت بلور در زون دگرسانی پتاسیک - فلیک (XPL, $\times 25$) ۱۶
- شکل ۱-۶: فنوکریستال پلاژیوکلاز با زونینگ نوسانی در توده گرانیتیوئیدی اینچه (XPL, $\times 25$) ۱۷
- شکل ۱-۷: بافت میکرولیتیک تا اینترگرانولار جریانی در توده ولکانیکی اکوز داغی (XPL, $\times 100$) ۱۸
- شکل ۱-۸: موقعیت نمونه‌های سطحی و گمانه‌های حفاری شده در منطقه ۱۹
- شکل ۱-۲: شکل کلی طبقه‌بندی داده‌ها با استفاده از SVM، X_1 و X_2 محور ویژگی می‌باشند. a) زمانی که داده‌ها به صورت کامل از هم تفکیک شده باشند، b) زمانی که داده‌ها به صورت کامل از هم تفکیک نشده باشند. ۲۶
- شکل ۲-۲: مثالی از KNN با $K=10$ ۳۱
- شکل ۲-۳: بخش‌های مختلف شبکه عصبی ۴۷
- شکل ۲-۴: ساختار شبکه عصبی MLP ۴۹
- شکل ۱-۳: نقشه زمین‌شناسی معدن مس پرفیری نوچون به همراه موقعیت آن در نقشه ایران ۶۷
- شکل ۲-۳: نقشه غلظت مس ۶۸
- شکل ۳-۳: نقشه پتانسیل معدنی حاصل از تلفیق داده‌های اکتشافی بوسیله SVM به همراه شماره گمانه‌ها ... ۶۹
- شکل ۳-۴: نقشه زمین‌شناسی و آلتراسیون منطقه کانی‌زایی مس پورفیری کوه‌پنج ۷۰
- شکل ۳-۵: نقشه پتانسیل مس (a) سطحی در افق ۲۷۲۰ متری و (b) عمقی در افق ۲۵۹۰ متری ۷۰
- شکل ۳-۶: نقشه‌های طبقه‌بندی شده مناطق آنومال پیش‌بینی شده توسط الگوریتم‌های مختلف (a) نتیجه الگوریتم آدابوست (b) نتیجه الگوریتم جنگل تصادفی (c) نتیجه الگوریتم ماشین‌های بردار پشتیبان ۷۱
- شکل ۳-۷: جداسازی آنومالی از زمینه مس با استفاده از مدل فرکتال ۷۲
- شکل ۳-۸: موقعیت جغرافیایی معدن طلای Sunrise Dam در غرب استرالیا. ۷۳
- شکل ۳-۹: نمودار میله‌ای غلظت عناصر طلا، آرسنیک، آنتیموان، روبیدیوم و کرم در طول گمانه UGD1944 (۳۰۷ متر حفاری) نمودارهای بالا اثر قطعه‌ای طلا (عدم پیوستگی) و همچنین ارتباط فضایی طلا با سایر عناصر ژئوشیمیایی را نشان می‌دهد. ۷۴

شکل ۳-۱۰: مقایسه مقدار طلا در گمانه UGD1944 در مقایسه با نتایج الگوریتم‌های سه گانه برای عناصر کروم، روبیدیم و آرسنیک. (a) طلا، (b) الگوریتم نزدیک‌ترین K همسایه (KNN) با $K=501$ ، (c) تخمین‌گر چگالی هسته‌ای (KDE) و (d) احتمال شرطی بیز (NB) ۷۵

شکل ۳-۱۱: مقایسه مقدار طلا در گمانه حفاری UGD1944 در مقایسه با نتایج الگوریتم‌های سه گانه برای عناصر کروم، روبیدیم و آنتیموان. (a) طلا، (b) الگوریتم نزدیک‌ترین K همسایه (KNN) با $K=501$ ، (c) تخمین‌گر چگالی هسته‌ای (KDE) و (d) احتمال شرطی بیز (NB) ۷۵

شکل ۳-۱۲: نقشه نقطه‌ای آنتیموان در مقابل روبیدیم که بوسیله پارامترهای مختلف رنگ‌گذاری شده اند، (a) مقدار عیار طلا، (b) تخمین‌گر چگالی هسته‌ای (KDE)، (c) احتمال شرطی بیز (NB)، (d) الگوریتم نزدیک‌ترین K همسایه (KNN) با $K=101$ ، (e) الگوریتم نزدیک‌ترین K همسایه (KNN) با $K=501$ ، (f) الگوریتم نزدیک‌ترین K همسایه (KNN) با $K=1001$ ۷۶

شکل ۳-۱۳: موقعیت محدوده مورد مطالعه در استان راجستان هند. دایره‌های سیاه رنگ نشانگر ذخایر فلزات پایه است. ۷۷

شکل ۳-۱۴: نقشه پتانسیل تولید شده با استفاده از طبقه‌بندی بیزین ساده، (A) نقشه پتانسیل در مقیاس خاکستری (B) نقشه پتانسیل دودویی تولید شده. مثلث‌های خاکستری ذخایر فلزات پایه را نشان می‌دهند... ۷۸

شکل ۳-۱۵: نقشه پتانسیل تولید شده با استفاده از طبقه‌بندی بیزین ساده تقویت شده، (A) نقشه پتانسیل در مقیاس خاکستری (B) نقشه پتانسیل دودویی تولید شده. مثلث‌های خاکستری ذخایر فلزات پایه را نشان می‌دهند. ۷۹

شکل ۳-۱۶: نقشه پتانسیل تولید شده با استفاده از طبقه‌بندی بیزین ساده انتخابی، (A) نقشه پتانسیل در مقیاس خاکستری (B) نقشه پتانسیل دودویی تولید شده. مثلث‌های خاکستری ذخایر فلزات پایه را نشان می‌دهند ۸۰

شکل ۳-۱۷: موقعیت گمانه‌ها با دوایر سیاه رنگ و نمونه‌های سطحی با مثلث‌های سبز رنگ در کانسار طلای ساری‌گونی و همچنین بلوک بندی سطحی داده‌ها با ابعاد 25×25 (طبقه A به رنگ قرمز) (آنومالی)، طبقه B به رنگ زرد (پتانسیل دار) و طبقه C به رنگ آبی (زمینه) ۸۲

شکل ۳-۱۸: نقشه پراکندگی وضعیت کانی‌زایی طلا و موقعیت گمانه‌ها در کانسار ساری - گونی (طبقه A به رنگ قرمز) (آنومالی)، طبقه B به رنگ زرد (پتانسیل دار) و طبقه C به رنگ آبی (زمینه) ۸۳

شکل ۳-۱۹: نقشه کنثوری طلا به دست آمده از داده‌های ژئوشیمیایی خاک (واحد اندازه گیری طلا در مقیاس میله‌ای ppb است) ۸۳

- شکل ۳-۲۰: نقشه پراکندگی مدل‌سازی طلا در سه طبقه به روش الف) بی‌زی، ب) نزدیک‌ترین K همسایه، ج) روش پنجره پارزن و د) روش بی‌زی ساده در منطقه ساری‌گونئی (ترکیب رنگی‌های انتخاب شده مشابه شکل ۳-۱۹ است)..... ۸۴
- شکل ۴-۱: طراحی تخمینگر با استفاده از کل داده‌ها برای عناصر ژئوشیمیایی..... ۸۹
- شکل ۴-۲: مدل طراحی تخمینگر شبکه عصبی MLP..... ۹۰
- شکل ۴-۳: دقت تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از کل داده‌ها..... ۹۰
- شکل ۴-۴: خطای تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از کل داده‌ها..... ۹۰
- شکل ۴-۵: طراحی تخمینگر با استفاده از داده‌های خوشه‌بندی شده برای عناصر ژئوشیمیایی..... ۹۲
- شکل ۴-۶: شاخص‌های اعتبارسنجی PC و CE..... ۹۲
- شکل ۴-۷: شاخص‌های اعتبارسنجی SC، S و XB..... ۹۳
- شکل ۴-۸: شاخص‌های اعتبارسنجی DI و ADI..... ۹۳
- شکل ۴-۹: پراکندگی نمونه‌های خوشه‌بندی شده به روش FCM به صورت سه بعدی در فضا..... ۹۴
- شکل ۴-۱۰: دقت تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از داده‌های خوشه‌بندی شده به روش FCM..... ۹۵
- شکل ۴-۱۱: خطای تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از داده‌های خوشه‌بندی شده به روش FCM..... ۹۵
- شکل ۴-۱۲: مقایسه دقت کل تخمینگرها در حالت استفاده از کل داده‌ها و داده‌های خوشه‌بندی شده به روش FCM..... ۹۶
- شکل ۴-۱۳: نمودار خوشه‌بندی عناصر ژئوشیمیایی (عناصر عناصر مرتبط با مس پورفیری)..... ۹۷
- شکل ۴-۱۴: نمودار رگرسیونی داده‌های واقعی و تخمینی عنصر Zn در حالت استفاده از کل داده‌ها. محور x داده‌های واقعی و محور y داده‌های تخمینی است. a) داده‌های آموزشی، b) داده‌های اعتبارسنجی، c) داده‌های آزمون و d) کل داده‌ها..... ۹۹
- شکل ۴-۱۵: نمودار رگرسیونی داده‌های واقعی و تخمینی عنصر Zn در حالت استفاده از داده‌های خوشه‌بندی شده با روش FCM. محور x داده‌های واقعی و محور y داده‌های تخمینی است. a) داده‌های آموزشی، b) داده‌های اعتبارسنجی، c) داده‌های آزمون و d) کل داده‌ها..... ۹۹
- شکل ۴-۱۶: مقایسه خطای تخمین در حالت استفاده از کل داده‌ها و حالت استفاده از الگوریتم خوشه‌بندی ۱۰۰..... ۱۰۰

شکل ۴-۱۷: نمودار تغییرات خطا و هیستوگرام پراکندگی مقادیر خطا عنصر Zn، (a1) نمودار خطا در حالت استفاده از کل داده‌ها و (a2) حالت استفاده از داده‌های خوشه‌بندی شده، (b1) هیستوگرام خطا در حالت استفاده از کل داده‌ها و (b2) حالت استفاده از داده‌های خوشه‌بندی شده ۱۰۱

شکل ۴-۱۸: پراکندگی نمونه‌های خوشه‌بندی شده به روش گوستافسون کسل به صورت سه بعدی در فضا... ۱۰۲

شکل ۴-۱۹: دقت تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از داده‌های خوشه‌بندی شده به روش GK ۱۰۳

شکل ۴-۲۰: خطای تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از داده‌های خوشه‌بندی شده به روش GK ۱۰۳

شکل ۴-۲۱: مقایسه دقت کل تخمینگرها در حالت استفاده از کل داده‌ها و داده‌های خوشه‌بندی شده به روش GK ۱۰۵

شکل ۴-۲۲: نمودار رگرسیونی داده‌های واقعی و تخمین زده شده. محور x داده‌های واقعی و محور y داده‌های تخمینی است. (a1) نمودار رگرسیونی عنصر Al در حالت استفاده از کل داده‌ها، (a2) نمودار رگرسیونی عنصر Al در حالت استفاده از داده‌های خوشه‌بندی شده، (b1) نمودار رگرسیونی عنصر Zn در حالت استفاده از کل داده‌ها و (b2) نمودار رگرسیونی عنصر Al در حالت استفاده از داده‌های خوشه‌بندی شده ۱۰۶

شکل ۴-۲۳: مقایسه نقشه غلظت مقادیر واقعی و مقادیر تخمینی در سطح، (c1) نقشه غلظت عنصر Al با مقادیر واقعی، (c2) نقشه غلظت عنصر Al با مقادیر تخمینی، (d1) نقشه غلظت عنصر Zn با مقادیر واقعی و (d2) نقشه غلظت عنصر Zn با مقادیر تخمینی، (e1) نقشه غلظت عنصر Cu با مقادیر واقعی و (e2) نقشه غلظت عنصر Cu با مقادیر تخمینی ۱۰۷

شکل ۴-۲۴: مقایسه خطای تخمین در حالت استفاده از کل داده‌ها و حالت استفاده از الگوریتم خوشه‌بندی گوستافسون کسل ۱۰۸

شکل ۴-۲۵: نمودار تغییرات خطا و هیستوگرام پراکندگی مقادیر خطا عنصر As، (f1) نمودار خطا در حالت استفاده از کل داده‌ها و (f2) حالت استفاده از داده‌های خوشه‌بندی شده، (g1) هیستوگرام خطا در حالت استفاده از کل داده‌ها و (g2) حالت استفاده از داده‌های خوشه‌بندی شده ۱۰۹

شکل ۴-۲۶: پراکندگی نمونه‌های خوشه‌بندی شده به روش گت جوا به صورت سه بعدی در فضا ۱۱۰

شکل ۴-۲۷: دقت تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از داده‌های خوشه‌بندی شده به روش GG ۱۱۱

شکل ۴-۲۸: خطای تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از داده‌های خوشه‌بندی شده به روش GG ۱۱۱

شکل ۴-۲۹: مقایسه دقت کل تخمینگرها در حالت استفاده از کل داده‌ها و داده‌های خوشه‌بندی شده به روش GG ۱۱۲

شکل ۴-۳۰: نمودار رگرسیونی داده‌های واقعی و تخمینی عنصر Zn در حالت استفاده از کل داده‌ها. محور x داده‌های واقعی و محور y داده‌های تخمینی است. (a) داده‌های آموزشی، (b) داده‌های اعتبارسنجی، (c) داده‌های آزمون و (d) کل داده‌ها. ۱۱۴

شکل ۴-۳۱: نمودار رگرسیونی داده‌های واقعی و تخمینی عنصر Zn در حالت استفاده از داده‌های خوشه‌بندی شده با روش GG. محور x داده‌های واقعی و محور y داده‌های تخمینی است. (a) داده‌های آموزشی، (b) داده‌های اعتبارسنجی، (c) داده‌های آزمون و (d) کل داده‌ها. ۱۱۴

شکل ۴-۳۲: مقایسه خطای تخمین در حالت استفاده از کل داده‌ها و حالت استفاده از الگوریتم خوشه‌بندی ۱۱۵

شکل ۴-۳۳: نمودار تغییرات خطا و هیستوگرام پراکندگی مقادیر خطا عنصر As، (f1) نمودار خطا در حالت استفاده از کل داده‌ها و (f2) حالت استفاده از داده‌های خوشه‌بندی شده، (g1) هیستوگرام خطا در حالت استفاده از کل داده‌ها و (g2) حالت استفاده از داده‌های خوشه‌بندی شده ۱۱۶

شکل ۴-۳۴: پراکندگی نمونه‌های خوشه‌بندی شده به روش K means به صورت سه بعدی در فضا ۱۱۷

شکل ۴-۳۵: دقت تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از داده‌های خوشه‌بندی شده به روش K means ۱۱۷

شکل ۴-۳۶: خطای تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از داده‌های خوشه‌بندی شده به روش K means ۱۱۸

شکل ۴-۳۷: مقایسه دقت کل تخمینگرها در حالت استفاده از کل داده‌ها و داده‌های خوشه‌بندی شده به روش K means ۱۲۰

شکل ۴-۳۸: نمودار رگرسیونی داده‌های واقعی و تخمینی عنصر Mo در حالت استفاده از کل داده‌ها. محور x داده‌های واقعی و محور y داده‌های تخمینی است. (a) داده‌های آموزشی، (b) داده‌های اعتبارسنجی، (c) داده‌های آزمون و (d) کل داده‌ها. ۱۲۱

شکل ۴-۳۹: نمودار رگرسیونی داده‌های واقعی و تخمینی عنصر Mo در حالت استفاده از داده‌های خوشه‌بندی شده با روش K means. محور x داده‌های واقعی و محور y داده‌های تخمینی است. (a) داده‌های آموزشی، (b) داده‌های اعتبارسنجی، (c) داده‌های آزمون و (d) کل داده‌ها. ۱۲۱

شکل ۴-۴۰: نمودار تغییرات خطا و هیستوگرام پراکندگی مقادیر خطا عنصر As، (f1) نمودار خطا در حالت استفاده از کل داده‌ها، (f2) نمودار خطا در حالت استفاده از داده‌های خوشه‌بندی شده، (g1) هیستوگرام خطا در حالت استفاده از کل داده‌ها و (g2) هیستوگرام خطا در حالت استفاده از داده‌های خوشه‌بندی شده ۱۲۲

شکل ۴-۴۱: مقایسه خطای تخمین در حالت استفاده از کل داده‌ها و حالت استفاده از الگوریتم خوشه‌بندی K	۱۲۳
..... means	۱۲۳
شکل ۴-۴۲: مقایسه دقت تخمینگر در حالت‌های مختلف	۱۲۵
.....	۱۲۵
شکل ۴-۴۳: مقایسه خطای تخمینگر در حالت‌های مختلف	۱۲۵
.....	۱۳۱
شکل ۵-۱: نمودار خوشه‌ای داده‌های سطحی و عمقی	۱۳۱
.....	۱۳۲
شکل ۵-۲: موقعیت پروفیل‌های ژئوشیمیایی و گمانه‌های اکتشافی بر روی نقشه غلظت مس در سطح	۱۳۲
.....	SOI-14 و SOI-12 ، SOI -07
شکل ۵-۳: تغییرات غلظت مس نسبت به عمق در طول گمانه‌های اکتشافی	SOI -07 ، SOI -12 و SOI-14
.....	۱۳۳
شکل ۵-۴: تغییرات غلظت مس نسبت به عمق در طول گمانه‌های اکتشافی	SOI -08 ، SOI -09 ، SOI-11 و SOI -10
.....	۱۳۳
شکل ۵-۵: مقاطع عرضی تغییرات عنصر مس در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی	۱۳۵
.....	۱۳۶
شکل ۵-۶: تغییرات غلظت مس در سطوح مختلف عمقی	۱۳۶
.....	۱۳۷
شکل ۵-۷: تغییرات غلظت آرسنیک در سطوح مختلف عمقی	۱۳۷
.....	شکل ۵-۸: مقاطع عرضی تغییرات عنصر آرسنیک در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی
.....	۱۳۸
شکل ۵-۹: تغییرات غلظت وانادیوم در سطوح مختلف عمقی	۱۳۹
.....	شکل ۵-۱۰: مقاطع عرضی تغییرات عنصر وانادیوم در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی
.....	۱۳۹
شکل ۵-۱۱: تغییرات غلظت کبالت در سطوح مختلف عمقی	۱۴۰
.....	شکل ۵-۱۲: مقاطع عرضی تغییرات عنصر کبالت در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی
.....	۱۴۱
شکل ۵-۱۳: تغییرات غلظت آلومینیوم در سطوح مختلف عمقی	۱۴۲
.....	شکل ۵-۱۴: مقاطع عرضی تغییرات عنصر آلومینیوم در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی
.....	۱۴۲
شکل ۵-۱۵: تغییرات غلظت آهن در سطوح مختلف عمقی	۱۴۴
.....	شکل ۵-۱۶: مقاطع عرضی تغییرات عنصر آهن در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی
.....	۱۴۴
شکل ۵-۱۷: تغییرات غلظت گوگرد در سطوح مختلف عمقی	۱۴۵

- شکل ۵-۱۸: مقاطع عرضی تغییرات عنصر گوگرد در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی ۱۴۵.....
- شکل ۵-۱۹: تغییرات غلظت روی در سطوح مختلف عمقی ۱۴۶.....
- شکل ۵-۲۰: مقاطع عرضی تغییرات عنصر روی در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی ۱۴۶.....
- شکل ۵-۲۱: تغییرات غلظت سرب در سطوح مختلف عمقی ۱۴۷.....
- شکل ۵-۲۲: مقاطع عرضی تغییرات عنصر سرب در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی ۱۴۷.....
- شکل ۵-۲۳: تغییرات غلظت مولیبدن در سطوح مختلف عمقی ۱۴۸.....
- شکل ۵-۲۴: مقاطع عرضی تغییرات عنصر مولیبدن در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی ۱۴۸.....
- شکل ۵-۲۵: مقاطع عرضی تغییرات عنصر مس در طول پروفیل ژئوشیمیایی EE' در سطوح مختلف عمقی. ۱۴۹.....
- شکل ۵-۲۶: مقاطع عرضی تغییرات عنصر آرسنیک در طول پروفیل ژئوشیمیایی EE' در سطوح مختلف عمقی ۱۵۰.....
- شکل ۵-۲۷: مقاطع عرضی تغییرات عنصر وانادیوم در طول پروفیل ژئوشیمیایی EE' در سطوح مختلف عمقی ۱۵۱.....
- شکل ۵-۲۸: مقاطع عرضی تغییرات عنصر کبالت در طول پروفیل ژئوشیمیایی EE' در سطوح مختلف عمقی ۱۵۱.....
- شکل ۵-۲۹: مقاطع عرضی تغییرات عنصر آلومینیوم در طول پروفیل ژئوشیمیایی EE' در سطوح مختلف عمقی ۱۵۲.....
- شکل ۵-۳۰: مقاطع عرضی تغییرات عنصر آهن در طول پروفیل ژئوشیمیایی EE' در سطوح مختلف عمقی. ۱۵۳.....
- شکل ۵-۳۱: مقاطع عرضی تغییرات عنصر گوگرد در طول پروفیل ژئوشیمیایی EE' در سطوح مختلف عمقی ۱۵۳.....
- شکل ۵-۳۲: مقاطع عرضی تغییرات عنصر سرب در طول پروفیل ژئوشیمیایی EE' در سطوح مختلف عمقی ۱۵۴.....
- شکل ۵-۳۳: مقاطع عرضی تغییرات عنصر روی در طول پروفیل ژئوشیمیایی EE' در سطوح مختلف عمقی. ۱۵۴.....
- شکل ۵-۳۴: مقاطع عرضی تغییرات عنصر مس در طول پروفیل ژئوشیمیایی BB' در سطوح مختلف عمقی. ۱۵۵.....
- شکل ۵-۳۵: مقاطع عرضی تغییرات عنصر آرسنیک در طول پروفیل ژئوشیمیایی BB' در سطوح مختلف عمقی ۱۵۶.....

- شکل ۵-۳۶: مقاطع عرضی تغییرات عنصر وانادیوم در طول پروفیل ژئوشیمیایی BB' در سطوح مختلف عمقی ۱۵۷.....
- شکل ۵-۳۷: مقاطع عرضی تغییرات عنصر مس در طول پروفیل ژئوشیمیایی CC' در سطوح مختلف عمقی. ۱۵۸.....
- شکل ۵-۳۸: مقاطع عرضی تغییرات عنصر آرسنیک در طول پروفیل ژئوشیمیایی CC' در سطوح مختلف عمقی ۱۵۹.....
- شکل ۵-۳۹: مقاطع عرضی تغییرات عنصر وانادیوم در طول پروفیل ژئوشیمیایی CC' در سطوح مختلف عمقی ۱۵۹.....
- شکل ۵-۴۰: مقاطع عرضی تغییرات عنصر مس در طول پروفیل ژئوشیمیایی DD' در سطوح مختلف عمقی ۱۶۰.....
- شکل ۵-۴۱: مقاطع عرضی تغییرات عنصر آرسنیک در طول پروفیل ژئوشیمیایی DD' در سطوح مختلف عمقی ۱۶۱.....
- شکل ۵-۴۲: مقاطع عرضی تغییرات عنصر وانادیوم در طول پروفیل ژئوشیمیایی DD' در سطوح مختلف عمقی ۱۶۱.....
- شکل ۵-۴۳: نقشه نسبت آرسنیک به وانادیوم ۱۶۲.....
- شکل ۵-۴۴: نمودار فرکتالی عنصر مس در منطقه ۱۶۴.....
- شکل ۵-۴۵: نقشه فرکتالی تغییرات غلظت مس در سطح ۱۶۴.....
- شکل ۵-۴۶: نقشه فرکتالی تغییرات غلظت مس در سطح ۱۶۵.....
- شکل ۵-۴۷: نمودار مقادیر ویژه مولفه‌های اصلی داده‌ها ۱۶۷.....
- شکل ۵-۴۸: استفاده از داده‌های خوشه‌بندی شده بوسیله الگوریتم FCM جهت طراحی طبقه‌بند ۱۶۸.....
- شکل ۵-۴۹: ماتریس دقت طبقه‌بندی داده‌های آزمون با استفاده از روش بیزین و استفاده از کل داده‌ها ۱۶۹.....
- شکل ۵-۵۰: ماتریس دقت طبقه‌بندی داده‌های آزمون با استفاده از روش بیزین و استفاده از داده‌های خوشه‌بندی شده با FCM ۱۷۰.....
- شکل ۵-۵۱: طبقه‌بندی غلظت مس در سطوح مختلف عمقی با استفاده از روش بیزین (a) سطح عمقی ۵۰ متر، (b) سطح عمقی ۱۰۰ متر، (c) سطح عمقی ۱۵۰ متر و (d) سطح عمقی ۲۰۰ متر. ۱۷۱.....
- شکل ۵-۵۲: طبقه‌بندی غلظت مس در سطوح مختلف عمقی با استفاده از روش بیزین (a) سطح عمقی ۲۵۰ متر، (b) سطح عمقی ۳۰۰ متر، (c) سطح عمقی ۳۵۰ متر و (d) سطح عمقی ۳۷۰ متر. ۱۷۲.....
- شکل ۵-۵۳: ماتریس دقت طبقه‌بندی داده‌های آزمون با استفاده از روش SVM و استفاده از کل داده‌ها ۱۷۳.....
- شکل ۵-۵۴: ماتریس دقت طبقه‌بندی داده‌های آزمون با استفاده از روش SVM و استفاده از داده‌های خوشه‌بندی شده با FCM ۱۷۴.....

شکل ۵-۵۵: طبقه‌بندی غلظت مس در سطوح مختلف عمقی با استفاده از روش SVM (a) سطح عمقی ۵۰ متر، (b) سطح عمقی ۱۰۰ متر، (c) سطح عمقی ۱۵۰ متر و (d) سطح عمقی ۲۰۰ متر. ۱۷۵

شکل ۵-۵۶: طبقه‌بندی غلظت مس در سطوح مختلف عمقی با استفاده از روش SVM (a) سطح عمقی ۲۵۰ متر، (b) سطح عمقی ۳۰۰ متر، (c) سطح عمقی ۳۵۰ متر و (d) سطح عمقی ۳۷۰ متر. ۱۷۵

شکل ۵-۵۷: ماتریس دقت طبقه‌بندی داده‌های آزمون با استفاده از روش KNN و استفاده از کل داده‌ها ۱۷۶

شکل ۵-۵۸: ماتریس دقت طبقه‌بندی داده‌های آزمون با استفاده از روش KNN و استفاده از داده‌های خوشه‌بندی شده با FCM ۱۷۷

شکل ۵-۵۹: طبقه‌بندی غلظت مس در سطوح مختلف عمقی با استفاده از روش KNN (a) سطح عمقی ۵۰ متر، (b) سطح عمقی ۱۰۰ متر، (c) سطح عمقی ۱۵۰ متر و (d) سطح عمقی ۲۰۰ متر. ۱۷۸

شکل ۵-۶۰: طبقه‌بندی غلظت مس در سطوح مختلف عمقی با استفاده از روش KNN (a) سطح عمقی ۲۵۰ متر، (b) سطح عمقی ۳۰۰ متر، (c) سطح عمقی ۳۵۰ متر و (d) سطح عمقی ۳۷۰ متر. ۱۷۹

شکل ۵-۶۱: طراحی طبقه‌بند با استفاده از شبکه عصبی MLP ۱۸۰

شکل ۵-۶۲: ماتریس دقت طبقه‌بندی داده‌های آموزش و آزمون با استفاده از روش ANN و استفاده از کل داده‌ها ۱۸۰

شکل ۵-۶۳: ماتریس دقت طبقه‌بندی داده‌های آزمون با استفاده از روش ANN و استفاده از داده‌های خوشه‌بندی شده با FCM ۱۸۱

شکل ۵-۶۴: طبقه‌بندی غلظت مس در سطوح مختلف عمقی با استفاده از روش ANN (a) سطح عمقی ۵۰ متر، (b) سطح عمقی ۱۰۰ متر، (c) سطح عمقی ۱۵۰ متر و (d) سطح عمقی ۲۰۰ متر. ۱۸۲

شکل ۵-۶۵: طبقه‌بندی غلظت مس در سطوح مختلف عمقی با استفاده از روش ANN (a) سطح عمقی ۲۵۰ متر، (b) سطح عمقی ۳۰۰ متر، (c) سطح عمقی ۳۵۰ متر و (d) سطح عمقی ۳۷۰ متر. ۱۸۲

شکل ۶-۱: نمایش نقاط برداشت الکتریکی با آرایش مستطیلی بر روی نقشه زمین‌شناسی ۱۸۷

شکل ۶-۲: نمایش تغییرات کلی بارپذیری برداست ژئوفیزیکی انجام شده بر روی نقشه زمین‌شناسی منطقه ۱۸۷

شکل ۶-۳: نمایش پروفیل‌های دوقطبی - دوقطبی DD00 و DD600 بر روی نقشه زمین‌شناسی. ۱۸۸

شکل ۶-۴: نتایج مدلسازی مقاومت ویژه (a)، بارپذیری (b) برای پروفیل قطبی - دوقطبی DD00 ۱۸۹

شکل ۶-۵: نتایج مدلسازی مقاومت ویژه (a)، بارپذیری (b) برای پروفیل دوقطبی - دوقطبی DD600 ۱۸۹

فهرست جدول‌ها

- جدول ۱-۱: عناصر آنالیز شده برای نمونه‌های سطحی ۱۹
- جدول ۲-۱: عناصر آنالیز شده برای نمونه‌های گمانه‌ها ۲۰
- جدول ۳-۱: عناصر مشترک نمونه‌های سنگ آنالیز شده سطحی و عمقی ۲۰
- جدول ۱-۳: میزان خطای الگوریتم آدابوست در ترکیب با دیگر الگوریتم‌های انتخاب ورودی منطقه کوه‌پنج ۷۲
- جدول ۲-۳: نتایج اعتبارسنجی ۸۰
- جدول ۳-۳: نتایج اعتبارسنجی نقشه‌های پتانسیل ۸۱
- جدول ۴-۳: نتایج طبقه‌بندی داده‌های آزمایشی به چهار روش نظریه بیزین ۸۵
- جدول ۱-۴: نتایج خوشه‌بندی داده‌ها با استفاده از الگوریتم خوشه‌بندی FCM ۹۴
- جدول ۲-۴: جدول دقت تخمین غلظت عناصر ژئوشیمیایی در حالت استفاده از کل داده‌ها و حالت استفاده از الگوریتم خوشه‌بندی FCM ۹۸
- جدول ۳-۴: نتایج خوشه‌بندی داده‌ها با استفاده از الگوریتم خوشه‌بندی GK ۱۰۲
- جدول ۴-۴: جدول دقت تخمین غلظت عناصر ژئوشیمیایی در حالت استفاده از کل داده‌ها و حالت استفاده از الگوریتم خوشه‌بندی GK ۱۰۵
- جدول ۵-۴: نتایج خوشه‌بندی داده‌ها با استفاده از الگوریتم خوشه‌بندی GG ۱۱۰
- جدول ۶-۴: جدول دقت تخمین غلظت عناصر ژئوشیمیایی در حالت استفاده از کل داده‌ها و حالت استفاده از الگوریتم خوشه‌بندی ۱۱۳
- جدول ۷-۴: نتایج خوشه‌بندی داده‌ها با استفاده از الگوریتم خوشه‌بندی K means ۱۱۶
- جدول ۸-۴: جدول دقت تخمین غلظت عناصر ژئوشیمیایی در حالت استفاده از کل داده‌ها و حالت استفاده از الگوریتم خوشه‌بندی K means ۱۱۹
- جدول ۹-۴: مقایسه دقت و خطای تخمینگر در حالت‌های مختلف ۱۲۴
- جدول ۱۰-۴: مقایسه حداقل و حداکثر داده‌های واقعی و تخمینی (ppm) ۱۲۶
- جدول ۱۱-۴: مقایسه میانه و میانگین داده‌های واقعی و تخمینی (ppm) ۱۲۷
- جدول ۱-۵: عناصر مهم در کانسارهای مس پورفیری ۱۳۴
- جدول ۲-۵: تغییرات فرکتالی غلظت مس ۱۶۳
- جدول ۳-۵: عناصر ژئوشیمیایی مورد استفاده برای طبقه‌بندی مس ۱۶۶
- جدول ۴-۵: داده‌های مورد استفاده برای طبقه‌بندی ۱۶۸
- جدول ۵-۵: نتایج طبقه‌بندی در حالت استفاده از کل داده‌ها ۱۸۳

جدول ۵-۶: نتایج طبقه‌بندی در حالت استفاده از داده‌های خوشه‌بندی شده با FCM..... ۱۸۳

فصل

۱- کلیات

۱-۱- مقدمه

بخشی از هدف اکتشافات ژئوشیمیایی، شناخت مناطق آنومالی، تعیین نوع کانی‌سازی احتمالی، معرفی زون‌های کانی‌سازی و به دست آوردن پاراژنهای کانی‌سازی است (McClenaghan et al., 2000). در این پژوهش تلاش می‌شود تا با استفاده از روش‌های شناسایی الگو نخست به صورت مستقل و سپس به صورت ترکیب با شبکه‌های عصبی یک مدل دقیق از کانسار ساخته شود، به نحوی که با کمترین اطلاعات در دسترس و کمترین ابهام، افق‌های کانی‌سازی را مدل نمود. برای این منظور، بهینه‌سازی ساختار شبکه عصبی جهت بهبود تخمین، جدایش آنومالی‌های تک‌عنصری و چندگانه، زون‌بندی کانسار و استخراج اطلاعات زمین‌شناسی کانسار با استفاده از تکنیک‌های شناسایی الگو از جمله اهداف این پژوهش است.

روش‌های متفاوتی جهت تخمین غلظت عناصر و مدل‌سازی ژئوشیمیایی کانسار وجود دارد که از جمله به روش‌های زمین‌آماری، فرکتالی و هوش مصنوعی می‌توان اشاره نمود. شبکه‌های عصبی مصنوعی از جمله روش‌هایی است که با پردازش داده‌ها، روابط نهفته را کشف و قادر به ارائه مدل‌های دقیقی است (Duda et al., 2002; Theodoridis and Kourtombas, 2010).

تاکنون در حوزه علوم زمین، شبکه‌های عصبی با جایگاه ویژه خود به گونه‌ای عمل نموده که بیشتر پژوهشگران سعی بر استفاده از آن به عنوان ابزاری قدرتمند دارند. در این زمینه می‌توان به بررسی‌های ژئوشیمیایی غلظت عناصر با استفاده از شبکه‌های عصبی نوع احتمالی (Singer, 2006)، طبقه‌بندی مواد آلی رسوبی (Weller et al., 2005)، تعیین شوری به دست آمده از تصاویر ذرات آلی رسوبی (Andrew et al., 2007)، کمی کردن داده‌های ژئوشیمیایی سنگ‌های آتشفشانی و طبقه‌بندی آن‌ها (Lacassie et al., 2006; Lacassie et al., 2004)، تعیین ویژگی‌های سفره آب‌زیرزمینی (Rizzo and Dougherty, 1994)، تعیین ویژگی‌های مواد ناخالص در معادن سنگ‌آهک (Koike and Matsuda, 2003)، تخمین ذخیره (Wu and Zhou, 1993) اشاره کرد.

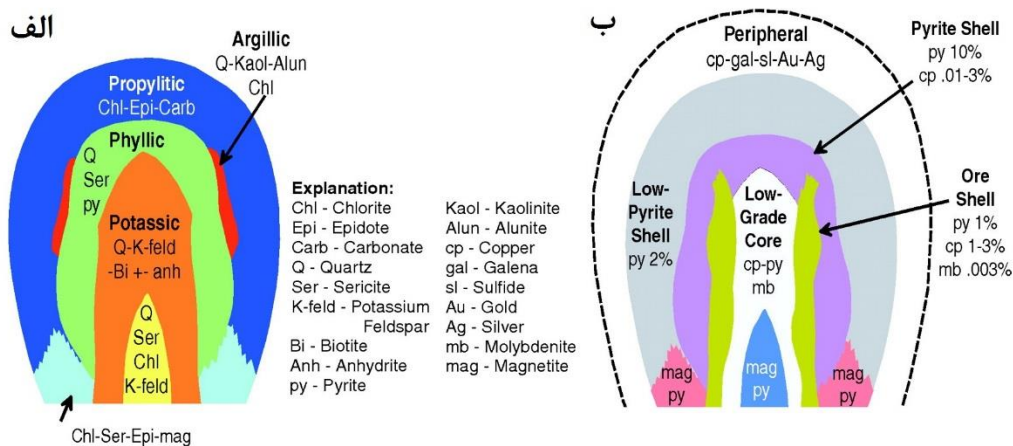
در مجموعه داده‌های عظیم، هنگامی که داده‌های درون مجموعه مورد مطالعه از نظم معقول پیروی می‌کنند، با استفاده از الگوریتم‌های شناسایی الگو این نظم قابل شناسایی و استفاده در مطالعات اکتشافی است. کانسارهای فلزی دارای تیپ‌های مختلف بوده و براساس مطالعات انجام شده، کانسارهای پورفیری به صورت ذاتی دارای یک نظم درونی هستند در این ذخایر افزایش و کاهش مقادیر عناصر مرتبط با کانی‌زایی از نظم خاصی تبعیت می‌کنند که شناسایی این نظم کمک فراوانی در انجام مراحل بعدی اکتشاف خواهد کرد. در این تحقیق تمرکز بر روی کانسارهای پورفیری متمرکز شده است.

۱-۲- کانسارهای مس پورفیری

کانسارهای مس پورفیری را می‌توان محصول سرد شدن سیستم‌های هیدروترمالی همراه با نفوذی‌های عمق پورفیری تشکیل شده در کمان‌های ماگمایی مرتبط با فرورانش دانست (Titley and Bean, 1981). بطور کلی مدل ژنتیکی کانسارهای هیدروترمالی را می‌توان به صورت دو مدل ژنتیکی مرتبط با فرآیندهای زمین‌شناسی و ژئوشیمیایی و مدل ژنتیکی مرتبط با فرآیندهای دگرشکلی پوسته‌ای و محیط‌های ژئودینامیکی بررسی نمود. مطالعات محققین نشان دهنده آن است که تشکیل و جایگیری کانسارهای مس پورفیری نه تنها تحت تأثیر فرآیندهای ماگمایی و هیدروترمالی بلکه تکتونیک ناحیه‌ای، محلی و رژیم‌های تکتونیک‌های زمان تشکیل قرار می‌گیرد (Richards *et al.*, 2001; Padilla *et al.*, 2001). بسیاری معتقدند که کانسارهای مس پورفیری، در مقایسه با سایر کانسارهای هیدروترمالی همزاد با توده‌های نفوذی است که در امتداد سیستم‌های گسلی امتداد لغز کمان‌های قاره‌ای و جزیره‌ای تشکیل می‌شوند. بنابراین به منظور بررسی جایگاه مناسب جایگیری نفوذی‌های پورفیری و کانسارهای مس مرتبط، شناسایی محل‌های تمرکز سیالات ماگمایی درون پوسته بسیار با اهمیت می‌باشد (Carranza, 2002). بطور کلی مهاجرت سیالات ماگمایی توسط شرایط محیطی و ترمودینامیک کنترل شده به طوری که سیالات عموماً از زون‌های پرفشار به مناطق کم فشار با درجه حرارت پایین‌تر مهاجرت می‌کنند. در چنین شرایطی، سیالات

ماگمایی بیش از آنکه متمرکز گردند، پراکنده می‌شوند. لذا بررسی هندسه و مکانیسم گسل‌های مرتبط با کنسارهای مس پورفیری، جهت مطالعه مکان‌های مناسب برای جایگیری توده‌های پورفیری می‌تواند از اهمیت بالایی برخوردار باشد (Zarasvandi, 2004; Zarasvandi et al., 2005). در این راستا زون‌های برشی، حوضه‌های کافتی مجزای از هم^۱ و انقطاع در امتداد گسل‌های امتداد لغز مکان‌هایی بسیار مناسب جهت نفوذ توده‌های تفریق یافته پورفیری و تشکیل کنسارهای مس پورفیری هستند (Carranza, 2002). اکثر ذخایر مس ایران به خصوص پورفیری‌ها در زون ولکانو - پلوتونیک ارومیه - دختر، ارسباران، طارم و بلوک لوت مرکزی واقع شده‌اند که ارتباط زمانی بسیار جالبی با گسل‌های امتداد لغز و توده‌های نفوذی گرانیتوئیدی میوسن در این زون دارند (قوامی ریایی و دارابی گلستان، ۱۳۹۳).

لاول و گیلبرت (۱۹۷۰)، توده معدنی سان مانوئل - کلامازو^۲ را توصیف و یافته‌های خویش را با ۲۷ کنسار مس پورفیری دیگر مقایسه و به عنوان مدل لاول - گیلبرت شناخته می‌شود، ارائه نمودند (شکل ۱-۱).



شکل ۱-۱: مدل ذخیره مس پورفیری لاول و گیلبرت شامل الف) مقطع عرضی مناطق آلتراسیون‌های هیدروترمال و ب) مواد معدنی همراه این مناطق (Mars et al., 2006)

1 Pull-apart basin
2 San Manuel, Kalamazoo

آن‌ها در این تحقیق ارزشمند و بنیادی نشان دادند که بهترین چهارچوب مرجعی که می‌تواند تمام سیمایهای دیگر این ذخایر را به هم ارتباط دهد، ماهیت و توزیع مناطق دگرسانی گرمابی سنگ دیواره است. بر اساس ادعای آن‌ها عموماً چهار منطقه دگرسانی وجود دارد. این مناطق بیشتر در اطراف استوک‌پورفیری به صورت مناطق هم‌محوری که پوسته‌هایی هم‌مرکز و اغلب ناکامل را می‌سازند متمرکز شده و غالباً در اکتشاف ذخایر مس‌پورفیری به عنوان یک راهنما مورد استفاده قرار می‌گیرد (شکل ۱-۱). این مناطق در مدل لاول - گیلبرت عبارت است از (Lowell and Guilbert, 1970؛ قوامی ریایی و دارایی گلستان، ۱۳۹۳):

الف) منطقه پتاسیک^۳: با تشکیل ارتوکلاز، بیوتیت ثانویه و مگنتیت یا تشکیل ارتوکلاز - بیوتیت - کلریت شناسایی می‌شود. سریسیت نیز ممکن است وجود داشته باشد. این کانی‌های ثانویه، جانشین ارتوکلاز، پلاژیوکلاز و کانی‌های مافیک اولیه توده نفوذی می‌شوند. انیدریت ممکن است در این منطقه غالب باشد.

ب) منطقه فیلیک^۴: فیلیک یا سریسیتی شدن نوعی دگرسانی است که بر اساس مجموعه کانی‌های کوارتز - سریسیت - پیریت مشخص می‌شود و معمولاً مقادیر کمی کلریت، ایلیت و روتیل نیز همراه آن است. سریسیتی شدن بر فلدسپارها و بیوتیت اولیه تاثیر می‌گذارد و دگرسانی بیوتیت و مقدار ناچیزی روتیل تولید می‌کند. این دگرسانی‌ها واکنش‌هایی سیلیسی‌زا هستند و به همین دلیل مقدار زیادی کوارتز تشکیل می‌شود (سیلیسی شدن). سطح تماس این منطقه با منطقه پتاسیک، سطح تماسی تدریجی با ضخامت بیش از ده‌ها متر است. منطقه فیلیک در صورت وجود، دارای بیشترین توسعه پیریت افشان و رگچه‌ای است.

3 Potassic zone
4 Phyllic zone

ج) منطقه آرژلیک⁵: این منطقه همیشه وجود ندارد. کانی‌های رسی، ویژگی این منطقه بوده و کائولن با نزدیک‌تر شدن به توده معدنی فراوان شده و با دور شدن از آن مونت‌موریونیت افزایش می‌یابد.

د) منطقه پروپیلیتیک⁶: این منطقه که خارجی‌ترین دگرسانی است، همواره وجود دارد. کلریت رایج‌ترین کانی این منطقه است. پیریت، کلسیت و اپیدوت نیز با آن همراه است. کانی‌های مافیک اولیه (بیوتیت و هورنبلند) به طور بخشی یا کامل به کلریت و کربنات تبدیل می‌شود و پلاژیوکلاز ممکن است بدون تغییر باقی بماند. این منطقه در طول صدها متر به سمت سنگ‌های احاطه کننده آن، محو می‌شود.

در این تحقیق هدف آن است که بر روی مثال‌های موردی از کانسارهای پورفیری متمرکز شویم و با شناسایی نظم حاکم بر کانی‌زایی در این کانسارها، با بهره‌گیری از الگوریتم‌های شناسایی‌الگو، بتوانیم تغییرات عیار کانسار در افق‌های مختلف عمقی را شناسایی و پیش‌بینی کنیم.

۱-۳- بیان مساله و ضرورت انجام تحقیق

به دلیل ناهمگنی در توزیع خواص فیزیکی و شیمیایی زمین، تخمین و مدل‌سازی این خواص توأم با عدم قطعیت بالایی است. تفکیک یک فضای ناهمگن به زیرفضاهای همگن، و مدل‌سازی مجزای هر کدام از این زیرفضاها راه‌حلی شناخته شده برای غلبه بر این مشکل است. نوع نگاه و تکنیکی که به وسیله آن اقدام به تقسیم فضا به زیرفضا می‌شود، در پاسخ نهایی بسیار اثرگذار است. با توجه به نوع و حجم داده‌های در دسترس و ماهیت غیرخطی داده‌های ژئوشیمیایی، روش شبکه عصبی چندلایه به دلیل توانایی شناسایی روابط غیرخطی برای تخمین پارامتر مورد نظر قابل استفاده است.

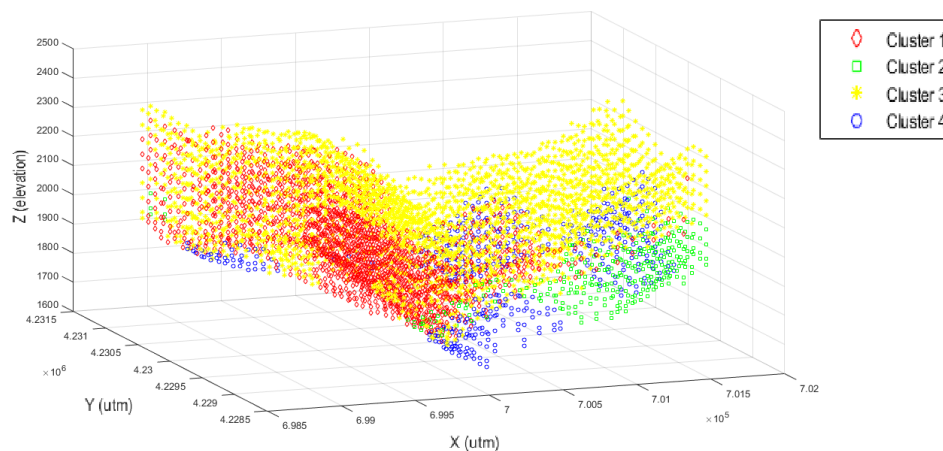
در مطالعات پیشین، اکثر مطالعات معطوف به بررسی‌های سطحی تغییرات غلظت عناصر بوده و کمتر این تغییرات در عمق مورد بررسی قرار داده شده است (Roshani et al., 2013; Han et al., 2015; Granian et al., 2016; Mohammadi Gonbadi et al., 2015; Shahi et al., 2016; et al., 2015; Harraz et al., 2015).

5 Argillic zone
6 Propylitic zone

مطالعاتی که به بررسی وجود یا عدم وجود کانی‌سازی در عمق پرداخته شده است صرفاً به معرفی مناطق پتانسیل‌دار در سطح اکتفا شده و نقشه پراکندگی هاله‌های محوری ژئوشیمیایی عناصر در سطوح مختلف عمقی بررسی نشده است. بنابراین، ضرورت ارائه روشی مبنی بر بررسی عمقی هاله‌های محوری عناصر مختلف و شناسایی کانی‌سازی پنهان در سطوح مختلف (بویژه زمانی که گمانه‌های محدودی حفر شده باشد) احساس می‌شود و در این رساله روشی برای حل این مشکل ارائه خواهد گردید.

در خوشه‌بندی و طبقه‌بندی هر چه تعداد پارامترهای اندازه‌گیری شده (ویژگی‌ها) در یک نقطه بیشتر باشد (افزایش بعد فضا) تفکیک نقاط آسان‌تر و هر چه تعداد نقاط اندازه‌گیری شده در یک محیط (در اینجا غلظت ژئوشیمیایی عناصر مرتبط با کانی‌سازی) بیشتر باشد ضریب اطمینان افزایش می‌یابد.

از الگوریتم‌های خوشه‌بندی جهت شکستن فضای کلی مجموعه داده‌ها به چندین زیرفضای همگن استفاده خواهد شود (شکل ۱-۲) و سپس انجام تخمین غلظت در هاله‌های ژئوشیمیایی با استفاده از روش شبکه عصبی در زیرفضاهای همگن صورت خواهد پذیرفت تا نتایج حاصل از تخمین غلظت عنصر یا عناصر دارای دقت و صحت بیشتری باشد.



شکل ۱-۲: تبدیل فضای ناهمگن کل داده‌ها به فضای همگن با استفاده از الگوریتم‌های خوشه‌بندی

پس از تخمین غلظت ژئوشیمیایی عناصر مختلف در سطوح مختلف عمقی، با انتخاب ویژگی‌های مناسب و مرتبط با کانی‌سازی مس پورفیری (عناصر ژئوشیمیایی) اقدام به طبقه‌بندی داده‌ها جهت شناسایی هاله‌های ژئوشیمیایی مس خواهد شد. در کانسارهای مس پورفیری بسته به جهت و حجم سیال کانه‌ساز، یک الگوی منطقه‌بندی ژئوشیمیایی در هاله‌های اولیه محوری وجود دارد. در این پژوهش جهت شناسایی این نظم، از الگوریتم‌های طبقه‌بندی مانند بیز، ماشین بردار پشتیبان و سایر الگوریتم‌های توانمند در این زمینه استفاده خواهد شد.

۱-۴- اهداف و سوالات تحقیق

در این پژوهش دو هدف وجود دارد: نخست مدل تخمین غلظت عناصر مرتبط با هاله‌های ژئوشیمیایی و کانی‌سازی و پیش‌بینی تغییرات غلظت عناصر مورد اکتشاف در افق‌های مختلف عمقی محدوده مورد مطالعه با استفاده از الگوریتم‌های خوشه‌بندی و تخمینگر شبکه عصبی چندلایه، دوم ارائه یک رویکرد جدید مبتنی بر طبقه‌بندی غلظت ژئوشیمیایی عناصر مرتبط با کانی‌سازی به منظور ارزیابی گسترش هاله‌های عنصری در افق‌های مختلف عمقی و شناسایی کانی‌سازی‌های پنهان در منطقه است. برای دستیابی به این مهم، اهداف زیر نیز مدنظر است:

الف) تقسیم فضای هاله‌های عناصر مرتبط با کانی‌سازی به زون‌ها و بخش‌های مختلف براساس دیدگاه‌های متفاوت و الگوریتم‌های مختلف خوشه‌بندی: بدیهی است که زون‌بندی در علوم‌زمین به صورت عام و در بهره‌برداری از معادن به صورت خاص بسیار کاربردی است. بنابراین از خروجی این بخش می‌توان برای تقسیم‌بندی داده‌ها در خوشه‌های مختلف و استفاده از خوشه‌ها برای افزایش دقت تخمین غلظت عناصر استفاده نمود.

ب) مدل‌سازی تغییرات عمقی غلظت عناصر در هاله‌های لیتوژئوشیمیایی اولیه محوری با استفاده از الگوریتم‌های طبقه‌بندی: اگر بپذیریم که در کانی‌سازی توسط سیال هیدروترمال بسته به ویژگی‌های ترمودینامیکی و شیمیایی خود سیال و ویژگی مکانی مسیر حرکتی آن، یک نظم قابل کشف در هاله‌های عناصر مرتبط با کانی‌سازی به وجود می‌آید، لذا تلاش خواهد شد تا تغییرات رفتاری کلاس مس را با استفاده از تغییرات غلظت سایر عناصر به عنوان ویژگی، در سطوح مختلف عمقی شناسایی شود.

به بیان دیگر در مرحله نخست تلاش می‌شود با استفاده از تکنیک‌های مختلف خوشه‌بندی و پارامترهای ورودی متفاوت و بهینه‌سازی تعداد خوشه‌ها به بهترین شکل ممکن در فضای آمار، محدوده هاله‌های اولیه محوری منطقه‌بندی شود و از نتایج حاصل از خوشه‌بندی جهت افزایش دقت تخمین غلظت ژئوشیمیایی عناصر استفاده کرد. سپس مدل گسترش هاله‌های محوری با استفاده از تحلیل‌های حاصل از الگوریتم‌های خوشه‌بندی و طبقه‌بندی مختلف تخمین زده شده و نهایتاً مدل تفسیر می‌شود.

سوالات اصلی تحقیق شامل موارد زیر خواهد بود:

الف) تخمین عمقی تغییرات غلظت عناصر هاله‌های ژئوشیمیایی در افق‌های مختلف از چه دقتی برخوردار است؟

ب) تغییرات غلظت عناصر مرتبط با کانی‌سازی چه ارتباطی با تغییرات حاصل از مجموع عناصر مورد مطالعه دارد؟

ج) آیا با استفاده از روش‌های شناسایی‌الگو امکان کشف ارتباط نمونه‌های سطحی و عمقی امکان‌پذیر است؟

۱-۵- فرض‌های تحقیق

➤ داده‌های ژئوشیمیایی مورد مطالعه در این تحقیق از دیدگاه زمین‌آماری، داده‌های مکانی هستند و نمونه‌ها متاثر از مناطق بالادست و ارتفاعات نبوده و می‌شود از نظر منطقی داده‌ها را درونیایی کرد.

➤ نمونه‌های برداشت شده از منطقه توانسته‌اند تغییرات غلظت در منطقه را به خوبی انعکاس دهند و این داده‌ها معرف خوبی از پراکندگی عناصر در منطقه هستند.

➤ اطلاعات و داده‌هایی که در مراحل اکتشافی تفصیلی بدست آمده و جهت اعتبار سنجی استفاده می‌شوند حائز اعتبار بوده و قابل اعتماد می‌باشد.

۱-۶- روش اجرای تحقیق

الف) با توجه به داده‌های در دسترس از کانسارهای مس پورفیری بررسی‌های زمین‌شناسی، اکتشافی و آماری صورت خواهد پذیرفت.

ب) براساس فاکتورهای اعتبارسنجی، تعداد خوشه‌ها بهینه‌سازی می‌شود. فاکتورهای بهینه‌سازی نیز شامل PC, CE, DI, ADI, S, Sc, Xb خواهد بود که نتایج‌شان براساس رأی اکثریت ترکیب خواهد شد و بهترین تعداد خوشه انتخاب خواهد گردید.

ج) خوشه‌بندی داده‌های سطحی و عمقی با استفاده از الگوریتم‌های مختلف خوشه‌بندی و تعداد خوشه بهینه بدست آمده از مرحله قبل انجام خواهد شد.

د) از نتایج مرحله قبل برای بهبود دقت تخمین‌گر عناصر ژئوشیمیایی استفاده خواهد شد. به این صورت که ابتدا تخمین با استفاده از کل داده‌ها و تخمین‌گر شبکه عصبی انجام خواهیم شد و سپس با انتخاب داده‌های خوشه‌بندی شده با روش‌های مختلف به عنوان ورودی شبکه عصبی سعی بر بهبود دقت تخمین خواهد شد و در نهایت بهترین تخمین‌گر طراحی و عناصر ژئوشیمیایی تخمین زده خواهد شد.

ه) در این مرحله با استفاده از الگوریتم‌های طبقه‌بندی پیشنهادی (شامل SVM^Y، KNN^A و بیزین و سایر الگوریتم‌های توانمند در این زمینه)، الگوی موجود در کل هاله‌های لیتوژئوشیمیایی اولیه محوری با

^Y Support vector machine

استفاده از آموزش الگوریتم‌ها بوسیله داده‌های آموزشی موجود (داده‌های سطحی و داده‌های گمانه‌ای) و سپس اعتبارسنجی نتایج با استفاده از داده‌های آموزشی بدست خواهد آمد و کلاس‌های مختلف هاله ژئوشیمیایی مس در داده‌های تخمین زده شده شناسایی خواهد شد.

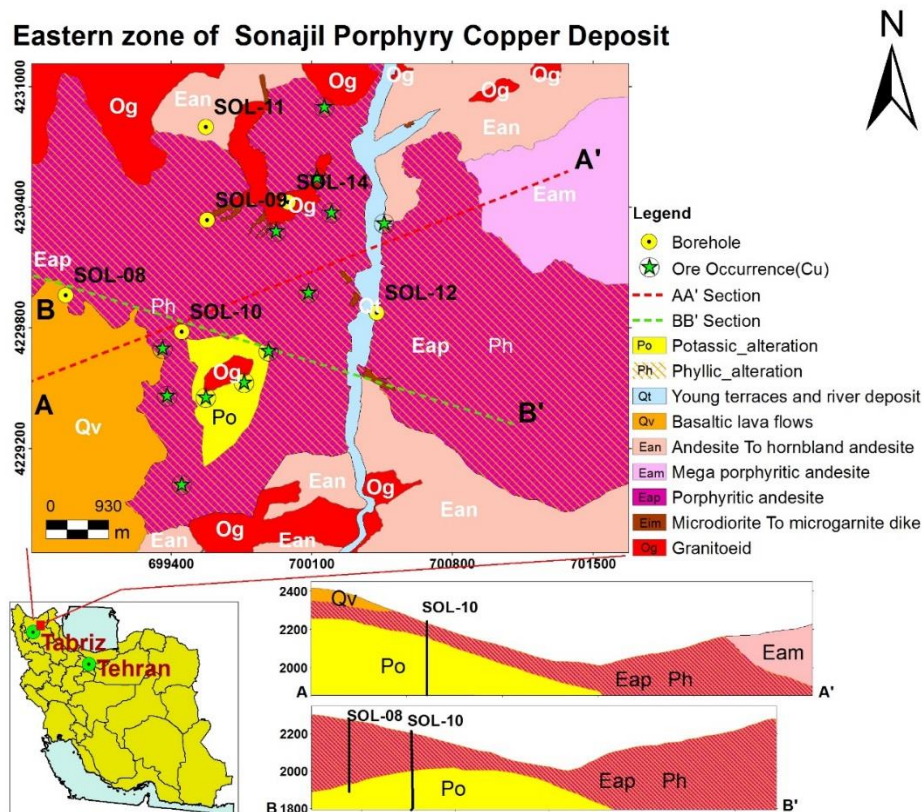
(و) با دستیابی به اطلاعات مرحله قبل، نتایج در فضای هاله‌های ژئوشیمیایی در سطوح مختلف عمقی به نقشه در خواهد آمد و نتایج مورد تفسیر قرار خواهند گرفت.

۷-۱- زمین‌شناسی منطقه

توده مس پورفیری سوناجیل در فاصله هوایی ۱۷ کیلومتری شرق هریس و ۴۵ کیلومتری جنوب شرق اهر از استان آذربایجان شرقی قرار گرفته است. این منطقه طبق تقسیمات زمین‌شناسی ایران در زون البرز غربی- آذربایجان قرار گرفته و بخشی از نوار آتشفشانی البرز- آذربایجان به سن ترشیری می‌باشد. وجود آلتراسیون‌های فراوان و سنگ‌های آندزیتی تا آندزیت مگاپرفیری (که خاستگاه مناسبی برای کانه‌سازی مس هستند)، منجر به پیشنهاد انجام عملیات اکتشافی در محدوده‌ای به وسعت تقریباً دو کیلومترمربع گردید.

توده پورفیری سوناجیل قدیمی‌ترین توده نفوذ در طی ماگماتیسم ترشیری در منطقه بوده و دارای دگرسانی شدید، ترکیب سنگ شناسی خاص و پیچیدگی زمین‌شناسی است. منطقه از پتانسیل کانی‌سازی مطلوبی بخصوص برای تیپ‌های مس - مولیبدن و طلا برخوردار می‌باشد. منطقه به علت فعالیت‌های ماگماتیسم گسترده، چه به لحاظ زمانی، مکانی و چه به لحاظ تنوع عمق سنگ‌ها، احتمال امید بخش بودن ذخیره پورفیری را افزایش می‌دهد. این امید بخشی بخصوص با توسعه زون‌های دگرسانی در منطقه افزایش پیدا می‌کند.

بررسی‌های صحرایی در محدوده کانسار سوناجیل نشان می‌دهد که سنگ‌های درونگیر توده پورفیری سوناجیل شامل تناوبی از نهشته‌های ولکانیک و ولکانو - کلاستیک ائوسن و سنگ‌های نفوذی ائوسن - الیگوسن می‌باشد که از قدیم به جدید شامل جریانات بازالتی (Qv)، آندزیت و آندزیت‌های هورنبلنددار (Ean)، آندزیت‌های مگاپورفیری (Eam)، آندزیت‌پورفیری (Eap)، میکرودیوریت و دایک‌های میکرودیوریتی (Eim) و گرانیتوئیدها (Og) می‌باشند. آندزیت پورفیری بخش وسیعی از محدوده مورد مطالعه را پوشش داده است. توده پورفیری سوناجیل دگرسانی‌های متنوعی از جمله پتاسیک و فلیک در سطح به نمایش گذاشته است (Hezarkhani, 2003). (شکل ۳-۱).



شکل ۳-۱: نقشه زمین‌شناسی شرق محدوده اکتشافی سوناجیل هریس (After Hezarkhani, 2003)
 در ادامه در بازه بین ائوسن تا کواترنر به شرح واحدهای سنگی متنوع موجود در منطقه می‌پردازیم.

۱-۷-۱- ائوسن

۱-۱-۷-۱- واحد آندزیتی ائوسن (Ean)

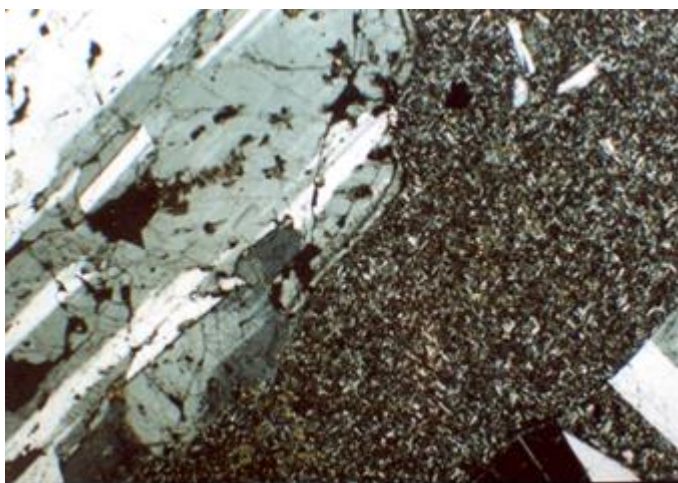
این واحد ولکانیک که به نظر می‌رسد قدیمی‌ترین واحد لیتولوژیک در این ناحیه می‌باشد به طور گسترده در بخش‌هایی از جنوب و شمال غرب منطقه دیده می‌شود. در بخش‌های جنوبی منطقه (شمال روستای آتمیان) حالت لایه‌بندی حاصل از جریانات گدازه‌ای در این سنگ‌ها دیده می‌شود. از نظر ماکروسکوپی این سنگ دارای بافت پورفیری است.

آپوفیزها و دایک‌هایی با ترکیب دیوریتی - گرانودیوریتی و حتی دیابازی در توده ولکانیک ائوسن این منطقه تزریق شده است. این واحد در مقایسه با واحدهای دیگر ائوسن در این منطقه آلتراسیون نسبتاً اندکی را متحمل شده است. این آلتراسیون‌ها عمدتاً از نوع پروپیلیتیک و تا حدی آرژیلیک است.

آلکالی فلدسپات‌ها شدیداً سریسیتی و آرژیلیتی شده‌اند و بافت این سنگ‌ها از نظر میکروسکوپی آفانیتیک تا پورفیری بوده است که در بعضی قسمت‌ها متحمل خردشدگی و شکستگی شده است. کانی‌های اصلی شامل پلاژیوکلاز، آلکالی فلدسپات، کوارتز و هورنبلند می‌باشد. پلاژیوکلازها بصورت فنوکریست درشت و میکروولیت تیغه‌ای ریز در سنگ‌های با بافت پورفیری دیده می‌شوند. در سنگ‌های ریزدانه این کانی به همراه اورتوکلاز دیده می‌شود. این کانی به صورت وسیع به سریسیت، کلسیت، تا حدی اپیدوت و کلریت تبدیل شده‌اند که به مقدار کم در بعضی نمونه‌ها دیده شده است. کوارتز اکثراً ثانویه بوده و اندازه آن در فضاهای خالی و شکستگی‌ها بزرگتر بوده و در متن سنگ ریزدانه می‌باشد. کانی فرعی آپاتیت می‌باشد که به صورت بلورهای ریز در متن کانی‌های دیگر مانند پلاژیوکلاز دیده می‌شود. کانی‌های ثانویه شامل کانی‌های زیر می‌باشد: کلریت حاصل تجزیه هورنبلند و پلاژیوکلاز، کلسیت حاصل تجزیه پلاژیوکلازها و هورنبلند، زئولیت‌ها، سریسیت حاصل تجزیه فلدسپات‌ها، کوارتز حاصل تجزیه اکثر کانی‌ها (صنایع مس ایران، ۱۳۸۰).

۱-۷-۱-۲- واحد آندزیت مگاپورفیری (Eam)

این واحد از نظر ترکیبی بیشتر دارای آندزیت تا تراکی آندزیت با بافت مگاپورفیری بوده و بلورهای درشت سفید یا صورتی رنگ پلاژیوکلاز در متن سبز رنگ پراکنده‌اند. این واحد از گسترده‌ترین لیتولوژی‌های منطقه می‌باشد که در اطراف روستای آتمیان، نرمیق و در منتهی‌الیه شمالی رودخانه اسماعیل‌کندی گسترده شده است. از نظر میکروسکوپی این سنگ‌ها دارای بافت مگاپورفیری می‌باشد. طول تیغه‌های پلاژیوکلاز گاهی به ۳ سانتی‌متر می‌رسد. در متن سنگ‌ها آثاری از پیریت‌های کاملاً اکسید شده دیده می‌شود. آلتراسیون هیدروترمال در این واحد شدید نبوده و آثاری از سریسیتی و آرژیلیتی شدن بلورهای درشت فلدسپات به چشم می‌خورد. از نظر میکروسکوپی بافت این سنگ پورفیری، میکروولیت پورفیری و مگاپورفیری می‌باشد (شکل ۱-۴). پلاژیوکلازها بسیار درشت‌البت و الیگوکلاز با ماکل‌البتیتی در آن مشاهده می‌شوند. این فنوکریست‌ها در بسیاری از موارد به کلسیت تبدیل شده‌اند. ماتریکس سنگ شامل پلاژیوکلاز، کلسیت ثانویه و آثاری از کانیهای فرومنیزین (احتمالاً بیوتیت) که به طور کامل به اکسید آهن تبدیل شده، می‌باشد. مقادیر زیادی تیغه آپاتیت در بین و داخل پلاژیوکلازها دیده می‌شود (صنایع مس ایران، ۱۳۸۰).



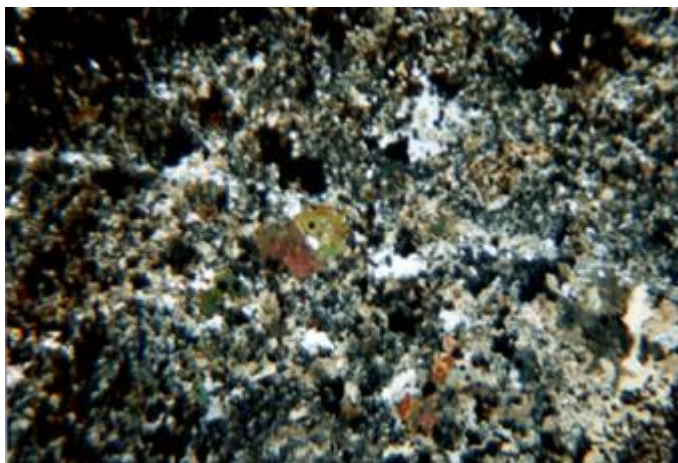
شکل ۱-۴: بافت میکروولیتی پورفیریک در آندزیت مگاپورفیرهای منطقه (XPL, $\times 25$)

۱-۷-۱-۳- واحد آندزیت پورفیری (Eap)

این واحد در بخش‌های وسیعی از مرکز محدوده رخنمون دارد. از نظر ماکروسکوپی این سنگ‌ها دارای بافت ریزدانه تا پورفیری می‌باشند. شدت آلتراسیون کاملاً متغیر است. شدیدترین آلتراسیون هیدروترمال و کانی‌سازی در این واحد رخ داده که در دره‌ای به نام آق‌یار و همچنین در ارتفاعات روستایی قره‌توپراخ و منتهی‌الیه شرقی محدوده مشاهده می‌شود. در این محدوده آلتراسیون‌های پتاسیک، آرژیلیک، فیلیک، سیلیسی و پروپیلیتیک قابل مشاهده می‌باشد (شکل ۱-۵).

دایک‌ها و آپوفیزهای فراوانی از توده نفوذی گرانیتوئیدی به داخل این واحد تزریق شده است. در بسیاری از موارد گدازه‌های بازالتی کواترنر این واحد را پوشانده است بطوریکه تشخیص کانی‌سازی و آلتراسیون در سطح مشکل است. توده‌های اکسید آهن در بالاترین سطوح این واحد در اطراف روستای جنگل و دره آق‌یار برونزد دارد. در این واحد شکستگی‌های فراوان و پرشدگی استوک‌ورکی از سیلیس و اکسید آهن دیده می‌شود. از نظر میکروسکوپی این سنگ‌ها دارای بافت پورفیری با فنوکریست‌های پلاژیوکلاز و میکروکریستال‌های پلاژیوکلاز در زمینه بوده است. کانی‌های اصلی به علت آلتراسیون شدید اغلب به کانی‌های ثانویه تبدیل شده‌اند. پلاژیوکلاز مهمترین کانی این سنگ است که عمدتاً به کلسیت، اپیدوت، کوارتز و کلریت تبدیل شده است. اکسیدهای آهن و کانی‌های ثانویه حاصل دگرسانی هورنبلند می‌باشند. کانی‌های ثانویه شامل تورمالین، کلسیت، اسکاپولیت، کوارتز، سیلیس و اکسیدهای آهن می‌باشد. مهمترین قسمت سیستم کانی‌سازی در این واحد متمرکز است کانی‌سازی، مس در این بخش بصورت رگچه‌ای و پراکنده رخ داده است. کانه‌های مس عبارت از کالکوپیریت، بورنیت، کالکوسیت، کولیت و مالاکیت سطحی است. پیریت نیز در زون کانی‌سازی به مقدار فراوان دیده می‌شود. هم رشدی بورنیت و کالکوپیریت، جانشینی پیریت‌ها توسط گوتیت و کالکوسیت از ویژگی‌های منطقه کانی‌سازی می‌باشد (صنایع مس ایران، ۱۳۸۰). کانی‌های فرعی موجود در سنگ شامل: روتیل، زیرکن، اسفن، تورمالین،

زئولیت، کلریت و کانی‌های رسی می‌باشد. دگرسانی پتاسیک، کوارتز، سربیسیت و آرژیلیک از مهمترین دگرسانی‌های این بخش می‌باشند.



شکل ۱-۵: بافت پورفیریک با خمیره درشت بلور در زون دگرسانی پتاسیک - فلیک (XPL, $\times 25$)

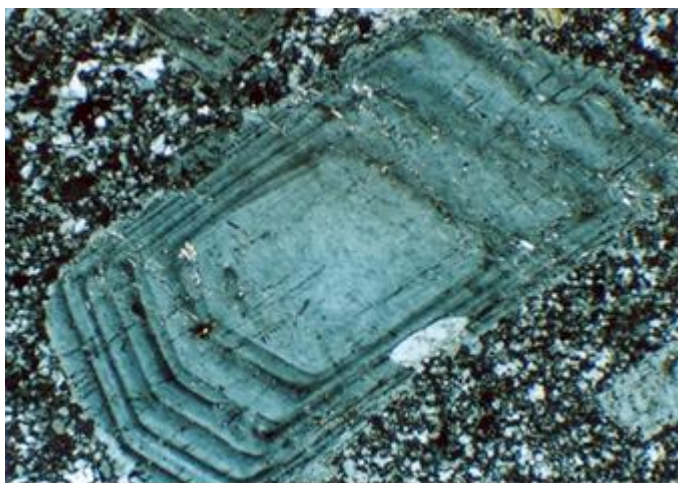
۱-۷-۲- الیگوسن

۱-۷-۲-۱- سنگ‌های گرانیتوئیدی الیگوسن (Og)

سنگ‌های گرانیتوئیدی یک استوک کوارتز مونزونیتی تا کوارتز مونزودیوریتی است که در راستای N-S و بطول تقریبی ۱۰ کیلومتر و عرض متوسط ۲ کیلومتر در قسمت وسیعی از شمال منطقه مورد مطالعه گسترش دارد. دگرسانی‌های خیلی وسیع در بخش جنوبی این توده حادث شده است.

آپوفیزهایی از این توده در سنگ‌های آندزیتی تزریق شده‌اند. و بعضی از آنها رسوبات ولکانیک ائوسن را قطع کرده‌اند. این توده در نهشته‌های ائوسن تزریق شده است و باعث دگرسانی، کانی‌سازی و شکستگی‌های شدید در آنها شده است. در خود توده نفوذی نیز فعالیت‌های تکتونیکی، سیلیسی شدن و خردشدگی شدید خصوصاً در دره منتهی به روستای اینچه مشاهده شده است. دگرسانی هیدروترمال در قسمت‌هایی از توده نفوذی دیده می‌شود. در جنوب دره اصلی جنگل برونزدهای مربوط به توده نفوذی اصلی با ترکیبی در حد مونزونیت تا مونزودیوریت دیده می‌شود. سنگ‌های این توده از نظر ماکروسکوپی

دارای بافت گرانولار است، این سنگ‌ها در ظاهر دگرسان بوده و آثاری از کلریت در شکستگی‌ها دیده می‌شود. به نظر می‌رسد در مرحله اول استوک نفوذی با رنگ تیره‌تر سپس واحد نفوذی روشن و در نهایت رگه سیلیسی نفوذ نموده است. دایک‌ها و رگه‌های وسیع سیلیسی در داخل توده نفوذی دیده می‌شود. که در اطراف این دایک‌ها کانی‌سازی مس بصورت ملاکیت دیده می‌شود. ترکیب سنگ‌شناسی این توده در حد آلکالی گرانیت تا هورنبلندگرانیت می‌باشد. این واحد نفوذی قسمتی از استوک اینچه است. از نظر میکروسکوپی بافت سنگ‌ها در این واحد گرانولار تا پورفیری بوده است در بافت گرانولار، فلدسپات و آمفیبول در اندازه درشت پراکنده‌اند و در بافت پورفیری پلاژیوکلاز و هورنبلند در زمینه‌ای از پلاژیوکلاز قرار دارد (شکل ۱-۶). کانی‌ها اصلی شامل پلاژیوکلاز سدیک که در اثر دگرسانی به سریسیت و اپیدوت تبدیل شده است. اورتوکلاز که تبدیل شدگی شدید به سریسیت و کانی‌های رسی نشان می‌دهد. آمفیبول (هورنبلند) که در اکثر موارد به مخلوطی از کلسیت، اپیدوت و کانی‌های اوپاک تبدیل شده است. کوارتز که در اکثر نمونه ۱۵ تا ۵ درصد دیده می‌شود. کانی‌های فرعی شامل آپاتیت، کلریت، اسفن و روتیل است. کلسیت، کوارتز، اپیدوت، کلریت، سریسیت و کانی‌های اوپاک، کانی‌های ثانویه را تشکیل می‌دهند که از آلتراسیون پلاژیوکلاز و هورنبلند بوجود آمده‌اند (صنایع مس ایران، ۱۳۸۰).

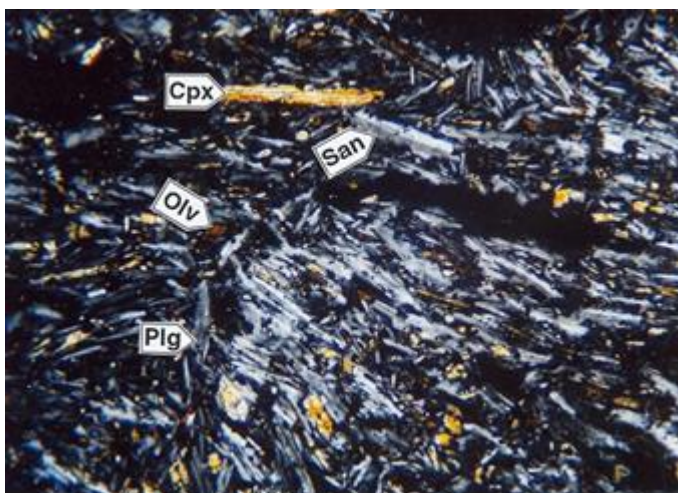


شکل ۱-۶: فنوکریستال پلاژیوکلاز با زونینگ نوسانی در توده گرانیتیوئیدی اینچه (XPL, $\times 25$)

۱-۷-۳- کوآترنر

۱-۷-۳-۱- واحد بازالتی کوآترنر (Qv)

این واحد جوانترین واحد سنگی منطقه بوده است که قسمت آتشفشانی منطقه را بنام کوه اکوزداغی با ارتفاع حدود ۲۶۰۰ متر با ترکیب آلكالی الیوین بازالت تا تراکی بازالت را تشکیل داده است. در این سنگ‌ها میکروولیت‌های پلاژیوکلاز به صورت جریان‌ی حدود ۹۰٪ حجم سنگ را تشکیل داده است. پلاژیوکلازها در بعضی موارد بصورت تیغه‌ای و دارای منطقه‌بندی دیده می‌شوند. آثاری از اولیوین و فضای خالی ناشی از آلتراسیون آن در متن سنگ قابل مشاهده است (شکل ۱-۷). واریزه‌هایی از این واحد در حجم گسترده بر روی واحدهای قدیمی‌تر قرار گرفته‌اند. این روانه‌های بازالتی دارای یک جبهه مشخص در قسمت‌های جنوبی بوده و قطعات بزرگ بازالتی حالت بالشی و گرد شده از خود نشان می‌دهند (صنایع مس ایران، ۱۳۸۰).



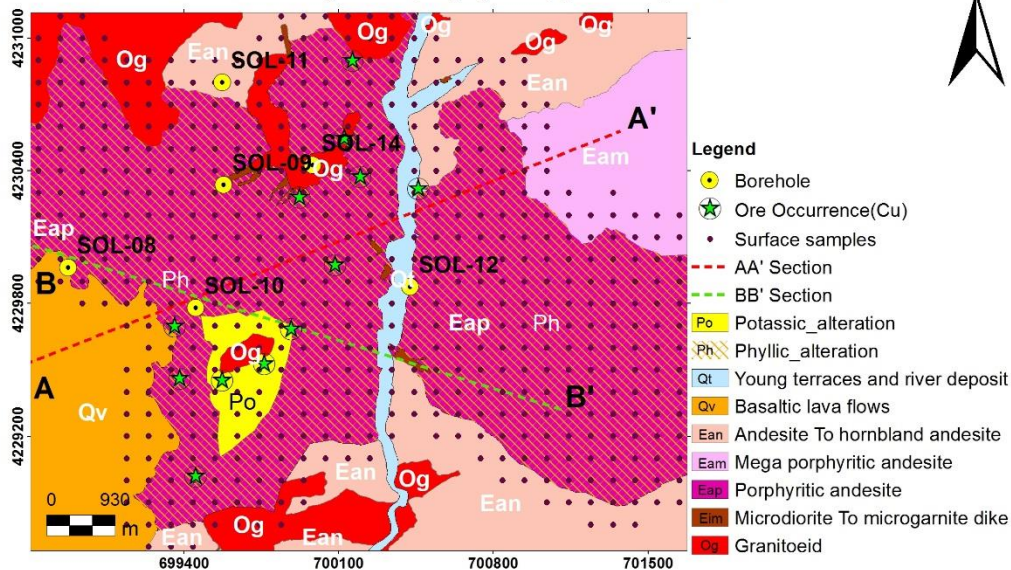
شکل ۱-۷: بافت میکروولیتیک تا اینترگرانولار جریان‌ی در توده ولکانیکی اکوز داغی (XPL, $\times 100$)

۱-۸- داده‌های ژئوشیمیایی استفاده شده

نمونه‌برداری سطحی سنگ در شبکه‌ای مربعی و با فاصله تقریبی ۱۰۰ متر به ۱۰۰ متر به تعداد ۵۶۲ از منطقه برداشت گردید. نمونه‌های مذکور شامل تنوعی از واحدهای سنگ‌شناسی موجود در منطقه می‌باشد. علاوه بر آن ۶

گمانه اکتشافی با ۲۴۶۵ متر مغزه‌گیری نیز حفاری شده است. نمونه‌برداری با فواصل ۲ متری در طول گمانه‌ها انجام شد که شامل ۱۱۹۳ نمونه می‌باشد. در مجموع تعداد ۱۷۵۵ نمونه برای این مطالعه مورد استفاده قرار گرفته است (شکل ۸-۱).

Eastern zone of Sonajil Porphyry Copper Deposit



شکل ۸-۱: موقعیت نمونه‌های سطحی و گمانه‌های حفاری شده در منطقه

نمونه‌های سطحی سنگ برداشت شده برای ۴۵ عنصر مورد آنالیز قرار گرفته‌اند که در (جدول ۱-۱) فهرست این عناصر آورده شده است. در این پروژه آنالیز شیمیایی نمونه‌ها در کشور استرالیا و آزمایشگاه AMDEL صورت‌پذیرفته است. روش اندازه‌گیری همه عنصر به جز طلا روش ICP-OES/MS بوده است. طلا به روش Fire Assay اندازه‌گیری شده است.

جدول ۱-۱: عناصر آنالیز شده برای نمونه‌های سطحی

Ag	Al	As	Au	B	Ba	Be	Bi	Ca	Cd	Ce	Co
Cr	Cs	Cu	Fe	Hg	K	La	Li	Mg	Mn	Mo	Na
Nb	Ni	P	Pb	Rb	Re	S	Sb	Sc	Sn	Sr	Te
Th	Ti	Tl	U	V	W	Y	Zn	Zr			

نمونه‌های برداشت شده از گمانه‌ها برای ۲۳ عنصر مورد آنالیز قرار گرفته‌اند که در جدول ۲-۱ فهرست این عناصر ارائه شده است.

جدول ۲-۱: عناصر آنالیز شده برای نمونه‌های گمانه‌ها

Ag	Al	B	Ba	Be	Ca	Co	Cu	Fe	K	La	Mg
Mn	Mo	Na	Ni	P	Pb	S	Sc	Sr	V	Zn	

نمونه‌های سطحی و گمانه‌ای به صورت جداگانه مورد آنالیز قرار گرفته‌اند و هر کدام در طیف خاصی از عناصر مورد آنالیز واقع شده‌اند. در این مطالعه ما نیازمند عناصری هستیم که تغییرات غلظت آنها هم در سطح و هم در عمق موجود باشد. عناصر مشترک در سطح و عمق، ۲۲ عنصر می‌باشد که برای این مطالعه انتخاب شده‌اند (جدول ۳-۱).

جدول ۳-۱: عناصر مشترک نمونه‌های سنگ آنالیز شده سطحی و عمقی

Ag	Al	Ba	Be	Ca	Co	Cu	Fe	K	La	Mg	Mn
Mo	Na	Ni	P	Pb	S	Sc	Sr	V	Zn		

۱-۹- ساختار پایان نامه

فصل اول، با یک مقدمه شروع و با بیان مسئله و سوال‌های اصلی تحقیق، ضرورت انجام تحقیق و بیان فرضیه‌های آن ادامه و با ذکر هدف مطالعه، روش انجام تحقیق، منطقه مطالعاتی و داده‌های اکتشافی پایان می‌یابد. در فصل دوم روش‌های مورد استفاده شامل الگوریتم‌های خوشه‌بندی و طبقه‌بندی، شاخص‌های اعتبارسنجی، شبکه عصبی و روش جداسازی آنومالی از زمینه فرکتالی عیار - مساحت توضیح داده خواهد شد. مروری بر مطالعات قبلی شامل تخمین عیار و ذخیره در کانسارهای معدنی، بهبود تخمین‌گرها، جدایش آنومالی از زمینه، الگوریتم‌های شناسایی الگو، کاربرد روش‌های پتانسیل‌یابی در مطالعات ژئوشیمیایی و مثال‌هایی کاربردی از روش‌های شناسایی الگو در فصل سوم آورده خواهد شد. در فصل

چهارم تخمین غلظت عناصر با استفاده از شبکه عصبی انجام شده و افزایش دقت تخمین با ترکیب الگوریتم‌های خوشه‌بندی و شبکه عصبی صورت گرفته و بهترین تخمینگر انتخاب خواهد گردید. در فصل پنجم بررسی تغییرات عمقی غلظت عناصر ژئوشیمیایی در سطوح مختلف عمقی و سپس مدل‌سازی تغییرات غلظت کلاس‌های مختلف مس در سطوح مختلف عمقی با ترکیبی از الگوریتم خوشه‌بندی FCM و الگوریتم‌های طبقه‌بندی ماشین‌بردار پشتیبان، K نزدیکترین همسایگی، بیزین و شبکه عصبی انجام و بهترین طبقه‌بندی انتخاب و مدل کانسار در عمق ارائه خواهد شد. در فصل ششم اعتبارسنجی بوسیله اطلاعات حفاری و ژئوفیزیکی انجام شده و نتیجه گیری و پیشنهادات ارائه می‌گردد.

فصل

۲- معرفی روش‌ها

۲-۱- مقدمه

شناسایی الگو یکی از شاخه‌های یادگیری ماشین بوده که در پی کشف ارتباط میان داده‌ها است. در این شاخه بسته به آگاهی از ماهیت داده‌ها از دو رویکرد خوشه‌بندی و طبقه‌بندی استفاده می‌شود. خوشه‌بندی یا آموزش بدون نظارت و طبقه‌بندی یا آموزش با نظارت تلاش می‌کنند تا الگوی بین داده‌های یک مجموعه را کشف کنند.

هرگاه ماهیت داده‌ها برای کاربر ناشناخته باشند و هدف رسیدن به پاسخ این سوال باشد که در این مجموعه داده چند گروه یا دسته داده‌ی وابسته به هم وجود دارد، از خوشه‌بندی استفاده می‌شود. برای مثال چنانچه در ژئوشیمی، یک مجموعه داده در اختیار باشد اما اینکه کدام داده رفتار آنومالی یا زمینه ماده معدنی و یا باطله بوده ممکن است برای کاربر ناشناخته باشد. برای رسیدن به پاسخ این پرسش می‌توان از خوشه‌بندی استفاده نمود و با تقسیم داده‌ها به دو یا چند گروه و سپس تفسیر هر گروه، ماهیت داده‌ها را معلوم نمود.

اگر یک مجموعه داده در اختیار باشد که در آن ماهیت بخشی از داده‌ها معلوم و ماهیت بخش دیگر نامعلوم باشد و هدف پاسخ به این سوال باشد که با استفاده از داده‌های معلوم ماهیت داده‌های مجهول، تعیین شود از طبقه‌بندی استفاده می‌گردد. فرض کنید در یک کانسار عملیات ژئوفیزیک صورت گرفته و چند ویژگی فیزیکی کانسار در تمام فضای کانسار اندازه‌گیری شده است. سپس چند گمانه حفر و در آن‌ها موقعیت بخش‌های پرعیار مشخص شده باشد. در این حالت دو سری داده در اختیار است. در موقعیت گمانه‌ها هم ویژگی‌های فیزیکی در اختیار است و هم موقعیت بخش پرعیار، حال سوال این است که ویژگی‌های فیزیکی با ماهیت معلوم، پرعیار هستند یا خیر؟ در بقیه فضای کانسار فقط ویژگی‌های فیزیکی در اختیار است. در اینجا با استفاده از طبقه‌بندی تلاش می‌شود تا الگوی ارتباطی بین ویژگی‌های فیزیکی و عیار استخراج گردد تا بتوان در تمام فضای کانسار موقعیت بخش‌های پرعیار را تعیین نمود.

بدیهی است، در صورت در اختیار بودن ماهیت بخشی از داده‌ها، طبقه‌بندی بر خوشه‌بندی اولویت دارد اما می‌توان موقعیت‌هایی را تصور نمود که در هر کدام می‌بایست از خوشه‌بندی یا طبقه‌بندی استفاده نمود. در این پژوهش برای تفکیک آنومالی یک و چندگانه از زمینه، زون‌بندی کانسار از منظر دیدگاه‌های مختلف و بهبود تخمین، از خوشه‌بندی استفاده می‌شود. در مقابل از طبقه‌بندی برای طراحی یک نوع مدل‌ساز و تخمین‌گر نیمه‌کمی سریع و با راندمان بالا، درواقع یک زون‌بندی کننده، جهت شناسایی بخش‌های با غلظت‌های مختلف استفاده خواهد شد.

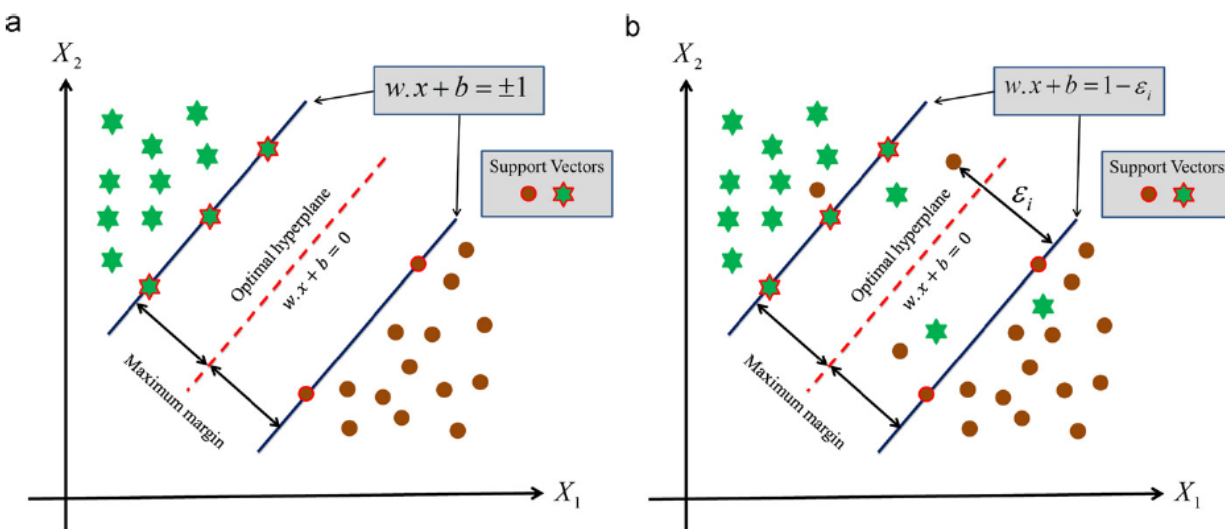
۲-۲- ماشین بردار پشتیبان (SVM)

الگوریتم ماشین بردار پشتیبان اولیه در سال ۱۹۶۳ توسط وپنیک^۹ ابداع شد و در طی سال‌های ۱۹۸۲ تا ۱۹۹۸ توسعه یافت و به علت ویژگی جذاب و کارایی مطلوب مورد استفاده قرار گرفت. (Solomatine and Ostfeld, 2008). این الگوریتم کاربرد چندین ساله‌ای در حل مشکلات طبقه‌بندی، رگرسیون و شناسایی الگوریتم‌های اکتشافی بر اساس تئوری یادگیری آماری دارد (Vapnik, 1995). از ویژگی‌های مهم SVM که یک الگوریتم یادگیری نظارت‌شده است، تعیین پارامترهای بهینه مدل می‌باشد (Bishop, 2006). SVM با استفاده از مجموعه بردارهای پشتیبان با برچسب کلاس‌های شناخته شده بوسیله آبر صفحه‌هایی که کلاس‌های مختلف را از هم جدا می‌کند، طراحی می‌شود. داده‌های بردارها به عنوان یک مجموعه آموزشی استفاده می‌شوند و هر داده بردار با مجموعه‌ای ویژگی‌های منحصر به فرد که بر اساس آن طبقه‌بندی شده، مشخص می‌شود. این روش در تئوری یادگیری ماشین به عنوان یادگیری نظارت شده شناخته می‌شود. در موارد پیچیده‌تر و غیرخطی، وظیفه الگوریتم یافتن ویژگی‌های ترکیبی برای افزایش بعد فضای ویژگی است، به نحوی که امکان طبقه‌بندی خطی داده‌ها با دقت مطلوب‌تر فراهم شود. به این ترتیب داده‌ها در یک فضای چند بعدی، به عنوان فضای ویژگی، طبقه‌بندی خواهند شد. توابع هسته‌ای

9 vapnik

مختلفی برای این انتقال مورد استفاده قرار می‌گیرند. هنگامی که یک طبقه‌بندی کننده بهینه مشخص شود، داده‌های جدید با اطلاعات طبقه‌بندی ناشناخته (داده‌های آزمون) می‌توانند به وسیله SVM آموزش دیده و بر اساس ویژگی‌های خود طبقه‌بندی شوند (Smirnoff et al., 2008).

برای توضیح بیشتر، ابتدا یک مجموعه دو کلاسه را در نظر می‌گیریم. مجموعه‌ای از داده‌های آموزشی متشکل از l بردار ویژگی $x_i \in R^n$ به شرطی که $i (= 1, 2, 3, \dots, n)$ تعداد نمونه‌ها باشد را در نظر بگیرید. داده‌ها در y_i طبقه‌بندی می‌شوند به طوری که یک کلاس برابر ۱ و کلاس دیگر برابر -۱ باشد ($y_i \in \{-1, 1\}$). اگر دو کلاس به صورت خطی از هم جدا شوند، آنگاه باید یک ابر صفحه ۱۰ طراحی شود که در ادامه معادلات آن ارائه می‌شود (Kavzoglu and Colkesen, 2009) (شکل ۱-۲-۲-۱):



شکل ۱-۲-۲: شکل کلی طبقه‌بندی داده‌ها با استفاده از SVM، X_1 و X_2 محور ویژگی می‌باشند. (a) زمانی که داده‌ها به صورت کامل از هم تفکیک شده باشند، (b) زمانی که داده‌ها به صورت کامل از هم تفکیک نشده باشند (Kavzoglu and Colkesen, 2009).

$$w \cdot x_i + b \leq -1 \quad \text{for } y_i = -1 \quad (1-2)$$

$$w \cdot x_i + b \geq +1 \quad \text{for } y_i = +1 \quad (2-2)$$

w : بردار نرمال است، که به ابرصفحه عمود است. x_i : یک بردار حقیقی p بعدی. b : عرض از مبدا.

که معادله (۱-۲) و (۲-۲) به صورت معادله (۳-۲) نیز بیان می‌شود:

$$y_i (w x_i + b) \geq 1, \quad i = 1, 2, \dots, n \quad (3-2)$$

y_i : به عبارتی قید یا برچسب کلاس‌های می‌باشد.

پس از این می‌توان ابر صفحه جداکننده را به عنوان یک تابع تصمیم به رسمیت شناخت،

$$f(x) = \text{sign}(wx + b) \quad (4-2)$$

که در آن SIGN یک تابع نشانه‌ای است که به صورت زیر تعریف شده است:

$$\text{sign}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases} \quad (5-2)$$

پارامترهای ابرصفحه جداکننده (b, w) به عنوان یک تابع تصمیم را می‌توان با حل تابع بهینه‌سازی زیر

بدست آورد:

$$\text{minimize } \tau(x) = \frac{1}{2} \|w\|^2 \quad (6-2)$$

منوط به

$$y_i (w x_i + b) \geq 1, \quad i = 1, \dots, l \quad (7-2)$$

راه‌حل این مشکل بهینه‌سازی:

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^l \alpha_i (y_i ((x_i w) + b) - 1) \quad (۸-۲)$$

$$\frac{\partial}{\partial w} L(w, b, \alpha) = 0 \quad (۹-۲)$$

$$\frac{\partial}{\partial b} L(w, b, \alpha) = 0 \quad (۱۰-۲)$$

$$\frac{\partial}{\partial \alpha} L(w, b, \alpha) = 0 \quad (۱۱-۲)$$

که در آن α_i ضرایب لاگرانژ است. تابع لاگرانژ با توجه به w و b معادله را به حداقل و با توجه به $\alpha_i > 0$ حداکثر می‌رساند. ضرایب لاگرانژ α_i توسط تابع بهینه‌سازی زیر تعیین می‌شوند:

$$\text{maximize } \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{ij=1}^l \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \quad (۱۲-۲)$$

منوط به:

$$\alpha_i \geq 0, \quad i = 1, \dots, l \quad \text{and} \quad \sum_{i=1}^l \alpha_i y_i = 0 \quad (۱۳-۲)$$

روش جداسازی بر اساس ابرصفحه بهینه بوسیله تابع هزینه زیر خواهد بود (Zuo and Carranza, 2011):

$$f(x) = \text{sign} \left(\sum_{i=1}^l y_i \alpha_i (x x_i) + b \right) \quad (۱۴-۲)$$

راه حل فوق را می‌توان به مسائلی که در فضای ویژگی از هم قابل تفکیک هستند اعمال کرد. به منظور بهبود طبقه‌بندی، مقدار جریمه‌ی^{۱۱} C برای خطاهای طبقه‌بندی نادرست و متغیرهای جاافتاده^{۱۲} مثبت ϵ_i مورد استفاده قرار گرفتند (شکل ۱-۲ - b). این متغیرها در قیدهای (۱-۲)، (۲-۲) به شرح زیر گنجانده شده‌اند (Huang et al., 2002).

$$\begin{aligned} wx_i + b &\geq 1 - \epsilon_i & \text{for } & y_i = +1 \\ wx_i + b &\leq 1 + \epsilon_i & \text{for } & y_i = -1 \end{aligned} \quad (۱۵-۲)$$

11 Penalty
12 Slack

$$\varepsilon_i \geq 0, \quad i = 1, 2, \dots, n$$

در نهایت، تابع زیر باید بهینه‌سازی شود:

$$\text{minimize } \tau(x) = \frac{1}{2} \|w\|^2 + C \left(\sum_{i=1}^l \varepsilon_i \right)^k \quad (16-2)$$

منوط به:

$$y_i((wx_i) + b) \geq 1 - \varepsilon_i \quad i = 1, \dots, l \quad (17-2)$$

C می‌تواند توسط فرآیند اعتبارسنجی متقابل تعیین شود. اگر $K=1$ باشد، روش بهینه‌سازی تابع بالا همانند کلاس داده‌های گسسته می‌باشد (Huang et al., 2002).

در مواردی که داده‌ها به صورت خطی از هم جدا نشوند، یک تابع طراحی $\phi(x)$ جهت به نقشه درآوردن داده‌ها از فضای اصلی x به فضای هیلبرت^{۱۳} X استفاده می‌شود. در تابع بهینه‌سازی SVM، اطلاعات ویژگی‌ها در داده‌های آموزشی به صورت ضرب نقطه‌ای به نظر می‌رسند $(x_i x_j)$ (توابع (۲-۱۲) و (۲-۱۳)) (Yang et al., 2008).

بنابراین، تابع کرنل^{۱۴} K به صورت زیر است:

$$K(x_i, x_j) = \langle \phi(x_i) \times \phi(x_j) \rangle \quad (18-2)$$

بنابراین ما فقط نیاز داریم که از K در برنامه آموزشی خود استفاده کنیم، بدون دانستن فرم صحیح و روشن از $\phi(x)$ (Huang et al., 2002). همین رویکرد را می‌توان با یک تابع تصمیم استفاده کرد. بنابراین، فرمول دوگانه این مشکل به شرح زیر است (El-Khoribi, 2008):

13 Hillbert
14 kernel function

$$\text{maximize} \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{ij=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (19-2)$$

منوط به:

$$\alpha_i \geq 0, \quad i=1, \dots, l \quad \text{and} \quad \sum_{i=1}^l \alpha_i y_i = 0 \quad (20-2)$$

لازم به ذکر است که کدنویسی الگوریتم SVM در محیط نرم افزار MATLAB با کمک گیری از ابزارهای موجود برای این الگوریتم در نرم افزار مذکور انجام خواهد گرفت.

۲-۳- طبقه بند نزدیک ترین k همسایگی

یکی از روش های طبقه بندی داده های با بُعد بالا، روش نزدیک ترین k همسایگی (KNN) است. منظور از متغیر K، اشاره به تعداد نقاط داده ای دارد که در همسایگی با نقطه ای قرار دارند که می خواهیم آن را طبقه بندی کنیم. این روش بر مشکلات ناشی از تراکم کم داده ها که از آن ها برای طبقه بندی یک نقطه جدید استفاده می شود غلبه می کند (Hechenbichler and Schliep, 2004).

این طبقه بندی، نمونه آزمون را متعلق به کلاسی می داند که بیشترین آرا را در بین نزدیک ترین k همسایگان آن که نرمالیزه شده به تعداد داده های کلاس معیار طبقه بندی هستند داشته باشد. برای بدست آوردن نزدیک ترین همسایگان یک نمونه، معمولاً از فاصله اقلیدسی طبق رابطه (۲۱-۲) استفاده می شود:

$$d_{\text{eucl}}(x, t) = \sqrt{\sum_{i=1}^m d_{\text{eucl}}^i(x, t)} \quad (21-2)$$

d_{eucl}^i فاصله دو نقطه x و t، اندازه پاره خطی است که آنها را به هم متصل می کند (xt) و موقعیت هر نقطه به طور کلی بصورت $a_i(x)$ و $a_i(t)$ نمایش داده می شود.

اگر مقادیر خصوصیات عددی و پیوسته باشد، d_{eucl}^i از رابطه (۲۲-۲) بدست می آید:

$$d^i_{eucl}(x, t) = \sum (a_i(x) - a_i(t))^2 \quad (22-2)$$

و اگر خصوصیات مقادیر اسمی داشته باشند، رابطه بالا به صورت (23-2) تغییر پیدا خواهد کرد:

$$d^i_{eucl}(x, t) = \begin{cases} 1 & \text{if } a_i(x) \neq a_i(t) \\ 0 & \text{if } a_i(x) = a_i(t) \end{cases} \quad (23-2)$$

طبقه‌بند KNN به دلیل قابلیت درک بالا و عدم نیاز به ایجاد فرضیه روی داده‌ها، روشی ساده و کاربردی

محسوب می‌شود. اما این طبقه‌بند، علاوه بر مزایایی که یاد شد، دارای دو مشکل به شرح زیر می‌باشد:

- اولین اشکالی که به این الگوریتم وارد است، تعیین مقدار K توسط کاربر است (شکل 2-2).

الگوریتم KNN را با $K=10$ در نظر بگیرید. نمونه آزمون با نقطه نشان داده شده است. در این

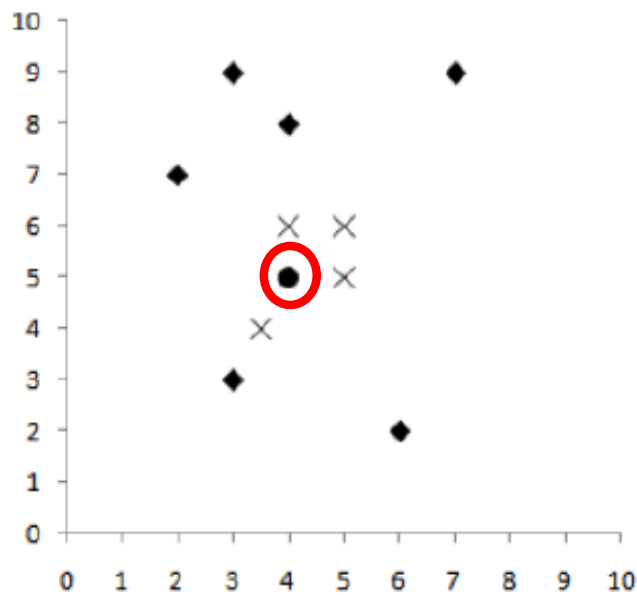
شکل دو کلاس وجود دارد که 4 همسایه متعلق به کلاس اول (با ضربدر نمایش داده شده) و 6

همسایه کلاس دوم، (که با لوزی نشان داده شده) است. با یک رای‌گیری ساده مشخص می‌شود که

این الگوریتم نمونه آزمون را متعلق به کلاس دوم می‌داند اگر چه با توجه به شکل 2-2 می‌توان

فهمید که نمونه آزمون به کلاس یک تعلق دارد. به هر حال دقت طبقه‌بندی با این روش بستگی

زیادی به تعداد همسایگان یعنی مقدار K دارد (Verbiest et al., 2012).



شکل 2-2: مثالی از KNN با $K=10$ (Verbiest et al., 2012)

- اما دومین اشکال این روش عبارت است از این که برای تخصیص برچسب، همه نزدیکترین همسایگان نمونه آزمون را با درجه اهمیت یکسانی در نظر می‌گیرد. برای حل این مشکل کلر^{۱۵} و همکارانش (Keller et al., 1985) از تئوری مجموعه‌های فازی استفاده کردند. آن‌ها بعد از به دست آوردن نزدیکترین K همسایه نمونه آزمون، نمونه آزمون را به کلاسی نسبت می‌دهند که مقدار عبارت رابطه (۲۴-۲) آن بیشترین مقدار باشد.

$$\frac{\sum_{x \in NN} R(x, t) C(x)}{\sum_{x \in NN} R(x, t)} \quad (24-2)$$

t مرکز کلاس C، x نمونه مورد مطالعه، شباهت بین x و t، R(x, t)، از رابطه (۲۵-۲) به دست می‌آید:

$$R(x, t) = \frac{1}{d_{eucl} \frac{2}{m-1}} \quad (25-2)$$

اگر x متعلق به کلاس C باشد مقدار C(x)=1 خواهد بود و در غیر این صورت برابر صفر است. m تعداد نمونه‌ها می‌باشد.

برای مثال در شکل ۲-۲ این روش در استراتژی رای‌گیری به چهار همسایه نمونه آزمون که متعلق به کلاس ۱ هستند وزن بیشتری اختصاص می‌دهد و لذا نمونه آزمون در کلاس ۱ طبقه‌بندی می‌شود.

۲-۴- طبقه‌بندی داده‌ها بر اساس تئوری بیزین

۲-۴-۱- مبانی علمی روش

در صورتی که C کلاس $\omega_1, \omega_2, \dots, \omega_c$ داشته باشیم و بخواهیم نمونه جدید x را به یکی از این کلاس‌ها نسبت دهیم، بایستی احتمالات شرطی $P(\omega_i|x)$ ($i=1, 2, \dots, c$) را محاسبه کنیم. بدیهی است که نمونه جدید x به کلاسی تعلق دارد که دارای بیشترین احتمال شرطی باشد. اما از آنجا که محاسبه

احتمال‌های شرطی به صورت مستقیم میسر نمی‌باشد، از تئوری بیزین به صورت معادله (۲۶-۲) استفاده می‌شود:

$$P(\omega_i | x) = \frac{P(\omega_i | x)P(\omega_i)}{P(x)} \quad (26-2)$$

که $P(x)$ احتمال غیرشرطی و $P(\omega_i)$ احتمال اولیه هر طبقه است. احتمال اولیه هر طبقه، از تقسیم تعداد نمونه‌های هر طبقه بر تعداد کل نمونه‌ها بدست می‌آید. بنابراین نمونه جدید x به طبقه‌ای تعلق خواهد گرفت که $g_i(x) = P(\omega_i|x)P(x)$ بالاتری داشته باشد (Theodoridis and Koutroumbas, 2009).

۲-۴-۲- روش طبقه‌بندی بیزین

در روش طبقه‌بندی بیزین، ابتدا یک تابع توزیع احتمال (فرضاً گوسی، بتا و غیره) به داده‌های هر طبقه برازش می‌شود. سپس به کمک داده‌های آموزشی پارامترهای این تابع چگالی احتمال تخمین می‌شود. برای مثال اگر توزیع داده‌ها نرمال فرض شوند، پارامترهای بردار میانگین و ماتریس کوواریانس داده‌ها به عنوان تابع چگالی گوسی برآورد می‌شود. بنابراین خواهیم داشت:

$$g_i(x) = \frac{1}{(2\pi)^{\frac{m}{2}} \left| \sum_i \frac{1}{2} \right|^{\frac{1}{2}}} \exp \left[\frac{1}{2} (x - \mu_i)^T \sum_i^{-1} (x - \mu_i) \right] \times P(\omega_i) \quad (27-2)$$

$i = 1, 2, \dots, c$

در این رابطه m تعداد متغیرها، μ_i و Σ_i به ترتیب بردار میانگین ($m \times 1$ بعدی) و ماتریس واریانس - کوواریانس ($m \times m$ بعدی) طبقه i ام هستند، که از روابط (۲۸-۲) و (۲۹-۲) به دست می‌آیند:

$$\mu_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ji} \quad (28-2)$$

$$\Sigma_i = \frac{1}{n_i} \sum_{j=1}^{n_i} (x_{ji} - \mu_i)(x_{ji} - \mu_i)^T \quad (29-2)$$

با تخمین پارامترها و استفاده از رابطه (۲-۲۷)، تعلق هر نمونه جدید به هر طبقه محاسبه می‌شود. با توجه به این دیدگاه، روش بیزین جزء روش‌های طبقه‌بندی پارامتری قرار می‌گیرد (Theodoridis and Koutroumbas, 2009 , Duda et al., 2002).

۲-۵ - خوشه بندی (FCM) Fuzzy C-Means

یکی از مهمترین و کاربردیترین الگوریتم‌های خوشه بندی، الگوریتم C میانگین می باشد. در این الگوریتم نمونه‌ها به C خوشه تقسیم می‌شوند و تعداد C مفروض فرض می‌شود. در نسخه فازی این الگوریتم نیز تعداد خوشه‌ها (c) مفروض است. الگوریتم FCM برای پارامترهای y_{ji} ($i = 1, \dots, n; j = 1, \dots, c$) حلی به کمک رابطه (۲-۳۰) دارد (تابع هدف) (Webb, 2002):

$$J_r = \sum_{i=1}^n \sum_{j=1}^c y_{ji}^r |x_i - m_j|^2 \quad (۲-۳۰)$$

که تحت شرایط زیر کمینه می‌شود (از روی y_{ji} می‌توان یک ماتریس y تعریف کرد که دارای n سطر و c ستون می‌باشد و مولفه‌های آن هر مقداری بین ۰ تا ۱ را می‌توانند اختیار کنند. اگر تمامی مولفه‌های ماتریس y بصورت ۰ و ۱ باشند الگوریتم مشابه C میانگین کلاسیک خواهد بود. با اینکه مولفه‌های ماتریس y می‌توانند هر مقداری بین ۰ تا ۱ را اختیار کنند اما مجموع مولفه‌های هر یک از ستون‌ها باید برابر ۱ باشد) (Webb, 2002):

$$\sum_{j=1}^c y_{ji} = 1 \quad 1 \leq i \leq n \quad (۲-۳۱)$$

$$y_{ji} \geq 0 \quad i = 1, \dots, n; \quad j = 1, \dots, g$$

معنای شرط این است که مجموع تعلق هر نمونه به C خوشه باید برابر ۱ باشد. پارامتر y_{ji} درجه تابع عضویت آامین شیء در آامین خوشه را نشان می‌دهد. پارامتر r یک عدد حقیقی است که فازی بودن

را کنترل می‌کند ($r \geq 1$). اگر $r=1$ قرار دهیم تابع هدف خوشه‌بندی c میانگین غیرفازی به دست می‌آید. پارامتر x_i نمونه i ام و پارامتر m_j مرکز j امین خوشه است و n تعداد نمونه‌ها (اشیاء) می‌باشد.

برای بدست آوردن فرمول‌های مربوط به y_{ji} و m_j باید تابع هدف تعریف شده را کمینه کنیم. با

استفاده از شرط فوق و برابر صفر قرار دادن مشتق تابع هدف خواهیم داشت: (Webb, 2002)

$$y_{ji} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{ik}} \right)^{\frac{2}{r-1}}} \quad (32-2)$$

$$m_j = \frac{\sum_{i=1}^n y_{ji}^r x_i}{\sum_{i=1}^n y_{ji}^r} \quad (33-2)$$

با استفاده از دو رابطه (32-2) و (33-2) محاسبه شده، الگوریتم خوشه‌بندی c میانگین فازی

بصورت زیر می‌باشد:

مراحل انجام الگوریتم:

الف) انتخاب r ($1 < r < \infty$) مقداردهی اولیه برای $(i = 1, \dots, n; j = 1, \dots, c)$

ب) مراکز خوشه‌ها محاسبه شوند (محاسبه m_j ها).

ج) محاسبه d_{ij} ، $i = 1, \dots, n; j = 1, \dots, c$ که $d_{ij} = |x_i - m_j|$.

د) محاسبه تابع عضویت: اگر $d_{il} = 0$ باشد برای مقادیری از l آنگاه $y_{li} = 1$ است و $y_{ji} = 0$

است برای همه $j \neq l$ ، در غیر این صورت با توجه به معادله (34-2) حل می‌شود.

ه) اگر همگرا نبود، به مرحله ب برگرد.

یکی از شرط‌های توقف الگوریتم زمانی است که تغییرات نسبی در مقادیر مرکز کوچک می‌شود

:(Webb, 2002)

$$D_z = \left\{ \sum_{j=1}^c |m_j(k) - m_j(k-1)|^2 \right\}^{\frac{1}{2}} < \varepsilon \quad (34-2)$$

که $m_j(k)$ مقدار ژامین مرکز روی k امین تکرار و ε حد مشخص شده است. شرط‌های توقف دیگر بر اساس مقادیر تابع عضویت، y_{ji} یا تابع هزینه، J_r می‌باشد. شرط دیگر بر اساس بهینه‌سازی محلی تابع هزینه که توسط سلیم و اسماعیل^{۱۶} (۱۹۸۶) مطرح شده‌اند، است. زمانی توقف داریم که:

$$\max_{1 \leq i \leq n} \alpha_i < \varepsilon \quad (35-2)$$

که

$$\alpha_i = \max_{1 \leq j \leq g} y_{ji}^{r-1} |x_i - m_j|^2 - \min_{1 \leq j \leq g} y_{ji}^{r-1} |x_i - m_j|^2 \quad (36-2)$$

۲-۶- الگوریتم گوستافسون کسل^{۱۷}

الگوریتم گوستافسون کسل به عنوان یکی از روش‌های خوشه‌بندی نرم، بر خلاف روش‌های همچون میانگین k ، که صرفاً توانایی تشخیص خوشه‌های کروی^{۱۸} را دارند، می‌تواند خوشه‌های خطی^{۱۹} و بیضوی^{۲۰} را شناسایی کند. در الگوریتم گوستافسون کسل هر خوشه براساس یک نقطه و یک ماتریس ایجاد می‌شود که در آن، نقطه بیانگر مرکز خوشه و ماتریس بیانگر کواریانس آن می‌باشد. بر این اساس الگوریتم گوستافسون کسل قابلیت تفکیک خوشه‌های بیضوی را نیز دارد و برخلاف اغلب روش‌های خوشه‌بندی محدودیتی در قبال ایجاد خوشه‌های دایره‌ای^{۲۱} ندارد (Lesot and Kruse, 2008).

16 Selim and Ismail

17 Gustafson kessel algorithm (GK)

18 Spherical clusters

19 Linear Clusters

20 Elliptical Clusters

21 Circular clusters

بسیاری از الگوریتم‌های خوشه‌بندی فازی مبتنی بر نمونه، مانند FCM، مبتنی بر بهینه‌سازی یک طرح هستند و می‌خواهند تابع مناسبی مانند J را که نشان‌دهنده خطای اتصالات خوشه‌ای داده‌ها می‌باشد را به حداقل برسانند (Gustafson et al., 1987).

$$J(V, U) = \sum_{i=1}^c \sum_{k=1}^N (u_{ik})^\beta d_{ik}^2 \quad (37-2)$$

که در آن:

u_{ik} درجه عضویت داده نقطه‌ای x_k به مدل الگوی i ام (مرکز خوشه)، $U=[u_{ij}]$ ماتریس تقسیم‌بندی با ابعاد $c \times N$ ، $V=[v_i]$ ، $c \times q$ ماتریس $c \times q$ نمونه‌های اولیه و d_{ik} فاصله بین داده نقطه‌ای K ام و نمونه i ام است.

پارامتر $\beta > 1$ وزن نمایی است که فازی‌شدن تقسیم‌بندی را کنترل می‌کند (تعیین‌کننده این است که خوشه‌ها چه مقدار باهم تداخل دارند).

در حالی که الگوریتم FCM تنها قادر به تشخیص خوشه‌های کروی است، الگوریتم گوستافسون-کسل می‌تواند خوشه‌های بیضوی را نیز شناسایی کند. الگوریتم گوستافسون-کسل، گسترش یافته الگوریتم استاندارد FCM با استفاده از نرم فاصله تطبیقی به منظور تشخیص خوشه‌ها از اشکال مختلف هندسی در مجموعه داده می‌باشد.

فاصله تطبیقی برای هر خوشه مجزا مطابق رابطه (38-2) بدست می‌آید و $S_i, i=1 \dots c$ (ماتریس نرم القایی هر خوشه) از رابطه (39-2) قابل محاسبه است.

$$d_{ik}^2 = \|x_k - v_i\|_{S_i}^2 = (x_k - v_i)^T S_i (x_k - v_i) \quad (38-2)$$

$$S_i = [\rho_i \det(F_i)]^{1/q} F_i^{-1} \quad (39-2)$$

که در آن q تعداد ویژگی داده‌های اولیه، ρ_i حجم خوشه نام است و F_i ماتریس کواریانس فازی محاسبه شده مطابق رابطه (۴۰-۲) است:

$$F_i = \frac{\sum_{k=1}^N (u_{ik})^\beta (x_k - v_i)(x_k - v_i)^T}{\sum_k (u_{ik})^\beta} \quad (40-2)$$

به حداقل رساندن تابع هدف $J(V, U)$ تحت محدودیت $\sum_{i=1}^c u_{ik} = 1$ با استفاده از یک الگوریتم تکرار شونده که مراکز خوشه و درجه عضویت را بهینه‌سازی می‌کند، طبق روابط (۴۱-۲) و (۴۲-۲) انجام می‌شود (Serir et al., 2012):

$$v_i = \frac{\sum_{k=1}^N (u_{ik})^\beta x_k}{\sum_k (u_{ik})^\beta}, i = 1 \dots c, k = 1 \dots N \quad (41-2)$$

و

$$u_{ik} = \frac{1}{\sum_{j=1}^c (d_{ik} / d_{jk})^{2/\beta-1}}, i = 1 \dots c, k = 1 \dots N \quad (42-2)$$

۷-۲ - الگوریتم گت‌جوا^{۲۲} (GG)

الگوریتم خوشه‌بندی فازی گت‌جوا (GG) یک فرمت از الگوریتم خوشه‌بندی فازی گوستافسون کسل (GK) است و از اندازه و تراکم خوشه‌ها برای کلاس‌بندی استفاده می‌کند (Hoppner et al., 1999). از این‌رو، بهترین رفتار را در مقابل ویژگی‌های نامنظم از خود نشان می‌دهد. تفسیر احتمالی خوشه‌بندی GG در معادله (۴۳-۲) نشان داده شده است:

$$P(X|\eta) = \sum_{i=1}^c P(X, \eta_i) = \sum_{i=1}^c P(\eta_i)P(X|\eta_i) \quad (43-2)$$

22 Gath-Geva (GG)

به طوریکه:

$$P(\underline{x}_j | \eta_i) = \frac{p_i}{(2\pi)^{\frac{p}{2}} \sqrt{|\Sigma_i|}} \exp \left[-\frac{1}{2} (\underline{x}_j - \underline{a}_i)' \Sigma_i^{-1} (\underline{x}_j - \underline{a}_i) \right] \quad (44-2)$$

ماتریس داده به صورت معادله (۴۵-۲) است:

$$X = [\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n], \underline{x}_j \in \mathbb{R}^p, j = 1, 2, \dots, n \quad (45-2)$$

ماتریس کواریانس خوشه i ، $\Sigma_i \in \mathbb{R}^{p \times p}$ می‌باشد که P تعداد ابعاد داده است، C تعداد خوشه‌ها است، $d^2 = (\underline{x}_j, \underline{a}_i)$ برای الگوریتم GG به طور غیر مستقیم متناسب با معادله (۳۵-۲) انتخاب می‌شود که احتمال خلفی^{۲۳} (حد آستانه‌ای^{۲۴}) تابع است.

فاصله کوچک به معنی احتمال بالا و فاصله بزرگ به معنی احتمال کم برای عضویت است. الگوریتم GG در به حداقل رساندن مجموع وزن مربع فاصله‌ی بین داده‌ها و مراکز خوشه تابع هدف بر اساس معادله (۴۶-۲) استوار است.

$$J_{GG}^m(V, U) = \sum_{i=1}^c \sum_{j=1}^n \mu_{ij}^\beta d^2(\underline{x}_j, \underline{a}_i) \quad (46-2)$$

شرایط لازم برای احتمال تقسیم‌بندی خوشه‌ها عبارتند از:

$$\beta \in [1, \infty); U = [\mu_{ij}]_{c \times n}; \mu_{ij} \in [0, 1], i = 1, 2, \dots, c, j = 1, 2, \dots, n \quad (47-2)$$

$$\sum_{i=1}^c \mu_{ij} = 1, j = 1, 2, \dots, n, 0 < \sum_{j=1}^n \mu_{ij} < n, i = 1, 2, \dots, c \quad (48-2)$$

23 Posterior probability

24 Likelihood

$$d^2(\underline{x}_j, a_i) = \frac{1}{P(\underline{x}_j | \eta_i)} = \frac{(2\pi)^{\frac{p}{2}} \sqrt{|\sum_i^{-1}|}}{p_i} \exp \left[\frac{1}{2} (\underline{x}_j - a_i)' \sum_i^{-1} (\underline{x}_j - a_i) \right] \quad (49-2)$$

۲-۷-۱- شاخص‌های اعتبارسنجی الگوریتم‌های خوشه‌بندی فازی

تعیین تعداد خوشه بهینه باید تا حد ممکن قابل تفسیر باشد. به منظور بررسی نتایج خوشه‌بندی از شاخص‌های اعتبارسنجی استفاده می‌شود که تعیین کننده فشردگی یا تراکم (اعضای هر خوشه باید تا جایی که ممکن است نزدیک به هم قرار بگیرند) و جدایش (خود خوشه‌ها باید به صورت گسترده از هم جدا باشند) نتایج خوشه‌بندی باشند. گروه اول (فشردگی) فقط از عضویت‌های خوشه‌بندی استفاده می‌کند در حالی که گروه دوم (جدایش) از عضویت‌های در ارتباط با خود داده‌ها استفاده می‌کند. در گروه اول، اغلب از ضریب تقسیم‌بندی^{۲۵} (PC) (Bezdek, 1981) و آنتروپی طبقه‌بندی^{۲۶} (CE) استفاده می‌شود. در گروه دوم، اغلب اعتبارسنجی شاخص زی - بنی^{۲۷} (XB) (Xie and Beni, 1991)، شاخص تقسیم‌بندی^{۲۸} (SC) (Bensaid et al., 1996) و شاخص جدایی^{۲۹} (S) (Bensaid et al., 1996) به کار بسته می‌شوند. در ادامه این پارامترها معرفی می‌شوند:

۲-۷-۱-۱- شاخص PC

شاخص PC (۲-۵۰) که توسط Bezdek (1981) ارائه شده است، "هم‌پوشانی ۳۰" بین خوشه‌ها را اندازه‌گیری می‌کند که مطابق رابطه (۲-۵۰) تعریف شده است:

$$PC = \frac{1}{N} \sum_{i=1}^m \sum_{k=1}^N \mu_{ik}^2 \quad (50-2)$$

که در آن N تعداد داده‌ها و μ_{ik} مقدار عضویت داده نقطه‌ای K در خوشه iام می‌باشد.

25 Partition co-efficient
 26 Classification entropy
 27 Xie-Beni
 28 Partition index
 29 Separation index
 30 Overlapping

۲-۷-۱-۲- شاخص CE

شاخص CE (۵۱-۲) مقدار فازی بودن^{۳۱} تقسیم‌بندی خوشه را اندازه‌گیری می‌کند که مطابق رابطه (۵۱-۲) تعریف می‌شود:

$$CE = \frac{1}{N} \sum_{i=1}^m \sum_{k=1}^N \mu_{ik} \log_a \mu_{ik} \quad (51-2)$$

هنگامی که خوشه‌های مختلف ارزیابی می‌شوند، نزدیک‌تر شدن شاخص PC به یک و شاخص CE به صفر، به عنوان بهترین خوشه در نظر گرفته می‌شوند. شاخص‌های PC و CE حساس به پارامتر β می‌باشند.

با نزدیک شدن شاخص PC به مقدار $1/\beta$ ، نتایج بدست آمده مبهم‌تر می‌شود. شاخص PC نقطه ضعیفی دارد و آن این است که مقدار آن با توجه به β کاهش می‌یابد و هیچ ارتباط مستقیمی با داده‌ها ندارد. با این حال، شاخص PC نمایش می‌دهد که خوشه‌ها تا چه مقدار با هم همپوشانی دارند. شاخص CE هم دارای معایب مشابه است. همانطور که تعداد خوشه افزایش می‌یابد، مقدار شاخص PC کاهش می‌یابد، در حالی که ارزش شاخص CE افزایش می‌یابد. به طور خلاصه، هر دو شاخص تخمین PC و CE، تا حدودی مشخص می‌کنند که تا چه حد خوشه‌ها باهم همپوشانی دارند.

۲-۷-۱-۳- شاخص زی - بنی XB

شاخص زی - بنی با هدف تعیین نسبت مقدار کل تغییرات درون خوشه‌ای و جدایی خوشه‌ها، بر اساس رابطه (۵۲-۲) تعریف می‌شود:

$$XB = \frac{\sum_{i=1}^m \sum_{k=1}^N \mu_{ik}^\beta \|x_k - u_i\|^2}{N \cdot \min_{i,k} \|x_k - u_i\|^2} \quad (52-2)$$

x_k معرف نمونه مورد بررسی و u_i معرف مرکز خوشه می‌باشد.

شاخص زی - بنی در خواص تراکم و جدایش متمرکز است. هرچه خوشه‌ها بیشتر از همدیگر تفکیک شوند، مقدار شاخص زی و بنی مقدار کمتری خواهد شد.

۲-۷-۱-۴- شاخص SC

شاخص SC نسبت مجموع تراکم به جدایش خوشه‌ها از یکدیگر است و مطابق معادله (۵۳-۲) تعریف می‌شود:

$$SC = \sum_{i=1}^m \frac{\sum_{j=1}^N \mu_{ij}^\beta \|x_j - u_i\|^2}{N_i \cdot \sum_{k=1}^m \|u_k - u_i\|^2} \quad (53-2)$$

هر چه مقدار شاخص SC کاهش یابد، بهترین خروجی خوشه‌بندی را خواهیم داشت. این مقدار به عنوان اعتبارسنجی زمانی مفید خواهد بود که خوشه‌بندی‌های مختلف، که دارای تعداد خوشه‌های یکسانی هستند باهم مقایسه شوند (Balasko et al., 2003).

۲-۷-۱-۵- شاخص جدایش S

شاخص جدایش S که در رابطه (۵۴-۲) تعریف شده است، از یک مقدار فاصله جدایش کمینه برای اعتبار تقسیم‌بندی استفاده می‌کند.

$$S = \frac{\sum_{i=1}^m \sum_{k=1}^N \mu_{ik}^2 \|x_k - u_i\|^2}{N \cdot \min_{i,k} \|u_i - u_k\|^2} \quad (54-2)$$

شاخص جدایش S، نشان دهنده یک تقسیم‌بندی بهینه معتبر است. بنابراین، شاخص‌های اعتبارسنجی، زمانی که پیش‌بینی تعداد خوشه‌ها ناشناخته است، مفید خواهند بود. مقادیر بیشتر شاخص S نشان دهنده جدایش بهتر خوشه‌ها از یکدیگر می‌باشد.

۲-۸- روش خوشه بندی K-means

این الگوریتم پارامتر K را به عنوان ورودی گرفته و مجموعه n شیء را به k خوشه افراز می‌کند. به طوری که سطح شباهت داخلی خوشه‌ها بالا بوده و سطح شباهت اشیاء بیرون خوشه‌ها پایین باشد. شباهت هر خوشه نسبت به متوسط اشیاء آن خوشه سنجیده شده که این متوسط مرکز خوشه نامیده می‌شود. این الگوریتم به صورت زیر کار می‌کند:

ورودی: K، تعداد خوشه‌ها است. پایگاه داده X شامل n شیء،

$$X = \{x_1, x_2, \dots, x_n\}, x_i \in R^m$$

خروجی: یک مجموعه از K خوشه که معیار مربع خطا را حداقل می‌کند.

الگوریتم:

الف) به صورت تصادفی k داده دلخواه را به عنوان مراکز خوشه‌های ابتدایی انتخاب می‌کنیم
 ب) هر داده را با توجه به نزدیکی آن به مراکز خوشه‌ها، به خوشه‌ها تخصیص می‌دهیم. معیار نزدیکی را فاصله اقلیدوسی (رابطه (۲-۵۵)) در نظر می‌گیریم:

$$d(x_i, m_j) = \|x_i - m_j\| = \sqrt{\sum_{k=1}^m (x_{ik} - m_{jk})^2} \quad (۲-۵۵)$$

ج) که در آن x_i ، i امین داده ورودی و m_j مرکز (میانگین) خوشه j ام است. داده فوق در خوشه ای قرار می‌گیرد که کمترین فاصله را با مرکز آن خوشه داشته باشد. مرکز خوشه نیز میانگین حسابی داده‌های آن خوشه می‌باشد.

د) مراکز خوشه‌ها را به روز می‌کنیم یعنی برای هر خوشه میانگین اعضای خوشه را به دست می‌آوریم.

ه) با توجه به مراکز جدید خوشه‌ها به مرحله ب برمی‌گردیم و فرآیند فوق را تا جایی ادامه می‌دهیم که هیچ تغییری در خوشه‌ها رخ ندهد (در این حالت الگوریتم پایان یافته است) (قنبری و همکاران ، ۱۳۸۹).

در واقع هدف کمینه کردن رابطه (۲-۵۶) که نشان دهنده فاصله درون خوشه‌ای است می‌باشد، برای مراکز خوشه‌ها $\{\mu_j, j = 1, \dots, k\}$ داریم (Webb, 2002):

$$\sum_{j=1}^k S_j \quad (۲-۵۶)$$

و) که مجموع مربعات برای خوشه j به صورت رابطه (۲-۵۷) است (Webb, 2002):

$$S_j = \sum_{i=1}^n z_{ji} |x_i - \mu_j|^2 \quad (۲-۵۷)$$

ز) که در آن $z_{ji} = 1$ است، اگر در گروه j باشد و در غیر این صورت، برابر صفر است. پارامتر μ_j برابر میانگین گروه j طبق رابطه (۲-۵۸) است (Webb, 2002).

$$\mu_j = \frac{1}{n_j} \sum_{i=1}^n z_{ji} x_i \quad (۲-۵۸)$$

ح) از مزایای این الگوریتم به سادگی آن می‌توان اشاره کرد، همچنین این الگوریتم همگرایی سریعی در رسیدن به جواب نهایی دارد و از معایب آن نیز می‌توان به این اشاره کرد که چون

در ابتدا تعداد خوشه ها به الگوریتم داده می شود بهینه بودن جواب به دست آمده را نمی توان تضمین کرد، زیرا تعداد خوشه ها به صورت دستی است و اشتباه در تعیین این عدد در جواب نهایی تاثیر زیادی دارد. همچنین کارایی الگوریتم بسیار وابسته به نحوه تعیین مراکز اولیه خوشه هاست (قنبری و همکاران، ۱۳۸۹).

۲-۸-۱- شاخص های اعتبارسنجی الگوریتم خوشه بندی K-means

۲-۸-۱-۱- شاخص دان (DI)

دان، شاخص دان^{۳۲} (DI) را در سال ۱۹۷۴ بعنوان شاخص اعتبارسنجی برای خوشه بندی های سخت^{۳۳} ارائه کرد. در این شاخص مجموعه داده با n داده و با X ویژگی $X = \{x_j; j=1, \dots, n\}$ درون $K = (C_1, C_2, \dots, C_k)$ خوشه دسته بندی شده اند. هر خوشه دارای مرکز v_i برابر با $V_i = \{i=1, 2, \dots, k\}$ می باشد. شاخص DI (۲-۵۹) بر اساس شناسایی خوشه های متراکم و مجزا از هم طبق معادله (۲-۵۹) تعریف شده است. بنابراین، نتیجه خوشه بندی به عنوان یک الگوریتم خوشه بندی سخت محاسبه می شود.

$$DI(c) = \min_{i \in c} \left[\min_{j \in c, i \neq j} \left\{ \frac{\min_{X \in c_i, Y \in c_j} d(x, y)}{\max_{k \in c} \left\{ \max_{x, y \in c} d(x, y) \right\}} \right\} \right] \quad (۲-۵۹)$$

که در آن d تابع فاصله است و C_i مجموعه عناصری است که به خوشه i اختصاص داده است. مقدار شاخص دان باید کمینه شود.

^{۳۲} Dunn's index
^{۳۳} Crisp clustering

۲-۸-۱-۲- شاخص دان جایگزین (ADI)

هدف از اصلاح شاخص اصلی دان جایگزین^{۳۴} (ADI) (معادله (۶۰-۲)) این بود زمانی که تابع عدم تشابه بین دو خوشه، $(\min_{x \in C_i, y \in C_j} d(x, y))$ بر اساس مقادیری تحت اثر مثلث نابرابر زیر رتبه‌بندی شده‌اند، محاسبات ساده‌تر شوند.

$$d(x, y) \geq |d(y, v_j) - d(x, v_j)| \quad (۶۰-۲)$$

که v_j مرکز خوشه J است.

$$ADI(c) = \min_{i \in c, i \neq j} \left\{ \frac{\min_{x_i \in c_i, x_j \in c_j} |d(y, v_j) - d(x_i, v_j)|}{\max_{k \in c} \{ \max_{x, y \in c} d(x, y) \}} \right\} \quad (۶۱-۲)$$

در حالت همپوشانی خوشه‌ها، به این دلیل که نتایج با روش تقسیم‌بندی سخت خوشه‌بندی می‌شوند نه فازی، مقادیر DI و ADI واقعا مناسب نمی‌باشند. کمترین مقادیر DI و ADI نشان دهنده بهترین تعداد خوشه بهینه است (Dunn, 1974).

۲-۹- شبکه عصبی مصنوعی

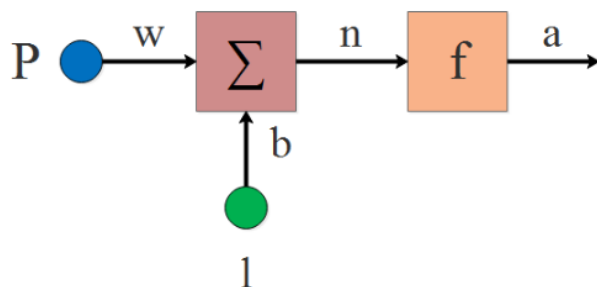
شبکه‌های عصبی مصنوعی با الهام از عملکرد نرون‌های عصبی عملیات پردازش داده را انجام داده و ابزاری قدرتمند جهت مدلسازی می‌باشند. مهمترین خصلت شبکه‌های عصبی یادگیری بر مبنای مثال‌های آموزشی می‌باشد. اگر فرآیند یادگیری شبکه‌ی عصبی به خوبی صورت پذیرد، این ویژگی بسیار مهم به ایجاد یک مدل محاسباتی بسیار کاربردی در حوزه‌های مختلف منجر خواهد شد بویژه جایی که فهم درستی از طبیعت مساله وجود نداشته باشد و یا اینکه رابطه‌ی مشخصی برای حل آنها پیشنهاد نشده

^{۳۴} Alternative Dunn's index

باشد. شبکه‌های عصبی مصنوعی نیازمند دستورات صریح و مشخصی نمی‌باشند و همانند مغز انسان با کسب تجربه، نتایج حاصله را تعمیم خواهند داد (Altrock, 1995).

شبکه‌های عصبی مصنوعی ساختاری لایه‌ای دارند و شامل لایه‌ی ورودی، مخفی و خروجی می‌باشند. داده‌ها در لایه‌ی ورودی قرار می‌گیرند و پس از پردازش خروجی نهایی بدست خواهد آمد. در شبکه‌های عصبی، لایه‌های ورودی کار خاصی انجام نمی‌دهد و صرفاً ورودی‌ها یا داده‌ها را در قسمت خروجی خود نشان می‌دهد. به جز لایه‌ی ورودی تمام لایه‌های دیگر در کار پردازش شرکت کرده و خروجی نهایی را تولید خواهند کرد. لایه‌های مابین لایه‌ی ورودی و لایه‌ی خروجی را لایه‌ی پنهان (داخلی یا میانی) می‌گویند (Altrock, 1995).

هر شبکه عصبی شامل پنج قسمت متعدد ماتریس‌های ورودی (P)، ماتریس اوزان (W)، تابع ترکیب (Σ)، تابع تحریک (f) و خروجی (a) می‌باشد (منهاج، ۱۳۸۷) (شکل ۲-۳).



شکل ۲-۳: بخش‌های مختلف شبکه عصبی

در شکل فوق کمیت‌های P و a به ترتیب ورودی و خروجی شبکه می‌باشد. میزان تاثیر ورودی P روی خروجی a به وسیله پارامتر وزن (w) تعیین می‌شود. ورودی دیگر یک مقدار ثابت ۱ است که در جمله بایاس^{۳۵} ضرب شده و سپس با wP جمع می‌شود. این حاصل جمع، ورودی خالص n برای تابع محرک f را تشکیل می‌دهد.

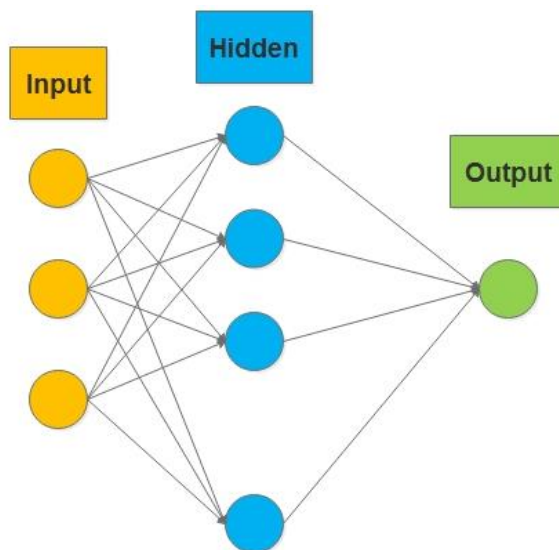
۳۵ - Bias

$$n = wp + b \quad (۶۲-۲)$$

سپس خروجی نرون به کمک معادله زیر به دست می‌آید:

$$a = f(wp + b) \quad (۶۳-۲)$$

در شبکه عصبی چند لایه پرسپترون محدوده تصمیم‌گیری حائز اهمیت بوده و با وجود اینکه شبکه عصبی پرسپترون چند لایه در بخشی یادگیری بسیار کند عمل کرده، ولی در صورت تمایز کافی ویژگی‌ها، بخشی یادگیری نیز سریع تر عمل خواهد نمود. یکی از مزایای شبکه عصبی چند لایه پرسپترون (شکل ۴-۲) نسبت به دسته‌کننده‌های آماری آن است که چندان متأثر از همبستگی ویژگی‌های ورودی نمی‌باشد. شبکه MLP به دنبال دسته‌بندی دو کلاس بوده و به دلیل تمایز مناسب ویژگی‌های استخراج شده با سرعت مناسبی در بخش آموزش عمل می‌کند و از یک سری لایه‌هایی تشکیل شده که نرون‌های آن به صورت موازی با هم عمل می‌کنند. هر نرون دارای P ورودی بوده که هر ورودی با انتخاب مناسب W وزن دار شده و جمع ورودی‌های وزن‌دار با بایاس ورودی تابع تحریک F را تشکیل می‌دهند. شبکه عصبی MLP از نوع شبکه‌های پیشخور بوده که از پر کاربردترین شبکه‌های عصبی می‌باشد. یکی دیگر از مزایای شبکه عصبی چند لایه پرسپترون نسبت به دیگر روش‌های آماری آن است که چندان متأثر از همبستگی ویژگی‌های ورودی نبوده است (کیا، ۱۳۸۷).



شکل ۲-۴: ساختار شبکه عصبی MLP

۲-۱۰- روش فرکتالی عیار - مساحت

هندسه فرکتالی روش جدیدی است که نسبت به هندسه اقلیدسی تفاوت زیادی دارد. در گذشته برای توصیف اشکال و ساختارهای منظم در طبیعت از هندسه اقلیدسی و روابط بین این ساختارها به عنوان ابزار قدرتمند ریاضی استفاده شده است. اما برای توصیف ساختارهای پیچیده‌تر مانند توزیع الگوهای ژئوشیمیایی، هندسه اقلیدسی نمی‌تواند مدلی را ارائه دهد. برای بیان این پیچیدگی‌ها الگوهای فرکتالی می‌توانند مناسب باشند. تئوری فرکتال اولین بار توسط ماندلبروت در سال ۱۹۸۳ مطرح شد و تاکنون به طور گسترده‌ای در علوم زمین به کار گرفته شده است. پژوهش‌های انجام شده در داده‌های ژئوشیمیایی ویژگی خودتشابهی را در این داده‌ها نشان می‌دهد (Zou and Wang, 2016). به طور کلی داده‌های ژئوشیمیایی رفتار مولتی فرکتالی دارند که این نشانگر میزان تغییرات در شرایط زمین‌شناسی، ژئوشیمیایی، دگرسانی، کانی‌سازی و غنی‌شدگی یک عنصر است.

روش‌های مبتنی بر هندسه فرکتال توزیع سطحی و فضایی داده‌ها و نیز شکل هندسی آنومالی‌ها را در نظر می‌گیرد. از این‌رو روش فرکتال و مولتی فرکتال مورد توجه قرار گرفته و برای تشخیص آنومالی‌ها به کار گرفته شده‌اند. چنگ در سال ۱۹۹۴ مدل فرکتال غلظت-مساحت را برای تفکیک زمینه از آنومالی ارائه

کرد. روش فرکتال عیار- مساحت مبتنی بر میزان مساحتی است که هر عیار خاص در منطقه مورد مطالعه اشغال کرده است. هر چه عیار عنصر افزایش یابد، میزان مساحت اشغال شده به وسیله آن کاهش می‌یابد. یکی از رایج‌ترین روش‌ها برای نمایش توزیع عیار یک عنصر در یک منطقه ترسیم نقشه کننتوری (منحنی- میزان) هم عیار عنصر مربوطه در منطقه مورد مطالعه است. اگر مقدار هر کننتور عیاری برابر با ρ در نظر گرفته شود، می‌توان یک معادله توانی به صورت زیر برای تمرکز مواد با خواص فرکتال ارائه نمود (Cheng *et al*, 1994)

$$A_{(>\rho)} \propto \rho^{-D} \quad (۶۴-۲)$$

مقدار D در حقیقت نمایانگر بعد فرکتالی مربوط به دامنه‌های متفاوت ρ را نشان می‌دهد. با ترسیم تغییرات مساحت در برابر عیار در نموداری لگاریتمی می‌توان بعد هر جامعه را از طریق شیب خط برازش شده به آن حساب نمود. نقاط شکست (تغییر شیب خط برازش شده) در این نمودار بیانگر گذر از جامعه‌ای به جامعه دیگر است، به گونه‌ای که علاوه بر جدا کردن زمینه می‌توان آنومالی‌های ممکن، احتمالی و قطعی یک عنصر و حتی در برخی موارد کانی‌سازی‌های اصلی و فرعی مربوط به آن عنصر را از یکدیگر جدا نمود (Li *et al.*, 2003).

پژوهش‌های زیادی در زمینه تشخیص آنومالی با روش فرکتال انجام شده است و روش‌های مختلفی بر اساس آن ارائه شد. از جمله می‌توان به روش‌های فرکتال غلظت- مساحت (Zou *et al.*, 2013; Daya, 2014)، غلظت- محیط (Daya, 2014)، غلظت- فاصله (Li *et al.*, 2003) و طیف توان- مساحت (Zuo *et al.*, 2013) اشاره کرد.

فصل

۳- مروری بر مطالعات پیشین

۳-۱- مقدمه

از طبقه‌بندی و خوشه‌بندی برای دستیابی به اهداف مختلفی در علوم زمین استفاده شده است. این الگوریتم‌ها در حوزه نفت به دلیل در اختیار بودن طیف وسیع و حجم زیادی از اطلاعات معمولاً به سادگی قابل اجرا و همراه با پاسخ مطلوب هستند. در حوزه معدن به دلیل محدودیت در حجم و تنوع اطلاعات استفاده از شناسایی الگو محدودتر و همواره همراه با ابهام بوده است. از خوشه‌بندی در جدایش آنومالی از زمینه و همچنین تهیه نقشه‌های سطحی کانسار به صورت گسترده استفاده شده است. از طبقه‌بندی نیز معمولاً در همین راستا استفاده شده است. در مطالعات گذشته تلاش شده تا با ورود اطلاعات در دسترس به الگوریتم، یک خروجی مطلوب گرفته شود که در موارد متعددی این استفاده مبهم و گاهی اوقات با ماهیت شناسایی الگو همخوانی نداشته است.

در این پژوهش تلاش می‌شود تا اولاً در حل یک سری مسائل از شناسایی الگو استفاده شود. در ادامه تلاش می‌شود تا الگوریتم‌ها را به سمت طراحی مدل‌های ساده، دقیق و با راندمان بالا هدایت کرد. از خوشه‌بندی نه تنها برای تفکیک آنومالی یک عنصر از زمینه و تهیه نقشه سطحی، بلکه برای جدایش آنومالی‌های چندعنصری استفاده خواهد شد. جهت بررسی منطقه‌بندی در تغییرات غلظت عناصر مختلف از خوشه‌بندی استفاده خواهد شد و سپس اثر استفاده از این رویکرد در بهبود تخمین غلظت عناصر بررسی خواهد شد. پس از این مرحله از ترکیب خوشه‌بندی و طبقه‌بندی جهت تخمین کلاس غلظت عناصر در افق‌های مختلف عمقی محدوده نمونه‌برداری سطحی استفاده خواهد شد.

۳-۲- تخمین عیار و ذخیره در کانسارهای معدنی

تخمین ذخیره یکی از مهمترین مشکلات و دغدغه‌های کارشناسان امور معادن است. علت این امر شاید وابستگی زیاد پروژه‌های معدنی به مقدار دقیق عیار و نیاز به داشتن دانش وسیعی از معدنکاری، برای تخمین عیار و سپس تناژ ماده معدنی باشد. اغلب به علت وجود گسل‌ها، دایک‌ها و شکستگی و ساختارهای

پیچیده، مدل‌سازی کانسار با مشکلات خاصی مواجه است. شایان ذکر است که یکی از پیچیده‌ترین و مشکل‌ترین مدل‌سازی‌ها، در ارتباط با مدل‌سازی شاخه‌های مختلف معدنی و زمین‌شناسی می‌باشد، دلیل این امر متفاوت بودن ساختار و بروز عوامل منحصربه‌فرد است (طهماسبی و هزارخانی، ۱۳۹۰).

در مراحل مختلف معدنکاری همواره با درجه‌ای از عدم قطعیت روبرو هستیم. برخی از این عدم قطعیت‌ها نظیر میزان ذخیره و عیار کانسار از تغییرات ذاتی عیار کانسار ناشی می‌شوند که به طور مستقیم بر شاخص‌های فنی و اقتصادی کانسار تأثیر می‌گذارند. از سوی دیگر هزینه‌های سنگین بخش اکتشاف اغلب باعث محدود شدن حجم عملیات اکتشافی می‌گردد که این خود لزوم به‌کارگیری روش‌های دقیق تخمین را ایجاب می‌نماید (حسنی پاک، ۱۳۸۲).

روش‌های گوناگونی برای تخمین عیار و ذخایر معدنی وجود دارد. این روش‌ها را می‌توان به روش‌های هندسی، مبتنی بر فاصله، زمین‌آماری و هوشمند تقسیم کرد. هر یک از این روش‌ها دارای محدودیت‌ها، مزایا و معایبی می‌باشند و بر یک سری فرضیات استوارند که انتخاب روش مناسب برای تخمین باید بر اساس این عوامل و نیز دقت و کیفیت مورد نظر انجام شود (حسنی پاک، ۱۳۸۴).

نخستین تجربه‌ها جهت به‌کارگیری روش‌های زمین‌آماری به مفهوم امروزی آن در محاسبات تخمین ذخیره از حدود ۹۰ سال پیش با شناسایی مقدماتی الگوهای توزیع طلا در معادن آفریقای جنوبی شروع شد. هوپر^{۳۶} و واترمایر^{۳۷} پیشگامان تئوری‌های زمین‌آماری بودند که در زمینه معدن طلا تحقیق کردند. اولین مقاله در این زمینه توسط واترمایر در سال ۱۹۱۹ منتشر شد که در آن لزوم به‌کارگیری میانگین‌وزنی به جای میانگین حسابی بیان شده بود. در ضمن شباهت بین مقادیر نمونه‌ها به عنوان تابعی از فاصله نمونه‌ها ارزیابی شده که این رابطه، پایه اصلی زمین‌آمار را تشکیل می‌دهد. این رابطه را اولین بار کریج^{۳۸} کارشناس معدن در آفریقای جنوبی در سال ۱۹۶۶ جهت ارزیابی معادن پیشنهاد کرد. به دنبال روند تکاملی روش‌های

36 Hooper
37 Watermeyer
38 Krige

آماري مورد استفاده در تخمين ذخاير معدني كه از سال‌ها قبل آغاز شده بود، ماترون³⁹ با انتشار مقاله‌اي در سال 1962 پايه‌هاي زمين‌آمار نوين را بنا نهاد (Ben-Jemaa et al., 1990).
با افزايش روزافزون حافظه و سرعت كامپيوترها، شاهد پيشرفت کاربردي سيستم‌هاي ديناميكي مبتني بر داده‌هاي تجربي در عرصه عمل هستيم. شبكه‌هاي عصبي هوشمند مصنوعي⁴⁰ مدل آزاد⁴¹ با الهام از عملکرد مغز انسان و واحدهاي پردازشگر آن بوجود آمده‌اند. اين مدل بر اين فرض استوار است كه همانند مغز بشر امكان يادگيري توسط واحدهاي عصبي (نرون) براي تخمين موجود مي‌باشد (Dowd and Sarac, 1994).

3-3- بهبود تخمين گرها

بررسي دقيق هاله‌هاي پراكندگي ژئوشيميائي عناصر در محدوده كاني‌سازي و انتخاب روش تخمين غلظت مي‌تواند در تصميم‌گيري‌هاي آينده پروژه‌هاي معدني موثر باشد (Goovaerts, 1997; Journel & Huijbregts, 1978). مدل پراكندگي ژئوشيميائي عناصر در مناطق كاني‌سازي با توجه به فرايندهاي زمين‌شناسي كه در تشكيل كانسارها دخالت دارند پيچيده است (Jalloh et al., 2016). ساختار پيچيده كانسارهاي معدني، مدل‌سازي پراكندگي غلظت عناصر ژئوشيميائي را با مشكل مواجه مي‌كند و مهندسين معدن و زمين‌شناسان، از دهه‌ها قبل به دنبال تكنيك‌هايي براي تخمين دقيق پراكندگي عناصر هستند (Goovaerts, 1997; Journel & Huijbregts, 1978; Rendu, 1979). روش‌ها و تكنيك‌هاي مختلفي براي تخمين غلظت عناصر ژئوشيميائي وجود دارد كه مي‌توان از آن جمله به روش‌هاي زمين‌آماري (Rendu, 1979, Hornik et al., 1989) و شبكه‌هاي عصبي مصنوعي (Samanta et al., 2002; Rooki et al., 2011; Koike & Matsuda, 2003; Jozanikohan et al., 2015) اشاره كرد. شبكه عصبي مصنوعي به دليل توانايي شناسايي روندهاي غيرخطي داده‌هاي ژئوشيميائي و شناخت روابط پيچيده‌ي غيرخطي

39 Matheron

40 Artificial Neural Networks (ANNs)

41 Model Free

موجود بین داده‌های ورودی و خروجی به ابزاری قدرتمند برای حل بسیاری از مشکلات محاسباتی تبدیل شده است. این روش به صورت گسترده‌ای در تخمین غلظت عناصر مورد استفاده قرار گرفته (Jafrasteh, 2013; Fathianpour, 2017; Nakhaei & Irannajad, 2013) و بر اساس ارتباط‌های پیچیده بین مختصات فضایی نمونه‌ها و عیار عنصر مورد بررسی، پراکندگی غلظت در محدوده مورد مطالعه شناسایی و مدل‌سازی می‌گردد (Guo, 2010). بر خلاف شبکه عصبی مصنوعی، بسیاری از روش‌های زمین‌آماري رایج مانند کریجینگ، تخمین‌گرهای خطی هستند (Hornik et al., 1989). در برخی از موارد که درجه توزیع و الگوهای روابط پیچیده است، روش‌های زمین‌آماري، قادر به دادن بهترین پاسخ نمی‌باشد و نیازمند روشی هستیم که بتواند روابط غیرخطی را شناسایی کند (Strebelle, 2002). علاوه بر این، در روش‌های زمین‌آماري مبتنی بر دونقطه، دقت کم است و این هم به دلیل برخی محدودیت‌ها و کاستی‌های این روش‌ها می‌باشد (Tahmasebi & Hezarkhani, 2012; Tahmasebi et al., 2012).

محققان مختلفی با تلفیق روش‌های مختلف با تخمین‌گر شبکه عصبی، دقت تخمین را افزایش و خطای آن را کاهش داده‌اند. مقایسه عملکرد شبکه عصبی به تنهایی و ترکیب با الگوریتم طبقه‌بندی آداپوست، نشان دهنده بهبود عملکرد تخمین غلظت عناصر ژئوشیمیایی است (Samanta et al., 2005) الگوریتم ژنتیک (GA) و روش لونبرگ - مارکوارت (LM) برای بهینه‌سازی وزن نمونه‌ها در مرحله آموزش سبب افزایش دقت تخمین شبکه عصبی می‌گردد (Mahmoudabadi et al., 2009). افزایش دقت تخمین با شبکه عصبی موجک (Xiao-li et al., 2009)، نرو فازی (Tahmasebi & Hezarkhani, 2010a)، منطق فازی (Tahmasebi & Hezarkhani, 2010b)، منطق فازی - الگوریتم ژنتیک (Tahmasebi & Hezarkhani, 2012; Allahkarami et al., 2017) انجام شده است.

۳-۴- جدایش آنومالی از زمینه

جدایش آنومالی از زمینه یک مسئله اساسی در ژئوشیمی اکتشافی است. در روش‌های آماری معمولاً غلظت عناصر شیمیایی در پوسته زمین از یک توزیع نرمال و در ناهنجاری‌های ژئوشیمیایی غالباً از توزیع لاگ نرمال پیروی می‌کند. یک ناهنجاری ژئوشیمیایی به عنوان منطقه‌ای که در آن غلظت عنصر خاصی بیشتر از مقدار حدآستانه‌ای خاص تعریف می‌شود (Li et al., 2003). یکی از روش‌های ارزیابی حدآستانه‌ای، استفاده از روش فرکتالی است. تئوری فرکتال از اواخر ۱۹۸۰ در مطالعات معدنی به کار گرفته شده است. Turcotte (1986) رابطه فراکتال بین عیار میانگین و ذخایر معدنی تجمعی را ارائه کرده است. منگ و ژائو (۱۹۹۱) به این نتیجه رسیدند که رفتار فراکتال در ساختارها و داده‌های زمین‌شناسی وجود دارد. چنگ و همکاران (۱۹۹۴) پیشنهاد کردند که توزیع داده‌های ژئوشیمیایی از ساختار فراکتالی در طبیعت پیروی می‌کنند. مطالعات متعدد علمی نشان می‌دهد که رابطه تجربی بین عیار و مساحت در طبیعت وجود دارد که به روش فرکتالی عیار - مساحت معروف است (Cheng, 1994; Sim et al., 1999; Wei and Pengda, 2002).

روش فرکتالی می‌تواند به عنوان یک روش مفید برای تجزیه و تحلیل توزیع ژئوشیمیایی عناصر استفاده شود. این روش به تدریج در حال تبدیل شدن به عنوان یک ابزار موثر و کارآمد برای تجزیه و تحلیل ساختار فضایی در داده‌های ژئوشیمیایی است (Pazand et al., 2011). تجزیه و تحلیل فراکتالی عیار - مساحت قادر به شناسایی کلاس ناهنجاری‌های ژئوشیمی است که برخی از آن‌ها در ارتباط با سطح ضعیفی از کانی‌سازی پنهان بوده و در نتیجه در تعریف اهداف اکتشافی بعدی و حفاری سیستماتیک مرحله اول مفید می‌باشند (Asadi et al., 2014). به هر حال، مناطق مختلف می‌توانند دارای ترکیبات سنگ متفاوت و همراه با فرآیندهای مختلف زمین‌شناسی باشند که در نتیجه مرزهای مختلف ژئوشیمیایی عناصر به وجود خواهد آمد. بنابراین، می‌توان از روش فرکتالی عیار - مساحت برای شناسایی

ناهنجاری‌های ژئوشیمیایی متفاوت حاصل از فرآیندهای مختلف زمین‌شناسی و سیالات هیدروترمالی استفاده کرد (Wang & Zuo, 2015).

۳-۵- الگوریتم‌های شناسایی الگو

تشخیص الگو^{۴۲} روشی است که شامل تکنیک‌های طبقه‌بندی اشیاء در تعدادی کلاس مختلف می‌باشد (Webb, 2002; Duda et al., 2012). از دهه ۱۹۷۰، تکنیک‌های تشخیص الگو برای شناسایی اطلاعات زمین‌شناسی، اقتصادی و کانی‌سازی پنهان در داده‌های ژئوشیمیایی و تعیین الگوی آنومالی و زمینه کار گرفته شده‌اند (Cheng, 2004). این روش‌ها در بررسی رابطه بین الگوهای ژئوشیمیایی ناحیه‌ای با ذخایر بزرگ معدنی نیز مورد استفاده واقع شده‌اند (Wang et al., 2008; Cheng, 2008).

در اکتشافات ژئوشیمیایی، کشف الگوهای ژئوشیمیایی غیرعادی یا آنومالی‌های ژئوشیمیایی مدنظر است (Ghavami Riabi et al., 2010)، که برای این منظور می‌بایست نسبت به شناخت مقدار زمینه عیار عناصر در محیط اقدام کرد. بدیهی است هر مقدار غلظت بالاتر از زمینه نشان دهنده یک آنومالی نیست. لذا برای تفکیک آنومالی، روش‌های آماری مختلفی وجود دارد که به عنوان مثال اگر ترکیبی از مقادیر یک گروه از عناصر معرف به جای مقدار یک عنصر خاص به کار گرفته شوند، هاله‌های ژئوشیمیایی در اطراف توده‌های کانسار بهتر مشخص می‌شوند (Beus and Grigorian, 1962). روش‌های مختلفی برای گروه‌بندی داده‌های ژئوشیمیایی و تفکیک آنومالی وجود دارد که تحت عنوان تجزیه و تحلیل خوشه‌ای یا کلاستر معرفی می‌شوند. به طور خاص از روش‌های کلاستری می‌توان به ارتباط بین‌گروهی^{۴۳}، ارتباط درون‌گروهی^{۴۴}، نزدیکترین همسایه^{۴۵}، دورترین همسایه^{۴۶}، خوشه‌بندی متمرکز و میانه اشاره کرد. بدیهی

۴۲ Pattern recognition

43 Between - groups linkage

44 Within - groups linkage

45 Nearest neighbor

است معرفی یکی از روش‌های خاص کلاستر به عنوان روشی که احتمالاً بیشترین توانمندی را در تفکیک آنومالی‌ها دارد، امکان‌پذیر نبوده و متناسب با ماهیت کانسار و داده‌ها می‌بایست از تکنیک خاصی استفاده نمود.

با توجه به این که داده‌های اکتشافی مختلف، دارای بخش‌های با ارزش متفاوت هستند که ممکن است همه آن‌ها برای اکتشاف یک ماده معدنی خاص دارای ارزش یکسان نباشند، لذا لازم است الگوهای شاهد کانی‌سازی، از مجموعه داده‌ها استخراج شود. نتیجه‌گیری بهترین الگوهای شاهد با استفاده از ارزش‌گذاری کلیه الگوها صورت می‌گیرد. بنابراین، ارزش‌گذاری بخش‌های مختلف یک نقشه برای جداسازی و نمایش شواهد با ارزش متفاوت (یا طبقه‌بندی) و همچنین تعیین میزان اهمیت حضور هر معیار اکتشافی، در هر موقعیت مکانی و برای هر نوع کانی‌سازی مشخص مورد جستجو، برای استفاده در مدل‌سازی مسئله ضروری است (Carranza, 2009).

تشخیص الگوی نظارت شده معمولاً شامل سه مرحله اصلی انتخاب‌ویژگی، طبقه‌بندی و ارزیابی است. نتیجه این روش به شدت تحت تاثیر مرحله طبقه‌بندی است که تاکنون موضوع پژوهش‌های متعددی بوده است. روش‌های مبتنی بر تئوری تصمیم‌گیری بیزین^{۴۷} (Friedman et al., 1997; Porwal et al., 2006)، روش‌های گروه^{۴۸} (Freund and Schapire, 1996; Dietterich, 2000; Nejadi et al., 2015)، درخت تصمیم‌گیری^{۴۹} (Breiman, 1996, 2001; Purwar et al., 2011)، روش‌های کرنل غیرخطی^{۵۰} (Schölkoph et al., 2000; Al-Anazi and Gates, 2010; Zuo and Carranza, 2011) و شبکه‌های عصبی^{۵۱} (Singer and Kouda, 1997; Porwal et al., 2003; Kashani et al., 2014) برخی از انواع

46 Neighbor furthest
47 Bayesian decision
48 Ensemble
49 Decision trees
50 Nonlinear kernel methods
51 Neural networks

الگوریتم‌های طبقه‌بندی نظارت‌شده است که در حل پژوهش‌های واقعی متعددی مورد استفاده قرار گرفته‌اند.

۳-۶- تاریخچه‌ای از کاربرد روش‌ها در ژئوشیمی

براساس مطالعات قبلی انجام شده، مشاهده می‌شود که اکثر مطالعات بر روی پتانسیل‌یابی سطحی و شناسایی آنومالی‌های ژئوشیمیایی در سطح استوار می‌باشند و مطالعاتی که منجر به تخمین تغییرات غلظت هاله‌های ژئوشیمیایی در افق‌های مختلف عمقی و در نقاط فاقد نمونه‌برداری با استفاده از داده‌های ژئوشیمیایی باشد مورد توجه قرار نگرفته است. لازم به ذکر است که ممکن است به صورت جزئی و در بخش کوچکی از مطالعات به این موضوع پرداخته شده باشد (Abedi et al., 2011, Roshani et al., 2013, Mohammadi Gonbadi et al., 2015, Geranian et al., 2016). در مواردی که هدف از مطالعات، شناسایی کانی‌سازی پنهان در عمق بوده است تنها از داده‌های گمانه‌ای برای برچسب‌زدن داده‌های سطحی استفاده شده است. نقشه‌های بدست آمده از این مرحله طبیعتاً تنها نشانگر تغییرات در سطح می‌باشد. در مناطقی که در سطح به‌عنوان مناطق پتانسیل‌دار شناسایی شده است، چنانچه منشاء مربوطه آلودگی زیست محیطی نباشد انتظار می‌رود که در عمق نیز کانی‌سازی اتفاق افتاده باشد.

مشکل تشخیص آنومالی‌ها، پیدا کردن الگوهای رفتاری است که یک سری داده از خود نشان می‌دهند (Carranza, 2009). روش‌های متداول تشخیص آنومالی‌های ژئوشیمیایی، با شیوه بدون نظارت شده عمل می‌کنند، به این معنی که از تمام خصوصیات و ویژگی‌ها در پردازش داده‌های ژئوشیمیایی استفاده نمی‌شود. این روش‌ها قادر به استفاده از یک پیش‌بینی کننده موثر اطلاعات (پیش‌بینی عمق کانی‌سازی) در پردازش داده‌های ژئوشیمیایی نمی‌باشند. الگوی تغییرات بین داده‌های سطحی و زیرسطحی، در مناطقی که داده‌های سطحی موجود و فاقد داده زیرسطحی است، به ما کمک خواهد کرد تا براساس این

روابط، تغییرپذیری غلظت عناصر را در عمق شناسایی کنیم. چنین رابطه‌ای را می‌توان با استفاده از الگوریتم‌های تشخیص الگوی نظارت‌شده به دست آورد (Mohammadi Gonbadi et al., 2015).

مطالعات متعددی در دهه اخیر انجام شده است که برای این کار از روش‌های مختلفی استفاده شده است. در سال ۲۰۱۰، کارانزا و صادقی از روش‌های آنالیز فرکتال و آنالیز فرای جهت تهیه نقشه‌های پیش‌بینی‌کننده پتانسیل‌های سطحی در منطقه اسکلفت^{۵۲} سوئد استفاده کردند (Carranza and Sadeghi, 2010). همچنین در همان سال، مطالعه‌ای مبنی بر تهیه نقشه پتانسیل پیش‌بینی‌کننده سطحی ذخایر طلا، با ترکیب اطلاعات مختلف ژئوشیمیایی و زمین‌شناسی در محیط GIS انجام شد. نتایج این مطالعه، شناسایی مناطق آنومال سطحی از لحاظ کانی‌سازی طلا بود که قبلاً شناسایی نشده بودند (Czarnota et al., 2010).

در سال ۲۰۱۲، عابدی و همکاران از روش تصمیم‌گیری چندمعیاره جهت تلفیق لایه‌های اطلاعاتی ژئوفیزیکی و ژئوشیمیایی برای اکتشاف مس پورفیری در منطقه ناچون استان کرمان استفاده کردند و نقشه پتانسیل سطحی مس را بدست آوردند (Abedi et al., 2012). در مطالعه‌ای که در ۲۰۱۳ توسط عابدی و همکاران در منطقه ناچون استان کرمان انجام شد، روش دانش‌محور جدید به نام فازی‌برتر^{۵۳} را برای تولید نقشه پتانسیل معدنی پیشنهاد شد و نقشه پتانسیل منطقه تهیه گردید (Abedi et al., 2013). روشی و همکاران در سال ۲۰۱۳ در منطقه کوه‌پنج، یک رویکرد جدید برای شناسایی آنومالی‌های ژئوشیمیایی با استفاده از آنالیز تمایز و مقادیر واقعی "آنومالی" و "زمینه" تعیین شده از مجموعه داده‌های سطحی و گمانه در نظر گرفته‌اند (Roshani et al., 2013). مطالعه‌ای توسط گرانیان و همکاران در سال ۲۰۱۴ بر روی کانسار تیپ جهانی طلای ساری‌گونئی^{۵۴} در غرب ایران انجام شد. هدف اصلی این

52 Skellefte

53 Fuzzy outranking

54 Sari Gunay

مطالعه استفاده از آنالیز رگرسیون چندمتغیره بر روی داده‌های لیتوژئوشیمیایی سطحی و گمانه‌ای ذخیره طلای ذکر شده به منظور تعریف مکان‌های مناسب برای حفاری‌های اکتشافی بیشتر در مناطق فاقد نمونه‌برداری عمقی بود (Geranian et al., 2014). در مطالعه دیگری که در ۲۰۱۴ نجفی و همکاران با استفاده از روش فازی فرآیند تحلیل سلسله مراتبی^{۵۵} (AHP)، روشی برای تهیه نقشه پتانسیل‌دار معدنی ارائه دادند که معمولاً برای اکتشاف ذخایر معدنی در سطح مورد استفاده قرار می‌گیرد، اطلاعات اکتشافی را تلفیق و نقشه پتانسیل معدنی تهیه کردند (Najafi et al., 2014).

گالیانو^{۵۶} و همکاران در سال ۲۰۱۵ در مطالعه‌ای که در منطقه معدنی طلای رودالکویلا^{۵۷} در جنوب اسپانیا انجام دادند. از چهار الگوریتم یادگیری ماشین از جمله شبکه‌های عصبی مصنوعی^{۵۸} (ANNs)، درختان رگرسیون^{۵۹} (RTS)، جنگل‌های تصادفی^{۶۰} (RF) و ماشین بردار پشتیبان (SVM) در تهیه نقشه پتانسیل معدنی استفاده و مناطق پتانسیل‌دار از لحاظ کانی‌سازی طلا را مشخص کردند (Galiano et al., 2015). مطالعه‌ای توسط ولادیمیر لیسیتسین^{۶۱} در سال ۲۰۱۵ تحت عنوان تجزیه و تحلیل مکانی الگوهای نقطه‌ای کانسار جهت کاربرد در اکتشاف کانسارها انجام شد و نقشه پتانسیل کانسارهای پنهان منطقه ارائه گردید (Lisitsin., 2015). مطالعه‌ای توسط چن^{۶۲} با عنوان تهیه نقشه پتانسیل معدنی با استفاده از ماشین بولتزن محدود در منطقه آلتای^{۶۳} در شمال سین‌کیانگ چین در سال ۲۰۱۵ انجام پذیرفت و نقشه‌های پتانسیل معدنی سطحی برای این منطقه تهیه گردید (Chen., 2015).

55 Fuzzy analytical hierarchy process

56 Galiano

57 Rodalquilar

58 Regression trees

59 Regression trees

60 Random forest

61 Vladimir Lisitsin

62 Chen

63 Altay

در ادامه‌ی سال ۲۰۱۵، شناسایی آنومالی‌های ژئوشیمیایی توسط الگوریتم‌های تشخیص‌الگوی نظارت شده در منطقه کوه‌پنج (Mohammadi Gonbadi et al., 2015)، تهیه نقشه پیش‌بینی کننده داده‌محور پتانسیل طلا، منطقه باگویو^{۶۴}، فیلیپین با استفاده از الگوریتم جنگل‌های تصادفی (Carranza and Laborte., 2015a)، مدل‌سازی پیش‌بینی کننده جنگل تصادفی مناطق پتانسیل‌دار کانسار مس پورفیری در منطقه ابرا، فیلیپین (Carranza and Laborte., 2015b)، استفاده از رگرسیون لجستیک^{۶۵} برای تهیه نقشه پیش‌بینی پتانسیل طلا در کمربند گرین‌استون جیانی^{۶۶} آفریقای جنوبی (Carranza et al., 2015) پاره‌ای از تحقیقات دیگر انجام شده بود.

یکی از روش‌های متداول برای تشخیص ناهنجاری‌ها در اکتشافات ژئوشیمیایی، روش تجزیه و تحلیل آنالیز تمایز^{۶۷} است. آنالیز تمایز (DA) یک روش آماری چند متغیره است که هر کدام از مشاهدات را در یک گروه خاص بر اساس متغیرها و گروه‌های از پیش تعریف شده طبقه‌بندی می‌کند. در مطالعه روشنی و همکاران (۲۰۱۳) رویکرد جدیدی برای شناسایی آنومالی ژئوشیمیایی با استفاده از DA و آنومالی و زمینه واقعی از پیش تعیین شده در نظر گرفته شده است. نمونه‌های آنومال و زمینه بر اساس وجود یا عدم وجود کانی‌سازی در عمق شناسایی می‌شوند. بنابراین، این روش به عنوان "رویکرد عینی"^{۶۸} معرفی شده است. به منظور طبقه‌بندی نمونه‌های ژئوشیمیایی سطحی به آنومالی و زمینه، از عیار مغزه‌های حفاری کانی‌سازی پورفیری مس کوه پنج استفاده شده است. اگر کانی‌سازی اثبات شده باشد، آنها به عنوان آنومالی طبقه‌بندی می‌شوند و اگر کانی‌سازی در گمانه تایید نشده باشد، به عنوان زمینه برچسب گذاری می‌شوند.

64 Baguio

65 Logistic regression

66 Giyani greenstone belt

67 Discriminant function analysis

68 Objective approach

برای رسیدن به توابع تمایز، تجزیه و تحلیل خطی گام به گام^{۶۹} (LDA) و تجزیه و تحلیل متقارن درجه دوم^{۷۰} (QDA) استفاده شده است (Roshani et al., 2013).

در این تحقیق داده‌های آزمایشی که برای تولید مدل‌ها استفاده شده، با دو روش LDA و QDA بطور کامل طبقه‌بندی شد و دقت اعتبارسنجی ۸۴٪ و ۷۴٪ به ترتیب برای LDA و QDA بدست آمد. نتایج این تحقیقات نشان داده اگر اطلاعات زمین شناسی و ژئوشیمی در دسترس منطقه هدف به کار گرفته شوند و مورد استفاده قرار گیرند، روش LDA می‌تواند بطور موثر به عنوان یک روش عینی برای شناسایی ناهنجاری ژئوشیمیایی به کار گرفته شود. همچنین نتایج نشان داده که تعریف ناهنجاری در اکتشافات ژئوشیمیایی می‌تواند به طرز قابل توجهی با استفاده از این روش بهبود یابد (Roshani et al., 2013).

تشخیص آنومالی ژئوشیمیایی مسئله‌ای مهم در اکتشاف مواد معدنی است. در دسترس بودن مجموعه داده‌های آموزشی که شامل نمونه‌های ژئوشیمیایی برچسب‌دار از کلاس‌های زمینه و آنومالی، ما را قادر می‌سازد که چارچوب شناخت الگو تحت نظارت را برای تشخیص آنومالی ژئوشیمیایی تعریف کنیم. بر این اساس محمدی گنبدی و همکاران (۲۰۱۵)، الگوریتم‌های مختلف طبقه‌بندی و انتخاب ویژگی‌ها را برای ساخت یک مدل پیش بینی شده و طبقه‌بندی نمونه‌های ژئوشیمیایی در کلاس‌های آنومالی و زمینه مورد استفاده قرار داده است. در این مطالعه، برخی از الگوریتم‌های انتخاب ویژگی و الگوریتم‌های طبقه‌بندی پیشرفته برای تشخیص آنومالی نظارت شده در منطقه پورفیری-کوه پنج استفاده شد (Mohammadi et al., 2015).

در ادامه، الگوریتم‌های آدابوست (ADB)، ماشین بردار پشتیبان (SVM) و جنگل‌های تصادفی (RF) با استفاده از نمونه‌های سنگی گمانه‌ای و سطحی از قسمت‌های حفاری شده منطقه مورد مطالعه، برای ایجاد

69 Stepwise Linear discriminant analysis
70 Quadratic discriminant analysis

یک نقشه طبقه‌بندی شده آنومالی (در قسمت‌هایی که حفاری انجام نشده بود) آموزش داده شد. نتایج نشان داد که الگوریتم‌های انتخاب ویژگی‌ها می‌توانند نقش مهمی در افزایش دقت و بهینه‌سازی طبقه‌بندی استفاده شده داشته باشند. از میان روش‌های طبقه‌بندی استفاده شده، آدابوست بهترین عملکرد را با کمترین خطا به میزان ۰/۰۶ به دست آورد. در همین حال، مقایسه نقشه طبقه‌بندی شده آدابوست با نقشه بدست آمده از روش فرکتال نشان داد که روش آدابوست در شناسایی مناطق پتانسیل‌دار موجود در منطقه موفق‌تر بود (Mohammadi Gonbadi et al., 2015).

هاله‌های عناصر فلزی در مناطق گسلی مرتبط با کانی‌سازی^{۷۱} (MEAHFZ) و ویژگی‌های ژئوشیمیایی آنها به عنوان روشی برای کشف ساختارهای پنهان ذخایر پلی‌متال سرب و روی کانسار کلینچانگ^{۷۲} چین در نظر گرفته شده‌اند. بررسی دقیق ناهنجاری‌های ژئوشیمیایی، کشف ذخایر پنهان شده به روش MEAHFZ موفقیت آمیز بوده و منجر به افزایش ۲ میلیون تن (Mt) در ذخایر فلزی سرب و روی شده است. اکتشاف موفقیت آمیز در ناحیه هوییز^{۷۳} نشان داد که روش MEAHFZ، روشی بالقوه برای افزایش ذخایر فلزی و یافتن ذخایر جدید معدنی در مناطق مجاور است (Han et al., 2015).

گرا دیان تغییرات ناهنجاری‌های اولیه، کنترل شده توسط ساختارها، می‌تواند جهت و گسترش کانسارهای پنهان را پیش‌بینی کند. در ضمن جهت جابجایی ناهنجاری‌ها در سطوح مختلف نشان دهنده جهت شیب کانسارهای پنهان در عمق است. بدین ترتیب، این روش نه تنها کانسارهای پنهان را اکتشاف می‌کند، بلکه اطلاعات مهمی در مورد ژنز کانسار فراهم می‌کند. این روش کم هزینه بوده و عملیات آن ساده است. لذا می‌تواند برای کاوش کانسارهای پنهانی که در عمق و در نواحی که ذخایر سرب و روی وجود دارد و کانسار به وضوح توسط ساختارها کنترل می‌شود مورد استفاده قرار گیرد (Han et al., 2015).

71 Metal–element association halos within mineralization-related fault zones

72 Qilinchang

73 Huize

گرانیان و همکاران (۲۰۱۵) رگرسیون چند متغیره بر روی داده‌های ژئوشیمیایی سطحی سنگ و گمانه ذخیره طلای اپی‌ترمال ساری گونئی در شمال غرب ایران به منظور مدل‌سازی کانسارهای زیرزمینی جهت حفاری بیشتر اعمال نمودند (Granian et al., 2015).

مدل‌های رگرسیون چندگانه سطحی، فاکتوریل، چندجمله‌ای و مدل رگرسیونی پاسخ سطحی^{۷۴} بر روی بخشی از داده‌های ژئوشیمیایی منطقه معدنی آموزش داده شد و برای ارزیابی دقت این مدل‌ها، از بخش دیگر داده‌ها که مورد استفاده قرار نگرفته بود کمک گرفته شد (Granian et al., 2015).

داده‌های ژئوشیمیایی ۳۱ عنصر در نمونه‌های سنگی ترانشه سطحی به عنوان متغیرهای مستقل مورد استفاده قرار گرفت و سه پارامتر یعنی میانگین غلظت^{۷۵}، مجموع غلظت^{۷۶} و پتانسیل مورد انتظار^{۷۷} در سلول‌های ۲۵ در ۲۵ متری که بوسیله درون‌یابی با روش کریجینگ با داده‌های گمانه‌ای بدست آمده بود به عنوان متغیرهای وابسته در نظر گرفته شد. برای هر سلول ۲۵ در ۲۵ متری، مجموع عیارهای بالای ۵ ppm در بلوک‌های عمودی به عنوان مجموع غلظت و ضرب مجموع غلظت در ضخامت بلوک به عنوان پتانسیل مورد انتظار در نظر گرفته شد (Granian et al., 2015).

تمام مدل‌های رگرسیون چند متغیره ضرایب تعیین شدت را برای سه پارامتر نشان دادند که در آن مدل رگرسیون سطحی بالاترین مقادیر را بدست آورد. نتایج رگرسیون سطحی به منظور مدل‌سازی کانی‌سازی طلای زیرزمینی در کانسار طلای ساری گونئی^{۷۸} برای طراحی حفاری‌های اضافی مورد استفاده قرار گرفت (Granian et al., 2015).

74 Response surface regression models

75 Average grade

76 Sum grade

77 Productivity

78 Sari Gunay

برای کشف کانسار پنهان، هراز^{۷۹} و همکاران (۲۰۱۵) ویژگی‌های کمی عناصر ردیاب هاله‌های کانی‌سازی منطقه را برای ۲۸۰ نمونه از سطح و سه سطح عمقی معدن زیرزمینی اعمال نمودند. از طریق تجزیه و تحلیل آماری چند متغیره (تجزیه و تحلیل فاکتوری) ۱۱ عنصر انتخاب شد. گسترش محوری (عمودی) هاله‌های اولیه در بالا و پایین کانسار برای تفسیر سطح فرسایش، تعیین شیب کانی‌سازی و همچنین بررسی اینکه آیا کانسار پنهان در معدن Atud امیدوار کننده است مورد استفاده قرار گرفتند. شاخص $\frac{Pb \cdot Cu}{U \cdot Zn}$ برای پیش‌بینی پتانسیل طلا در سطوح مختلف عمقی محاسبه و معرفی شد. علاوه بر این شاخص زونالیه $\frac{Pb}{U}$ شاخص مناسبی برای تعیین میزان فرسایش در معدن طلای Atud تشخیص داده شد. استفاده از نقشه شدت شاخص زونالیه ژئوشیمیایی در سطح، سطح از فرسایش آنومالی ژئوشیمیایی و همچنین کاهش محتوای طلا با افزایش عمق در سطوح مختلف زیرزمینی، ارائه مدل پیش‌بینی را برای ارزیابی مدل کانسار در عمق فراهم می‌کند (Harraz et al., 2015).

شاهی و همکاران (۲۰۱۶) داده‌های ژئوشیمیایی سطحی در حوزه فضایی را با استفاده از تبدیل فوریه دو بعدی به دامنه فرکانس منتقل کرد. تجزیه و تحلیل داده‌های ژئوشیمیایی سطحی در دامنه فرکانس به اطلاعات اکتشافی منجر شده است که ممکن است در حوزه فضایی داده‌های ژئوشیمیایی قابل دستیابی نباشد. در این تحقیق، دامنه فرکانس داده‌های ژئوشیمیایی سطحی برای شناخت الگوهای ژئوشیمیایی پیچیده مربوط به واحدهای سنگی مورد تجزیه و تحلیل قرار گرفته است. به منظور پیش‌بینی تغییرات کانی‌سازی در عمق و شناسایی ذخایر معدنی پنهان، روش جدید ضریب فرکانسی^{۸۰} (FCM) پیشنهاد شده و بر روی کانسار پنهان مس - مولیبدن پورفیری ظفرغند مورد استفاده قرار گرفته است. روش پیشنهادی به طور مطلوب ارتباط بین فرکانس‌های مختلف در نقشه توزیع ژئوشیمیایی سطحی و عمق‌های مختلف کانسار را نشان می‌دهد. نتایج حاصل از استفاده از روش پیشنهادی در یک بررسی واقعی، بهبود قابل

79 Harraz

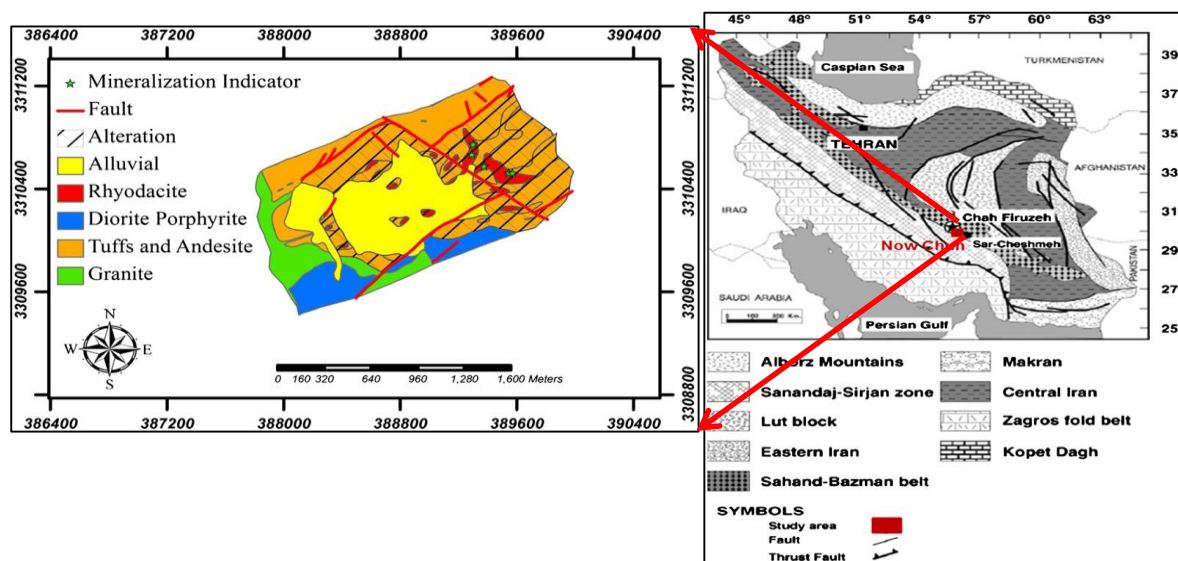
80 Frequency Coefficients Method

توجهی در مقایسه با نتایج به دست آمده از حوزه فضایی داده‌های ژئوشیمیایی نشان می‌دهد. روش معرفی شده به عنوان یک تکنیک تشخیص الگو، بدون حفاری اکتشافی، تعیین روند کانه‌زایی در عمق و تمایز بین کانی‌سازی پنهان و مناطق کانی‌سازی پراکنده را آشکار نمود (Shahi et al., 2016).

۷-۳- مثال‌هایی از کاربرد روش‌های شناسایی الگو

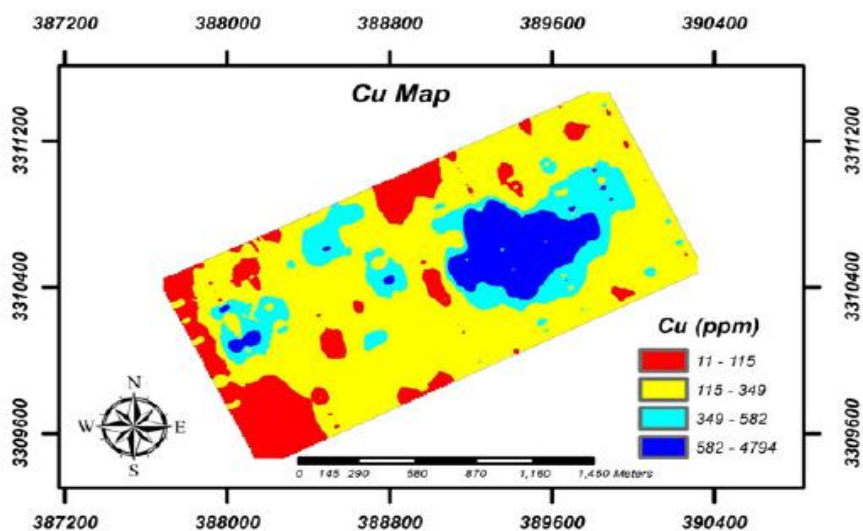
۳-۷-۱- کاربرد ماشین‌های بردار پشتیبان

در مقاله زیر اشاره‌ای به روش طبقه‌بندی نظارت شده ماشین‌بردار پشتیبان (SVM) برای تهیه نقشه پتانسیل معدنی کانسار مس پورفیری شده است. برای تهیه نقشه پتانسیل معدنی ذخیره مس پورفیری نوچون واقع در استان کرمان (شکل ۳-۱)، لایه‌های مختلف اطلاعاتی زمین‌شناسی، ژئوفیزیک و ژئوشیمی جمع‌آوری شده است.

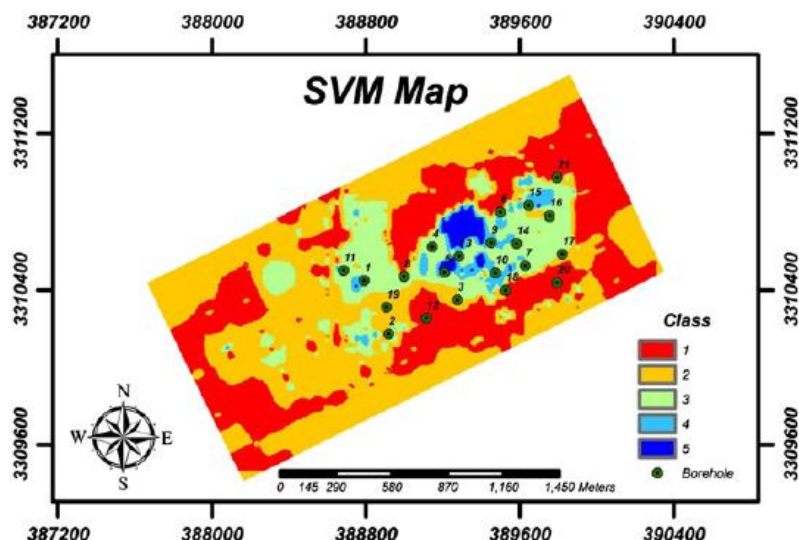


شکل ۳-۱: نقشه زمین‌شناسی معدن مس پورفیری نوچون به همراه موقعیت آن در نقشه ایران (Abedi et al., 2012).

روش SVM که یک روش مبتنی بر تشخیص الگو است، نرخ صحیح طبقه‌بندی را برای داده‌های بیست و یک گمانه که در پنج کلاس طبقه‌بندی شده‌اند ۵۲/۳۸ درصد نشان داد. نتایج حاصل از این مطالعه قابلیت SVM به عنوان یک ابزار الگوریتم یادگیری نظارت‌شده برای پیش‌بینی نقشه‌های پتانسل معدنی نشان می‌دهد. در این تحقیق، ابتدا غلظت مس (شکل ۲-۳) به نقشه درآورده شد و سپس نقشه حاصل از طبقه‌بندی داده‌ها با استفاده از SVM (شکل ۳-۳) تهیه شده است. با توجه به نقشه غلظت مس و نقشه بدست آمده از الگوریتم SVM، مشاهده می‌شود که در نقشه تهیه شده توسط SVM، جزئیات و پیوستگی بهتری از مناطق آنومال و هاله قابل شناسایی است. در مطالعات دقیق، طبقه‌بندی چندگانه مناطق معدنی منجر به افزایش وضوح نقشه‌های پتانسیل معدنی و همچنین کاهش ریسک عملیات حفاری می‌شود (Abedi et al., 2012).

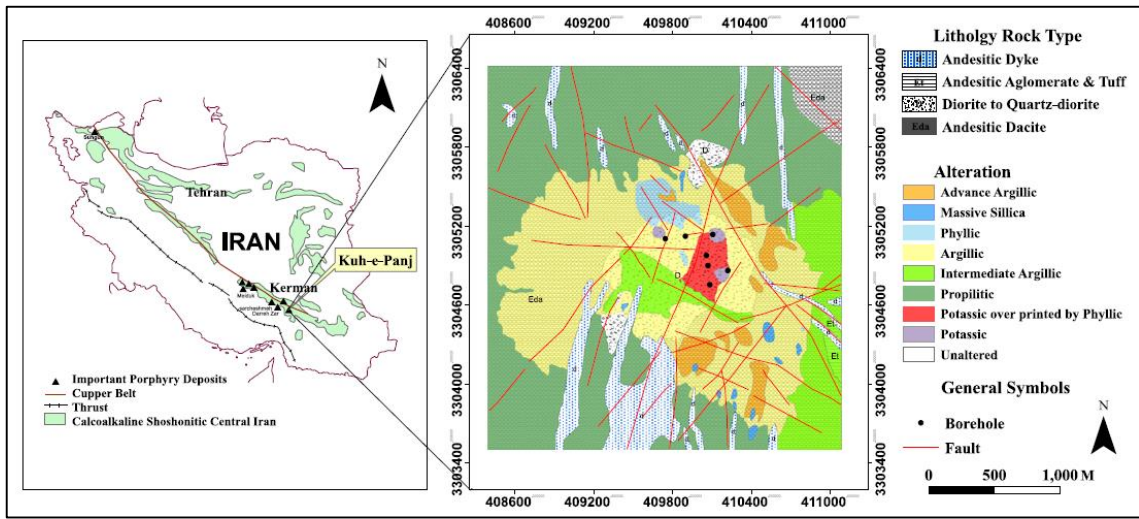


شکل ۲-۳: نقشه غلظت مس (Abedi et al., 2012)

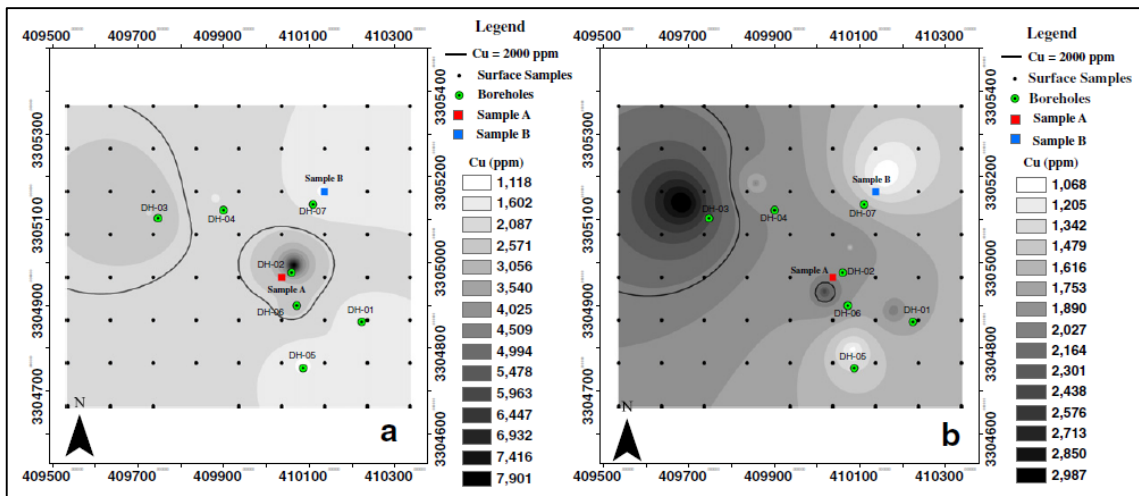


شکل ۳-۳: نقشه پتانسیل معدنی حاصل از تلفیق داده‌های اکتشافی بوسیله SVM به همراه شماره گمانه‌ها (Abedi et al., 2012)

در مطالعه‌ای که در معدن مس پورفیری منطقه کوه‌پنج توسط محمدی گنبدی و همکاران صورت گرفته، از برخی از الگوریتم‌های طبقه‌بندی جهت تشخیص آنومالی‌های نظارت شده استفاده گردیده است. با توجه به نقشه زمین‌شناسی و آلتراسیون‌های مختلف موجود در محدوده مورد مطالعه (شکل ۳-۴) انواع آلتراسیون‌های آرژیلیتیک و پروپلیتیک بیشترین گستردگی را در محدوده مورد مطالعه نشان می‌دهند. نقشه سطحی غلظت مس با استفاده از نمونه‌های سطحی ژئوشیمیایی منطقه رسم شده است (شکل ۳-۵ a) و نقشه عمقی غلظت مس با بهره‌گیری از نتایج آنالیز مغزه‌های حفاری و با درونیابی کردن غلظت مس بین گمانه‌ها با استفاده از روش کریجینگ شاخص بدست آمده است (شکل ۳-۵ b). با بررسی نقشه‌های زمین‌شناسی، آلتراسیون‌ها و نقشه‌های ژئوشیمیایی سطحی و عمقی، مشاهده می‌شود که غنی‌شدگی مس در زون آرژیلیتیک و پروپلیتیک صورت پذیرفته است (Mohammadi Gonbadi et al., 2015).

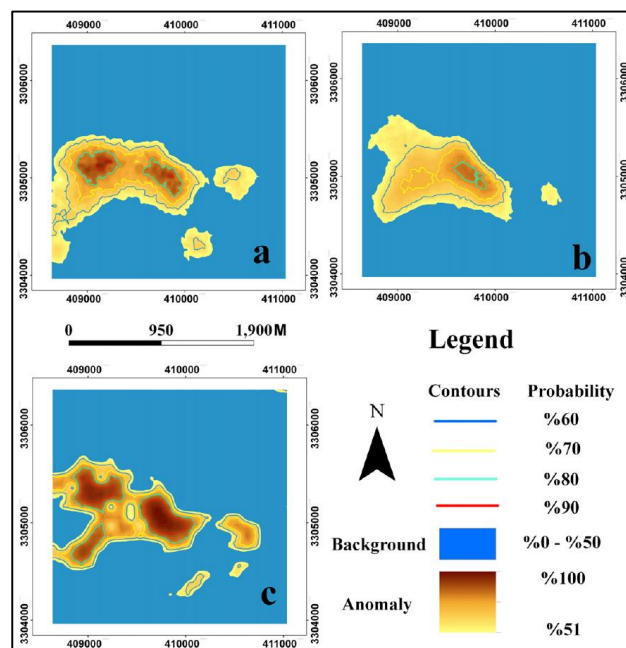


شکل ۳-۴: نقشه زمین‌شناسی و آتراسیون منطقه کانی‌زایی مس پورفیری کوه‌پنج (Mohammadi Gonbadi et al., 2015)



شکل ۳-۵: نقشه پتانسیل مس (a) سطحی در افق ۲۷۲۰ متری و (b) عمقی در افق ۲۵۹۰ متری (Mohammadi Gonbadi et al., 2015)

الگوریتم‌های آدا‌بوست^{۸۱} (ADB)، ماشین‌های بردار پشتیبان (SVM) و جنگل‌های تصادفی^{۸۲} (RF) با استفاده از نتایج آنالیزهای نمونه‌های سنگی سطحی و نمونه‌های حاصل از گمانه‌های حفر شده در منطقه مورد مطالعه برای ایجاد یک نقشه طبقه‌بندی شده آموزش‌دیده، که این نقشه مناطق آنومال را در قسمتهایی از محدوده مورد مطالعه که حفاری انجام نشده بود را نشان می‌دهد (شکل ۳-۶). نتایج نشان می‌دهد که الگوریتم‌های انتخاب ویژگی می‌توانند نقش مهمی در افزایش دقت و تعمیم‌توانایی طبقه‌بندی استفاده شده داشته باشند (Mohammadi Gonbadi et al., 2015).



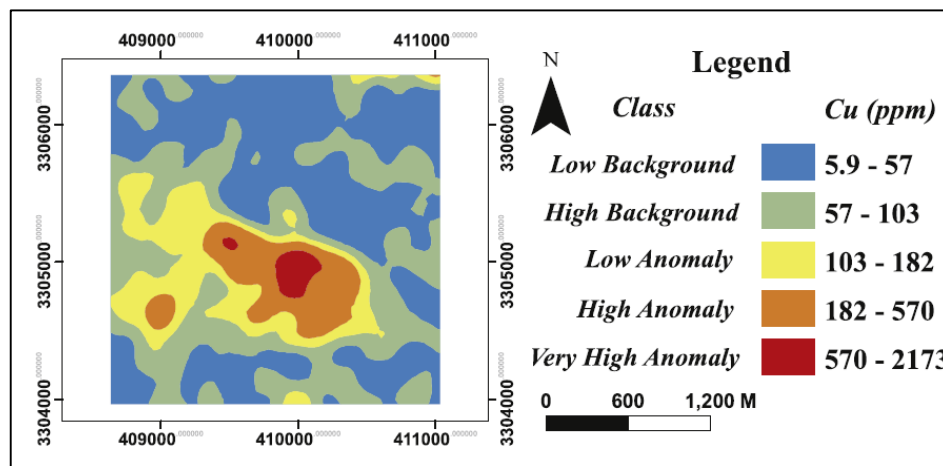
شکل ۳-۶: نقشه‌های طبقه‌بندی شده مناطق آنومال پیش‌بینی شده توسط الگوریتم‌های مختلف (a) نتیجه الگوریتم آدا‌بوست (b) نتیجه الگوریتم جنگل تصادفی (c) نتیجه الگوریتم ماشین‌های بردار پشتیبان (Mohammadi Gonbadi et al., 2015).

در بین روش‌های طبقه‌بندی، آدا‌بوست با میزان خطای ۰/۰۶ بهترین نتیجه را حاصل کرد (جدول ۳-۱).

جدول ۳-۱: میزان خطای الگوریتم آدابوست در ترکیب با دیگر الگوریتم‌های انتخاب ورودی منطقه کوه‌پنج (Mohammadi et al., 2015)

LOO			Error Probability			Classifier	Feature Selection	
T	BG	AN	T	BG	AN		Selected Features	Method
0.29	0.38	0.22	0	0	0	ADB	8 transformed feature	PCA
0.32	0.38	0.28	0	0	0	ADB	Cr, Cu, K, Mo, W, Zr	CFS + GA
0.32	0.38	0.28	0	0	0	ADB	Cr, K, Mo, Zr	CFS + Float
0.06	0.15	0	0	0	0	ADB	Ag, Al, As, Cu, Fe, Li, Mn, Mo, Sb, Sr, Th, V, W	WRP + GA
0.22	0.23	0.22	0	0	0	ADB	Ag, Al, Au, Ce, Co, Cr, Cu, Fe, Li, Mo, Sr, Te	WRP + Float
0.35	0.38	0.33	0	0	0	ADB	-	-

در همین حال، مقایسه نقشه‌های طبقه‌بندی شده با استفاده از آدابوست با نقشه پراکندگی غلظت عنصر Cu منطقه با روش فرکتال (شکل ۳-۷) و با توجه به خطای قابل قبول آدابوست که در حدود ۰/۰۶ نشان می‌دهد، نقشه آدابوست دارای تشخیص بالا و همچنین آینده‌نگری بهتری نسبت به فرکتال می‌باشد.

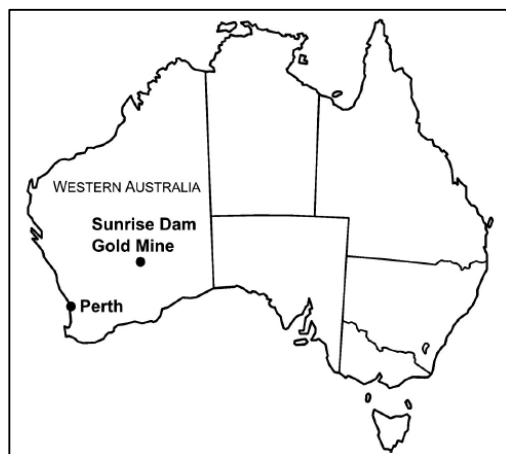


شکل ۳-۷: جداسازی آنومالی از زمینه مس با استفاده از مدل فرکتال (Mohammadi Gonbadi et al., 2015).

۲-۷-۳- کاربرد طبقه‌بندی KNN

ذخایر طلا با کانی‌سازی‌های رگه‌ای هیدروترمالی معمولاً به صورت قطعه‌ای هستند، به این معنی که توزیع فضایی طلا به صورت ناپیوسته است (اثر قطعه‌ای^{۸۳} زمین‌شناسی) (Dominy et al., 2003) و این باعث می‌شود که کار زمین‌شناسان و مهندسی‌ن در شناسایی و ایجاد مرزهای پیوسته کانسار در حفاری مغزه‌گیری و اتصال دستی مرزهای کانسار در گمانه‌ها به هم و ساخت مدل سه بعدی کانسار، دشوار گردد (Hill et al., 2014). زمین‌شناسان و مهندسی‌ن در تلاش هستند تا با ترکیب اطلاعات اکتشافی، مشکل جداسازی افق‌های پیوسته کانی‌سازی و غیرکانی‌سازی را حل کنند (Dominy et al., 2003).

معدن طلای سانرایس‌دم^{۸۴} استرالیا، یک ذخیره طلای در سطح جهانی دارای اثر قطعه‌ای بسیار بالا است (شکل ۸-۳). داده‌های ژئوشیمیایی چندعنصری این معدن، به منظور بهبود پیش‌بینی مناطق کانسار، جمع‌آوری شده است. عناصر پاراژنز^{۸۵} مناسب از این مجموعه داده‌ها انتخاب شدند، به ویژه عناصری که ارتباط فضایی خوبی با کانی‌سازی طلا نشان می‌دهند و در مقابل، همانند طلا توزیع قطعه‌ای نداشتند (عناصری مانند Cr و Rb، Sb).

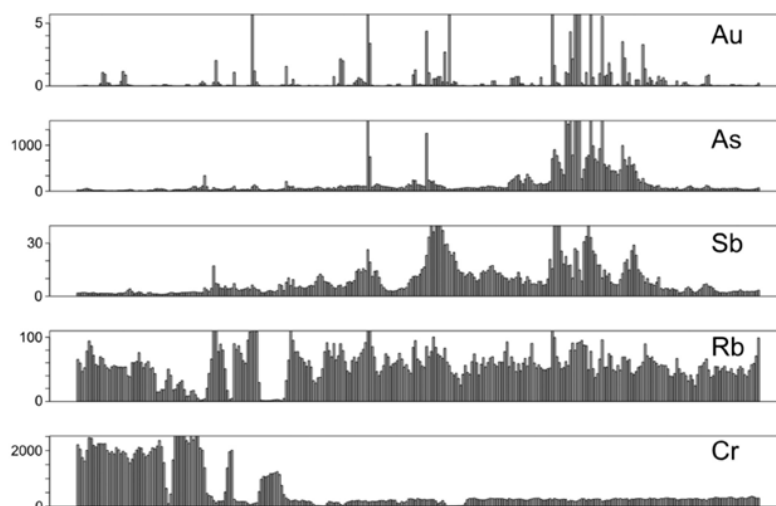


شکل ۸-۳: موقعیت جغرافیایی معدن طلای Sunrise Dam در غرب استرالیا (Hill et al., 2014).

83 Nugget effect
84 Sunrise Dam
85 Proxy

در این مطالعات، یک رویکردی جهت حل مشکل همبستگی بین ارزش طلا و سایر عناصر ژئوشیمیایی ارائه شده است. از سه روش نزدیک‌ترین همسایه K (KNN)، تخمین‌گر چگالی هسته‌ای^{۸۶} (KDE) و احتمال شرطی بیز^{۸۷} (NB) جهت بهبود طبقه‌بندی عناصر ژئوشیمیایی مرتبط با کانی‌سازی طلا و جداسازی مناطق آنومال و غیرآنومال استفاده شده است.

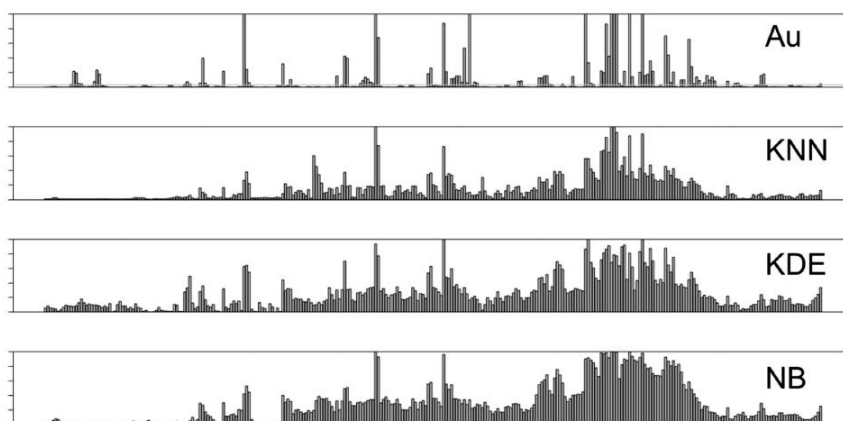
گمانه UGD1944 انتخاب گردید و غلظت عناصر طلا، آرسنیک، آنتیموان، روبیدیوم و کرم در طول گمانه به صورت نمودار میله‌ای رسم شد (شکل ۳-۹).



شکل ۳-۹: نمودار میله‌ای غلظت عناصر طلا، آرسنیک، آنتیموان، روبیدیوم و کرم در طول گمانه UGD1944 (۳۰۷ متر حفاری) نمودارهای بالا اثر قطعه‌ای طلا (عدم پیوستگی) و همچنین ارتباط فضایی طلا با سایر عناصر ژئوشیمیایی را نشان می‌دهد (Hill et al., 2014).

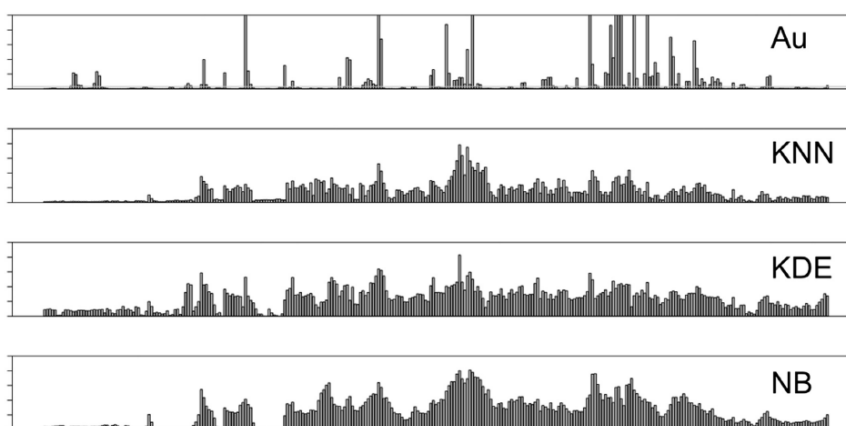
نمودار میله‌ای ترکیبی عناصر کروم، روبیدیوم و آرسنیک در مقابل طلا با استفاده از سه الگوریتم ذکر شده در قسمت اول رسم شده است (شکل ۳-۱۰).

86 Kernel density estimator
87 Bayes conditional probability



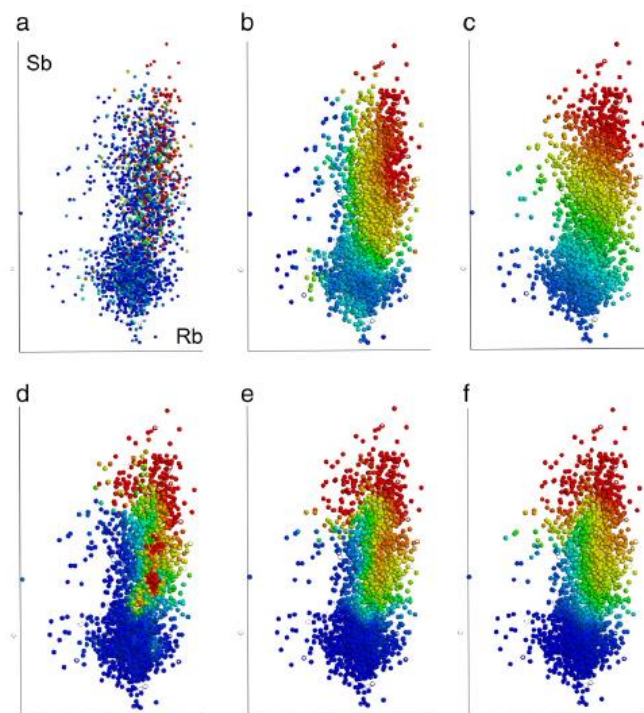
شکل ۳-۱۰: مقایسه مقدار طلا در گمانه UGD1944 در مقایسه با نتایج الگوریتم‌های سه گانه برای عناصر کروم، روبیدیم و آرسنیک. (a) طلا، (b) الگوریتم نزدیک‌ترین K همسایه (KNN) با $K=501$ ، (c) تخمین گر چگالی هسته‌ای (KDE) و (d) احتمال شرطی بیز (NB) (Hill et al., 2014).

در مرحله بعدی هنگامی که عنصر آنتیموان به جای آرسنیک جایگزین شد نتایج نسبت به قبل بهبود یافت و دلیل آن هم همبستگی فضایی بیشتر آنتیموان با عنصر طلا در منطقه کانه‌زایی فوق می‌باشد (شکل ۳-۱۱) (Hill et al., 2014).



شکل ۳-۱۱: مقایسه مقدار طلا در گمانه حفاری UGD1944 در مقایسه با نتایج الگوریتم‌های سه گانه برای عناصر کروم، روبیدیم و آنتیموان. (a) طلا، (b) الگوریتم نزدیک‌ترین K همسایه (KNN) با $K=501$ ، (c) تخمین گر چگالی هسته‌ای (KDE) و (d) احتمال شرطی بیز (NB) (Hill et al., 2014).

پس از نرمال‌سازی، داده‌های آنتیموان و روبیدیم، نمودار نقطه‌ای آنتیموان در مقابل روبیدیم رسم گردید و براساس مقدار عیار طلا و همچنین الگوریتم‌های سه گانه ذکر شده، کلاس‌بندی صورت گرفت (نتایج در شکل ۳-۱۲ ارائه شده است).

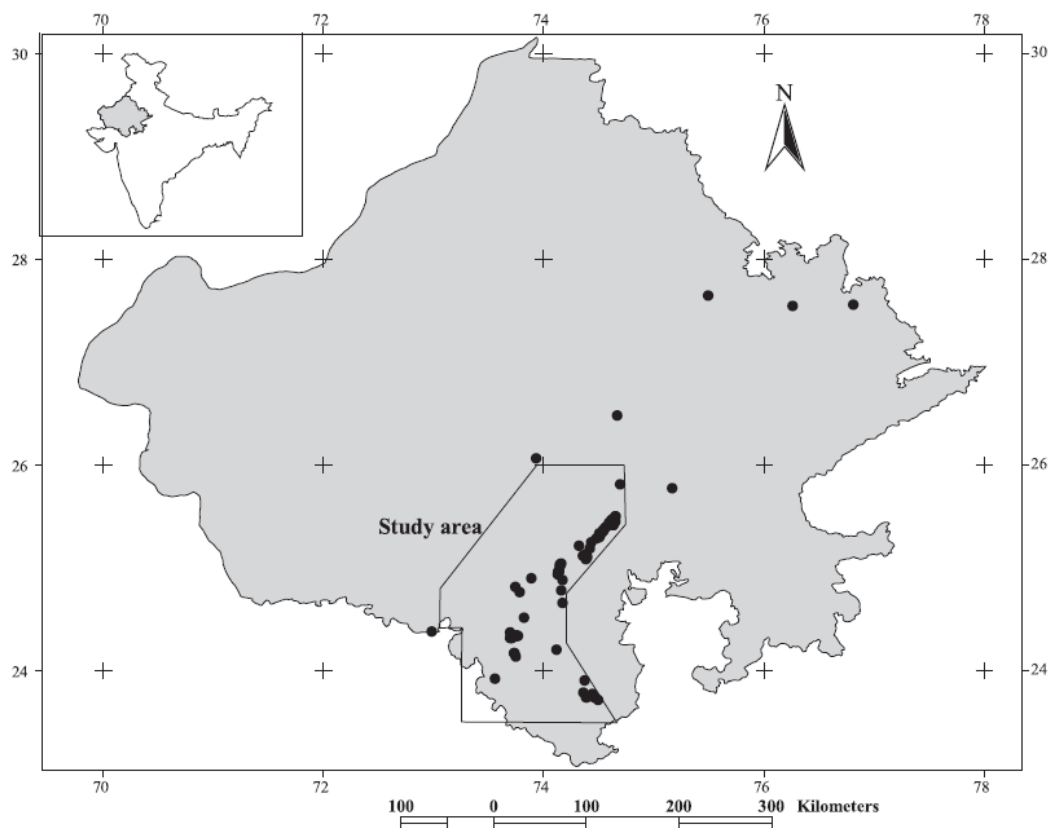


شکل ۳-۱۲: نقشه نقطه‌ای آنتیموان در مقابل روبیدیم که بوسیله پارامترهای مختلف رنگ‌گذاری شده اند، (a) مقدار عیار طلا، (b) تخمین گر چگالی هسته‌ای (KDE)، (c) احتمال شرطی بیز (NB)، (d) الگوریتم نزدیک‌ترین K همسایه (KNN) با $K=101$ ، (e) الگوریتم نزدیک‌ترین K همسایه (KNN) با $K=501$ ، (f) الگوریتم نزدیک‌ترین K همسایه (KNN) با $K=1001$ (Hill et al., 2014).

بر اساس مطالعات انجام شده مشاهده می‌شود که نتایج حاصل از طبقه‌بندی با الگوریتم نزدیک‌ترین K همسایه با تغییرات مقدار عددی K نتایج متفاوتی نشان می‌دهد، به طوری که واریانس‌های بدست آمده برای مقادیر K، ۱۱، ۱۰۱، ۵۰۱ و ۱۰۰۱ به ترتیب برابر با ۱۱، ۲/۱، ۰/۷ و ۰/۵ بود که نشان از بهتر شدن طبقه‌بندی با افزایش مقدار K می‌باشد (Hill et al., 2014).

۳-۷-۳- کاربرد روش‌های بیزین

در مطالعه‌ای که توسط پوروال^{۸۸} و همکاران (۲۰۰۶) بر روی ذخایر فلزات پایه در هند انجام پذیرفت، از سه طبقه‌بندی بیزی برای تهیه نقشه‌های پتانسیل معدنی استفاده شد که عبارتند از: الف) طبقه‌بندی ساده بیزی بر اساس فرض استقلال مشروط کامل الگوهای پیش‌بینی ورودی. ب) طبقه‌بندی بیزی ساده تقویت شده که وابستگی مشروط در میان الگوهای پیش‌بینی ورودی را به رسمیت می‌شناسد و ج) طبقه‌بندی ساده انتخابی که تنها از الگوهای پیش‌بینی مشروط مستقل استفاده می‌کند.

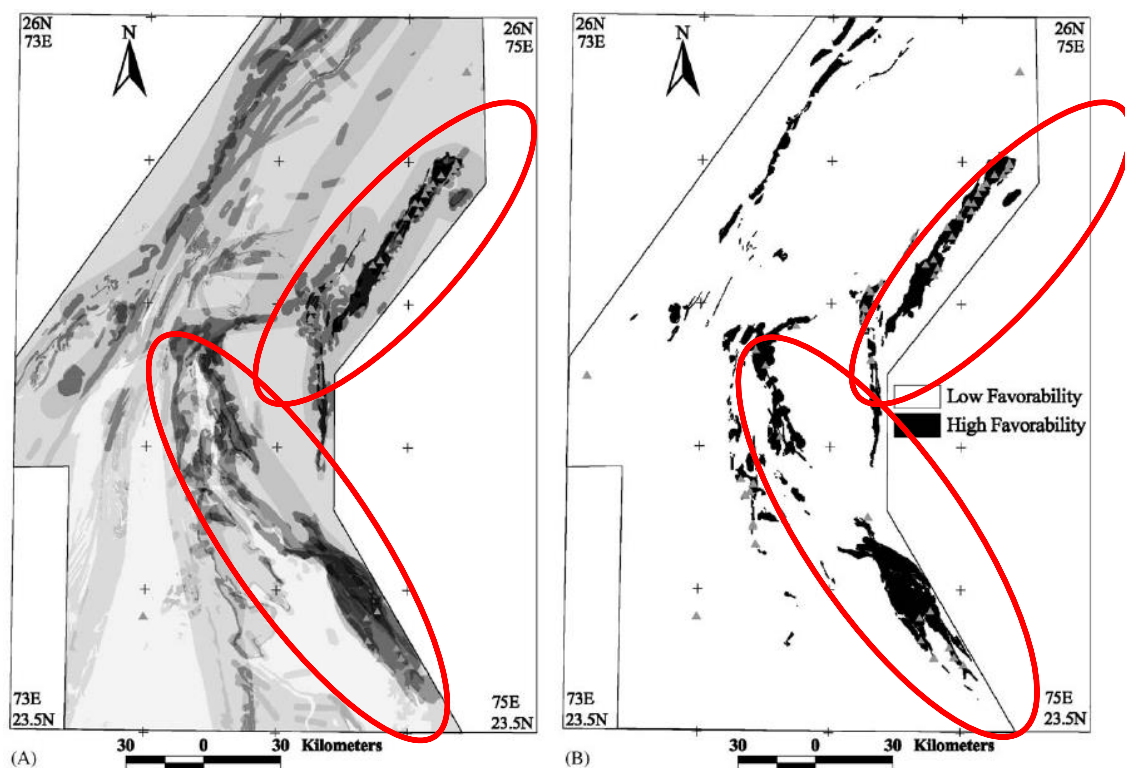


شکل ۳-۱۳: موقعیت محدوده مورد مطالعه در استان راجستان هند. دایره‌های سیاه نشانگر ذخایر فلزات پایه است (Porwal et al., 2006).

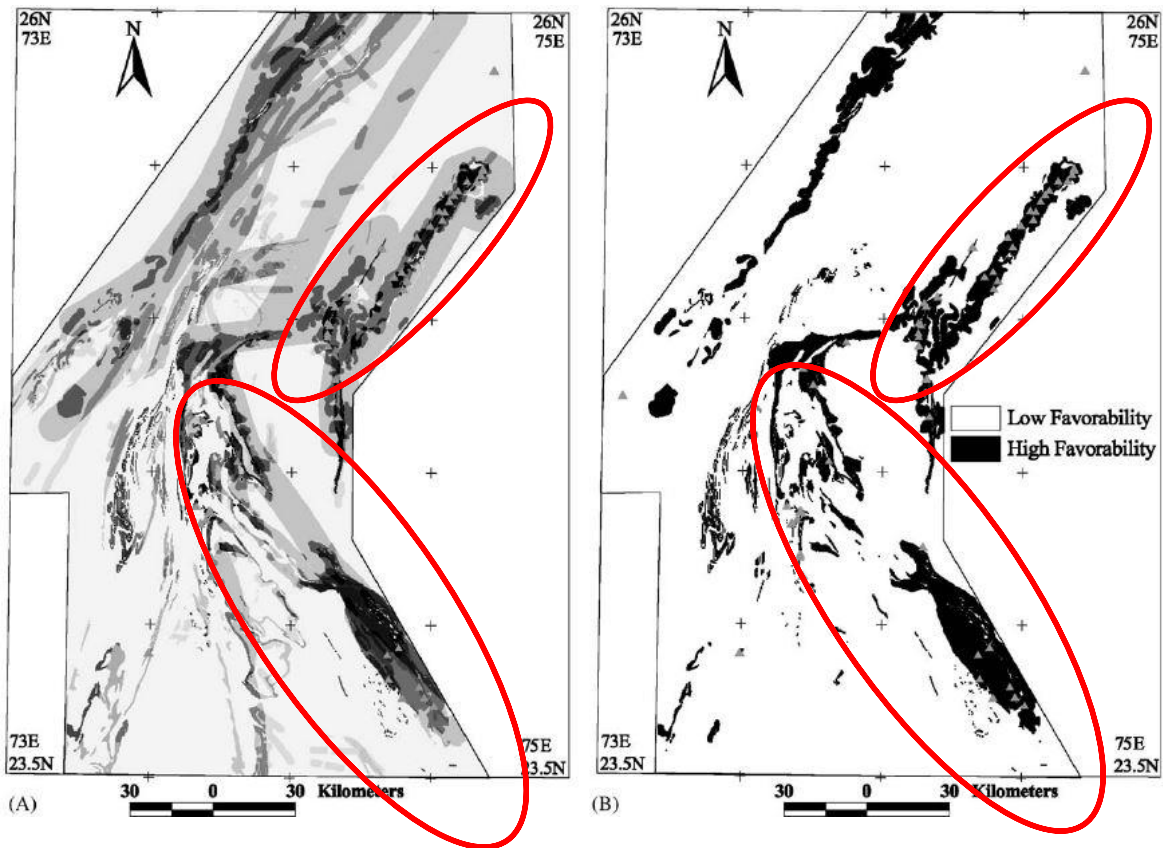
خروجی طبقه‌بندی‌های آموزش دیده، تعلق یافتن یک بردار ورودی را برای کلاس‌های حاوی کانی‌سازی و غیرکانی‌سازی مشخص می‌کند.

در این مطالعه، بردارهای ویژگی شناخته شده برای ۵۴ نمونه از ذخایر فلزات پایه موجود در منطقه و تعداد ۵۴ مورد از بردارهای ویژگی ذخایر غیرمعدنی که در ارتباط با کانسارهای فلزی شناخته شده بودند جمع آوری شده است که در مجموع تعداد داده‌های مورد مطالعه برابر با ۱۰۸ نمونه شد. در تهیه نقشه پتانسیل معدنی، متغیرها به صورت باینری "کانی‌سازی" و "غیرکانی‌سازی" برچسب‌گذاری شدند (شکل ۳-۱۳).

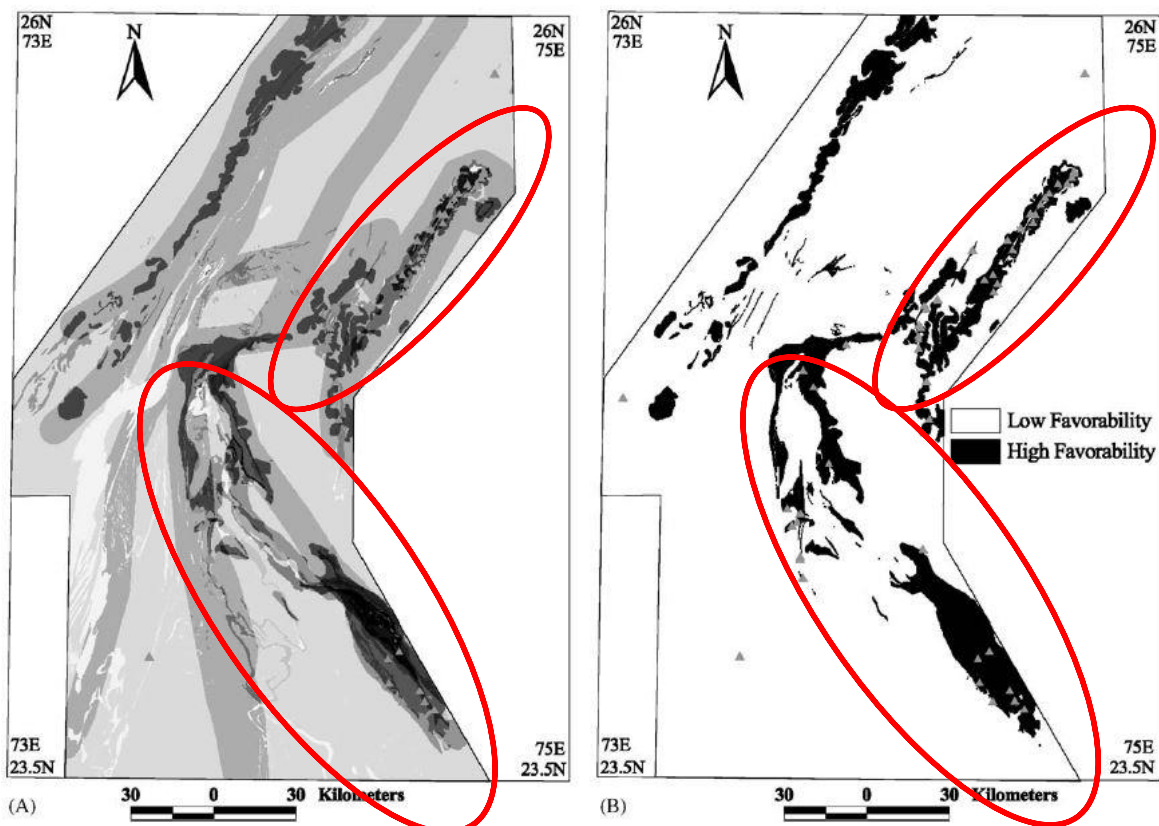
نتایج حاصل از طبقه‌بندی‌های مختلف، طبقه‌بندی ساده بیزی (شکل ۳-۱۴)، طبقه‌بندی بیزی ساده تقویت شده (شکل ۳-۱۵) و طبقه‌بندی ساده‌انتخابی (شکل ۳-۱۶) قابل مشاهده است.



شکل ۳-۱۴: نقشه پتانسیل تولید شده با استفاده از طبقه‌بندی بیزی ساده، (A) نقشه پتانسیل در مقیاس خاکستری (B) نقشه پتانسیل دودویی تولید شده. مثلث‌های خاکستری ذخایر فلزات پایه را نشان می‌دهند (Porwal et al., 2006).



شکل ۳-۱۵: نقشه پتانسیل تولید شده با استفاده از طبقه‌بندی بیزین ساده تقویت‌شده ، (A) نقشه پتانسیل در مقیاس خاکستری (B) نقشه پتانسیل دودوئی تولید شده. مثلث‌های خاکستری ذخایر فلزات پایه را نشان می‌دهند (Porwal et al., 2006).



شکل ۳-۱۶: نقشه پتانسیل تولید شده با استفاده از طبقه‌بندی بی‌زین ساده انتخابی ، (A) نقشه پتانسیل در مقیاس خاکستری (B) نقشه پتانسیل دودوئی تولید شده. مثلث‌های خاکستری ذخایر فلزات پایه را نشان می‌دهند (Porwal et al., 2006).

جدول ۳-۲ نشان می‌دهد که روش طبقه‌بندی ساده تقویت شده بهترین عملکرد را نشان می‌دهد و بعد از آن روش بی‌زین ساده و سپس روش بی‌زین انتخابی عملکرد مناسب‌تری نشان می‌دهند (Porwal et al., 2006).

جدول ۳-۲: نتایج اعتبارسنجی (Porwal et al., 2006)

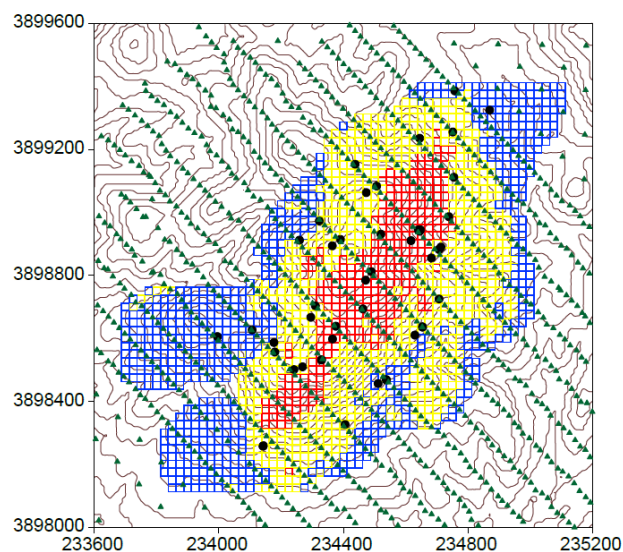
% of correctly classified deposits	Classifier
86.8	Naive
88.7	Augmented naive
83	Selective naive

جدول ۳-۳ نشان می‌دهد که ۷ درصد از کل ۳۴۰۰ کیلومتر مربع منطقه مطالعاتی که توسط روش بیزی ساده به عنوان منطقه پتانسیل دار شناسایی شده است، شامل ۸۸/۹ درصد از ذخایر شناخته شده منطقه است. روش بیزی ساده تقویت شده ۱۱/۳ درصد از کل منطقه مورد مطالعه را به عنوان محدوده پرتانسیل شناسایی کرده است که شامل ۹۲/۶ درصد از ذخایر شناخته شده منطقه می‌باشد و روش بیزی ساده انتخابی، در حالی ۱۱/۲ درصد از کل منطقه مطالعاتی را به عنوان منطقه پتانسیل دار معرفی می‌کند که شامل ۸۳/۳ درصد از ذخایر شناخته شده منطقه می‌باشد. نتایج جدول ۳-۳ تایید کننده نتایج حاصل از جدول ۲-۳ می‌باشد.

جدول ۳-۳: نتایج اعتبارسنجی نقشه‌های پتانسیل (Porwal et al., 2006)

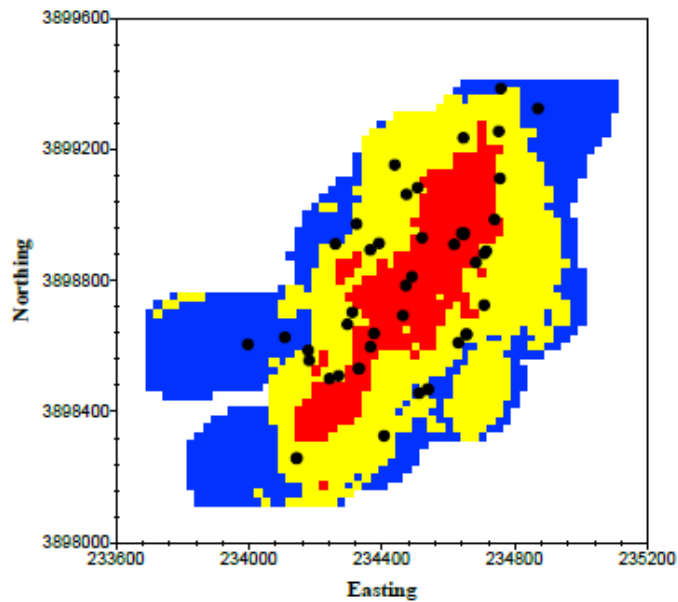
Percent of deposits (total deposits 54)	Percent of study area (total area: 34,000 Km ²)	Zone	Classifier
88.9	7.1	High favorability	Naive
11.1	92.9	Low favorability	
92.6	11.3	High favorability	Augmented naive
7.4	88.7	Low favorability	
83.3	11.2	High favorability	Selective naive
16.7	88.8	Low favorability	

گرانیان و همکاران (۱۳۹۲)، طبقه‌بندی داده‌های اکتشافی ژئوشیمیایی بر روی داده‌های گمانه‌ای و سطحی معدن طلای ساری‌گونی که شامل ۴۶ گمانه حفاری با آنالیز نمونه‌های از هر یک متر، و ۱۷۲۷ نمونه خاک از سطح که توسط ICP برای ۴۷ عنصر آنالیز شده انجام دادند (شکل ۳-۱۷).

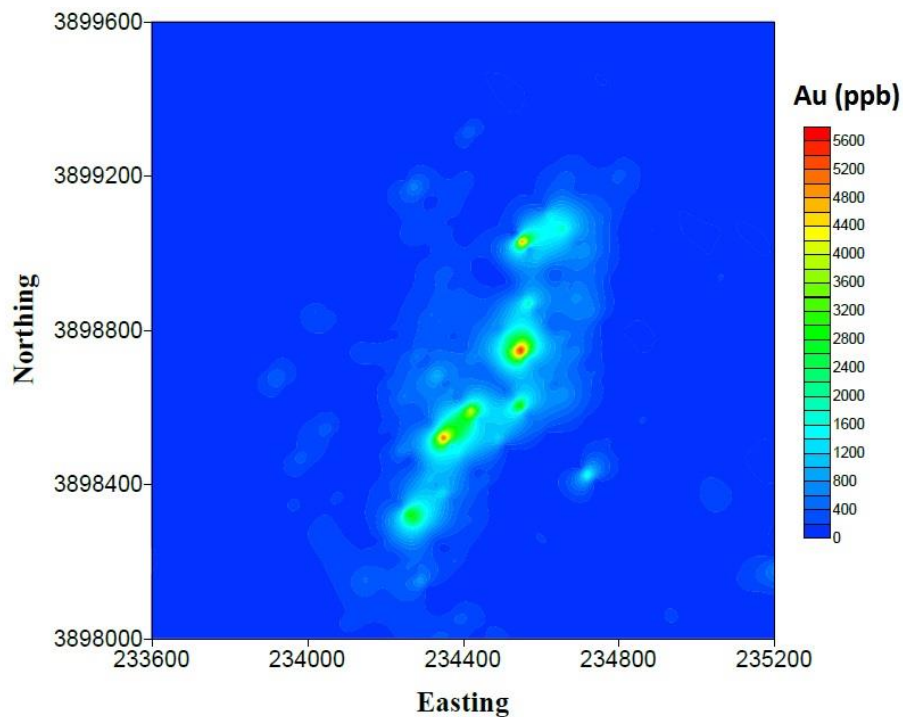


شکل ۳-۱۷: موقعیت گمانه‌ها با دوایر سیاه رنگ و نمونه‌های سطحی با مثلث‌های سبز رنگ در کانسار طلای ساری‌گونی و همچنین بلوک بندی سطحی داده‌ها با ابعاد 25×25 (طبقه A به رنگ قرمز (آنومالی)، طبقه B به رنگ زرد (پتانسیل دار) و طبقه C به رنگ آبی (زمینه)) (گرانیا و همکاران، ۱۳۹۲)

در این مطالعه با استفاده از داده‌های ژئوشیمیایی محیط خاکی و به کمک چهار روش طبقه بندی بیز، نزدیک‌ترین K همسایه، پنجره پارزن و بیزی ساده که همگی براساس تئوری بیزین بنا نهاده شده‌اند، موقعیت کانی‌زایی طلا در این کانسار مدل‌سازی شده است. پارامترهای شاخص تولید برای سلول‌های به ابعاد 25×25 متر در سطح زمین تعریف شده اند (شکل ۳-۱۸ و شکل ۳-۱۹).

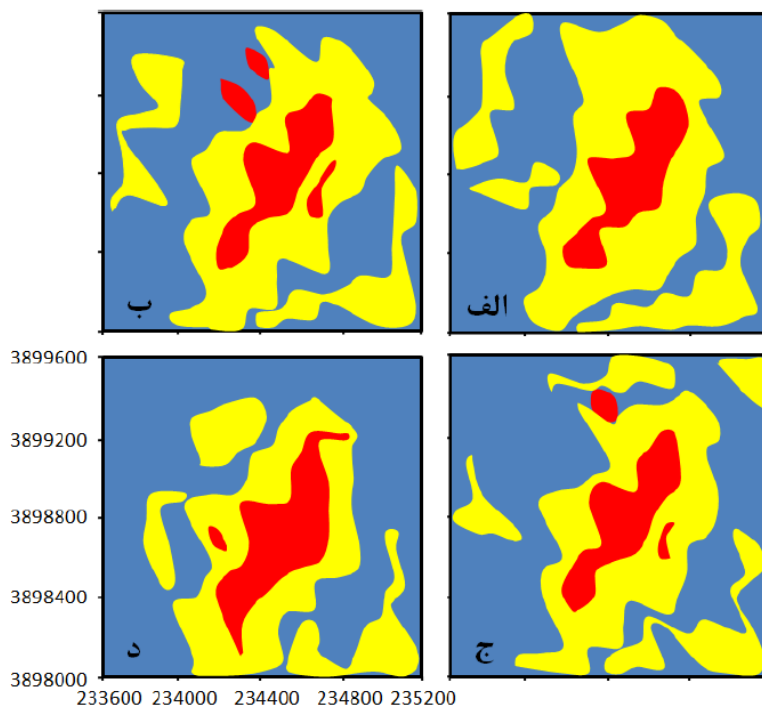


شکل ۳-۱۸: نقشه پراکندگی وضعیت کانی‌زایی طلا و موقعیت گمانه‌ها در کانسار ساری - گونئی (طبقه A به رنگ قرمز(آنومالی)، طبقه B به رنگ زرد(پتانسیل دار) و طبقه C به رنگ آبی(زمینه)) (گرانپان و همکاران، ۱۳۹۲)



شکل ۳-۱۹: نقشه کنتری طلا به دست آمده از داده‌های ژئوشیمیایی خاک (واحد اندازه گیری طلا در مقیاس میله‌ای ppb است) (گرانپان و همکاران، ۱۳۹۲)

۶۵ درصد داده‌ها به عنوان داده‌های آموزشی و ۳۵ درصد داده‌ها به عنوان داده‌های آزمون به کار رفته و با استفاده از روش‌های مذکور، مدل کانی‌زایی برای این منطقه بدست آمده است (شکل ۳-۲۰).



شکل ۳-۲۰: نقشه پراکندگی مدل‌سازی طلا در سه طبقه به روش الف) بی‌زی، ب) نزدیک‌ترین K همسایه، ج) روش پنجره پارزن و د) روش بی‌زی ساده در منطقه ساری‌گونئی (ترکیب رنگی‌های انتخاب شده مشابه شکل ۳-۱۹ است) (گرانیان و همکاران، ۱۳۹۲).

نتایج پیش‌بینی شاخص تولید توسط مقادیر طلای داده‌های ژئوشیمیایی محیط‌خاکی نشان می‌دهد که فقط امکان تعیین مقادیر با شاخص تولید بالا همراه با کمی جابجایی ممکن است. در حالی که به کمک روش‌های طبقه‌بندی بی‌زین افزون بر تخمین همه مناطق کانی‌زایی شده، اعتبار تخمین نیز در حد قابل قبولی افزایش می‌یابد. بنابراین استفاده از گمانه‌های حفر شده و ارتباط دادن آن‌ها با داده‌های ژئوشیمیایی سطحی به روش‌های طبقه‌بندی می‌تواند در موقعیت‌یابی گمانه‌های تکمیلی بهتر از نتایج تعیین آنومالی طلا عمل کند.

هر چهار روش طبقه‌بندی مورد استفاده قادر به تفکیک نواحی با کیفیت کانی‌زایی مختلف طلا در ساری - گونئی هستند. اما روش طبقه‌بندی بیزی با درستی ۷۲/۶ درصد و روش پنجره پارزن با درستی ۷۰/۴ درصد در طبقه‌بندی داده‌های آزمایشی و تعیین دقیق‌تر شکل و وسعت نواحی مختلف کانی‌زایی نسبت به دو روش دیگر عملکرد بهتری داشته‌اند (جدول ۳-۴) (گرانیان و همکاران، ۱۳۹۲).

جدول ۳-۴: نتایج طبقه‌بندی داده‌های آزمایشی به چهار روش نظریه بیزین

روش بیزی ساده			روش پنجره پارزن			روش نزدیک‌ترین K همسایه			روش بیزی			گروه طبقه‌بندی شده
گروه اولیه			گروه اولیه			گروه اولیه			گروه اولیه			
C	B	A	C	B	A	C	B	A	C	B	A	
۱	۱۷	۳۵	۰	۱۳	۳۴	۱	۱۴	۳۳	۱	۱۲	۳۵	A
۱۰	۳۵	۱۰	۹	۴۷	۱۳	۱۰	۴۲	۱۲	۸	۴۹	۱۲	B
۱۲	۱۳	۲	۱۴	۵	۰	۱۲	۹	۲	۱۴	۴	۰	C
%۵۲/۲	%۵۳/۸	%۷۴/۵	%۶۰/۹	%۷۲/۳	%۷۲/۳	%۵۲/۲	%۶۴/۶	%۷۰/۱۲	%۶۰/۹	%۷۵/۴	%۷۴/۵	صحت طبقه‌بندی هر گروه
%۶۰/۱۷			%۷۰/۱۴			%۶۴/۴			%۷۲/۶			صحت طبقه‌بندی هر روش

فصل

۴- بهبود دقت تخمین غلظت عناصر ژئوشیمیایی

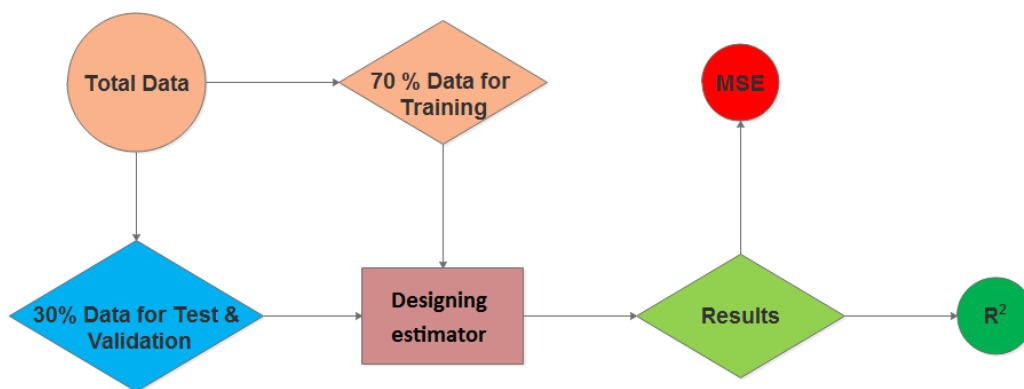
یکی از مشکلات اساسی در تحلیل‌های اکتشافی ذخایر معدنی، کمبود داده‌های اکتشافی است. تلاش مهندسان و محققان بر این است تا بتوانند بیشترین اطلاعات را از کمترین داده‌های موجود کسب نمایند. برای این منظور با استفاده از روش‌های تحلیل داده‌ها، سعی در شناسایی خصوصیات و ویژگی‌های مناطقی از محدوده مورد اکتشاف هستیم که دارای اطلاعات اکتشافی ناقص و در مواردی عدم وجود اطلاعات هستند. روش‌های شناسایی الگو از جمله روش‌هایی هستند که توانایی شناخت الگوهای پنهان موجود در داده‌های اکتشافی را داشته و می‌توان الگوی مناطق مشاهده شده را شناسایی و به مناطق مشاهده نشده تعمیم داد. شبکه عصبی چند لایه یکی از روش‌های تشخیص الگو است که توانایی شناسایی روابط غیرخطی موجود بین داده‌های ورودی و خروجی را دارد و با استفاده از این روش می‌توان غلظت عناصر ژئوشیمیایی را در مناطقی که فاقد اطلاعات هستند تخمین زد. نکته‌ای که باید به آن توجه داشت، دقت تخمین غلظت عناصر بوده که هر چه دقت تخمین بیشتر باشد اطلاعات بدست آمده دارای کیفیت قابل قبول‌تری خواهند بود. سه پارامتر خطا، دقت و صحت در تخمین داده‌ها از اهمیت بالایی برخوردار هستند. خطا، تفاوت مقدار اندازه‌گیری شده با مقدار واقعی، دقت، نزدیک بودن مقادیر اندازه‌گیری شده به هم و صحت، نزدیک بودن مقادیر اندازه‌گیری شده به مقدار واقعی هستند.

برای افزایش دقت تخمین عناصر می‌توان از خوشه‌بندی داده‌های اکتشافی استفاده کرد. داده‌های با بیشترین همبستگی بین نمونه‌ای درون خوشه‌های مشابه قرار می‌گیرند و الگوریتم‌های خوشه‌بندی توانایی شناسایی این خوشه‌ها را دارند. استفاده از داده‌های خوشه‌بندی شده در آموزش تخمینگر شبکه عصبی باعث می‌شود داده‌های ورودی و خروجی تخمینگر ارتباط بهتری برقرار کنند و در نتیجه احتمال افزایش دقت تخمین وجود دارد. این شرایط در مطالعات قبلی در نظر گرفته نشده است و نتایج بدست آمده از دقت پایینی برخوردار بوده است. در این مطالعه، برای ارزیابی تعداد خوشه بهینه داده‌های موجود از

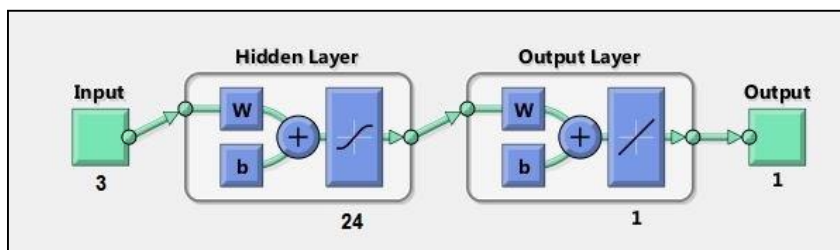
شاخص‌های اعتبارسنجی کمک گرفته شده است. با استفاده از الگوریتم‌های خوشه‌بندی، خوشه‌بندی انجام و تخمینگر شبکه عصبی بوسیله داده‌های خوشه‌بندی شده طراحی گردید و برای تخمینگر غلظت عناصر ژئوشیمیایی از آن‌ها استفاده شده است.

۲-۴- طراحی تخمینگر بر اساس کل داده‌ها برای عناصر ژئوشیمیایی

تخمینگر مورد استفاده برای تخمین غلظت عناصر ژئوشیمیایی شبکه عصبی پرسپترون چندلایه (MLP) دارای یک لایه پنهان و ۲۴ نرون، که با سعی و خطا انتخاب شده، می‌باشد. تابع آموزش این شبکه لونیگ - مارکوات (LM)، به دلیل عملکرد بهتر، انتخاب شده است. کل داده‌های موجود دارای تعداد ۱۷۵۵ نمونه با ۲۲ مؤلفه (غلظت عناصر ژئوشیمیایی) بود که از این تعداد ۷۰ درصد به عنوان آموزش و ۳۰ درصد به عنوان آزمون انتخاب گردید. مؤلفه‌های ورودی شبکه شامل X ، Y و Z بود که مختصات نقاط نمونه‌برداری است و مؤلفه خروجی شبکه، غلظت هر کدام از عناصر ژئوشیمیایی برای نمونه‌های مختلف است. برای بررسی دقت تخمینگر از ضریب تعیین (R^2) و برای بررسی خطای تخمین از میانگین توان دوم خطا (یا به اختصار MSE) استفاده شده است (شکل ۱-۴ و شکل ۲-۴).

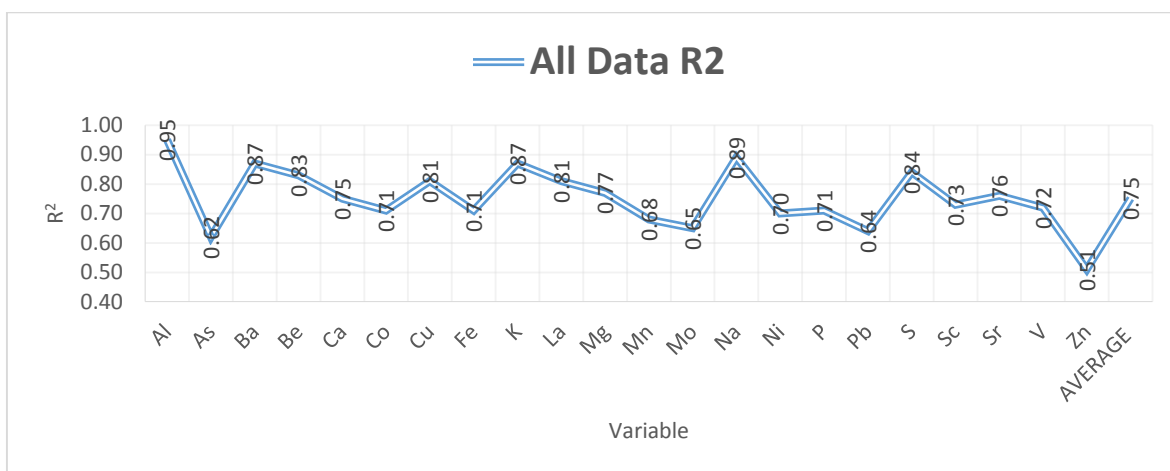


شکل ۱-۴: طراحی تخمینگر با استفاده از کل داده‌ها برای عناصر ژئوشیمیایی

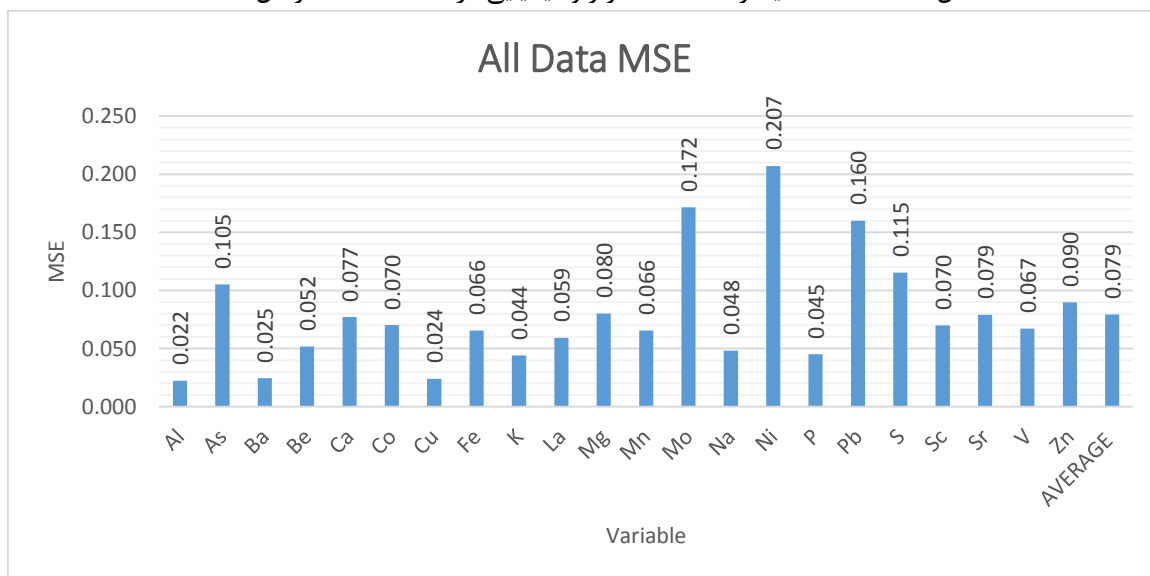


شکل ۴-۲: مدل طراحی تخمینگر شبکه عصبی MLP

دقت و خطای تخمینگر بدست آمده از تخمینگرهای تمامی عناصر ژئوشیمیایی در حالت استفاده از کل داده‌ها در شکل ۴-۳ و شکل ۴-۴ نشان داده شده است.



شکل ۴-۳: دقت تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از کل داده‌ها

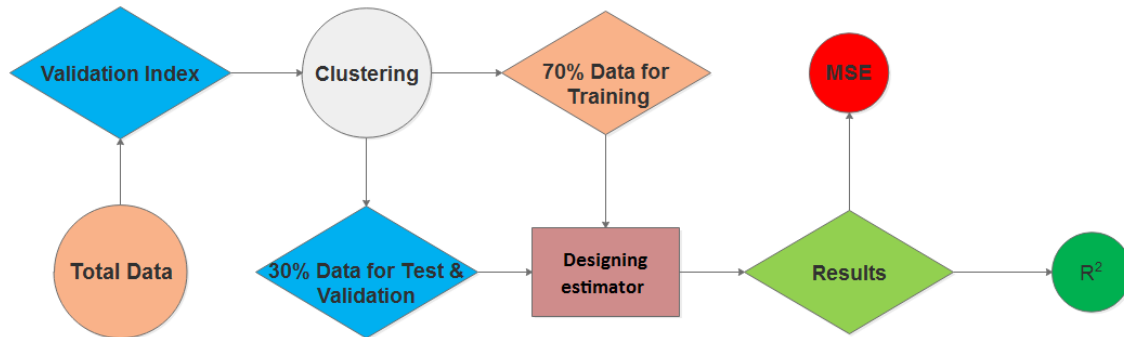


شکل ۴-۴: خطای تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از کل داده‌ها

بررسی نتایج نشان می‌دهد که بیشترین دقت مربوط به عنصر آلومینیوم با ۹۵ درصد و کمترین دقت مربوط به عنصر روی با ۵۱ درصد است. میانگین دقت تخمین‌گرهای ۲۲ عنصر برابر ۷۵ درصد است که دقت خوبی است اما دقت بعضی از عناصر مانند Zn, As, Pb, Mo و Mn که به ترتیب برابر ۰/۵۱، ۰/۶۲، ۰/۶۴، ۰/۶۵ و ۰/۶۸ هستند که دقت کمی محسوب می‌شوند. مطالعات ژئوشیمیایی بر اساس این دقت‌ها نتایج قابل قبولی ارائه نخواهد داد بنابراین باید با استفاده از یک سری راهکارها دقت را افزایش دهیم که برای این کار از الگوریتم‌های خوشه‌بندی K means, FCM, GK و GG استفاده خواهد شد. خطای میانگین تخمین نیز برابر با ۰/۰۷۹ است. در مراحل بعد سعی خواهد شد خطا کاهش یابد.

۳-۴- طراحی تخمینگر بر اساس داده‌های خوشه‌بندی شده برای عناصر ژئوشیمیایی

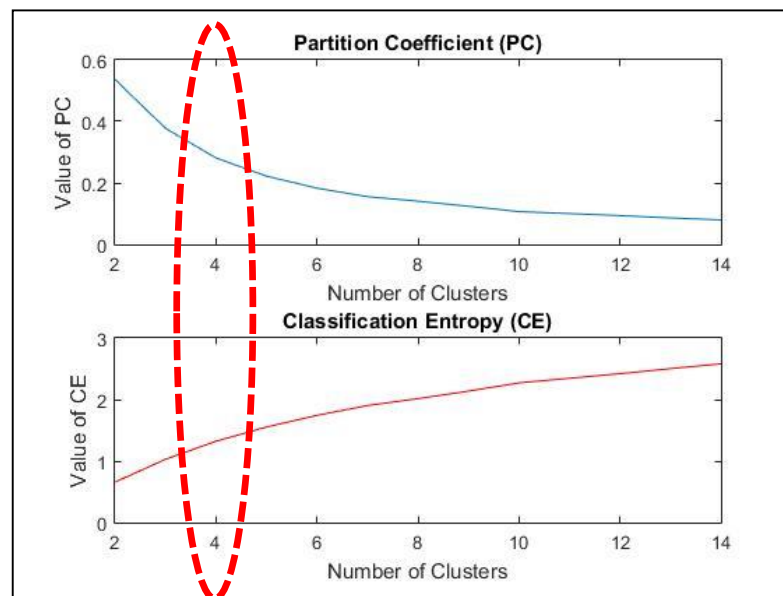
در مجموعه داده‌های ژئوشیمیایی احتمال وجود یک سری داده‌هایی است که عملکرد کل مجموعه را در تحلیل‌های اکتشافی دچار اشتباه می‌کند. با تقسیم‌بندی داده‌ها در خوشه‌های جداگانه، می‌توانیم بهترین خوشه‌ها را شناسایی کنیم و در تحلیل‌ها مورد استفاده قرار دهیم. قبل از خوشه‌بندی باید تعداد خوشه‌های بهینه را شناسایی کنیم که برای این کار از شاخص‌های اعتبارسنجی استفاده شده (شکل ۴-۶ تا شکل ۴-۸) و سپس از الگوریتم‌های خوشه‌بندی استفاده کرده و اعضای هر خوشه را شناسایی گردیده است. با تقسیم‌بندی داده‌ها به ۷۰ درصد آموزش و ۳۰ درصد آزمون، اقدام به طراحی تخمینگر پرداخته شده (شکل ۴-۵) تا بتوان بر اساس دقت و خطای تخمین، بهترین مجموعه داده را برای طراحی تخمینگر غلظت عناصر ژئوشیمیایی انتخاب نمود.



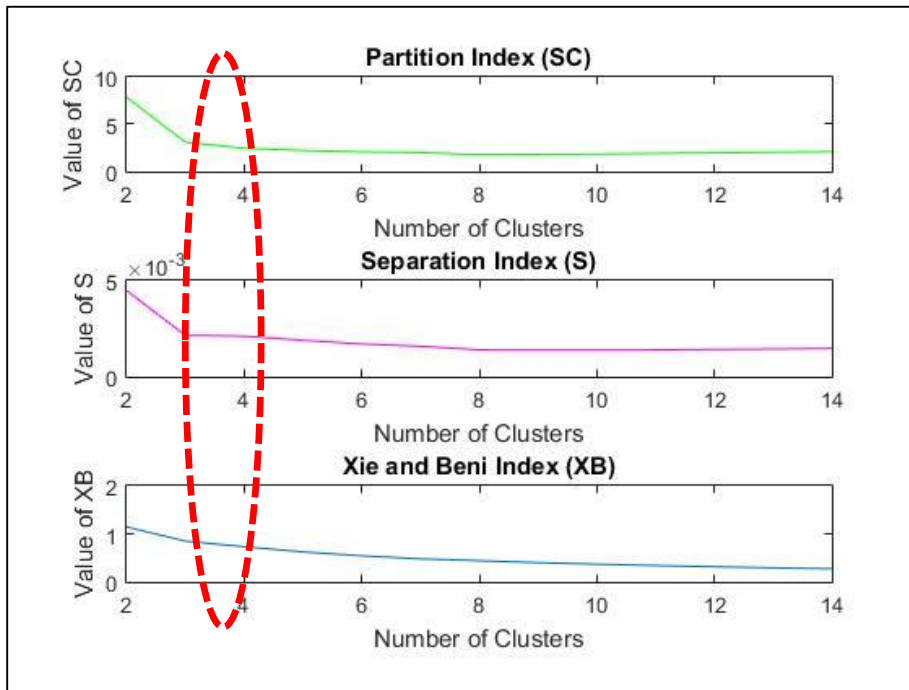
شکل ۴-۵: طراحی تخمینگر با استفاده از داده‌های خوشه‌بندی شده برای عناصر ژئوشیمیایی

۴-۳-۱- تعیین تعداد خوشه بهینه

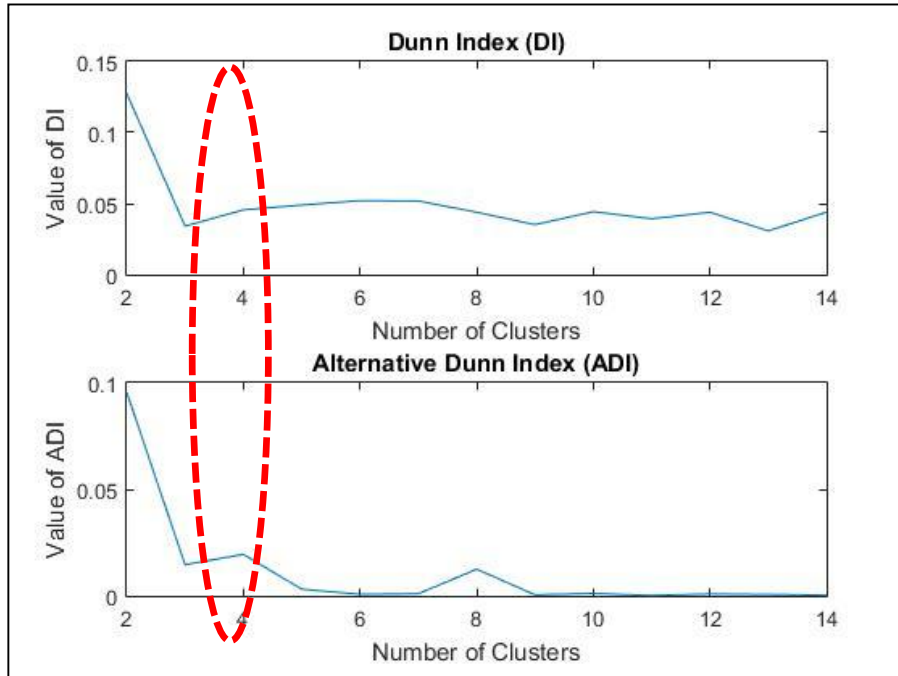
با توجه به شاخص‌های تعیین تعداد خوشه بهینه، مشاهده می‌شود در شاخص PC و CE تعداد خوشه بهینه تقریباً برابر با ۵ خوشه است و در دیگر شاخص‌ها تعداد خوشه‌های بهینه ۳ دیده می‌شود. بهینه‌ترین نتیجه، اشتراک بین تمام شاخص‌ها است که تعداد خوشه ۴ را نشان می‌دهد. لازم به ذکر است که با تعداد خوشه‌های ۳، ۴، ۵ و ۶ خوشه‌بندی انجام گردید و نتایج نشان داد که تعداد خوشه ۴ بهترین انتخاب برای خوشه‌بندی داده‌ها است بنابراین تعداد خوشه‌های بهینه برای الگوریتم‌های خوشه‌بندی برابر ۴ خوشه در نظر گرفته می‌شود (شکل ۴-۶، شکل ۴-۷ و شکل ۴-۸).



شکل ۴-۶: شاخص‌های اعتبارسنجی PC و CE



شکل ۴-۷: شاخص های اعتبارسنجی SC، S و XB



شکل ۴-۸: شاخص های اعتبارسنجی DI و ADI

اکنون داده‌ها را که شامل ۱۷۵۵ نمونه با ۲۲ مولفه است (جدول ۱-۳) با استفاده از الگوریتم‌های خوشه‌بندی K means، FCM، GK و GG خوشه‌بندی شده است.

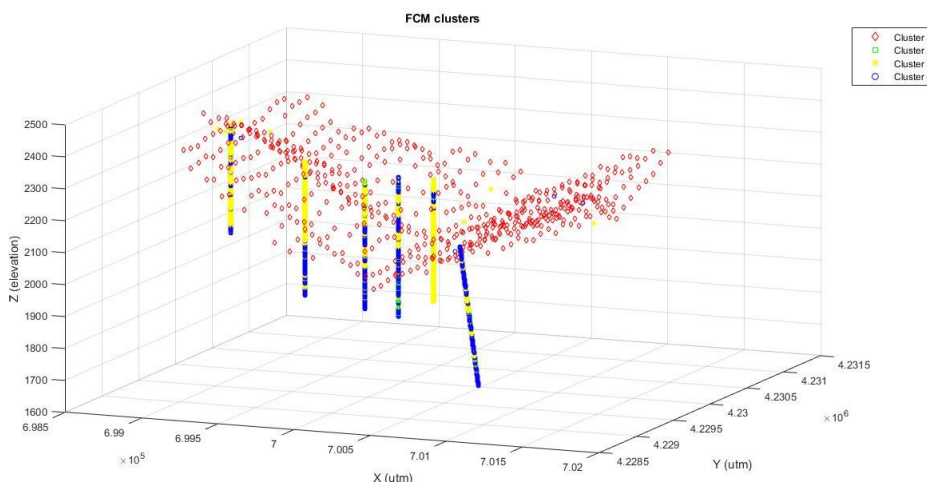
۴-۳-۲- افزایش دقت تخمین با الگوریتم خوشه‌بندی FCM

ابتدا اقدام به خوشه‌بندی داده‌های موجود با استفاده از الگوریتم خوشه‌بندی FCM می‌نماییم و سپس بر اساس شکل ۴-۵ اقدام به طراحی تخمینگر شبکه عصبی با استفاده از داده‌های خوشه‌بندی می‌کنیم. نتایج خوشه‌بندی با FCM به این صورت است که در خوشه اول ۵۳۵ داده، خوشه دوم ۴۴ داده، خوشه سوم ۶۴۳ داده و در خوشه چهارم ۵۳۳ داده قرار گرفته است. لازم به ذکر است که از داده‌های خوشه دوم به دلیل کم بودن تعداد نمونه‌ها در تحلیل‌ها استفاده نشده است (جدول ۱-۴).

جدول ۱-۴: نتایج خوشه‌بندی داده‌ها با استفاده از الگوریتم خوشه‌بندی FCM

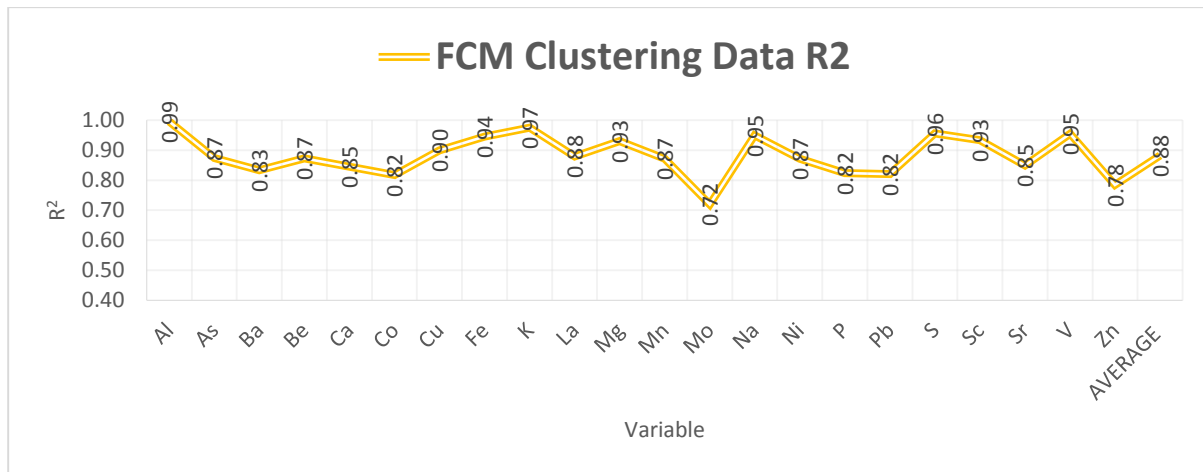
تعداد داده در خوشه‌های مختلف				تعداد کل داده‌ها	الگوریتم خوشه‌بندی
خوشه چهارم	خوشه سوم	خوشه دوم	خوشه اول	۱۷۵۵	Fuzzy c-means
۵۳۳	۶۴۳	۴۴	۵۳۵		

برای درک بهتر نحوه جدایش خوشه‌ها از یکدیگر، پراکندگی خوشه‌ها در فضای سه بعدی محورهای مختصات (شکل ۴-۹) نمایش شده است.

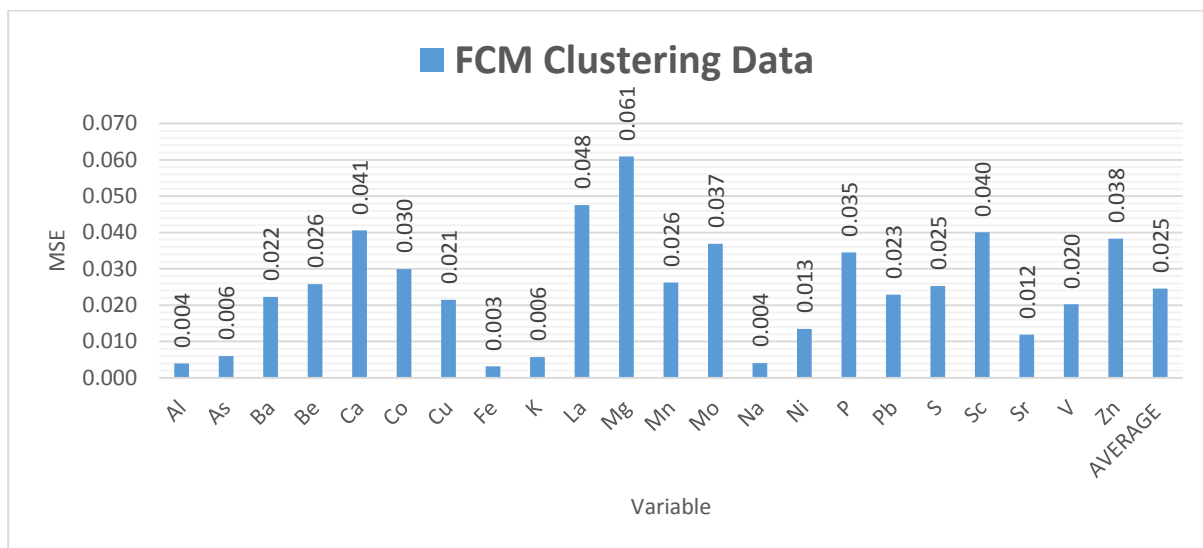


شکل ۴-۹: پراکندگی نمونه‌های خوشه‌بندی شده به روش FCM به صورت سه بعدی در فضا

دقت بدست آمده از تخمینگرهای عناصر ژئوشیمیایی در حالت استفاده از الگوریتم خوشه‌بندی FCM (شکل ۴-۱۰) نشان می‌دهد که عملکرد تخمینگرها در این حالت افزایش قابل توجهی داشته‌اند و از دقت ۷۵ درصدی با افزایشی ۱۳ درصدی به دقت ۸۸ درصد در میانگین دقت تخمینگرها رسیده است.



شکل ۴-۱۰: دقت تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از داده‌های خوشه‌بندی شده به روش FCM
 خطای تخمین حاصل از استفاده از داده‌های خوشه‌بندی شده با استفاده از روش خوشه‌بندی FCM (شکل ۴-۱۱) نشان می‌دهد که میانگین خطای تخمین عناصر ژئوشیمیایی ۰/۰۲۵ است که در مقایسه به حالت استفاده از کل داده‌ها که برابر با ۰/۰۷۹ بود سه برابر کاهش نشان می‌دهد.

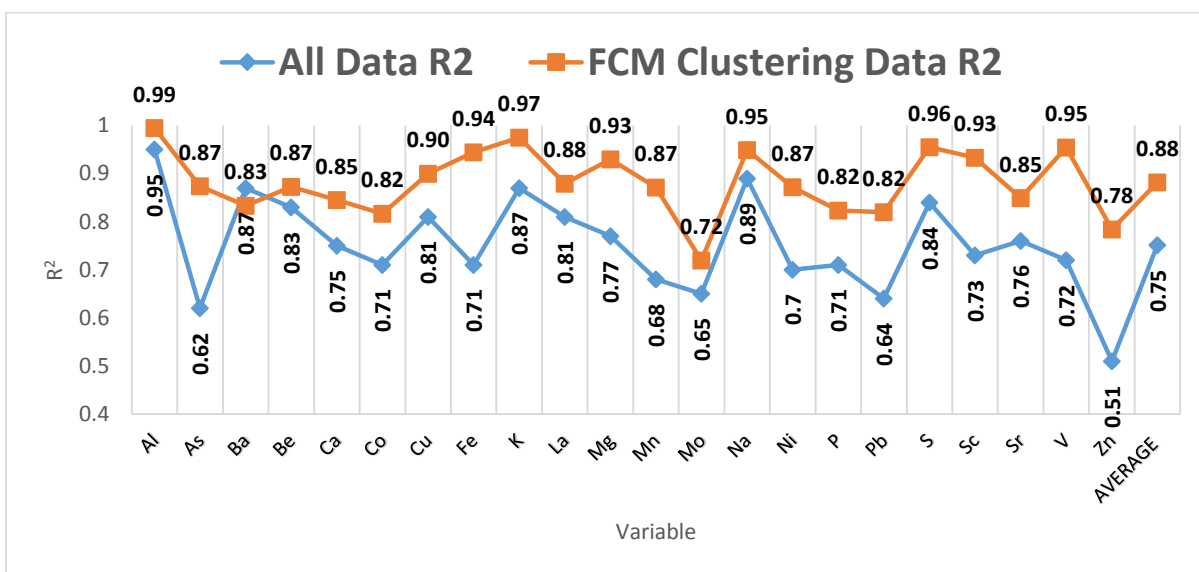


شکل ۴-۱۱: خطای تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از داده‌های خوشه‌بندی شده به روش FCM

۴-۳-۱- مقایسه دقت تخمینگر با حالت کلی

در حالت استفاده از الگوریتم خوشه‌بندی داده‌ها به خوشه‌هایی با همبستگی درونی بیشتری تقسیم‌بندی می‌شوند و انتظار داریم نتایج حاصل دارای دقت بالاتری نسبت به حالت استفاده از کل داده‌ها باشد. طراحی تخمینگرها در حالت استفاده از داده‌های خوشه‌بندی شده نشان از افزایش دقت میانگین ۱۳ درصدی نسبت به حالت قبل دارد به طوری که دقت میانگین از ۷۵ درصد به ۸۸ درصد افزایش پیدا کرده است (شکل ۴-۱۲).

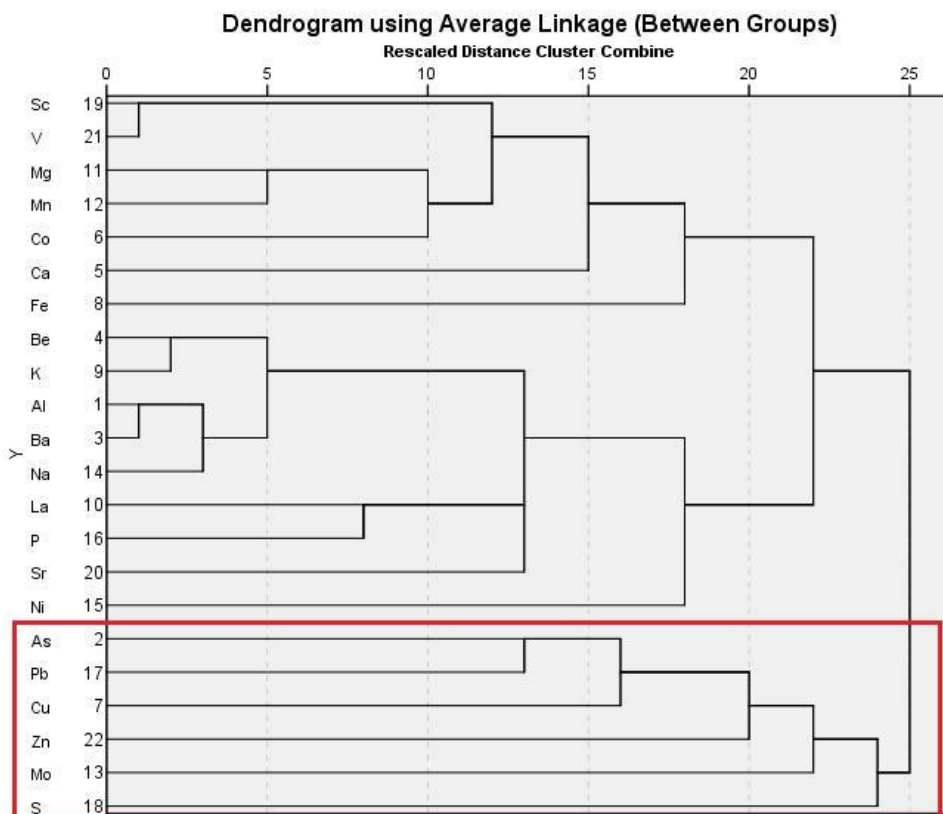
مقایسه نتایج نشان می‌دهد که بیشترین افزایش دقت در عنصر روی اتفاق افتاده است و با افزایشی ۲۷ درصدی از ۰/۵۱ به ۰/۷۸ رسیده است. عناصری مانند As با افزایش ۲۵ درصدی و V و Fe با افزایش ۲۳ درصدی دقت در رده‌های بعدی قرار گرفته‌اند (شکل ۴-۱۲).



شکل ۴-۱۲: مقایسه دقت کل تخمینگرها در حالت استفاده از کل داده‌ها و داده‌های خوشه‌بندی شده به روش FCM

در تحلیل‌های ژئوشیمیایی عناصری مورد توجه هستند که ارتباط خوبی با عنصر کانی‌ساز داشته باشند و به نوعی در ارتباط مستقیم با کانی‌سازی باشند. برای شناسایی این عناصر از تحلیل خوشه‌ای استفاده شده

که نشان دهنده جدایش عناصر As,Pb,Cu,Zn,Mo,S بعنوان عناصر مرتبط با کانی سازی مس پورفیری در یک خوشه می باشند (شکل ۴-۱۳).



شکل ۴-۱۳: نمودار خوشه بندی عناصر ژئوشیمیایی (عناصر مرتبط با مس پورفیری)

بررسی نتایج بدست آمده نشان می دهد که دقت تخمین عناصر Zn, As, Pb, Mo, Cu و S در حالت استفاده از کل داده ها به ترتیب دارای دقت های تخمین ۰/۵۱، ۰/۶۲، ۰/۶۴، ۰/۶۵، ۰/۸۱ و ۰/۸۴ بودند که در حالت استفاده از داده های خوشه بندی شده به روش FCM به ترتیب به ۰/۷۸، ۰/۸۷، ۰/۸۲، ۰/۷۲، ۰/۹۰ و ۰/۹۶ تغییر کرده اند که افزایش چشمگیری در دقت تخمین این عناصر مشاهده می شود. دقت تخمینگرها برای ۲۲ عنصر ژئوشیمیایی در حالت استفاده از کل داده ها و حالت استفاده از داده های خوشه بندی شده به روش FCM در جدول ۴-۲ آمده است.

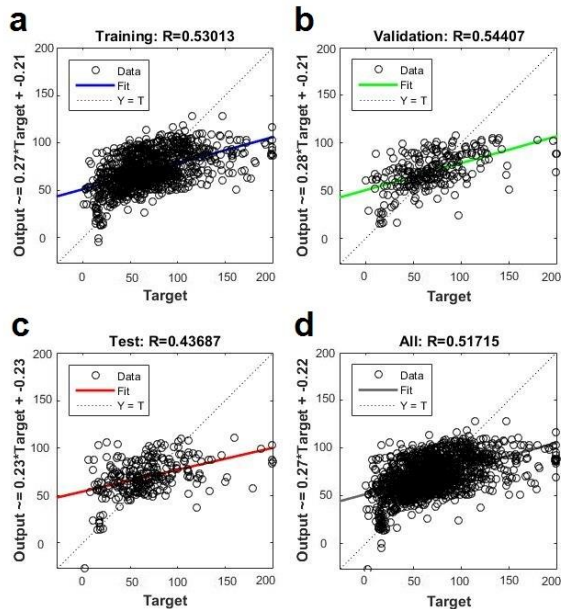
جدول ۴-۲: جدول دقت تخمین غلظت عناصر ژئوشیمیایی در حالت استفاده از کل داده‌ها و حالت استفاده از الگوریتم

خوشه‌بندی FCM

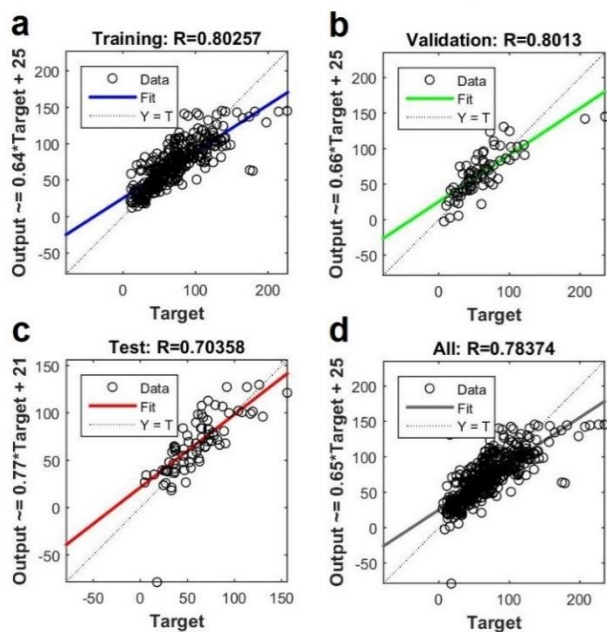
FCM Clustering method					Total Data				
Test R ²	Validation R ²	Training R ²	All R ²	Variable	Test R ²	Validation R ²	Training R ²	All R ²	Variable
0.99	0.998	0.99	0.99	Al	0.94	0.94	0.96	0.95	Al
0.76	0.71	0.89	0.87	As	0.6	0.57	0.64	0.62	As
0.88	0.83	0.82	0.83	Ba	0.89	0.85	0.86	0.87	Ba
0.82	0.83	0.89	0.87	Be	0.82	0.82	0.83	0.83	Be
0.8	0.79	0.87	0.85	Ca	0.67	0.76	0.77	0.75	Ca
0.82	0.83	0.82	0.82	Co	0.71	0.68	0.72	0.71	Co
0.91	0.83	0.92	0.9	Cu	0.86	0.82	0.81	0.81	Cu
0.83	0.99	0.89	0.94	Fe	0.7	0.72	0.71	0.71	Fe
0.97	0.97	0.98	0.97	K	0.85	0.86	0.87	0.87	K
0.8	0.89	0.9	0.88	La	0.74	0.78	0.82	0.81	La
0.91	0.96	0.93	0.93	Mg	0.72	0.77	0.78	0.77	Mg
0.81	0.84	0.89	0.87	Mn	0.64	0.6	0.7	0.68	Mn
0.77	0.77	0.71	0.72	Mo	0.64	0.62	0.66	0.65	Mo
0.9	0.81	0.97	0.95	Na	0.89	0.88	0.89	0.89	Na
0.89	0.8	0.91	0.87	Ni	0.69	0.7	0.7	0.7	Ni
0.81	0.84	0.82	0.82	P	0.65	0.73	0.72	0.71	P
0.81	0.82	0.83	0.82	Pb	0.6	0.62	0.65	0.64	Pb
0.94	0.95	0.96	0.96	S	0.82	0.86	0.84	0.84	S
0.93	0.93	0.94	0.93	Sc	0.74	0.73	0.73	0.73	Sc
0.71	0.73	0.9	0.85	Sr	0.8	0.71	0.76	0.76	Sr
0.95	0.96	0.96	0.95	V	0.71	0.76	0.71	0.72	V
0.7	0.8	0.8	0.78	Zn	0.43	0.54	0.53	0.51	Zn
0.85	0.86	0.89	0.88	Average	0.73	0.74	0.76	0.75	Average

هنگامی که خوشه‌بندی را بر روی داده‌های ژئوشیمیایی انجام می‌دهیم، هماهنگی داده‌ها درون خوشه‌ها افزایش پیدا می‌کند و باعث می‌شود شناسایی الگوی رفتاری بین پارامترهای ورودی و خروجی با دقت بیشتری انجام شود. نتیجه‌ای که از اعمال الگوریتم خوشه‌بندی FCM بر روی داده‌ها حاصل می‌شود علاوه بر افزایش دقت تخمین، صحت تخمین (انطباق شیب خط برازش شده به داده‌ها با خط $y=x$ ، هرچه انطباق بیشتری داشته باشد دقت و صحت بیشتر است) نیز افزایش پیدا می‌کند و مقادیر تخمینی به مقادیر واقعی نزدیکتر می‌شوند. برای مشاهده تاثیر خوشه‌بندی بر افزایش دقت و صحت (مثال - عنصر

Zn با کمترین دقت، نمودار رگرسیونی داده‌های واقعی و تخمینی در حالت استفاده از کل داده‌ها و داده‌های خوشه‌بندی شده در شکل ۴-۱۴ و شکل ۴-۱۵ آورده شده است.



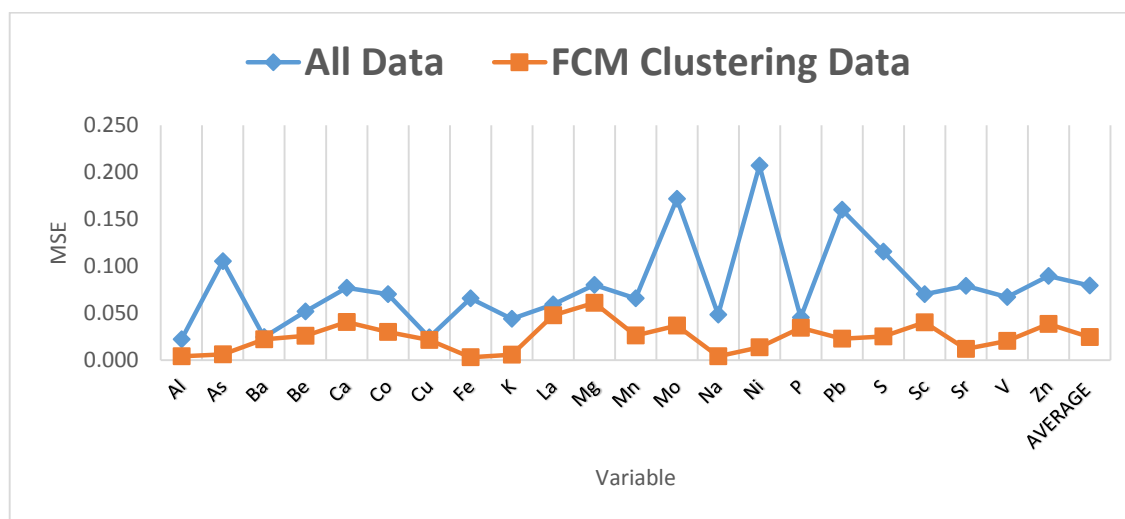
شکل ۴-۱۴: نمودار رگرسیونی داده‌های واقعی و تخمینی عنصر Zn در حالت استفاده از کل داده‌ها. محور x داده‌های واقعی و محور y داده‌های تخمینی است. (a) داده‌های آموزشی، (b) داده‌های اعتبارسنجی، (c) داده‌های آزمون و (d) کل داده‌ها.



شکل ۴-۱۵: نمودار رگرسیونی داده‌های واقعی و تخمینی عنصر Zn در حالت استفاده از داده‌های خوشه‌بندی شده با روش FCM. محور x داده‌های واقعی و محور y داده‌های تخمینی است. (a) داده‌های آموزشی، (b) داده‌های اعتبارسنجی، (c) داده‌های آزمون و (d) کل داده‌ها.

۴-۳-۲- مقایسه خطای تخمینگر با حالت کلی

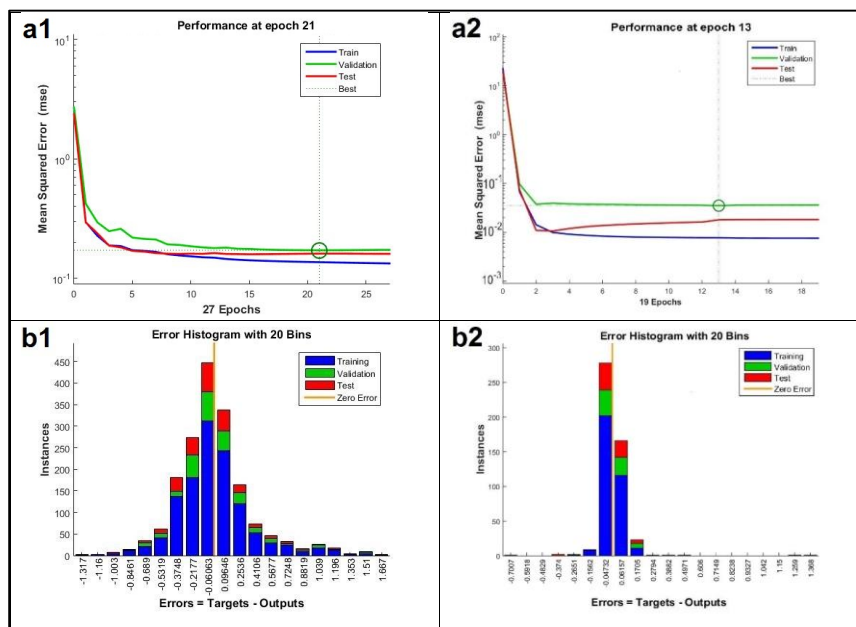
افزایش دقت تخمین زمانی ارزشمند است که تخمین انجام شده دارای کمترین میزان خطا بوده و تلاش بر کاهش میزان خطا است. در این مطالعه از میانگین توان دوم خطا (MSE) برای کنترل میزان خطای تخمینگر استفاده شده است و داده‌های استفاده شده در باز ۱- و ۱ استاندارد سازی شده‌اند تا عملکرد تخمینگر به بهترین شکل ممکن باشد. با توجه به نتایج بدست آمده، میانگین خطای تخمینگرها در حالت استفاده از کل داده‌ها برای طراحی تخمینگر برابر ۰/۰۷۹ است که این مقدار در حالت استفاده از داده‌های خوشه‌بندی شده به روش FCM کاهش چشمگیری نشان می‌دهد و با کاهشی ۳ برابری به ۰/۰۲۵ رسیده است که نشانگر آن است که استفاده از الگوریتم خوشه‌بندی علاوه بر افزایش دقت تخمین، خطای تخمین را نیز کاهش می‌دهد. این امر کارایی استفاده از روش‌های خوشه‌بندی را در تخمین غلظت عناصر ژئوشیمیایی نشان می‌دهد. نتایج خطای تخمینگر عناصر ژئوشیمیایی برای هر دو حالت ذکر شده، در شکل ۴-۱۶ آمده است.



شکل ۴-۱۶: مقایسه خطای تخمین در حالت استفاده از کل داده‌ها و حالت استفاده از الگوریتم خوشه‌بندی

بررسی نمودار خطاها نشان می‌دهد که کاهش خطا برای عناصر As، Mo، Ni، Pb، S و Zn به صورت چشمگیری اتفاق افتاده است و برای بقیه عناصر ژئوشیمیایی این کاهش در حد قابل قبولی انجام شده

است. از مواردی که می‌توان برای کنترل عملکرد تخمینگر استفاده کرد، نمودار کارایی ۸۹ و هیستوگرام پراکندگی ۹۰ خطا است. در شکل ۴-۱۷ این نمودارها برای عنصر Zn نشان می‌دهد که هنگام استفاده از داده‌های خوشه‌بندی شده، خطای تخمین با تعداد مراحل پردازش کمتر نسبت به حالت استفاده از کل داده‌ها کاهش خوبی داشته است به عبارتی دیگر زمان محاسبات کاهش یافته و خطا کمتر شده است (شکل ۴-۱۷ a1 و a2). در هیستوگرام خطا، هرچه هیستوگرام دارای پراکندگی کمتری باشد نشان می‌دهد که عملکرد تخمینگر بهتر است. از نتایج بدست آمده مشاهده می‌شود که هیستوگرام خطا در حالت استفاده از داده‌های خوشه‌بندی شده نسبت به حالت قبل دارای پراکندگی کمتر است (شکل ۴-۱۷ b1 و b2). به طور کلی خطای تخمین در تمامی عناصر هنگام استفاده از الگوریتم خوشه‌بندی FCM کاهش یافته است.



شکل ۴-۱۷: نمودار تغییرات خطا و هیستوگرام پراکندگی مقادیر خطا عنصر Zn. (a1) نمودار خطا در حالت استفاده از کل داده‌ها و (a2) حالت استفاده از داده‌های خوشه‌بندی شده، (b1) هیستوگرام خطا در حالت استفاده از کل داده‌ها و (b2) حالت استفاده از داده‌های خوشه‌بندی شده

۸۹ performance

۹۰ Error Histogram

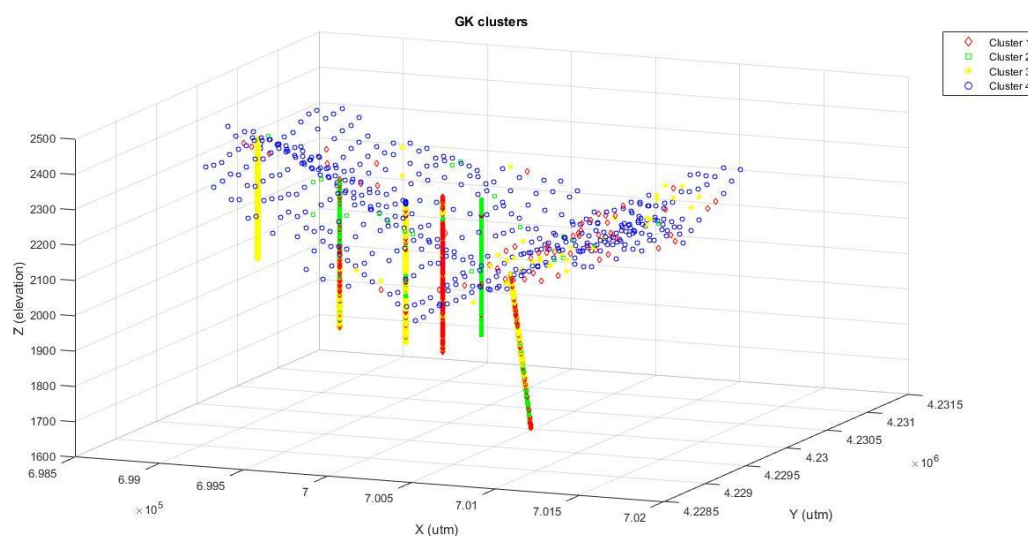
۳-۳-۴- افزایش دقت تخمین با الگوریتم خوشه‌بندی GK

در این مرحله خوشه‌بندی را با استفاده از الگوریتم خوشه‌بندی گوستافسون کسل و با تعداد ۴ خوشه تعیین شده با استفاده از شاخص‌های اعتبارسنجی انجام گردیده است. در خوشه اول ۴۱۶ داده، در خوشه دوم ۳۶۹ داده، در خوشه سوم ۵۳۲ داده و در خوشه چهارم ۴۳۸ داده قرار گرفته است. (جدول ۳-۴)

جدول ۳-۴: نتایج خوشه‌بندی داده‌ها با استفاده از الگوریتم خوشه‌بندی GK

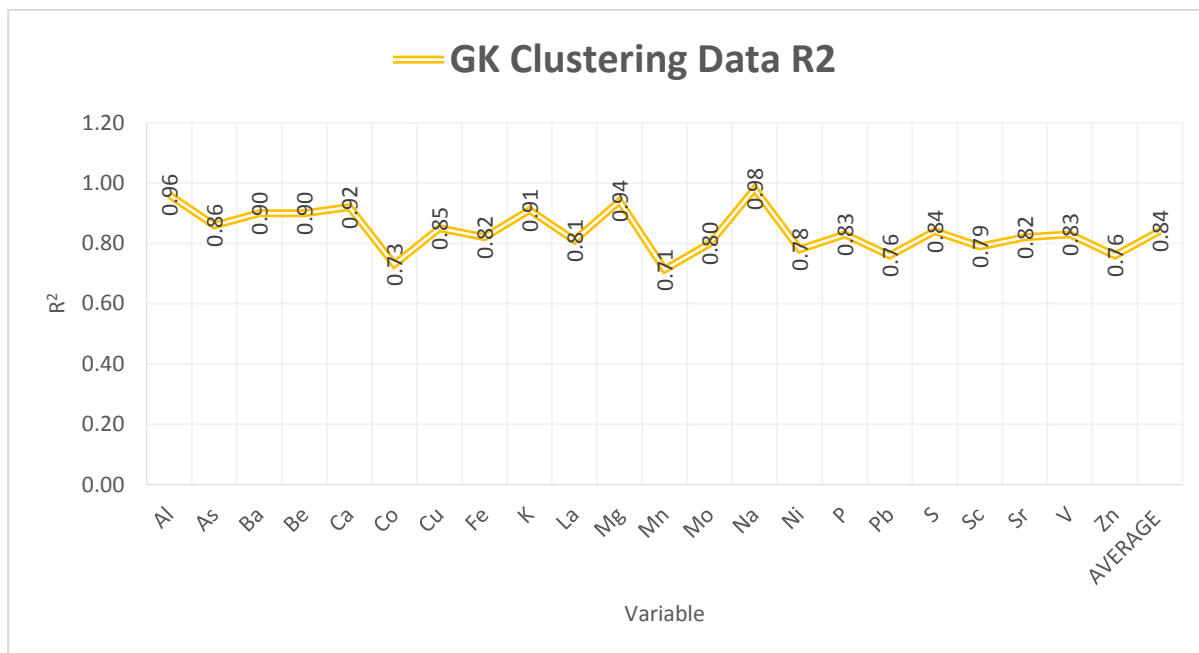
تعداد داده در خوشه‌های مختلف				تعداد کل داده‌ها	الگوریتم خوشه‌بندی
خوشه اول	خوشه دوم	خوشه سوم	خوشه چهارم	۱۷۵۵	Gustafson kessel
۴۱۶	۳۶۹	۵۳۲	۴۳۸		

پراکندگی خوشه‌ها در فضای سه بعدی محورهای مختصات در شکل ۴-۱۸ نمایان شده است.

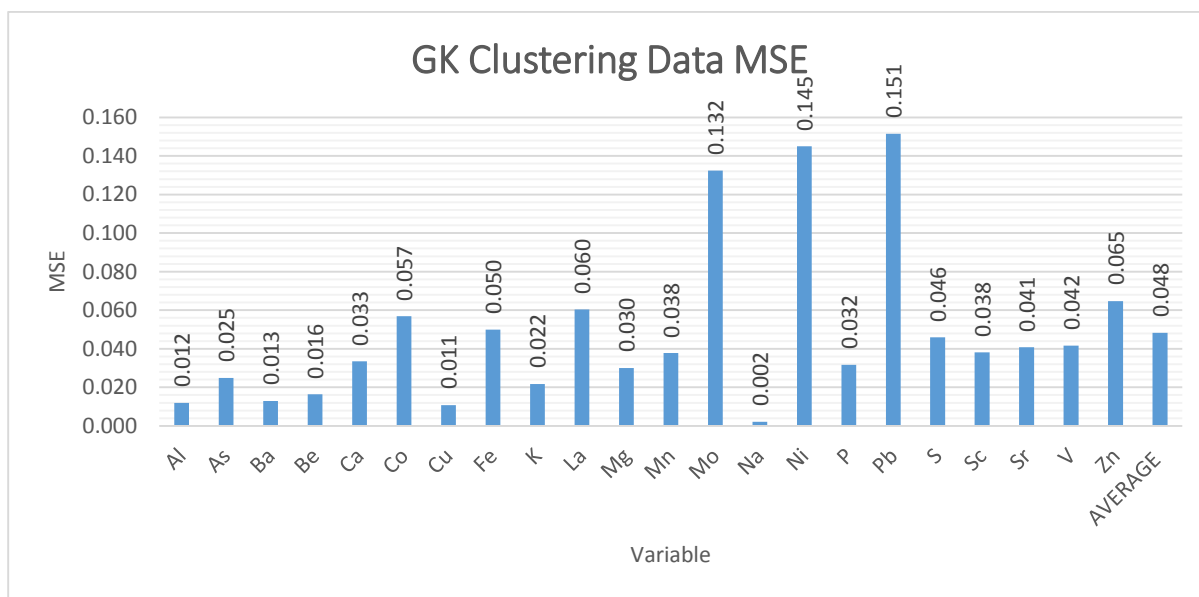


شکل ۴-۱۸: پراکندگی نمونه‌های خوشه‌بندی شده به روش گوستافسون کسل به صورت سه بعدی در فضا

در این مرحله، ورودی تخمینگر شبکه عصبی از داده‌های خوشه‌بندی شده توسط الگوریتم گوستافسون کسل انتخاب گردید و تخمینگر طراحی گردید. نتایج حاصل شده از طراحی تخمینگر در حالت استفاده از داده‌های خوشه‌بندی شده به روش GK برای عناصر ژئوشیمیایی در شکل ۴-۱۹ و شکل ۴-۲۰ آمده است.



شکل ۴-۱۹: دقت تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از داده‌های خوشه‌بندی شده به روش GK



شکل ۴-۲۰: خطای تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از داده‌های خوشه‌بندی شده به روش GK

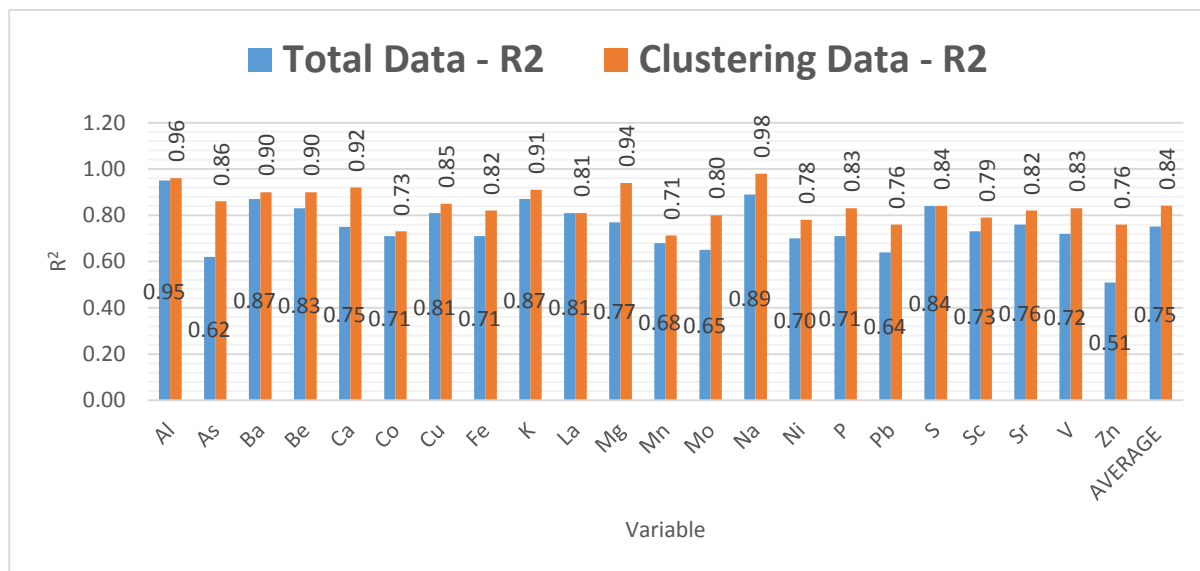
نتایج بدست آمده نشان می‌دهد که دقت تخمینگر به طور کلی افزایش یافته است و دقت میانگین با افزایش ۹ درصدی به ۸۴ درصد رسیده است. خطای تخمین حاصل از استفاده از داده‌های خوشه‌بندی

شده با استفاده از روش خوشه‌بندی GK در شکل ۴-۲۰ نشان می‌دهد که میانگین خطای تخمین عناصر ژئوشیمیایی ۰/۰۴۸ است که در مقایسه به حالت استفاده از کل داده‌ها که برابر با ۰/۰۷۹ بود کاهش نشان می‌دهد.

۴-۳-۳-۱- مقایسه دقت تخمینگر با حالت کلی

مقایسه نتایج حاصل از دو حالت طراحی تخمینگرها در حالت استفاده از کل داده‌ها و استفاده از داده‌های خوشه‌بندی شده برای تخمین غلظت عناصر مختلف نشان می‌دهد که دقت تخمینگرها در حالت استفاده از داده‌های خوشه‌بندی شده به میزان قابل توجهی افزایش پیدا کرده و دقت میانگین تخمینگرها از ۷۵ درصد به ۸۴ درصد افزایش می‌یابد (شکل ۴-۲۱).

به طوریکه عناصری مانند Zn، As، Pb، Mo و Mn که به ترتیب دارای دقت‌های پایین ۰/۵۱، ۰/۶۲، ۰/۶۴، ۰/۶۵ و ۰/۶۸ در حالت استفاده از کل داده‌ها بوده، در حالت استفاده از داده‌های خوشه‌بندی شده به ترتیب به ۰/۷۶، ۰/۸۶، ۰/۷۶، ۰/۸۰ و ۰/۷۱ افزایش یافته است. دقت عنصری مانند روی که در حالت اول عملاً غیر قابل قبول بود، در حالت استفاده از داده‌های خوشه‌بندی شده (با افزایش دقت ۲۵ درصدی) به ۷۶ درصد رسیده است که مورد قبول می‌باشد. نتایج تخمینگرهای عناصر مختلف در حالت استفاده از کل داده‌ها و حالت استفاده از داده‌های خوشه‌بندی شده به روش گوستافسون کسل در جدول ۴-۴ قابل مشاهده است.

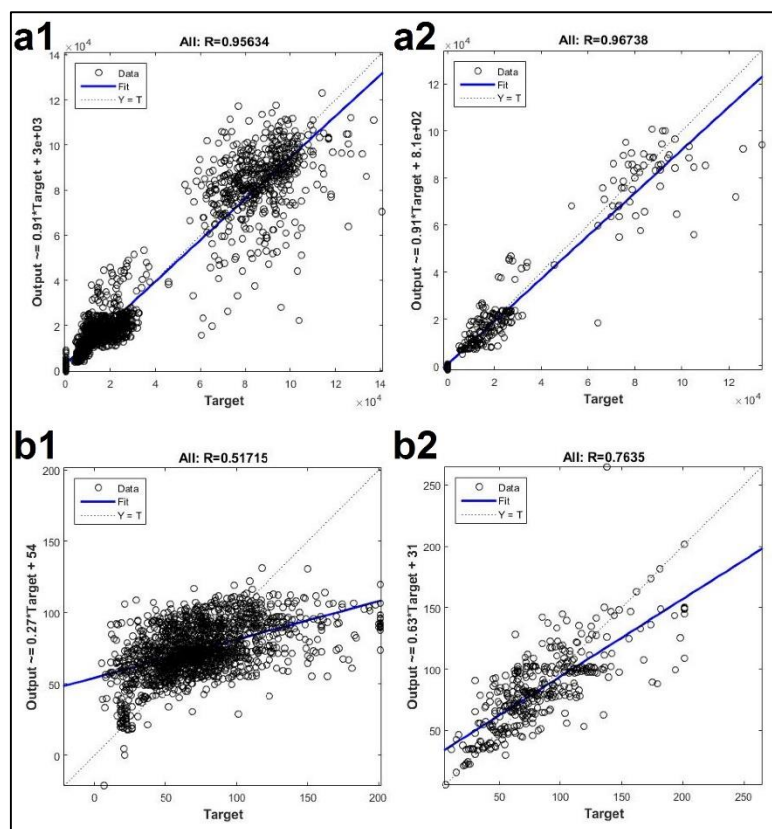


شکل ۴-۲۱: مقایسه دقت کل تخمینگرها در حالت استفاده از کل داده‌ها و داده‌های خوشه‌بندی شده به روش GK
 جدول ۴-۴: جدول دقت تخمین غلظت عناصر ژئوشیمیایی در حالت استفاده از کل داده‌ها و حالت استفاده از الگوریتم

خوشه‌بندی GK

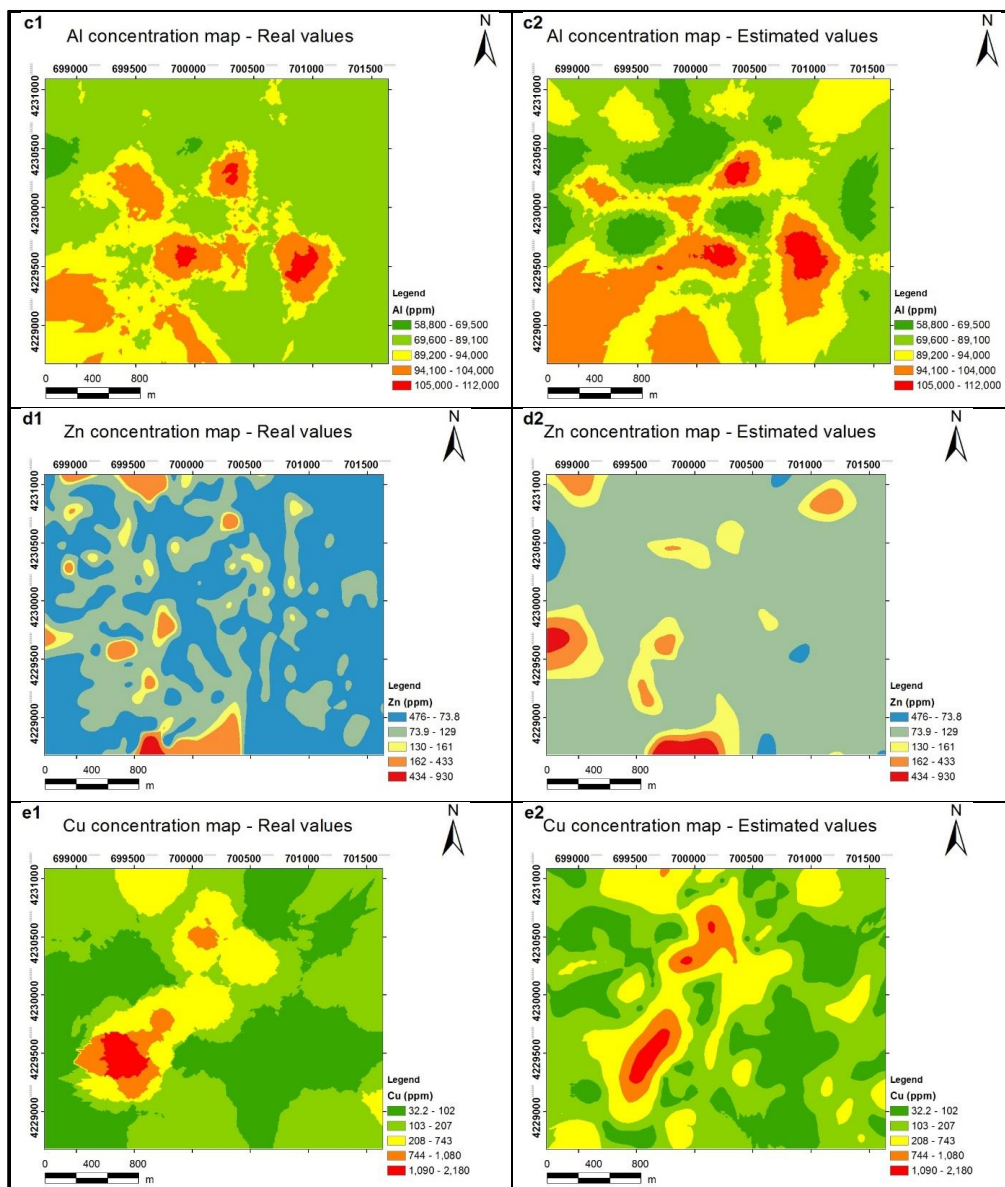
Gustafson kessel Clustering method					Total Data				Variable
Test R ²	Validation R ²	Training R ²	All R ²	Variable	Test R ²	Validation R ²	Training R ²	All R ²	
0.96	0.96	0.96	0.96	Al	0.94	0.94	0.96	0.95	Al
0.82	0.82	0.88	0.86	As	0.6	0.57	0.64	0.62	As
0.93	0.91	0.9	0.9	Ba	0.89	0.85	0.86	0.87	Ba
0.87	0.88	0.91	0.9	Be	0.82	0.82	0.83	0.83	Be
0.86	0.88	0.95	0.92	Ca	0.67	0.76	0.77	0.75	Ca
0.74	0.67	0.74	0.73	Co	0.71	0.68	0.72	0.71	Co
0.9	0.82	0.85	0.85	Cu	0.86	0.82	0.81	0.81	Cu
0.78	0.79	0.84	0.82	Fe	0.7	0.72	0.71	0.71	Fe
0.89	0.91	0.92	0.91	K	0.85	0.86	0.87	0.87	K
0.83	0.73	0.82	0.81	La	0.74	0.78	0.82	0.81	La
0.86	0.92	0.96	0.94	Mg	0.72	0.77	0.78	0.77	Mg
0.62	0.8	0.72	0.71	Mn	0.64	0.6	0.7	0.68	Mn
0.75	0.79	0.81	0.8	Mo	0.64	0.62	0.66	0.65	Mo
0.96	0.96	0.99	0.98	Na	0.89	0.88	0.89	0.89	Na
0.76	0.73	0.8	0.78	Ni	0.69	0.7	0.7	0.7	Ni
0.7	0.82	0.86	0.83	P	0.65	0.73	0.72	0.71	P
0.76	0.73	0.76	0.76	Pb	0.6	0.62	0.65	0.64	Pb
0.92	0.86	0.81	0.84	S	0.82	0.86	0.84	0.84	S
0.75	0.76	0.81	0.79	Sc	0.74	0.73	0.73	0.73	Sc
0.88	0.8	0.82	0.82	Sr	0.8	0.71	0.76	0.76	Sr
0.74	0.88	0.86	0.83	V	0.71	0.76	0.71	0.72	V
0.61	0.77	0.8	0.76	Zn	0.43	0.54	0.53	0.51	Zn
0.81	0.83	0.85	0.84	Average	0.73	0.74	0.76	0.75	Average

استفاده از الگوریتم خوشه‌بندی علاوه بر اینکه باعث افزایش دقت تخمین غلظت عناصر گردیده، صحت تخمین را نیز افزایش می‌دهد. برای درک بیشتر مطلب، نمودار رگرسیونی تخمین‌گرهای Al و Zn که به ترتیب دارای بیشترین و کمترین دقت تخمین بودند در شکل ۴-۲۲ آورده شده است. هنگام استفاده از الگوریتم خوشه‌بندی، به جهت اینکه داده‌های درون خوشه‌های دارای هماهنگی بیشتری نسبت به هم هستند، علاوه بر افزایش دقت تخمین، باعث نزدیک شدن شیب خط نمودار رگرسیونی به نمودار $y=x$ شده و صحت را افزایش داده است.



شکل ۴-۲۲: نمودار رگرسیونی واقعی و تخمین زده شده. محور x داده‌های واقعی و محور y داده‌های تخمینی است. (a1) نمودار رگرسیونی عنصر Al در حالت استفاده از کل داده‌ها، (a2) نمودار رگرسیونی عنصر Al در حالت استفاده از داده‌های خوشه‌بندی شده، (b1) نمودار رگرسیونی عنصر Zn در حالت استفاده از کل داده‌ها و (b2) نمودار رگرسیونی عنصر Al در حالت استفاده از داده‌های خوشه‌بندی شده

مقایسه نقشه‌های غلظت مقادیر واقعی و تخمینی نشانگر آن است که نتایج بدست آمده از تخمین، انطباق خوبی با واقعیت دارند. برای مثال، نقشه‌ی سطحی برای غلظت مقادیر واقعی و تخمینی عناصر Al، Zn و Cu رسم گردید (شکل ۴-۲۳). با بررسی نقشه‌های تخمینی بدست آمده از مجموعه داده‌های سطحی و عمقی می‌توان الگوی فضایی پراکندگی عناصر ژئوشیمیایی را شناسایی کرد.

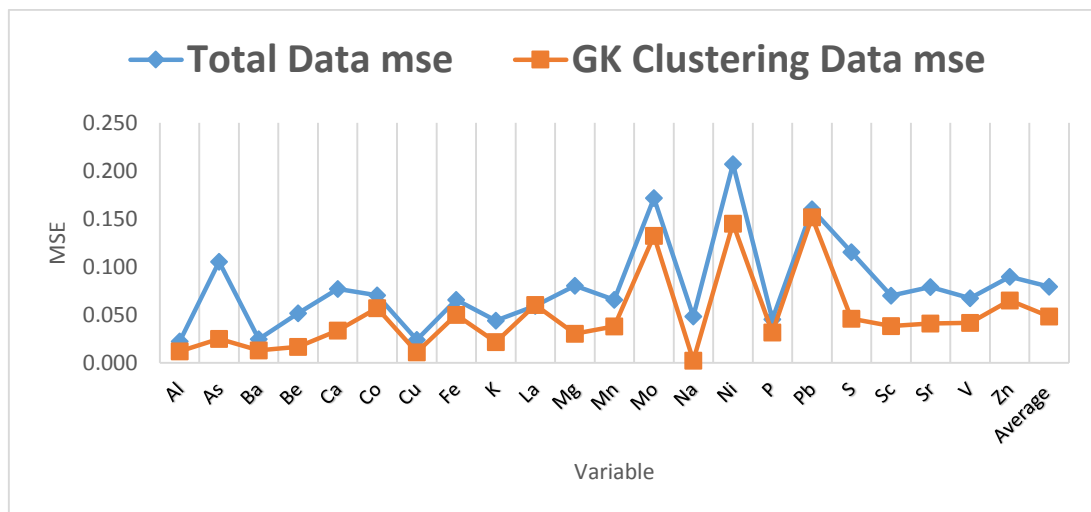


شکل ۴-۲۳: مقایسه نقشه غلظت مقادیر واقعی و مقادیر تخمینی در سطح، (c1) نقشه غلظت عنصر Al با مقادیر واقعی، (c2) نقشه غلظت عنصر Al با مقادیر تخمینی، (d1) نقشه غلظت عنصر Zn با مقادیر واقعی و (d2) نقشه غلظت عنصر Zn با مقادیر تخمینی، (e1) نقشه غلظت عنصر Cu با مقادیر واقعی و (e2) نقشه غلظت عنصر Cu با مقادیر تخمینی

۴-۳-۲- مقایسه خطای تخمینگر با حالت کلی

خطای تخمین نیز به عنوان یکی از پارامترها در بررسی عملکرد تخمینگرها مورد توجه بود و سعی بر آن است که میزان خطای تخمین به کمترین مقدار ممکن برسد. بازه داده‌های استانداردسازی شده مورد استفاده برای طراحی تخمینگرها در حالات مختلف بین ۱- تا ۱ است تا مقیاس تغییرات داده‌ها یکسان باشد و تخمینگر دچار اشتباه نگردد.

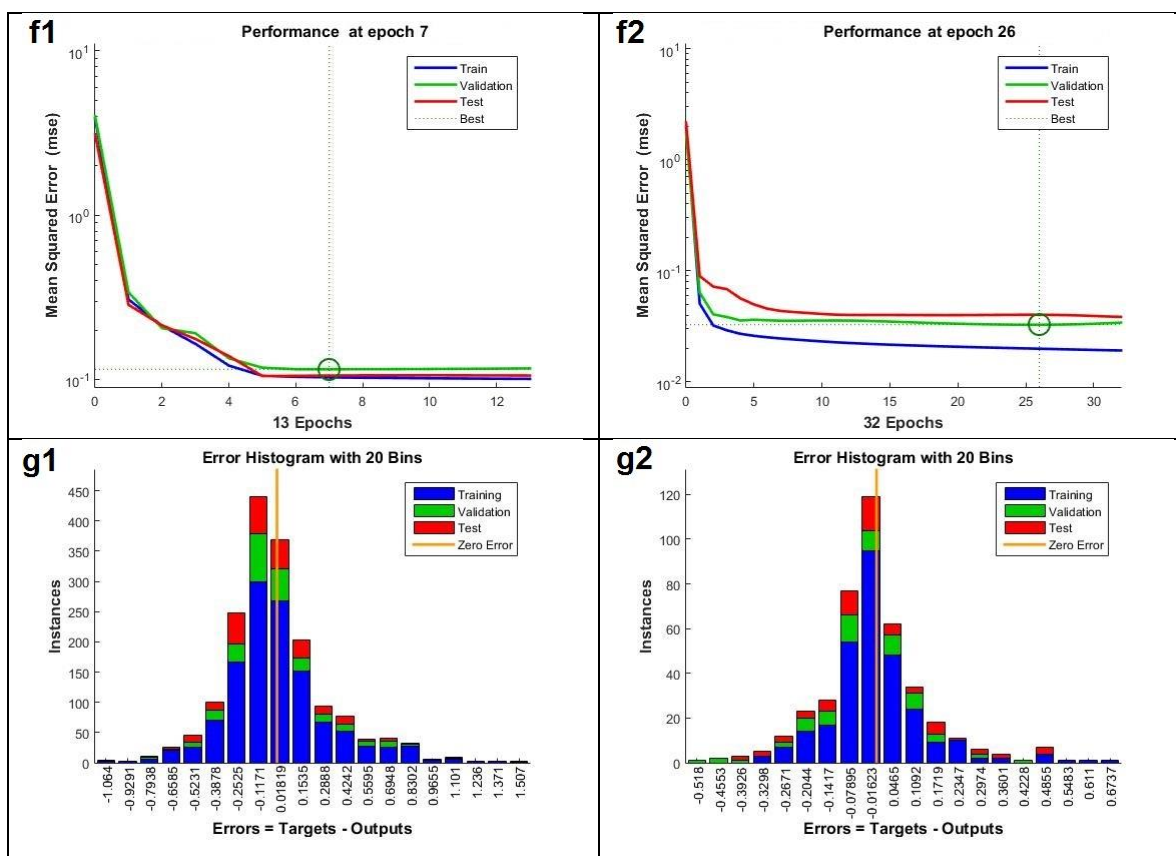
برای بررسی خطای تخمین از میانگین توان دوم خطا (یا به اختصار MSE) استفاده شده است. با توجه به نتایج بدست آمده از حالات مختلف طراحی تخمین مشاهده می‌شود که میزان میانگین خطا در حالتی که از کل داده‌ها برای طراحی تخمینگر استفاده می‌شود ۰/۰۷۹ است که در مقایسه با زمانی که از الگوریتم خوشه‌بندی استفاده شود، مقدار خطای بالا است. مقدار میانگین خطای تخمینگرهای ۲۲ عنصر مورد مطالعه در حالت استفاده از الگوریتم خوشه‌بندی GK به ۰/۰۴۸ کاهش یافته که نشانگر آن است که استفاده از الگوریتم خوشه‌بندی علاوه بر افزایش دقت تخمین، خطای تخمین را نیز کاهش می‌دهد. این امر کارایی استفاده از روش‌های خوشه‌بندی را نشان می‌دهد (شکل ۴-۲۴).



شکل ۴-۲۴: مقایسه خطای تخمین در حالت استفاده از کل داده‌ها و حالت استفاده از الگوریتم خوشه‌بندی گوستافسون کسل

بررسی نمودار خطاها نشان دهنده کاهش خطا از ۰/۰۲۲ به ۰/۰۱۲ در AI، از ۰/۱۰۵ به ۰/۰۲۵ در AS، از ۰/۱۱۵ به ۰/۰۴۶ در S بوده است.

نمودار کارایی و هیستوگرام پراکندگی مقادیر خطا برای عنصر AS (شکل ۴-۲۵) نشان می‌دهد که عملکرد تخمینگر برای داده‌های خوشه‌بندی شده با کاهش مقادیر خطا و متقارن سازی هیستوگرام همراه است. به طور کلی خطای تخمین در تمامی عناصر هنگام استفاده از الگوریتم خوشه‌بندی کاهش یافته است.



شکل ۴-۲۵: نمودار تغییرات خطا و هیستوگرام پراکندگی مقادیر خطا عنصر AS، (f1) نمودار خطا در حالت استفاده از کل داده‌ها و (f2) حالت استفاده از داده‌های خوشه‌بندی شده، (g1) هیستوگرام خطا در حالت استفاده از کل داده‌ها و (g2) حالت استفاده از داده‌های خوشه‌بندی شده

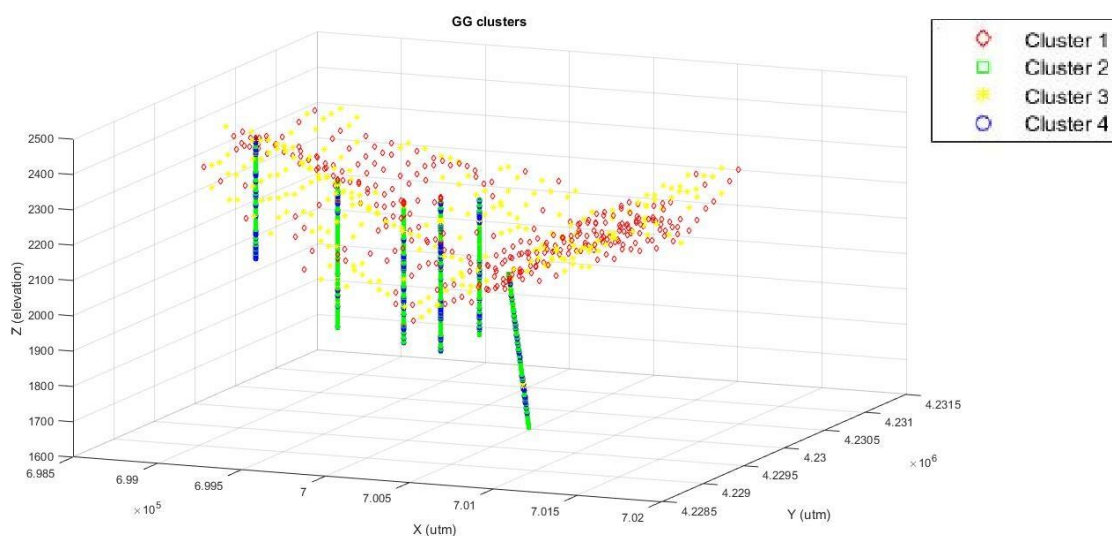
۴-۳-۴- افزایش دقت تخمین با الگوریتم خوشه‌بندی GG

در ادامه خوشه‌بندی با استفاده از الگوریتم خوشه‌بندی گت جوا و با تعداد ۴ خوشه حاصل از شاخص‌های اعتبارسنجی انجام شده است. در خوشه اول ۳۴۲ داده، خوشه دوم ۸۶۱، خوشه سوم ۲۷۹ و در خوشه چهارم ۲۷۶ داده قرار گرفته است (جدول ۴-۵).

جدول ۴-۵: نتایج خوشه‌بندی داده‌ها با استفاده از الگوریتم خوشه‌بندی GG

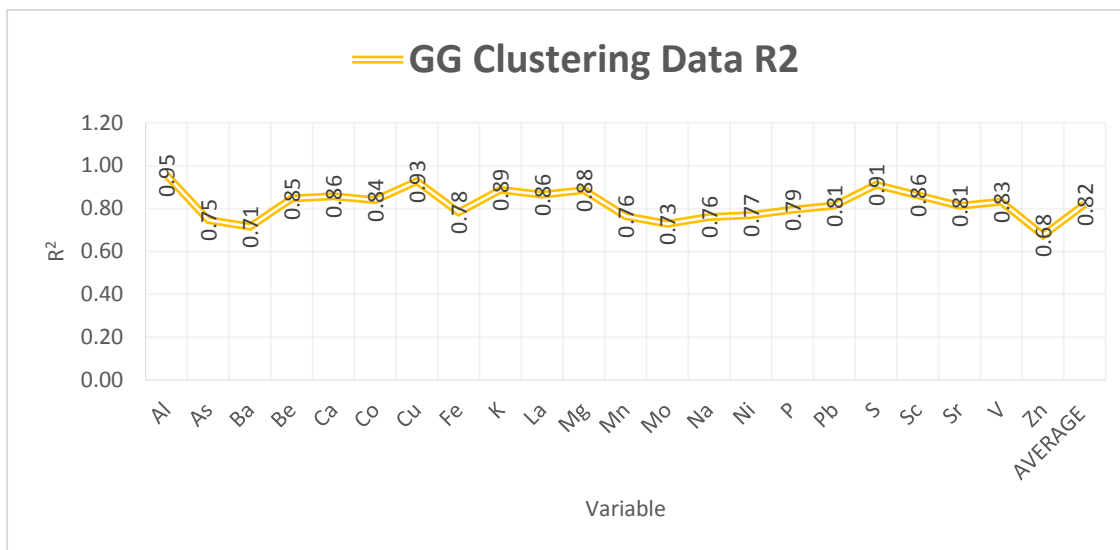
تعداد داده در خوشه‌های مختلف				تعداد کل داده‌ها	الگوریتم خوشه‌بندی
خوشه اول	خوشه دوم	خوشه سوم	خوشه چهارم	۱۷۵۵	Gath - Geva
۳۴۲	۸۶۱	۲۷۹	۲۷۶		

پراکندگی خوشه‌ها در فضای سه بعدی محورهای مختصات در شکل ۴-۲۶ نمایان شده است.

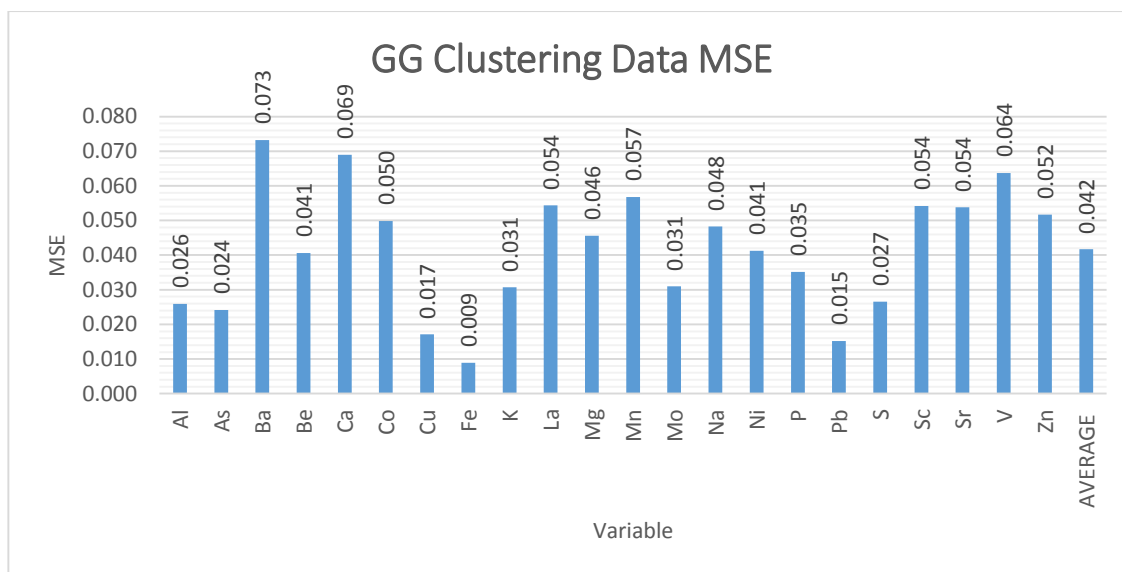


شکل ۴-۲۶: پراکندگی نمونه‌های خوشه‌بندی شده به روش گت جوا به صورت سه بعدی در فضا

در این مرحله، ورودی تخمینگر شبکه عصبی از داده‌های خوشه‌بندی شده توسط الگوریتم گت جوا انتخاب و تخمینگر طراحی گردیده است. نتایج حاصل شده از طراحی تخمینگر در حالت استفاده از داده‌های خوشه‌بندی شده به روش GG برای عناصر ژئوشیمیایی در شکل ۴-۲۷ و شکل ۴-۲۸ آمده است.



شکل ۴-۲۷: دقت تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از داده‌های خوشه‌بندی شده به روش GG نتایج بدست آمده نشان می‌دهد که دقت تخمینگر به طور کلی افزایش یافته است و دقت میانگین با افزایش ۷ درصدی به ۸۲ درصد رسیده است. خطای تخمین حاصل از استفاده از داده‌های خوشه‌بندی شده با استفاده از روش خوشه‌بندی GG در شکل ۴-۲۸ نشان می‌دهد که میانگین خطای تخمین عناصر ژئوشیمیایی ۰/۰۴۲ است که در مقایسه به حالت استفاده از کل داده‌ها که برابر با ۰/۰۷۹ بود کاهش نشان می‌دهد.

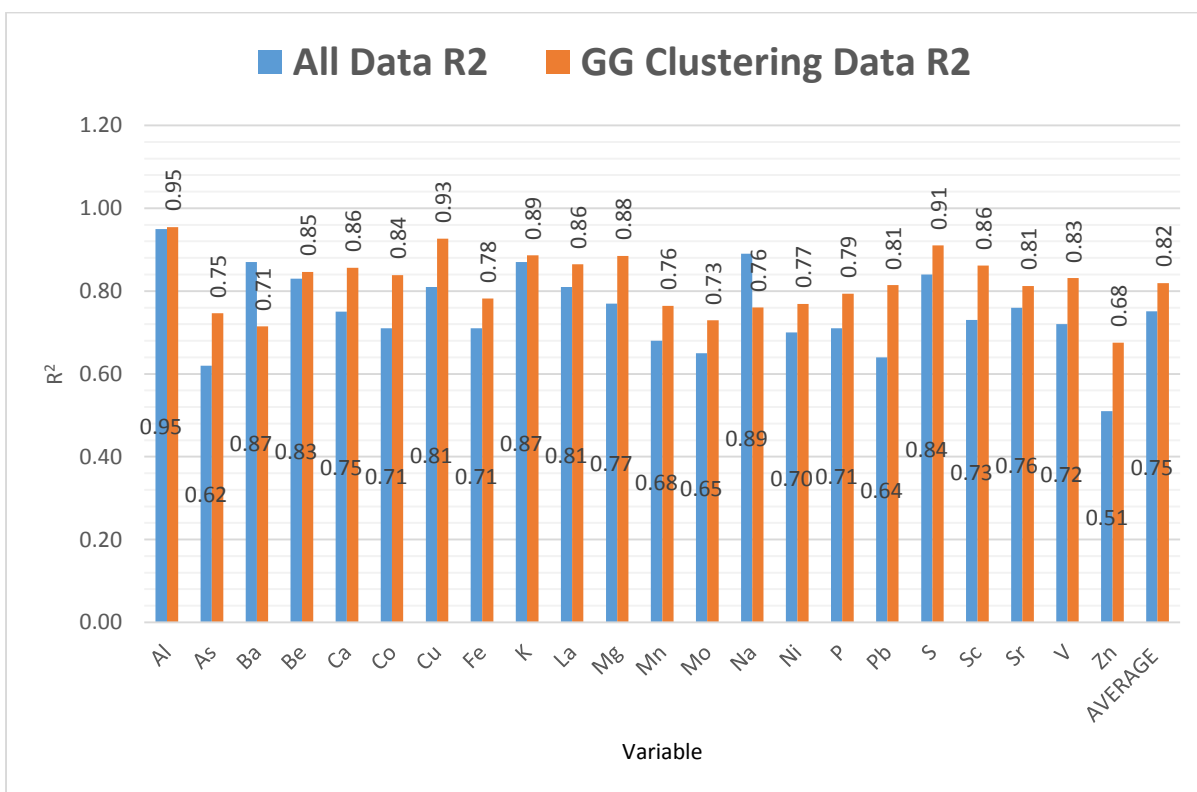


شکل ۴-۲۸: خطای تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از داده‌های خوشه‌بندی شده به روش GG

۴-۳-۴-۱- مقایسه دقت تخمینگر با حالت کلی

مقایسه نتایج حاصل از دو حالت طراحی تخمینگرها در حالت استفاده از کل داده‌ها و استفاده از داده‌های خوشه‌بندی شده برای تخمین غلظت عناصر مختلف نشان می‌دهد که دقت تخمینگرها در حالت استفاده از داده‌های خوشه‌بندی شده افزایش پیدا کرده و دقت میانگین تخمینگرها از ۷۵ درصد به ۸۲ درصد افزایش می‌یابد (شکل ۴-۲۹).

عناصری مانند Zn, As, Pb, Mo و Mn که به ترتیب دارای دقت‌های پایین ۰/۵۱، ۰/۶۴، ۰/۶۵ و ۰/۶۸ در حالت استفاده از کل داده‌ها بوده، در حالت استفاده از داده‌های خوشه‌بندی شده به ترتیب به ۰/۶۸، ۰/۷۵، ۰/۸۱ و ۰/۷۳ افزایش یافته است. نتایج تخمینگرهای عناصر مختلف در حالت استفاده از کل داده‌ها و حالت استفاده از داده‌های خوشه‌بندی شده به روش گوستافسون کسل در جدول ۴-۶ قابل مشاهده است.

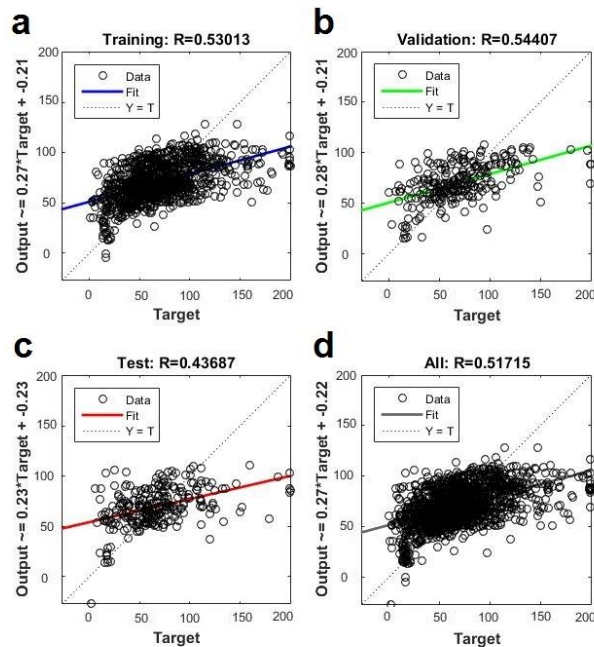


شکل ۴-۲۹: مقایسه دقت کل تخمینگرها در حالت استفاده از کل داده‌ها و داده‌های خوشه‌بندی شده به روش GG

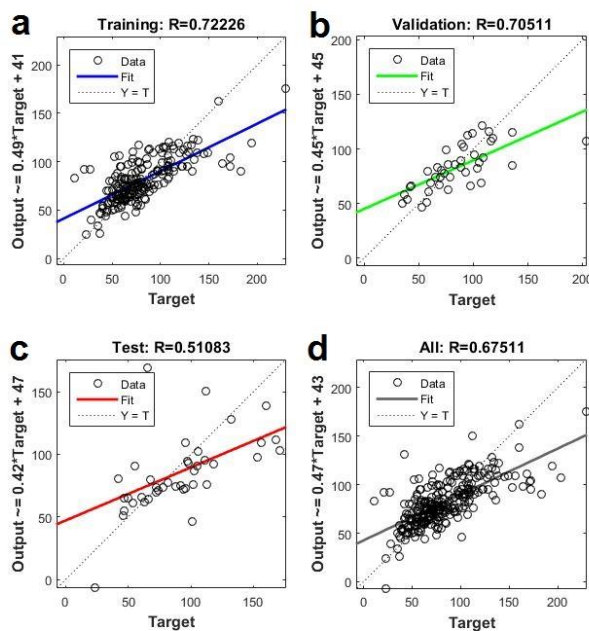
جدول ۴-۶: جدول دقت تخمین غلظت عناصر ژئوشیمیایی در حالت استفاده از کل داده‌ها و حالت استفاده از الگوریتم خوشه‌بندی

Gustafson kessel Clustering method					All Data				
Test R ²	Validation R ²	Training R ²	All R ²	Variable	Test R ²	Validation R ²	Training R ²	All Data R ²	Variable
0.92	0.91	0.97	0.95	Al	0.94	0.94	0.96	0.95	Al
0.73	0.75	0.75	0.75	As	0.6	0.57	0.64	0.62	As
0.7	0.68	0.73	0.71	Ba	0.89	0.85	0.86	0.87	Ba
0.81	0.85	0.86	0.85	Be	0.82	0.82	0.83	0.83	Be
0.82	0.91	0.85	0.86	Ca	0.67	0.76	0.77	0.75	Ca
0.84	0.83	0.84	0.84	Co	0.71	0.68	0.72	0.71	Co
0.95	0.93	0.92	0.93	Cu	0.86	0.82	0.81	0.81	Cu
0.71	0.92	0.76	0.78	Fe	0.7	0.72	0.71	0.71	Fe
0.87	0.91	0.89	0.89	K	0.85	0.86	0.87	0.87	K
0.72	0.86	0.9	0.86	La	0.74	0.78	0.82	0.81	La
0.86	0.86	0.89	0.88	Mg	0.72	0.77	0.78	0.77	Mg
0.75	0.73	0.77	0.76	Mn	0.64	0.6	0.7	0.68	Mn
0.81	0.7	0.72	0.73	Mo	0.64	0.62	0.66	0.65	Mo
0.75	0.8	0.75	0.76	Na	0.89	0.88	0.89	0.89	Na
0.81	0.73	0.77	0.77	Ni	0.69	0.7	0.7	0.7	Ni
0.78	0.72	0.81	0.79	P	0.65	0.73	0.72	0.71	P
0.84	0.82	0.81	0.81	Pb	0.6	0.62	0.65	0.64	Pb
0.92	0.94	0.9	0.91	S	0.82	0.86	0.84	0.84	S
0.85	0.84	0.87	0.86	Sc	0.74	0.73	0.73	0.73	Sc
0.74	0.85	0.82	0.81	Sr	0.8	0.71	0.76	0.76	Sr
0.83	0.82	0.83	0.83	V	0.71	0.76	0.71	0.72	V
0.51	0.71	0.72	0.68	Zn	0.43	0.54	0.53	0.51	Zn
0.8	0.82	0.82	0.82	AVERAG E	0.73	0.74	0.76	0.75	AVERAG E

در استفاده از الگوریتم خوشه‌بندی علاوه بر اینکه باعث افزایش دقت تخمین، صحت تخمین نیز افزایش پیدا می‌کند و مقادیر تخمینی به مقادیر واقعی نزدیکتر می‌شوند. برای مشاهده تاثیر خوشه‌بندی بر افزایش دقت و صحت (مثال - عنصر Zn با کمترین دقت)، نمودار رگرسیونی داده‌های واقعی و تخمینی در حالت استفاده از کل داده‌ها و داده‌های خوشه‌بندی شده در شکل ۴-۳۰ و شکل ۴-۳۱ آورده شده است.



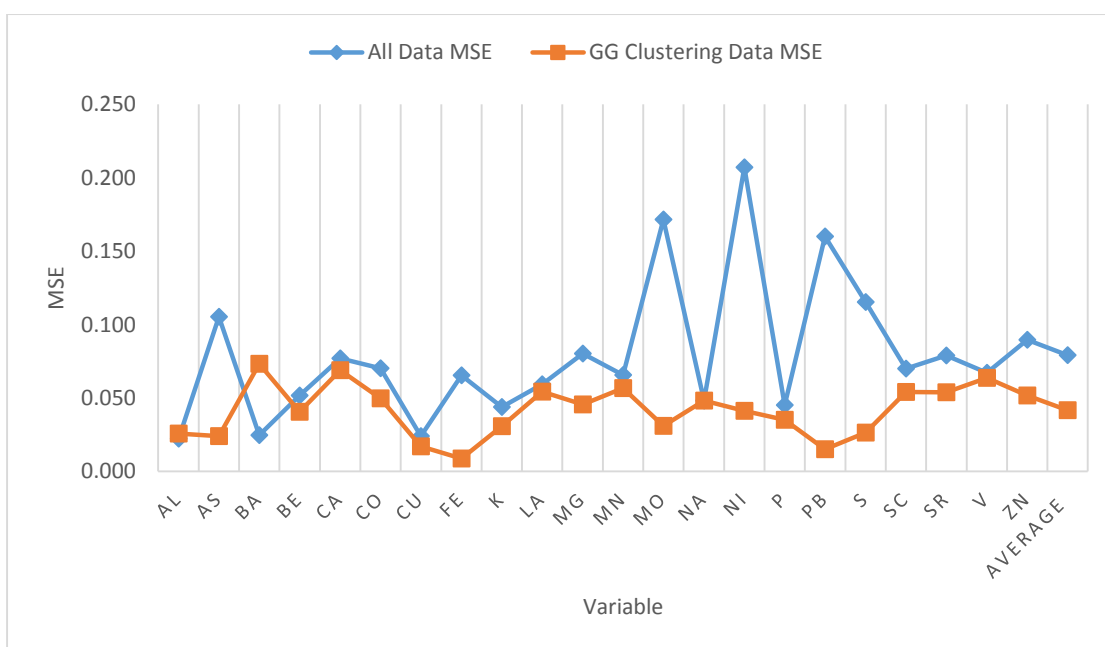
شکل ۴-۳: نمودار رگرسیونی داده‌های واقعی و تخمینی عنصر Zn در حالت استفاده از کل داده‌ها. محور x داده‌های واقعی و محور y داده‌های تخمینی است. (a) داده‌های آموزشی، (b) داده‌های اعتبارسنجی، (c) داده‌های آزمون و (d) کل داده‌ها.



شکل ۴-۳: نمودار رگرسیونی داده‌های واقعی و تخمینی عنصر GG در حالت استفاده از داده‌های خوشه‌بندی شده با روش GG. محور x داده‌های واقعی و محور y داده‌های تخمینی است. (a) داده‌های آموزشی، (b) داده‌های اعتبارسنجی، (c) داده‌های آزمون و (d) کل داده‌ها.

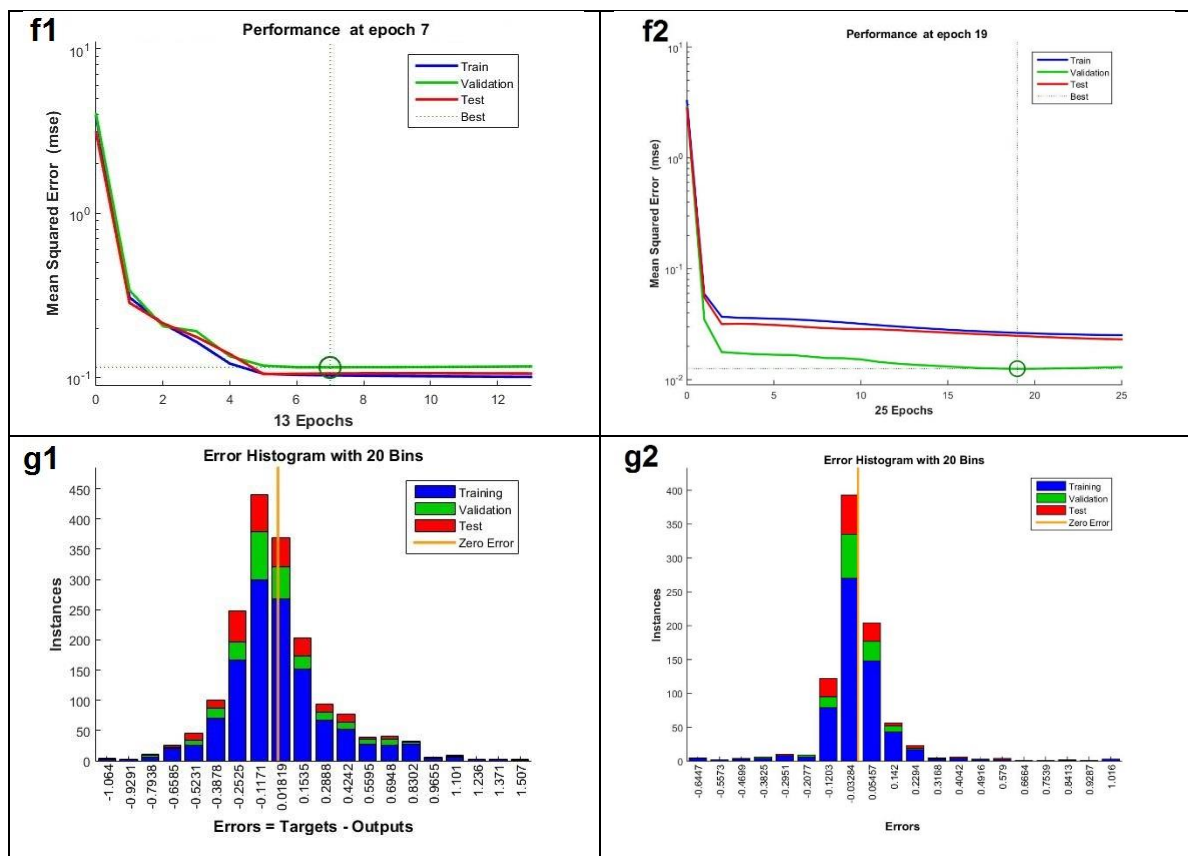
۴-۳-۲- مقایسه خطای تخمینگر با حالت کلی

مقدار میانگین خطای تخمینگرهای ۲۲ عنصر مورد مطالعه در حالت استفاده از الگوریتم خوشه‌بندی GG به ۰/۰۴۲ کاهش یافته که نشانگر آن است که استفاده از الگوریتم خوشه‌بندی علاوه بر افزایش دقت تخمین، خطای تخمین را نیز کاهش می‌دهد. این امر کارایی استفاده از روش‌های خوشه‌بندی را نشان می‌دهد. نتایج خطای تخمینگر عناصر ژئوشیمیایی در شکل ۴-۳۲ آمده است.



شکل ۴-۳۲: مقایسه خطای تخمین در حالت استفاده از کل داده‌ها و حالت استفاده از الگوریتم خوشه‌بندی

نمودار کارایی و هیستوگرام پراکندگی مقادیر خطا برای عنصر As در شکل ۴-۳۳ نشان می‌دهد که عملکرد تخمینگر برای داده‌های خوشه‌بندی شده با کاهش مقادیر خطا و متقارن سازی هیستوگرام همراه است. به طور کلی خطای تخمین در تمامی عناصر هنگام استفاده از الگوریتم خوشه‌بندی کاهش یافته است.



شکل ۴-۳: نمودار تغییرات خطا و هیستوگرام پراکندگی مقادیر خطا عنصر (f1, As) نمودار خطا در حالت استفاده از کل داده‌ها و (f2) حالت استفاده از داده‌های خوشه‌بندی شده، (g1) هیستوگرام خطا در حالت استفاده از کل داده‌ها و (g2) حالت استفاده از داده‌های خوشه‌بندی شده

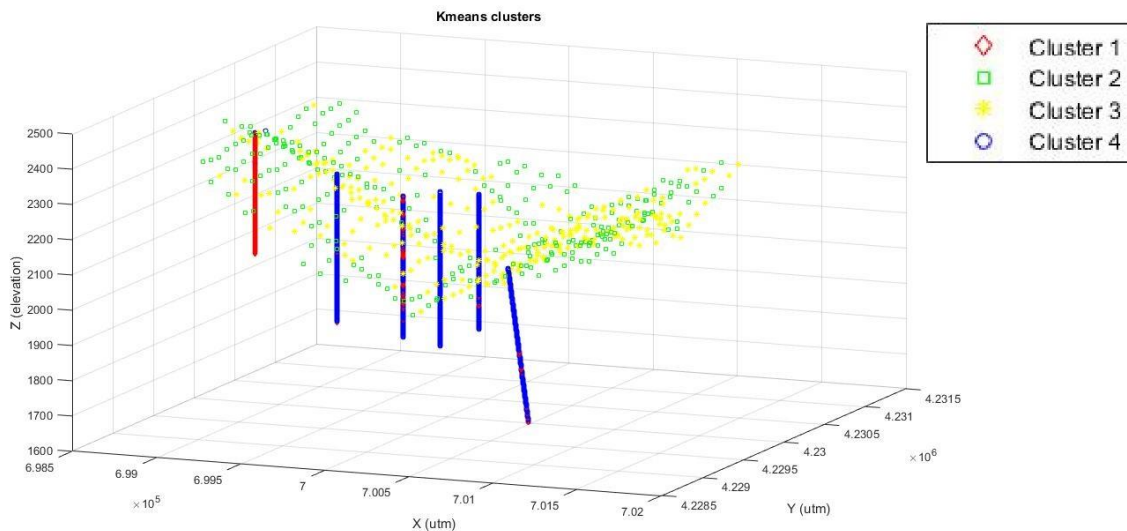
۴-۳-۵- افزایش دقت تخمین با الگوریتم خوشه‌بندی K means

خوشه‌بندی با استفاده از الگوریتم خوشه‌بندی میانگین K و با تعداد ۴ خوشه تعیین شده بوسیله شاخص‌های اعتبارسنجی انجام گردید. در خوشه اول ۲۵۳ داده، خوشه دوم ۲۶۲، خوشه سوم ۲۸۴ و در خوشه چهارم ۹۵۷ داده قرار گرفته است (جدول ۴-۷).

جدول ۴-۷: نتایج خوشه‌بندی داده‌ها با استفاده از الگوریتم خوشه‌بندی K means

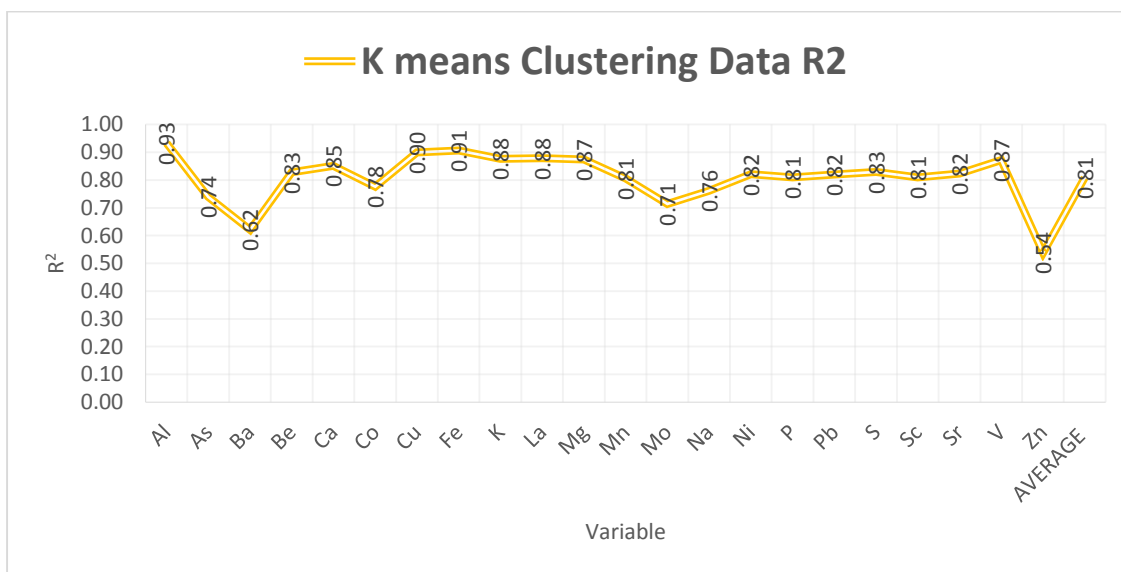
تعداد داده در خوشه‌های مختلف				تعداد کل داده‌ها	الگوریتم خوشه‌بندی
خوشه اول	خوشه دوم	خوشه سوم	خوشه چهارم		
۲۵۳	۲۶۲	۲۸۴	۹۵۷	۱۷۵۵	K means

پراکندگی خوشه‌ها در فضای سه بعدی محورهای مختصات در شکل ۴-۳۴ نمایش داده شده است.



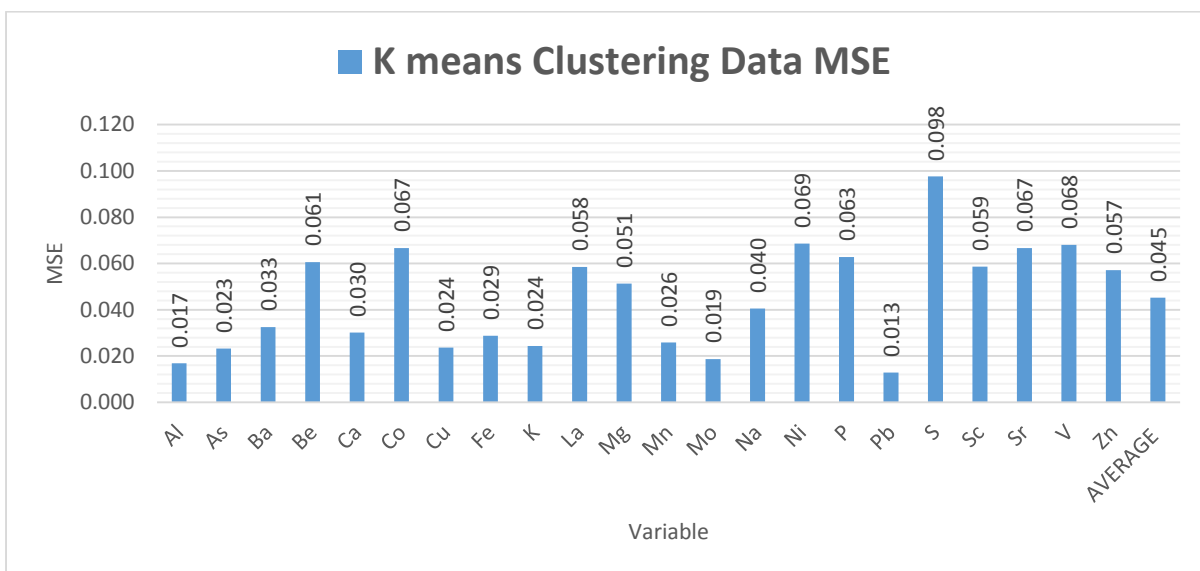
شکل ۴-۳۴: پراکندگی نمونه‌های خوشه‌بندی شده به روش K means به صورت سه بعدی در فضا

در این مرحله، ورودی تخمینگر شبکه عصبی از داده‌های خوشه‌بندی شده توسط الگوریتم K means انتخاب و تخمینگر طراحی شده است. نتایج حاصل از طراحی تخمینگر در حالت استفاده از داده‌های خوشه‌بندی شده به روش K means برای عناصر ژئوشیمیایی در شکل ۴-۳۵ و شکل ۴-۳۶ آمده است.



شکل ۴-۳۵: دقت تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از داده‌های خوشه‌بندی شده به روش K means

نتایج بدست آمده نشان می‌دهد که دقت تخمینگر به طور کلی افزایش یافته است و دقت میانگین با افزایش ۶ درصدی به ۸۱ درصد رسیده است. خطای تخمین حاصل از استفاده از داده‌های خوشه‌بندی شده با استفاده از روش خوشه‌بندی K means در شکل ۴-۳۶ نشان می‌دهد که میانگین خطای تخمین عناصر ژئوشیمیایی ۰/۰۴۵ است که در مقایسه به حالت استفاده از کل داده‌ها که برابر با ۰/۰۷۹ بود کاهش نشان می‌دهد.



شکل ۴-۳۶: خطای تخمینگر غلظت عناصر ژئوشیمیایی در حالت استفاده از داده‌های خوشه‌بندی شده به روش K means

۴-۳-۵-۱- مقایسه دقت تخمینگر با حالت کلی

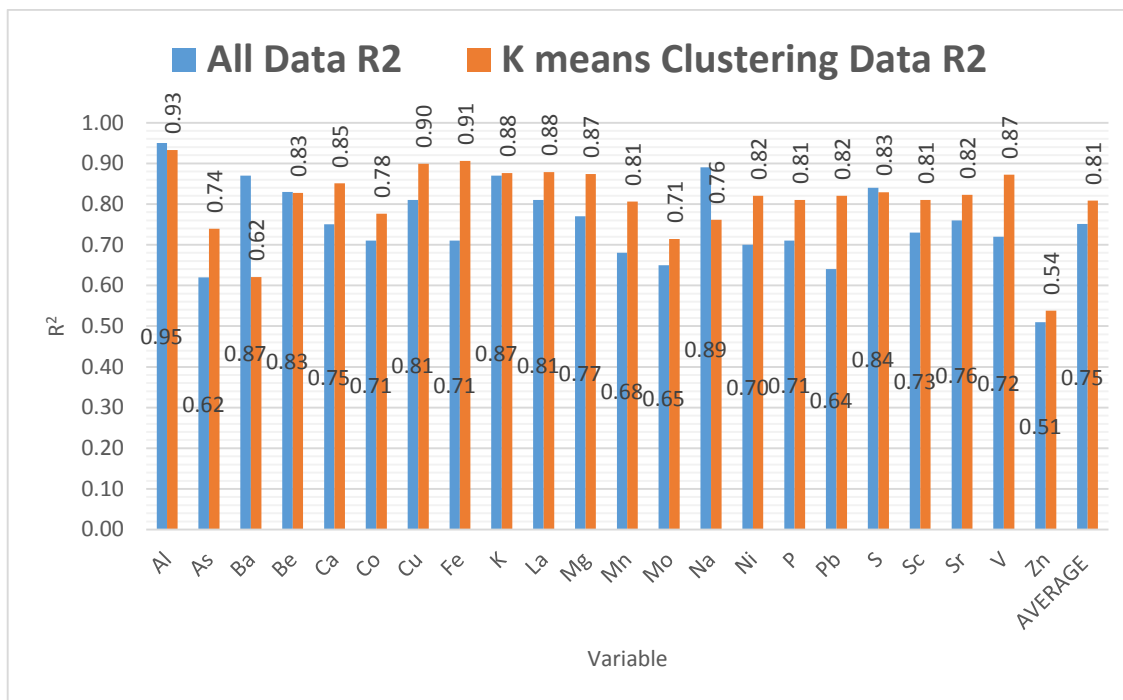
مقایسه نتایج حاصل از دو حالت طراحی تخمینگرها در حالت استفاده از کل داده‌ها و استفاده از داده‌های خوشه‌بندی شده برای تخمین غلظت عناصر مختلف نشان می‌دهد که دقت تخمینگرها در حالت استفاده از داده‌های خوشه‌بندی شده افزایش یافته و دقت میانگین تخمینگرها از ۷۵ درصد به ۸۱ درصد افزایش یافته است (شکل ۴-۳۷).

عناصری مانند Zn, As, Mo, Pb و Mn که به ترتیب دارای دقت‌های پایین ۰/۵۱, ۰/۶۲, ۰/۶۴, ۰/۶۵ و ۰/۶۸ در حالت استفاده از کل داده‌ها بوده، در حالت استفاده از داده‌های خوشه‌بندی شده به ترتیب به

۰/۵۴، ۰/۷۴، ۰/۸۲، ۰/۷۱ و ۰/۸۱ افزایش یافته است. این امر نشان می‌دهد که بر خلاف سایر عناصر، عنصر روی افزایش قابل قبولی ارایه نکرده است. نتایج تخمین‌گرهای عناصر مختلف در حالت استفاده از کل داده‌ها و حالت استفاده از داده‌های خوشه‌بندی شده به روش K means در جدول ۴-۸ قابل مشاهده است.

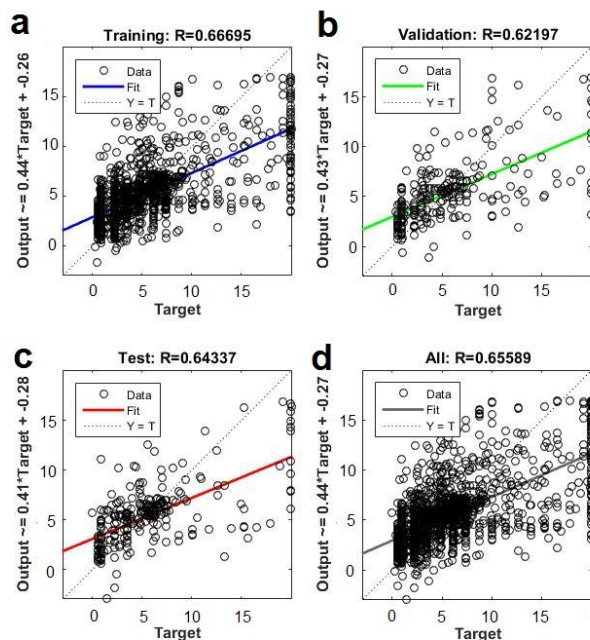
جدول ۴-۸: جدول دقت تخمین غلظت عناصر ژئوشیمیایی در حالت استفاده از کل داده‌ها و حالت استفاده از الگوریتم خوشه‌بندی K means

K means Clustering Method					All Data				
Test R ²	Validation R ²	Training R ²	All R ²	Variable	Test R ²	Validation R ²	Training R ²	All R ²	Variable
0.93	0.93	0.93	0.93	Al	0.94	0.94	0.96	0.95	Al
0.7	0.73	0.75	0.74	As	0.6	0.57	0.64	0.62	As
0.69	0.67	0.6	0.62	Ba	0.89	0.85	0.86	0.87	Ba
0.64	0.84	0.87	0.83	Be	0.82	0.82	0.83	0.83	Be
0.82	0.82	0.86	0.85	Ca	0.67	0.76	0.77	0.75	Ca
0.73	0.73	0.79	0.78	Co	0.71	0.68	0.72	0.71	Co
0.88	0.92	0.9	0.9	Cu	0.86	0.82	0.81	0.81	Cu
0.87	0.96	0.9	0.91	Fe	0.7	0.72	0.71	0.71	Fe
0.86	0.88	0.88	0.88	K	0.85	0.86	0.87	0.87	K
0.72	0.81	0.93	0.88	La	0.74	0.78	0.82	0.81	La
0.89	0.88	0.87	0.87	Mg	0.72	0.77	0.78	0.77	Mg
0.79	0.76	0.82	0.81	Mn	0.64	0.6	0.7	0.68	Mn
0.78	0.63	0.83	0.71	Mo	0.64	0.62	0.66	0.65	Mo
0.73	0.69	0.78	0.76	Na	0.89	0.88	0.89	0.89	Na
0.73	0.88	0.83	0.82	Ni	0.69	0.7	0.7	0.7	Ni
0.68	0.82	0.84	0.81	P	0.65	0.73	0.72	0.71	P
0.89	0.83	0.8	0.82	Pb	0.6	0.62	0.65	0.64	Pb
0.83	0.82	0.83	0.83	S	0.82	0.86	0.84	0.84	S
0.84	0.74	0.81	0.81	Sc	0.74	0.73	0.73	0.73	Sc
0.76	0.85	0.84	0.82	Sr	0.8	0.71	0.76	0.76	Sr
0.71	0.81	0.93	0.87	V	0.71	0.76	0.71	0.72	V
0.6	0.36	0.56	0.54	Zn	0.43	0.54	0.53	0.51	Zn
0.78	0.79	0.83	0.81	Average	0.73	0.74	0.76	0.75	Average

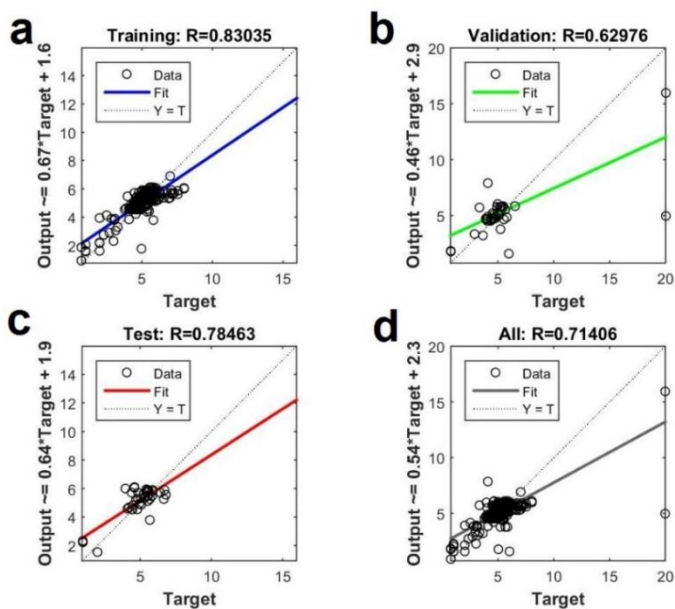


شکل ۴-۳۷: مقایسه دقت کل تخمینگرها در حالت استفاده از کل داده‌ها و داده‌های خوشه‌بندی شده به روش K means

در استفاده از الگوریتم خوشه‌بندی علاوه بر افزایش دقت تخمین، صحت تخمین نیز افزایش پیدا می‌کند و مقادیر تخمینی به مقادیر واقعی نزدیکتر می‌شوند. برای مشاهده تاثیر خوشه‌بندی بر افزایش دقت و صحت (مثال - عنصر Mo با افزایش دقت از ۶۵ درصد به ۷۳ درصد)، نمودار رگرسیونی داده‌های واقعی و تخمینی در حالت استفاده از کل داده‌ها و داده‌های خوشه‌بندی شده در شکل ۴-۳۸ و شکل ۴-۳۹ آورده شده است. بررسی نتایج نشان می‌دهد که استفاده از داده‌های خوشه‌بندی شده، همبستگی بین داده‌های ورودی و خروجی تخمینگر را افزایش داده است و دقت تخمین افزایش یافته است.



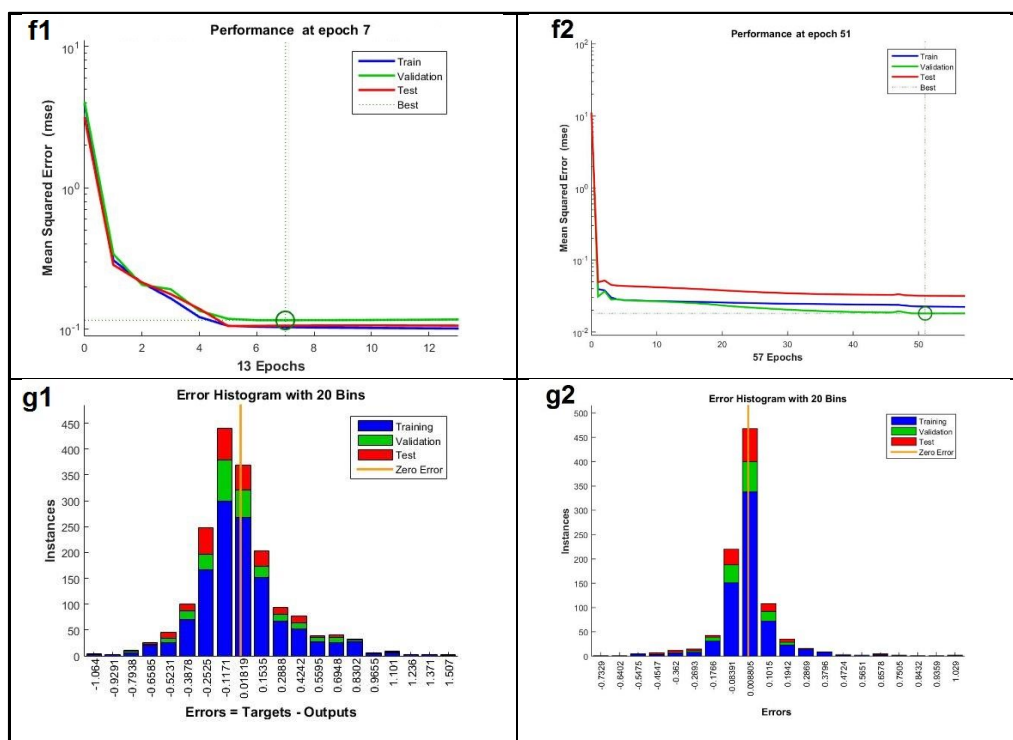
شکل ۴-۳۸: نمودار رگرسیونی داده‌های واقعی و تخمینی عنصر Mo در حالت استفاده از کل داده‌ها. محور x داده‌های واقعی و محور y داده‌های تخمینی است. (a) داده‌های آموزشی، (b) داده‌های اعتبارسنجی، (c) داده‌های آزمون و (d) کل داده‌ها.



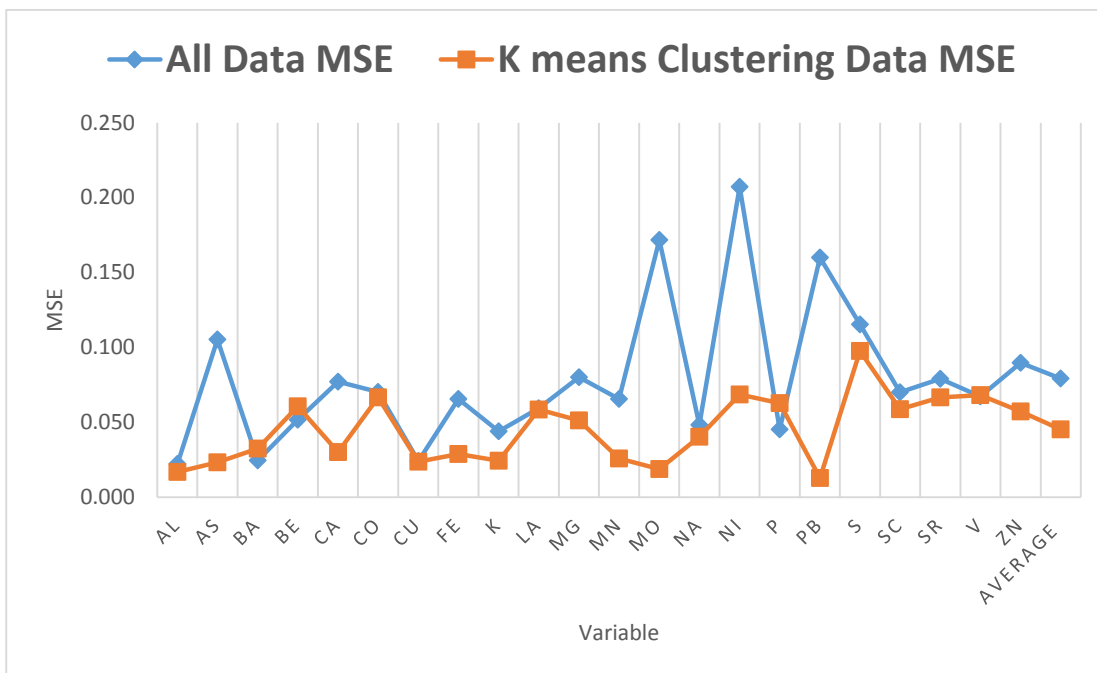
شکل ۴-۳۹: نمودار رگرسیونی داده‌های واقعی و تخمینی عنصر Mo در حالت استفاده از داده‌های خوشه‌بندی شده با روش K means. محور x داده‌های واقعی و محور y داده‌های تخمینی است. (a) داده‌های آموزشی، (b) داده‌های اعتبارسنجی، (c) داده‌های آزمون و (d) کل داده‌ها.

۴-۳-۵-۲- مقایسه خطای تخمینگر با حالت کلی

مقدار میانگین خطای تخمینگرهای ۲۲ عنصر مورد مطالعه در حالت استفاده از الگوریتم خوشه‌بندی K means به ۰/۰۴۵ کاهش یافته که نشانگر آن است که استفاده از الگوریتم خوشه‌بندی K means علاوه بر افزایش دقت تخمین، خطای تخمین را نیز کاهش می‌دهد. این امر کارایی استفاده از روش‌های خوشه‌بندی را نشان می‌دهد. نتایج خطای تخمینگر عناصر ژئوشیمیایی در شکل ۴-۴۱ آمده است. نمودار کارایی و هیستوگرام پراکندگی مقادیر خطا برای عنصر As در شکل ۴-۴۰ نشان می‌دهد که عملکرد تخمینگر برای داده‌های خوشه‌بندی شده با کاهش مقادیر خطا و متقارن سازی هیستوگرام همراه است. به طور کلی خطای تخمین در تمامی عناصر هنگام استفاده از الگوریتم خوشه‌بندی کاهش یافته است.



شکل ۴-۴۰: نمودار تغییرات خطا و هیستوگرام پراکندگی مقادیر خطا عنصر As، f1) نمودار خطا در حالت استفاده از کل داده‌ها، f2) نمودار خطا در حالت استفاده از داده‌های خوشه‌بندی شده، g1) هیستوگرام خطا در حالت استفاده از کل داده‌ها و g2) هیستوگرام خطا در حالت استفاده از داده‌های خوشه‌بندی شده



شکل ۴-۴: مقایسه خطای تخمین در حالت استفاده از کل داده‌ها و حالت استفاده از الگوریتم خوشه‌بندی K means

۴-۴- انتخاب بهترین تخمینگر

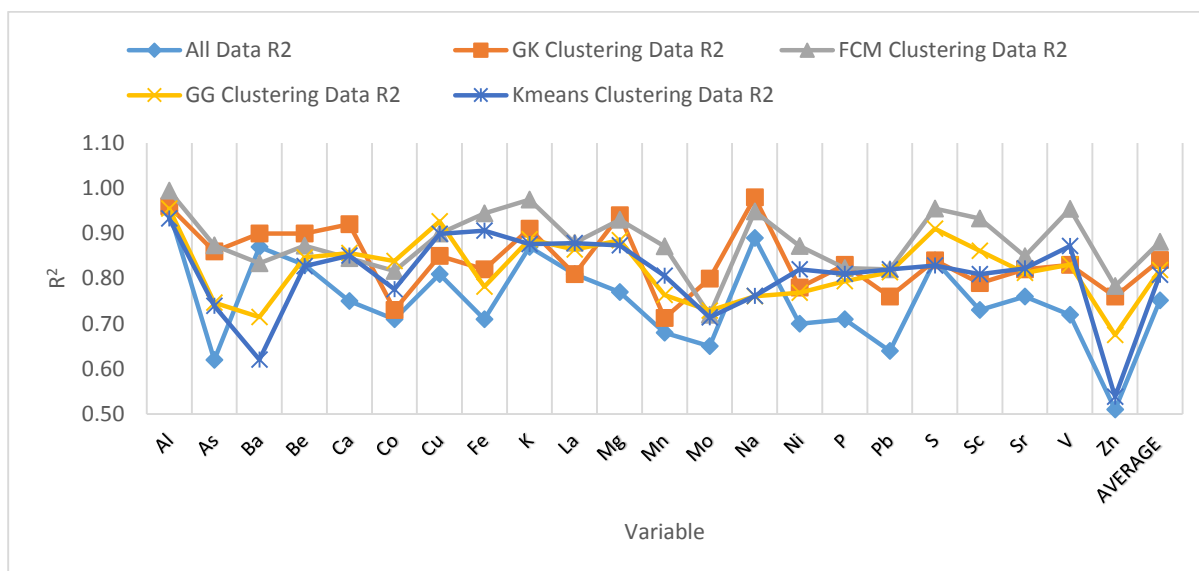
استفاده از تخمینگر روشی است که بوسیله آن می‌توانیم اطلاعات جامعه مشاهده شده را به جامعه مشاهده نشده تعمیم دهیم. انتخاب روش تخمین و همچنین پارامترهای ورودی تخمین ارتباط مستقیمی با دقت و صحت داده‌های تخمینی دارند. بنابراین شبکه عصبی به جهت داشتن ویژگی شناسایی روابط پیچیده بین داده‌های ورودی و خروجی بهترین روش تخمین برای داده‌های ژئوشیمیایی است و برای انتخاب بهترین ورودی‌ها برای تخمینگر شبکه عصبی، ابتدا داده‌ها به چهار روش FCM، GK، GG و K means خوشه‌بندی شده است. سپس تخمینگر شبکه عصبی با استفاده از داده‌های خوشه‌بندی طراحی و نتایج بدست آمده نشان داد که استفاده از داده‌های خوشه‌بندی شده همراه با افزایش دقت و کاهش خطای تخمین نسبت به زمانی است که از کل داده‌ها برای ورودی تخمینگر استفاده نماییم.

جدول ۴-۹: مقایسه دقت و خطای تخمینگر در حالت‌های مختلف

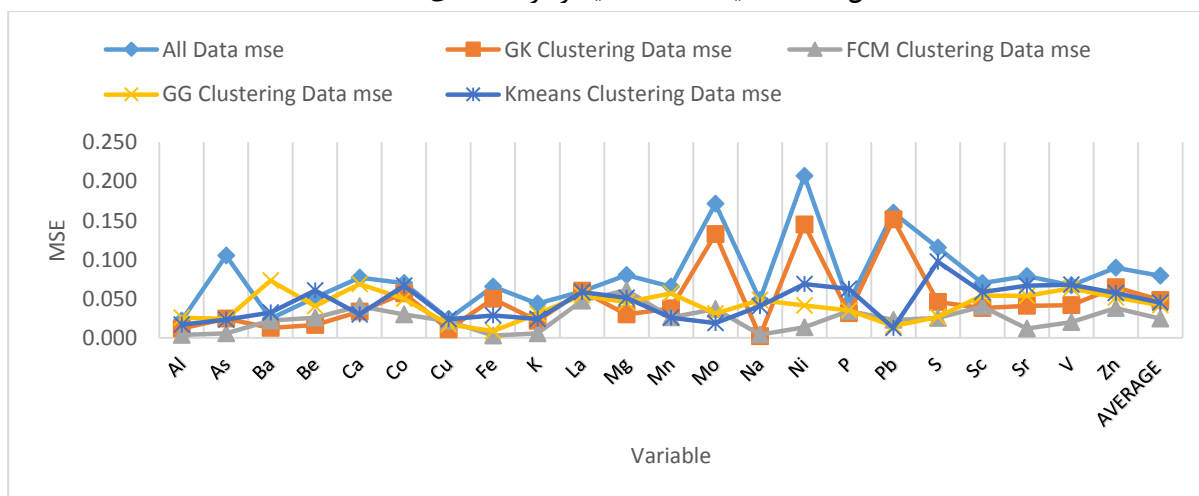
تخمینگر	داده‌های ورودی	تعداد عنصر	میانگین دقت (R^2)	میانگین خطا (MSE)
شبکه عصبی mlp	کل داده‌ها	۲۲	۷۵ درصد	۰/۰۷۹
شبکه عصبی mlp	داده‌های خوشه‌بندی شده به روش FCM	۲۲	۸۸ درصد	۰/۰۲۵
شبکه عصبی mlp	داده‌های خوشه‌بندی شده به روش GK	۲۲	۸۴ درصد	۰/۰۴۸
شبکه عصبی mlp	داده‌های خوشه‌بندی شده به روش GG	۲۲	۸۲ درصد	۰/۰۴۲
شبکه عصبی mlp	داده‌های خوشه‌بندی شده به روش K means	۲۲	۸۱ درصد	۰/۰۴۵

نتایج بدست آمده نشان می‌دهد که استفاده ترکیبی از روش‌های خوشه‌بندی و شبکه عصبی برای تخمین غلظت عناصر، باعث افزایش دقت تخمین و همچنین کاهش خطای تخمین می‌گردد. مقایسه حالت مختلف طراحی تخمینگر (جدول ۴-۹) نشانگر آن است که دقت تخمین در حالت استفاده از الگوریتم خوشه‌بندی FCM با افزایش ۱۳ درصدی از ۷۵ درصد در حالت استفاده از کل داده‌ها به ۸۸ درصد رسیده است و بهترین دقت را در مقایسه با دیگر روش‌ها ارایه می‌کند و همچنین میانگین خطای تخمین در حالت استفاده از این الگوریتم از ۰/۰۷۹ در حالت استفاده از کل داده‌ها به ۰/۰۲۵ کاهش پیدا کرده است و نسبت به سایر روش‌ها کاهش خطای بیشتری نشان داده است.

با مقایسه دقت و خطای بدست آمده از روش‌های مختلف ترکیبی تخمین غلظت عناصر، نتیجه‌گیری می‌شود که ترکیب روش خوشه‌بندی FCM با شبکه عصبی چندلایه بهترین دقت و کمترین خطا را حاصل می‌کند (شکل ۴-۴۲ و شکل ۴-۴۳). بنابراین ورودی تخمینگر شبکه عصبی را از داده‌های خوشه‌بندی شده به روش FCM انتخاب نموده و غلظت عناصر ژئوشیمیایی را در مقاطع مختلف عمقی تخمین می‌زنیم و در مرحله بعدی که شامل طبقه‌بندی غلظت عنصر کانی‌ساز مس در سطوح مختلف عمقی است را با استفاده از داده‌های تخمینی بدست آمده در فصل بعد انجام می‌دهیم.



شکل ۴-۴۲: مقایسه دقت تخمینگر در حالت‌های مختلف



شکل ۴-۴۳: مقایسه خطای تخمینگر در حالت‌های مختلف

۴-۴-۱- بررسی کیفیت داده‌های تخمینی

برای بررسی داده‌های تخمین زده شده از لحاظ کیفیت و همچنین اریب بودن تخمینگر مقایسه‌ای بین داده‌های واقعی و داده‌های تخمینی انجام می‌دهیم. پس از طراحی تخمینگر، داده‌های سطحی را که شامل ۵۶۲ نمونه با ۲۲ عنصر می‌باشند را تخمین می‌زنیم تا بتوانیم کیفیت داده‌های تخمینی را با داده‌های واقعی ارزیابی نماییم. برای اینکار حداقل، حداکثر، میانه و میانگین غلظت عناصر را در داده‌های واقعی و تخمینی مقایسه می‌کنیم.

جدول ۴-۱۰: مقایسه حداقل و حداکثر داده‌های واقعی و تخمینی (ppm)

Estimated Data Max	Real Data Max	Estimated Data Min	Real Data Min	Variable
123290	141000	25273	27800	Al
452	623	1	0.9	As
2985	3590	146	135	Ba
3.8	4.8	0.27	0.3	Be
95972	110000	540	477	Ca
25	32.9	0.8	0.2	Co
5307	6020	2	7.1	Cu
125510	147000	11668	12900	Fe
45946	59500	643	502	K
54.1	60	10.5	7.5	La
25451	27100	154.4	130	Mg
2797.4	3010	15.6	14	Mn
51.7	66.2	1	0.3	Mo
46346	45100	440	219	Na
51.5	71	0.5	0.75	Ni
3254	3220	490	208	P
260	227	1.15	3.4	Pb
29929	33800	25.41	30	S
20.72	25	3.3	2	Sc
2757	3440	49	60.4	Sr
275.4	349	16.8	13	V
331.8	290	7.5	5.4	Zn

جدول ۴-۱۱: مقایسه میانه و میانگین داده‌های واقعی و تخمینی (ppm)

Estimated Data Average	Real Data Average	Estimated Data Median	Real Data Median	Variable
83424.38	86618.1	85521.7	86200	Al
60.15	66.45	58.99	50	As
879.29	931.65	884	911	Ba
1.68	1.76	1.56	1.6	Be
21557.42	21324.45	20212.75	19200	Ca
11.46	11.38	11.45	10.7	Co
258.29	199.52	179.89	94.4	Cu
41036.4	41863.99	40265.67	39500	Fe
21848.75	23381.51	21089.56	22800	K
29.95	30.47	30.02	29	La
9276.36	8998.61	8906.21	7730	Mg
703.02	694.4	655.51	624	Mn
4.92	5.6	4.7	3.8	Mo
19507.04	20353.1	18462.41	21000	Na
9.02	12.26	8.7	7	Ni
1594.09	1612.25	1579.87	1560	P
18.41	23.07	17.11	15.8	Pb
2817.42	1655.81	2090.93	2210	S
9.29	9.56	9	9	Sc
652.03	724.81	642.78	611	Sr
126.78	128.99	123.85	111	V
74.12	74.75	73.84	73.1	Zn

بررسی حداقل، حداکثر، میانه و میانگین داده‌های واقعی (جدول ۴-۱۰ و جدول ۴-۱۱) نشان می‌دهد که داده‌های تخمینی از دقت قابل قبولی برخوردار هستند و در ادامه کار از این داده‌ها استفاده خواهد شد.

فصل

۵- نقشه غلظت عناصر در سطوح مختلف عمقی و

مدل کانسار

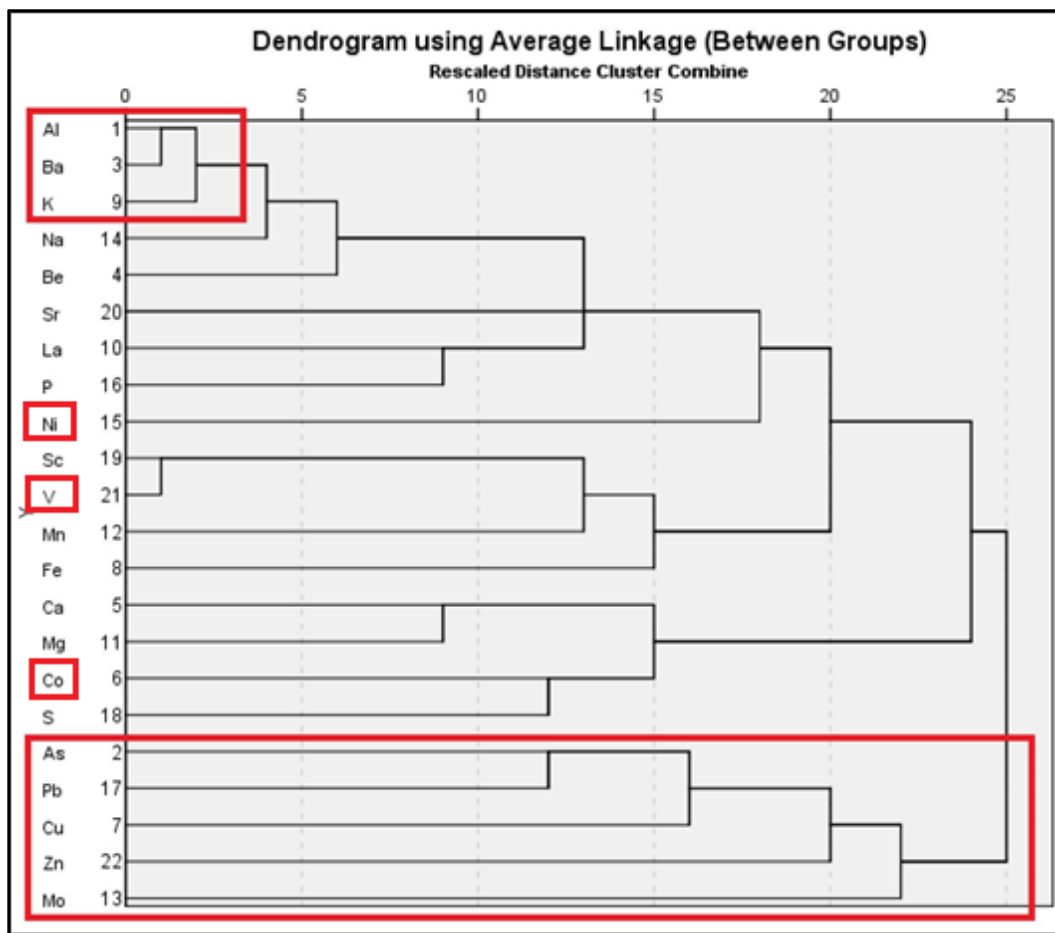
۵-۱- مقدمه

بر اساس نتایج بدست آمده در فصل قبل، ترکیب الگوریتم خوشه‌بندی FCM و تخمینگر شبکه عصبی چندلایه، بهترین دقت را در تخمین غلظت عناصر ژئوشیمیایی حاصل نمود. بنابراین با طراحی تخمینگر، تخمین غلظت عناصر ژئوشیمیایی در سطوح مختلف عمقی انجام شده است. نقشه غلظت و پروفیل‌های ژئوشیمیایی عناصر مرتبط با کانی‌سازی مس پورفیری در سطوح مختلف عمقی ترسیم و مورد بررسی قرار گرفت. هدف نهایی تخمین غلظت عناصر در سطوح مختلف عمقی، رسیدن به خصوصیات ژئوشیمیایی عناصر و بویژه مدل‌سازی کلاس تغییرات غلظت مس در سطوح مختلف عمقی است. برای این منظور از روش‌های شناسایی الگو جهت طبقه‌بندی داده‌ها مانند روش بیزین، K نزدیکترین همسایگی، ماشین بردار پشتیبان و شبکه عصبی که به صورت نظارت شده عمل می‌کنند استفاده شده است. برای این منظور ۷۰ درصد داده‌ها برای آموزش و ۳۰ درصد داده‌ها برای آزمون استفاده شد. برای شناسایی کلاس‌های مس داده‌های آموزشی و داده‌های آزمون، از روش فرکتالی عیار - مساحت استفاده شده است. برای کاهش ابعاد ویژگی‌های ورودی از روش تحلیل مولفه‌های اصلی استفاده شده است. پس از انجام طبقه‌بندی با روش‌های مختلف، نتایج مقایسه شده و در نهایت بهترین روش برای مدل‌سازی تغییرات غلظت مس در سطوح مختلف عمقی معرفی گردیده است.

۵-۲- عناصر تحت و فوق کانساری در سطوح مختلف عمقی

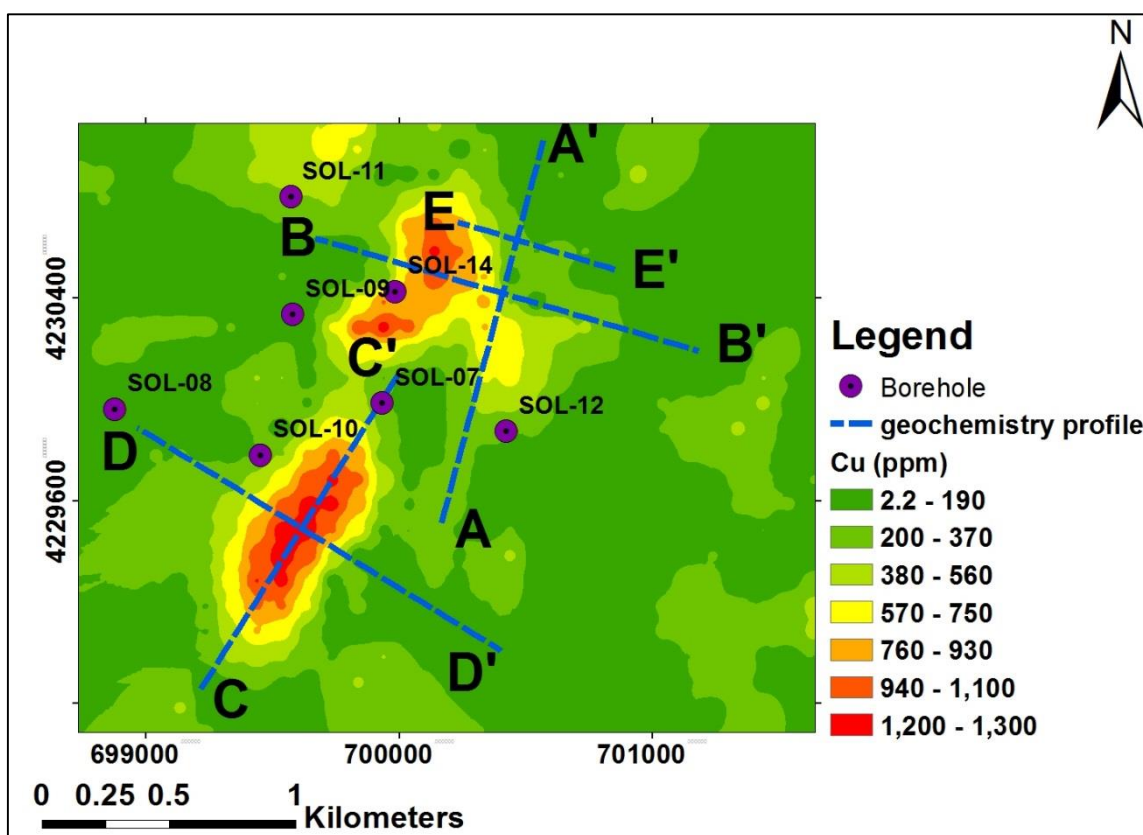
در کانسارهای مس پورفیری عناصری را به عنوان عناصر فوق کانساری، کانی‌ساز و تحت کانساری در نظر می‌گیرند. با بررسی تغییرات غلظت عناصر و نسبت عناصر فوق کانسار به تحت کانسار موقعیت کانسار پنهان و وجود یا عدم وجود کانی‌سازی پنهان در محدوده مورد مطالعه مورد بررسی قرار می‌گیرد. در این تحقیق این کار با بهره‌گیری از الگوریتم‌های شناسایی الگو صورت خواهد پذیرفت تا ارزیابی از وضعیت کانی‌سازی در عمق ارائه نماید. بررسی نمودار خوشه‌ای داده‌های تخمینی نشان داد که عناصر مرتبط با

کانی‌سازی مس (Pb,Cu,Zn,Mo) در ارتباط نزدیک با As در یک خوشه قرار گرفته‌اند. عناصری مانند Al و K که احتمالاً مرتبط با آلتراسیون‌های آرژیلیک و سرسیتیک هستند کنار هم درون یک خوشه از سایر عناصر جدا شده و عناصری مانند Ni، Co و V که نوعی معرف درجه حرارت‌های بالاتر (تحت‌کانساری کانسارهای پورفیری) جدا از هم و درون خوشه‌های مجزا قرار گرفته‌اند که مورد تحلیل قرار خواهند گرفت (شکل ۱-۵).

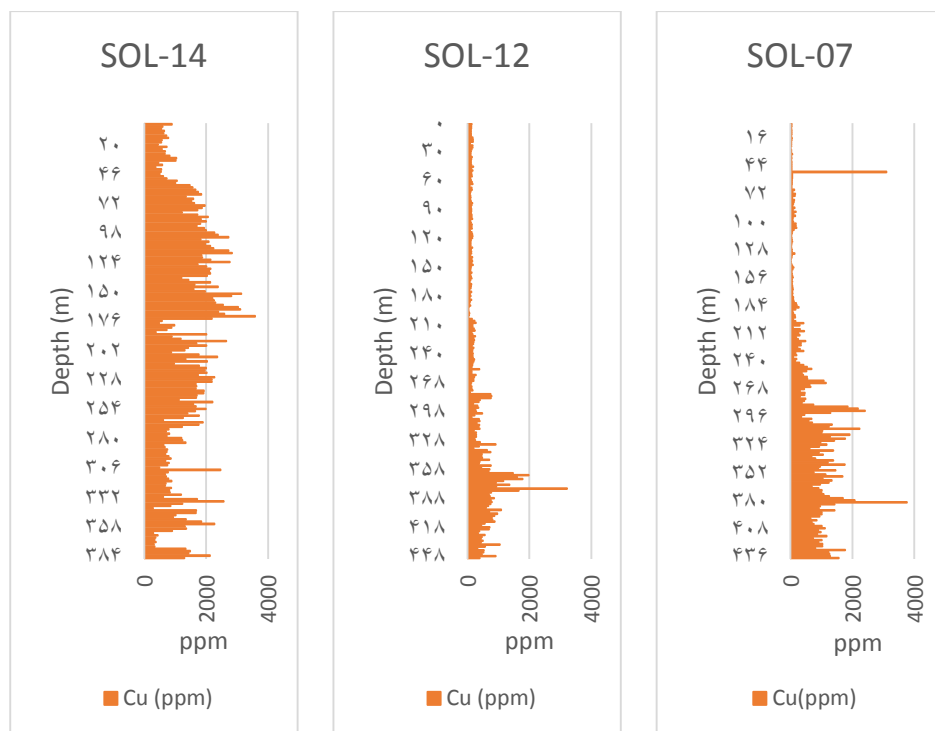


شکل ۱-۵: نمودار خوشه‌ای داده‌های سطحی و عمقی

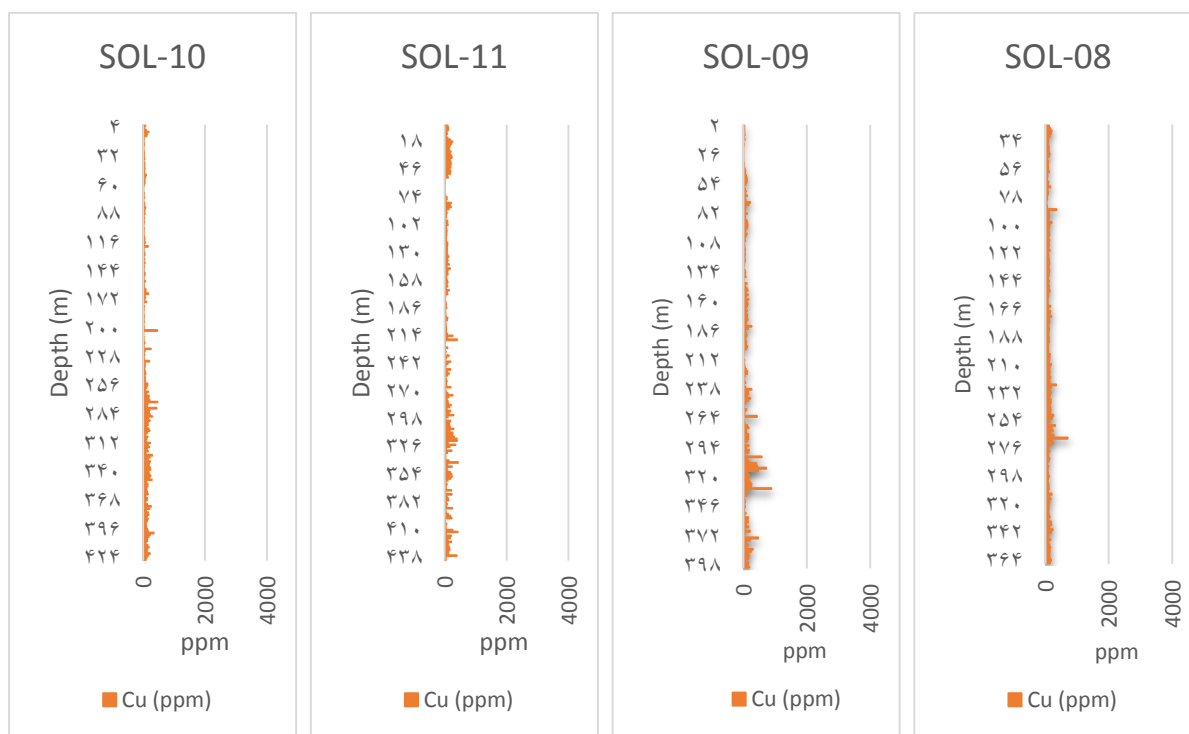
در این تحقیق، به منظور ارزیابی دقیق‌تری از عناصر مرتبط با کانی‌سازی مس، تحت‌کانساری و فوق‌کانساری، اقدام به ترسیم مقاطع عرضی ژئوشیمیایی در جهات و اعماق مختلف گردیده است. پس از بررسی تغییرات غلظت عناصر در سطوح مختلف عمقی، عناصر مهم شناسایی و مورد تحلیل قرار گرفتند. پروفیل‌های AA', BB' و EE' بر روی یکی از مناطق آنومال و CC' و DD' بر روی آنومالی سطحی دیگر از مس در نظر قرار گرفته است (شکل ۵-۲). موقعیت پروفیل‌های AA', BB' با پروفیل‌های ژئوفیزیکی نیز همخوانی دارند. بر روی این نقشه موقعیت گمانه‌های اکتشافی نیز نمایش داده شده است. برای اطمینان بیشتر از وضعیت کانی‌سازی پروفیل قائم گمانه‌های اکتشافی نسبت به عمق عنصر Cu ترسیم شده است (شکل ۵-۳ و شکل ۵-۴).



شکل ۵-۲: موقعیت پروفیل‌های ژئوشیمیایی و گمانه‌های اکتشافی بر روی نقشه غلظت مس در سطح



شکل ۳-۵: تغییرات غلظت مس نسبت به عمق در طول گمانه‌های اکتشافی SOL-14 و SOL-12 ، SOL-07



شکل ۴-۵: تغییرات غلظت مس نسبت به عمق در طول گمانه‌های اکتشافی SOL-10 و SOL-11, SOL-09, SOL-08

نقشه سطحی غلظت مس دو منطقه کانی‌سازی در جنوب‌غربی (نزدیک گمانه SOI-10) و شمال‌مرکزی (مجاورت گمانه SOI-14 در مرکز آن قرار گرفته) را نشان می‌دهد. بررسی تغییرات غلظت مس در گمانه‌ها نشان می‌دهد که گمانه‌های SOL-07، SOL-12 و SOL-14 به ترتیب تقریباً در عمق ۲۵۰ متر، ۳۰۰ متر و ۵۰ متر با مس بالای ۳ درصد (کانی‌سازی) برخورد نموده‌اند (شکل ۳-۵) و لذا حضور کانی‌سازی سطحی (گمانه SOI-14) و یا عمقی (گمانه‌های SOI-12 و SOI-07) می‌باشند. گمانه‌های SOI-08، SOI-09، SOI-11 و SOI-10 به دلیل غلظت پایین مس در عمق، نشان دهنده عدم کانی‌سازی هستند (شکل ۴-۵). با انطباق اطلاعات سطحی و گمانه‌ها مشخص می‌گردد که در بخش شمالی منطقه، کانی‌سازی از عمق ۵۰ متری به بعد شدت می‌یابد، در بخش مرکزی (مجاورت گمانه SOI-07) کانی‌سازی از عمق حدود ۳۰۰ متری شروع می‌شود که در حرکت به سمت غربی منطقه با عدم وجود کانی‌سازی مواجه خواهیم بود.

بر طبق اطلاعات اولیه بدست آمده از سطح فرسایش منطقه اقدام به طراحی پنج پروفیل اکتشافی ژئوشیمیایی 'AA'، 'BB'، 'CC'، 'DD' و 'EE' گردید (شکل ۲-۵) و در ادامه تغییرات غلظت عناصر مهم که در کانسارهای مس پورفیری دارای اهمیت هستند (جدول ۱-۵) در سطوح مختلف عمقی مورد بررسی قرار گرفت و عناصر فوق‌کانساری، تحت‌کانساری، عناصر نشان دهنده آلتراسیون‌ها و عناصر مرتبط با کانی‌سازی مشخص گردید.

جدول ۱-۵: عناصر مهم در کانسارهای مس پورفیری

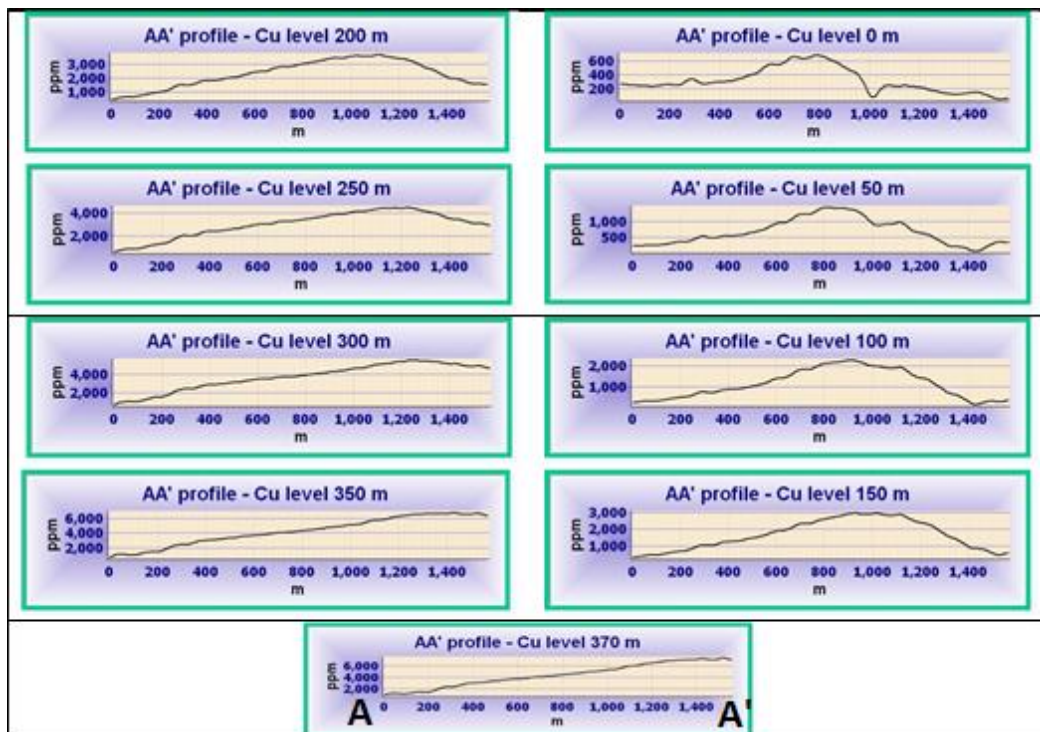
Mg	K	Fe	Cu	Co	Ca	As	Al
	Zn	V	Sr	S	Pb	Ni	Mo

۱-۲-۵- پروفیل ژئوشیمیایی 'AA'

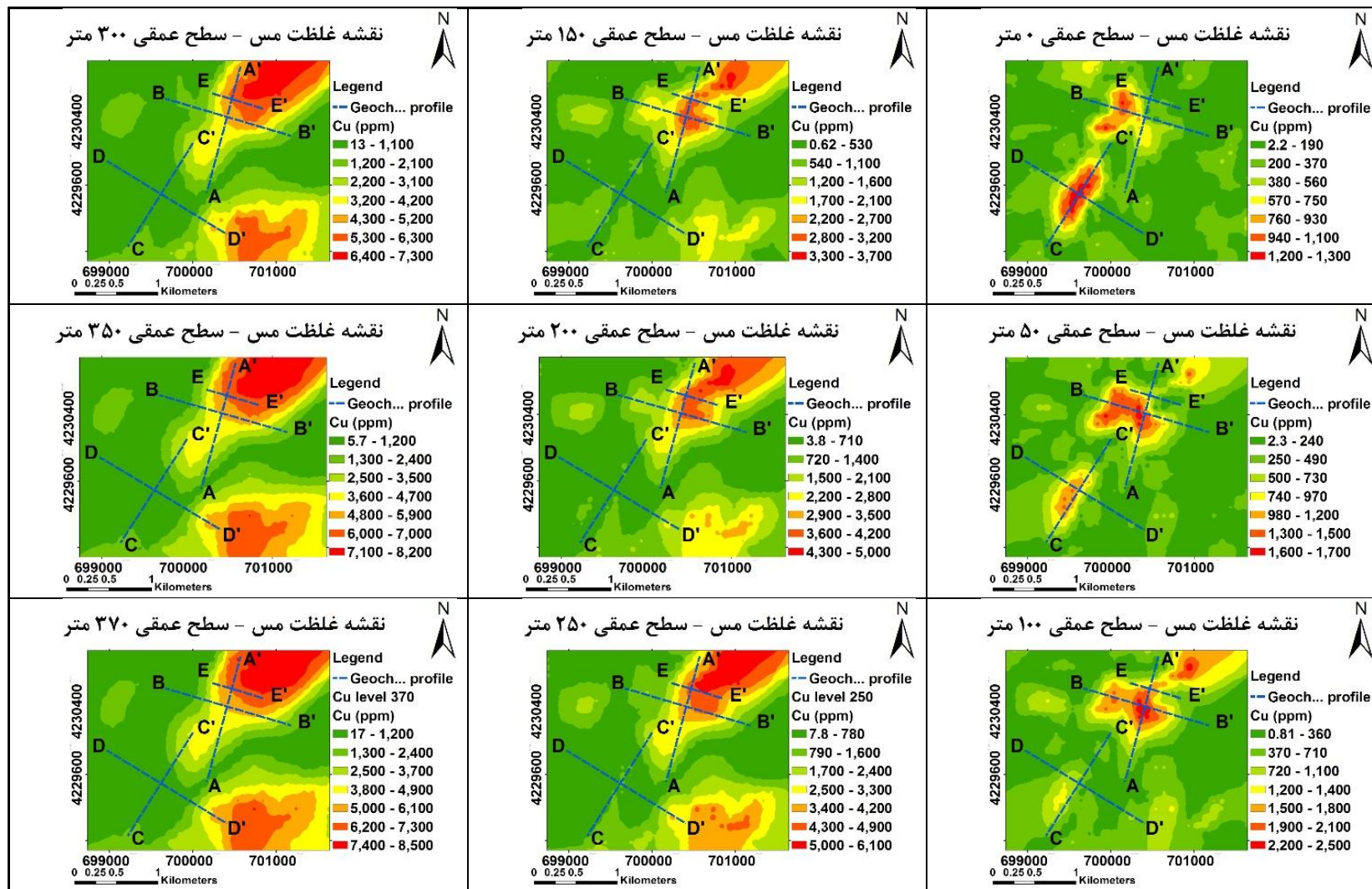
پس از بررسی تغییرات عناصر ژئوشیمیایی در سطوح مختلف عمقی و در راستای پروفیل‌های ژئوشیمیایی، نتایج بدست آمده نشان داد که بعضی از عناصر رفتارهای وابسته به عنصر کانی‌ساز یعنی مس دارند که با

افزایش و کاهش غلظت مس در سطح و عمق، افزایش و کاهش نشان می‌دهند. آرسنیک رفتاری فوق‌کانساری، وانادیوم رفتاری تحت‌کانساری، و عناصر آهن و گوگرد رفتار پاراژن با مس نشان دادند. عنصر آلومینیوم نیز با کاهش از سطح به عمق در شناسایی کانی‌های رسی و آلتراسیون آرژیلیک مناسب تشخیص داده شد. برای نمایش بهتر تغییرات غلظت عناصر، نقشه تغییرات غلظت به همراه پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی برای عناصر انتخاب شده آورده شده است.

نقشه تغییرات غلظت مس در سطوح مختلف عمقی در شکل ۵-۶ و شکل ۵-۵ نشان می‌دهد که آنومالی قسمت جنوب‌غربی این پروفیل بعد از عمق ۵۰ متری از بین می‌رود و کانی‌سازی ادامه ندارد. کانی‌سازی قسمت شمالی منطقه در عمق نیز ادامه پیدا می‌کند و متمایل به سمت شمال‌شرقی محدوده می‌شود. از عمق ۲۰۰ متری یک آنومالی مس در بخش جنوب‌شرقی نمایان شده که در عمق نیز ادامه یافته که وسعت گسترش آن نسبت به آنومالی شمالی کمتر است.

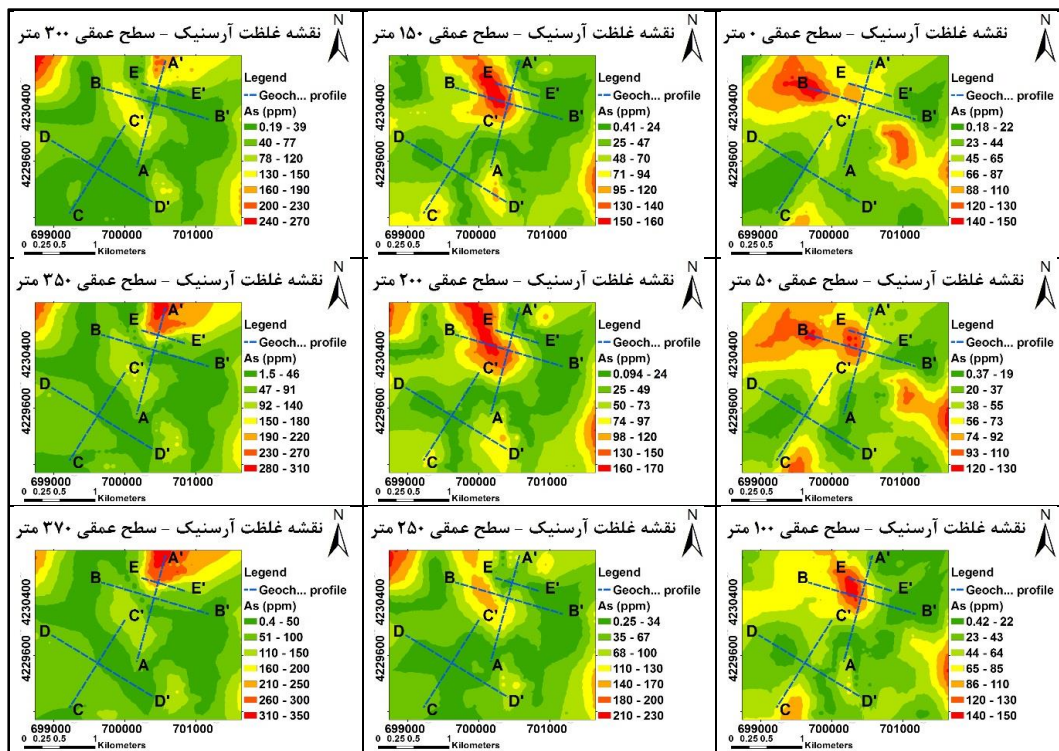


شکل ۵-۵: مقاطع عرضی تغییرات عنصر مس در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی



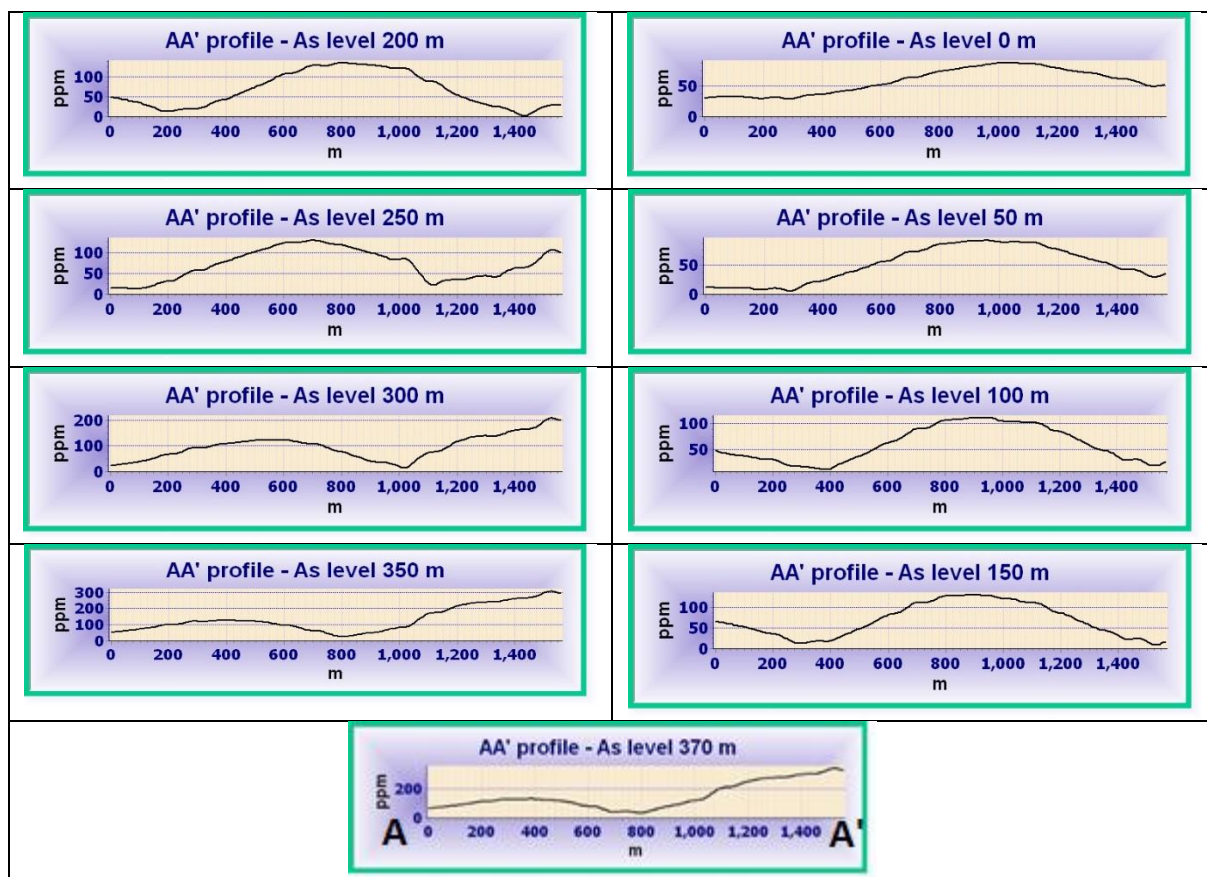
شکل ۵-۶: تغییرات غلظت مس در سطوح مختلف عمقی

تغییرات غلظت آرسنیک در سطوح مختلف عمقی شکل ۵-۷ و شکل ۵-۸ در نقشه بخش سطحی سه منطقه دارای غلظت بالا در قسمت شمالی - شمال غربی، جنوبی و شرقی مشاهده می‌شود که شدت غلظت و گسترش محدوده آنومال این عنصر در قسمت جنوبی و شرقی نسبت به قسمت شمالی کمتر است. شاید بتوان آنومالی آرسنیک را به عنوان هاله‌های اطراف آنومالی سطحی مس تصور نمود که با جابجایی موقعیت مناطق آنومال مس، آن‌ها نیز جابجا می‌شوند. از عمق ۱۵۰ متر در بخش جنوبی ناحیه، آنومالی کوچکی از آرسنیک تا عمق ۲۵۰ متر گسترش و در عمق ۳۰۰ متر ناپدید می‌شود. این آنومالی احتمالاً در بخش فوق‌کانساری آنومالی مس می‌تواند باشد که در این بخش از عمق ۲۰۰ متری شروع می‌شود. هاله As در عمق ۱۵۰ متر تا ۲۵۰ متر در بخش مرکزی پروفیل AA' به حداکثر میزان خود رسیده است (شکل ۵-۸). در اعماق ۳۰۰ متری به بعد، در بخش A' آنومالی آن شدت می‌یابد که می‌تواند ناشی از جهت یافتگی کانی‌سازی به سمت شمال شرقی باشد (شکل ۵-۷).



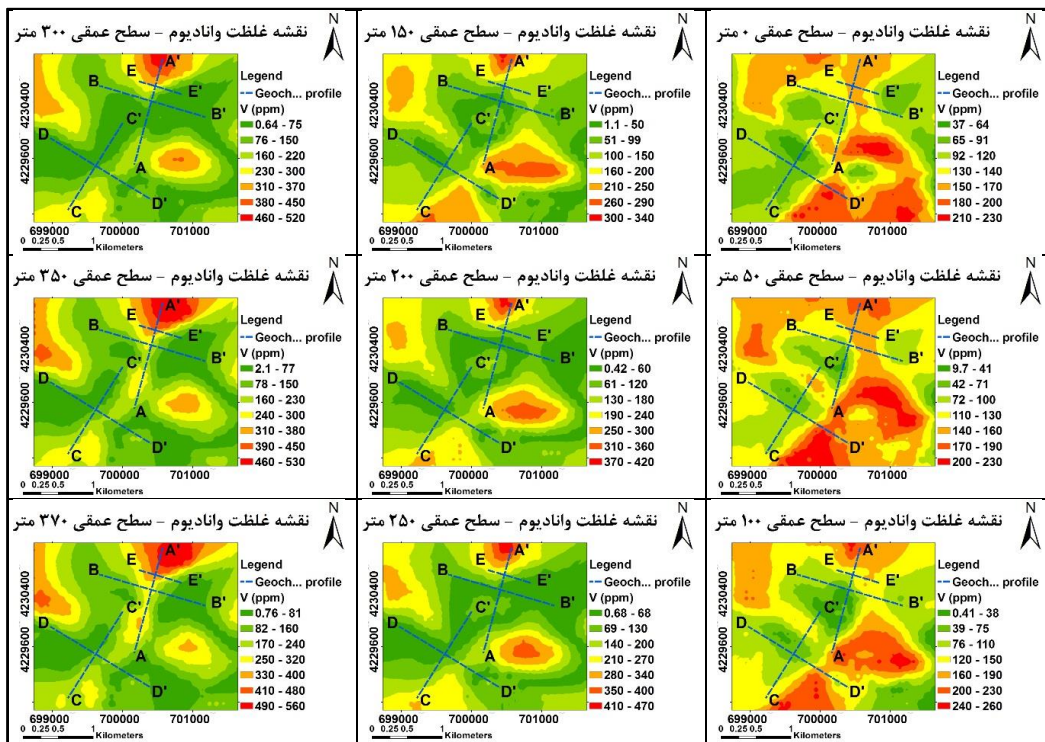
شکل ۵-۷: تغییرات غلظت آرسنیک در سطوح مختلف عمقی

این رفتار As، آنرا با ویژگی فوق کانساری بودن بیشتر معرفی می کند.

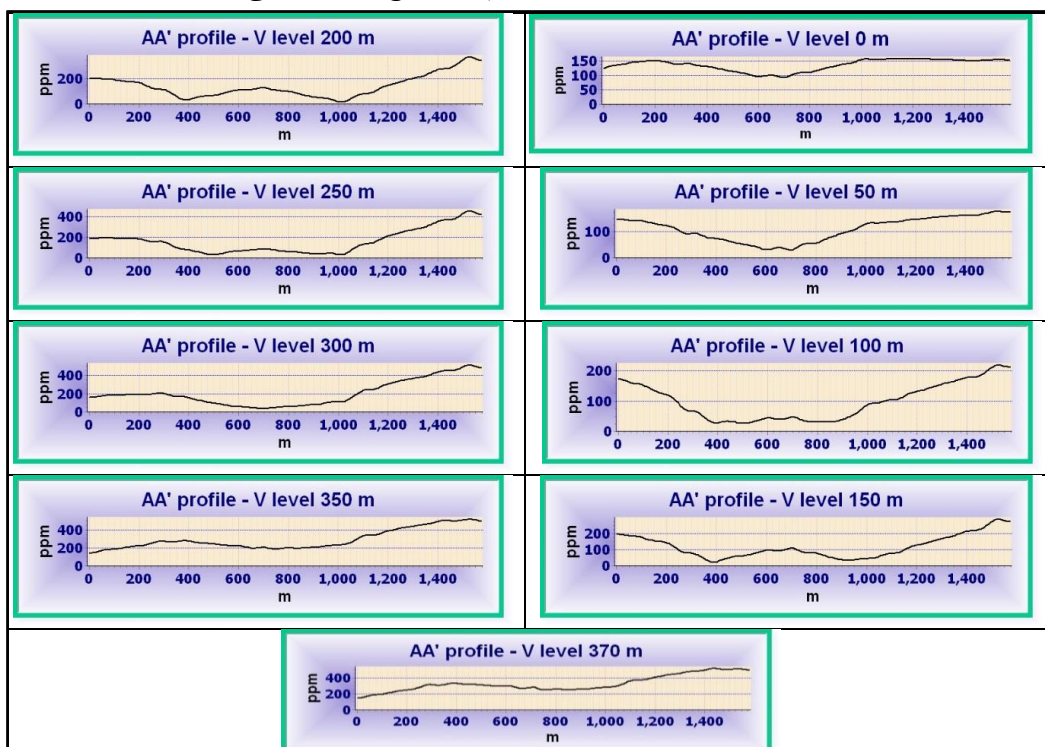


شکل ۵-۸: مقاطع عرضی تغییرات عنصر آرسنیک در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی

عنصر وانادیوم در مناطق مجاور که غلظت مس بالاست در سطح آنومالی نشان می دهد (مقایسه شکل ۵-۶ با شکل ۵-۹). این امر بویژه در خصوص آنومالی مسی که از عمق ۱۵۰ متر به بعد در بخش جنوبی ناحیه شروع می شود قابل ملاحظه است. این عنصر در مناطقی که در سطح غلظت مس کمی بالاست، پایین است (شکل ۵-۱۰) این امر به نوعی رفتار عنصر وانادیوم را با تحت کانساری تایید نموده که میزان غلظت آن در بدنه اطراف کانی سازی با رسیدن به کانسار افزایش نشان می دهد (شکل ۵-۹ و شکل ۵-۱۰).

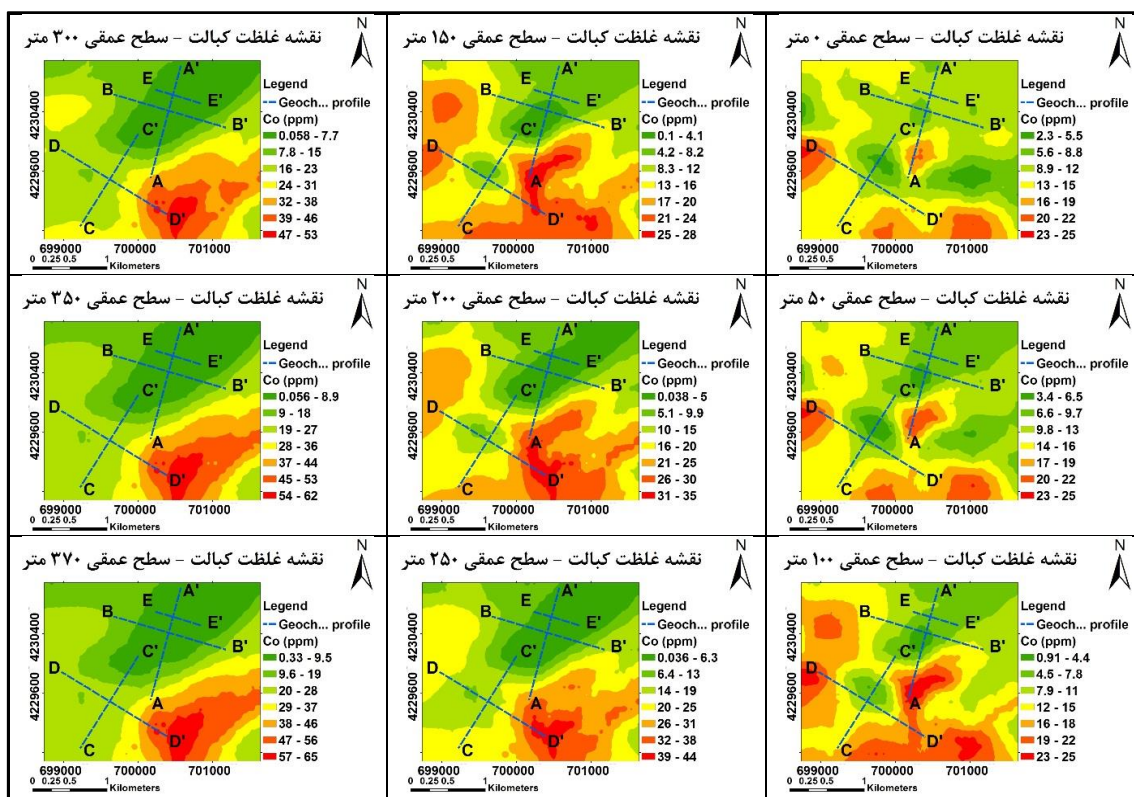


شکل ۵-۹: تغییرات غلظت وانادیوم در سطوح مختلف عمقی

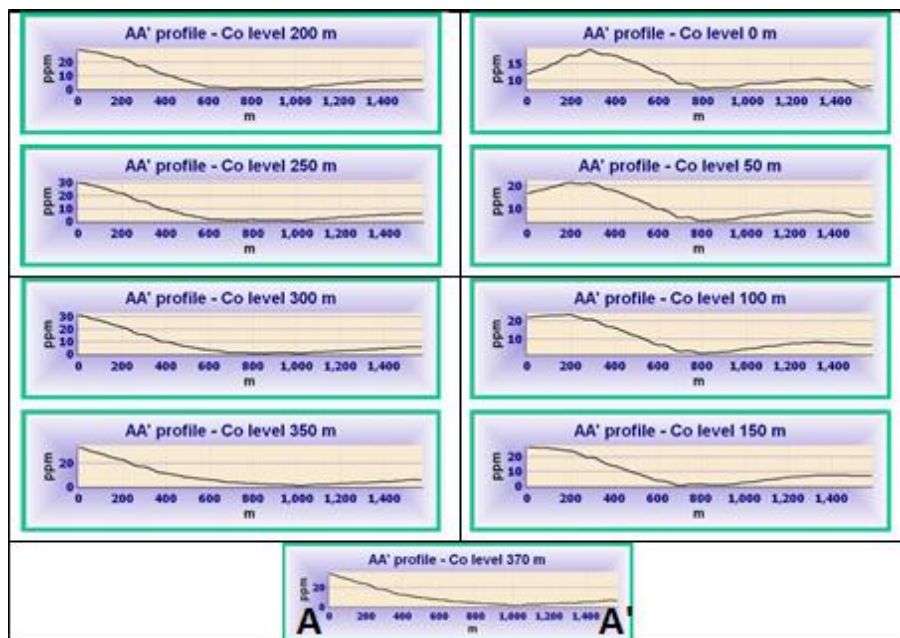


شکل ۵-۱۰: مقاطع عرضی تغییرات عنصر وانادیوم در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی

نقشه‌های تغییرات غلظت کبالت (شکل ۵-۱۱ و شکل ۵-۱۲) نشان دهنده آن است که این عنصر در مناطقی که کانی‌سازی مس شروع می‌شود افزایش آن آغاز می‌شود (انطباق مناطق آنومالی اقتصادی مس با آنومالی‌های کبالت). این امر به خوبی در خصوص منطقه کانی‌سازی شمال - شمال‌شرقی و جنوب از عمق ۱۰۰ متر به پایین قابل مشاهده است. بنابراین در کنار عنصر وانادیوم می‌توان از کبالت هم به عنوان عنصر تحت‌کانساری نام برد.

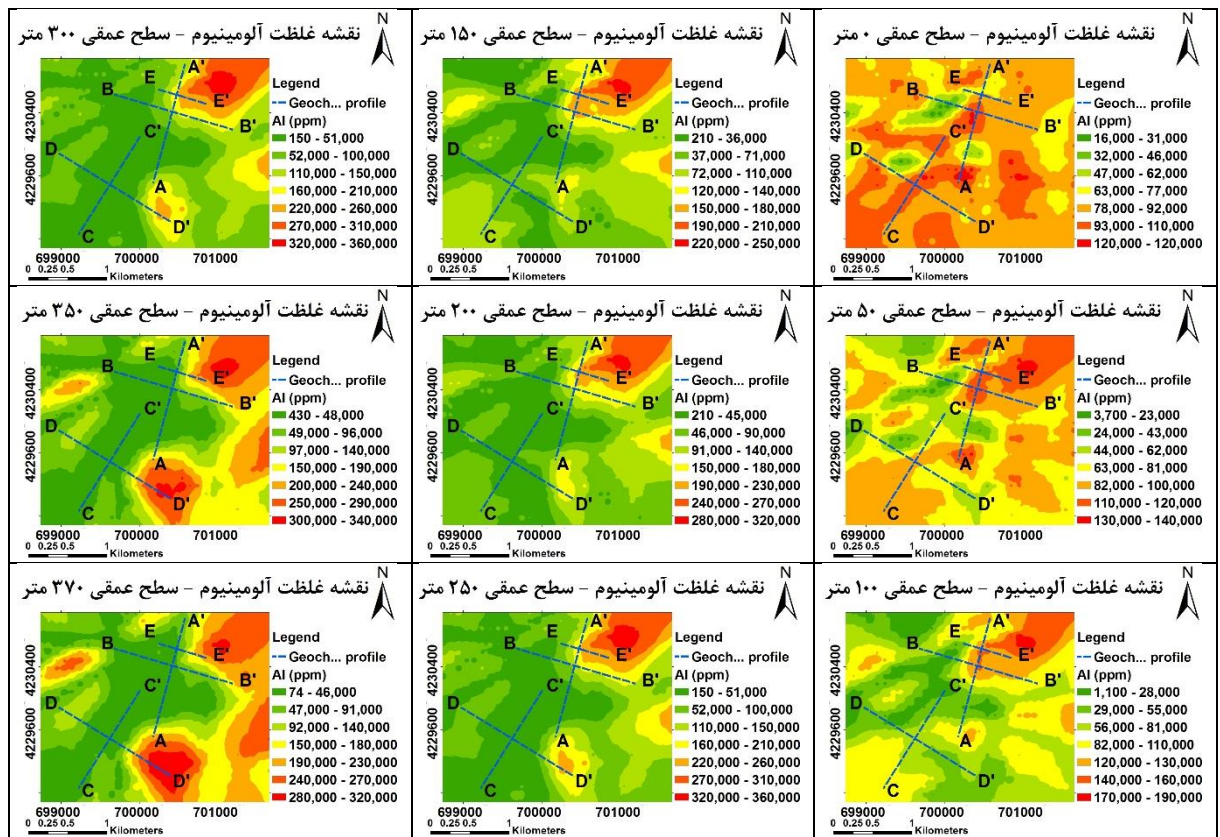


شکل ۵-۱۱: تغییرات غلظت کبالت در سطوح مختلف عمقی

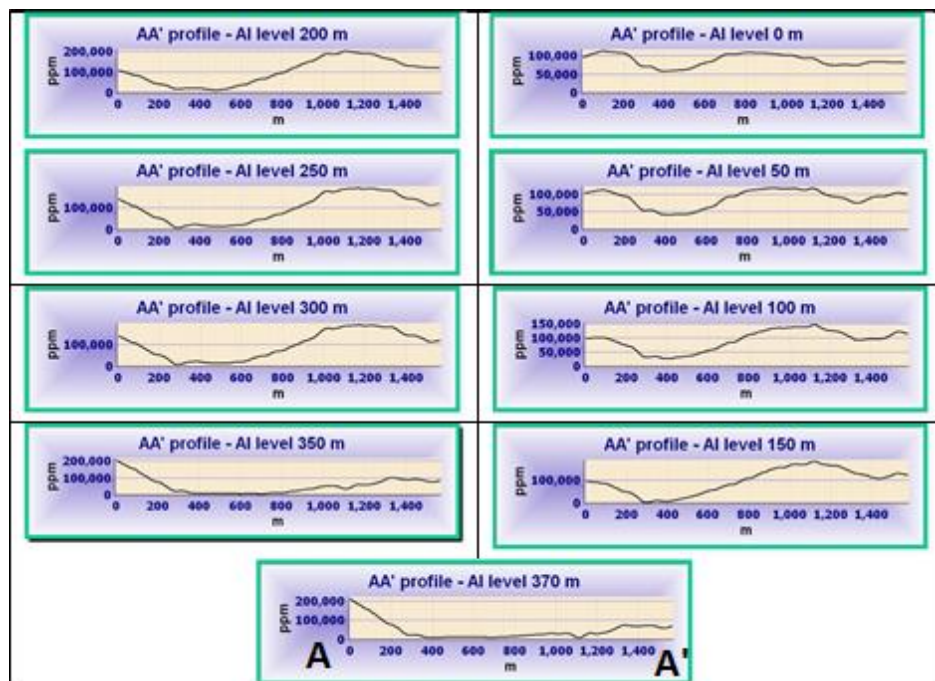


شکل ۵-۱۲: مقاطع عرضی تغییرات عنصر کبالت در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی

آلتراسیون کانی‌های رسی یکی از راهنماهای اکتشاف ذخایر مس پورفیری است. آلتراسیون آرژلیک در بالاترین قسمت مدل کانسارهای مس پورفیری در حاشیه خارجی منطقه سرسیتی شدن قرار می‌گیرد (شکل ۱-۱) مدل‌های ذخیره کانسار مس پورفیری گویای تمرکز ذخیره در منطقه سرسیتیک و پتاسیک است. بسته به سطح فرسایش منطقه، غلظت آلومینیوم در این دگرسانی تغییر کرده و با نزدیک شدن به توده کانسار، گسترش آلومینیوم کمتر می‌شود. در نقشه‌های غلظت آلومینیوم برای سطوح مختلف عمقی (شکل ۵-۱۳ و شکل ۵-۱۴) مشاهده می‌گردد که در مناطق کانه‌سازی مس با مناطق آنومال AI احتمالاً کانی‌های رسی بالا محصور می‌شود. نمونه شاخص این امر آنومالی مس در بخش جنوبی منطقه است که بوسیله حلقه‌ای از آنومالی‌های AI از عمق ۱۵۰ در بخش جنوب‌شرقی شروع شده و به مرور وسعت و گسترده‌تر می‌شود که بیشترین گسترش آن در عمق ۳۷۰ متری است که شاید گواهی از یک کانی‌سازی دیگر در عمق در این بخش از منطقه باشد.

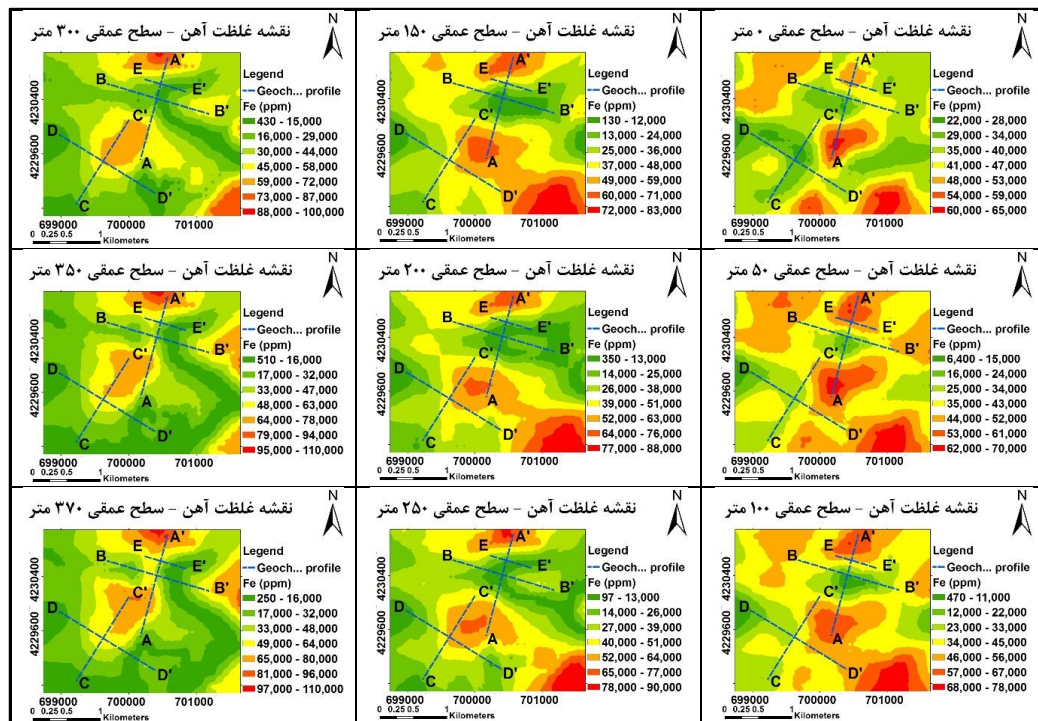


شکل ۵-۱۳: تغییرات غلظت آلومینیوم در سطوح مختلف عمقی

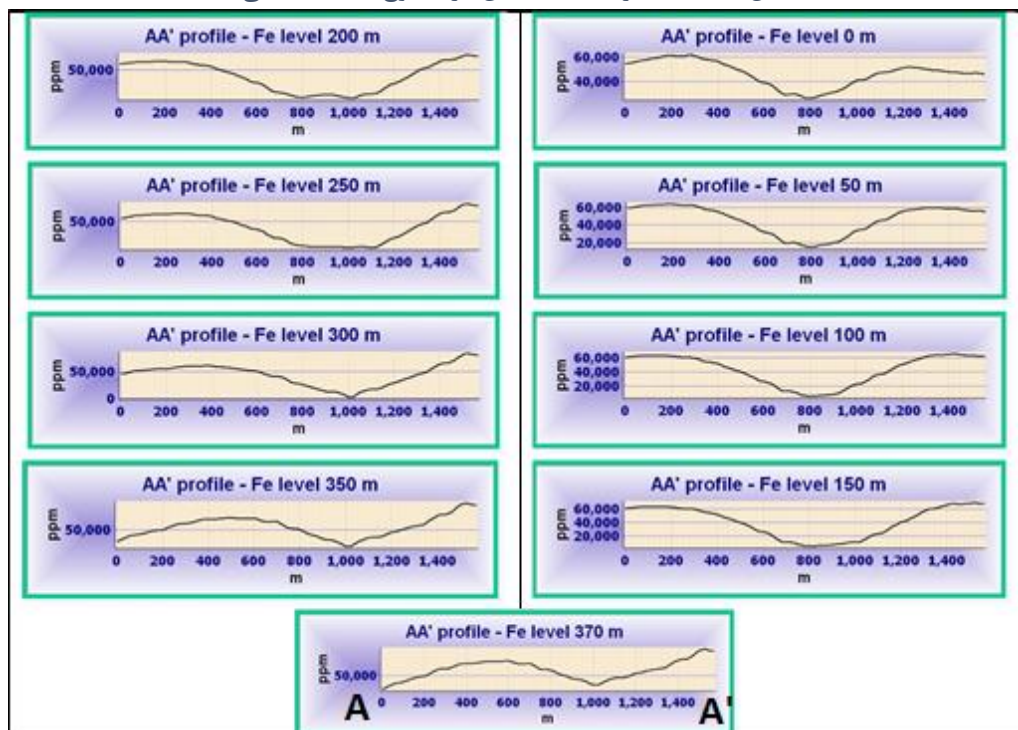


شکل ۵-۱۴: مقاطع عرضی تغییرات عنصر آلومینیوم در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی

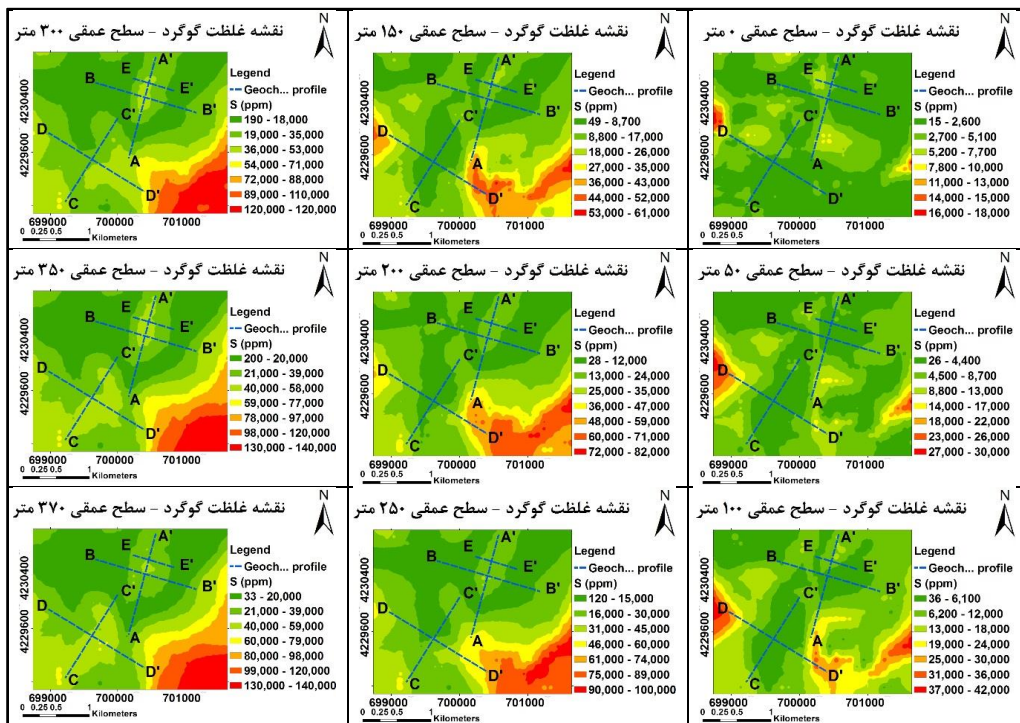
عنصر آهن به عنوان یکی از عناصری است که همراه کانی پیریت در دگرسانی فیلیک یا مگنتیت، کالکوپیریت و بورنیت در گرسانی پتاسیک کانسارهای مس پورفیری می‌تواند باشد و مناطقی با غلظت آنومال نشان دهد. در بررسی نقشه‌های این عنصر (شکل ۵-۱۵ و شکل ۵-۱۶)، غلظت بالای ۴ درصد از عمق ۵۰ متری به بعد با مناطق آنومال مس در بخش شمالی منطقه و از اعماق ۱۵۰ متری تا ۲۵۰ متری با آنومالی مس در بخش جنوب - جنوب‌شرقی منطقه هم‌خوانی نشان می‌دهد. گوگرد به عنوان یکی از عناصر کانه‌ساز کانسارهای پورفیری یا در منطقه دگرسانی سرسیتیک از اهمیت بالایی برخوردار است (شکل ۱-۱). در بررسی نقشه‌های این عنصر (شکل ۵-۱۷ و شکل ۵-۱۸) گوگرد از ۱۰۰ متری به بعد هم‌خوانی معنی‌داری با آنومالی مس در بخش جنوب - جنوب‌شرقی منطقه نشان داده و این همبستگی در عمق ادامه دارد. پیریت، کالکوپیریت و یا بورنیت از کانه‌های آهن و گوگرد دار در بخش کانه‌سازی ذخایر مس پورفیری هستند. لذا مناطق آنومال آهن و گوگرد در منطقه و سطوح عمقی مختلف اینگونه شاید تفسیر شود که آنومالی مس در بخش شمال-شمال‌شرقی منطقه به دلیل نزدیک سطح همراه با گوسن و فرآیندهای سوپرژن قرار گرفته است. این در حالی است که آنومالی بخش جنوب - جنوب‌شرقی منطقه که عمقی است همراه با سولفور مس حاوی بورنیت، کالکوپیریت و مگنتیت باشد.



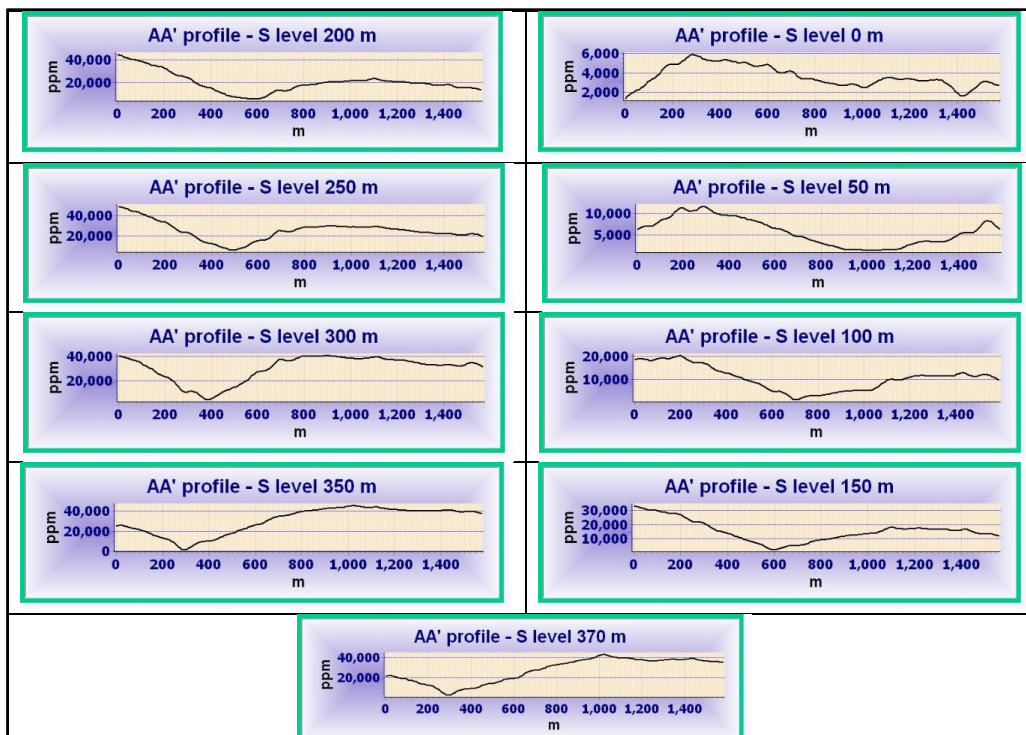
شکل ۵-۱۵: تغییرات غلظت آهن در سطوح مختلف عمقی



شکل ۵-۱۶: مقاطع عرضی تغییرات عنصر آهن در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی



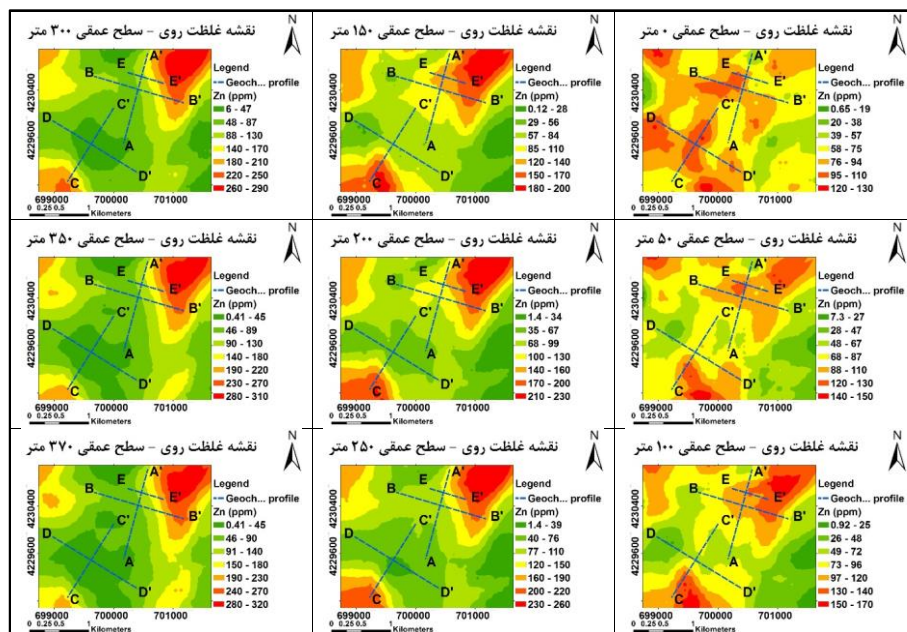
شکل ۵-۱۷: تغییرات غلظت گوگرد در سطوح مختلف عمقی



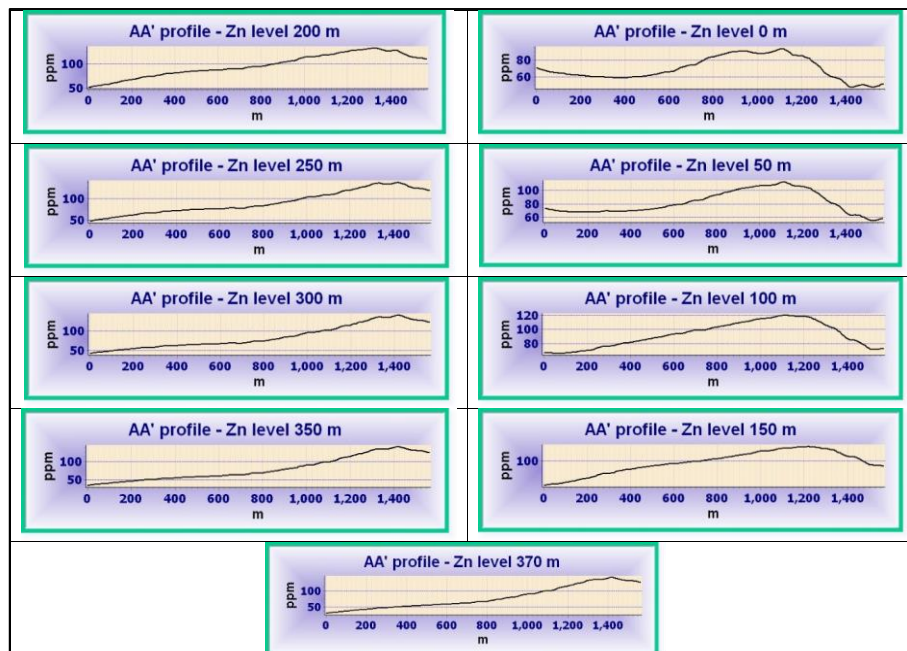
شکل ۵-۱۸: مقاطع عرضی تغییرات عنصر گوگرد در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی

مناطق آنومال روی در حاشیه محدوده کانی‌سازی مس در منطقه مشاهده می‌شود (شکل ۵-۱۹ و شکل

۵-۲۰) که این امر با توجه به تیپ پورفیری ذخیره مس منطقه قابل پیش‌بینی است.

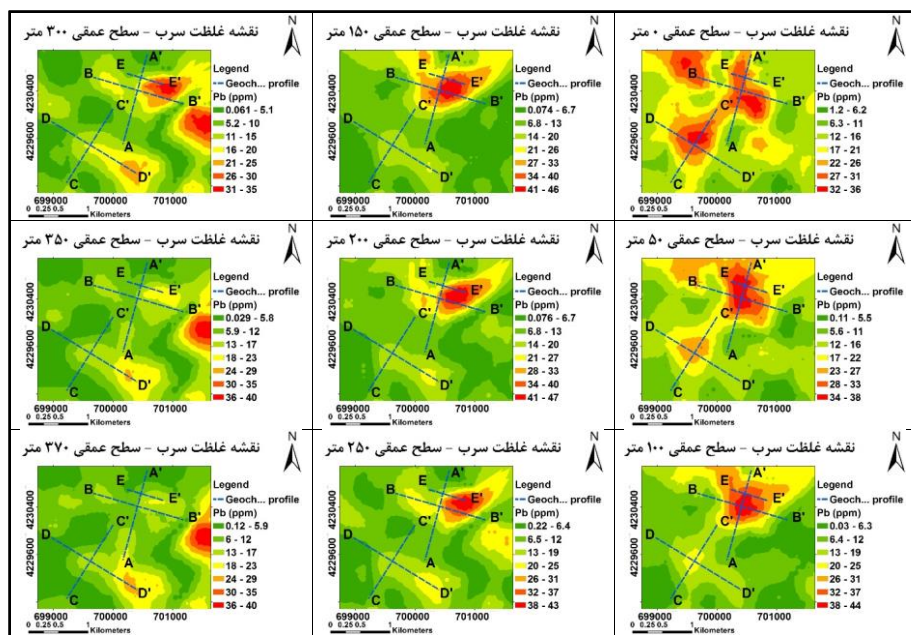


شکل ۵-۱۹: تغییرات غلظت روی در سطوح مختلف عمقی

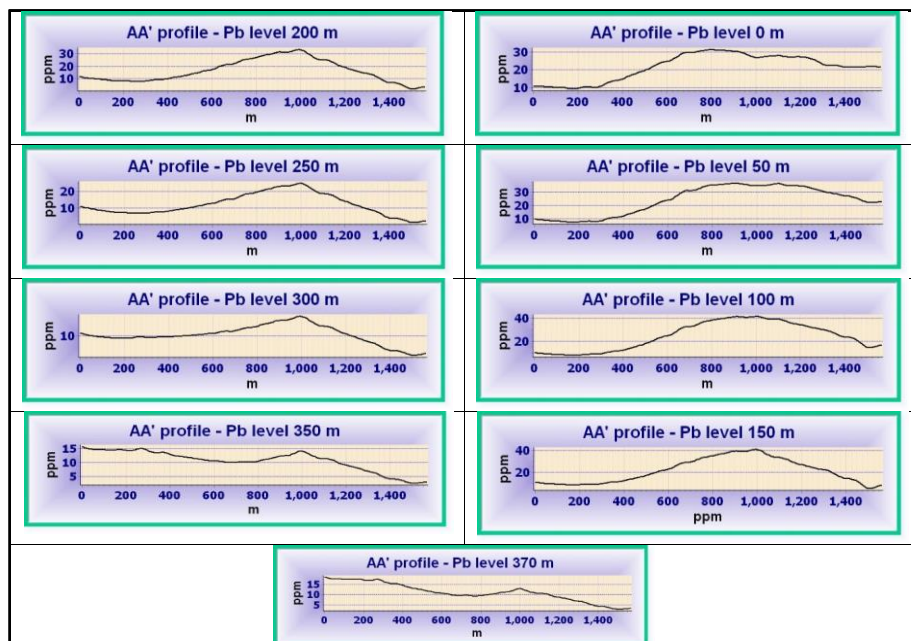


شکل ۵-۲۰: مقاطع عرضی تغییرات عنصر روی در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی

مناطق آنومال سرب به عنوان یکی از عناصر حاشیه دور مناطق کانی‌سازی مس در منطقه مورد مطالعه می‌باشد (شکل ۲۱-۵ و شکل ۲۲-۵) که این امر مشابه عنصر Zn، برای عنصر سرب نیز با توجه به تیپ پورفیری ذخیره قابل پیش‌بینی است.

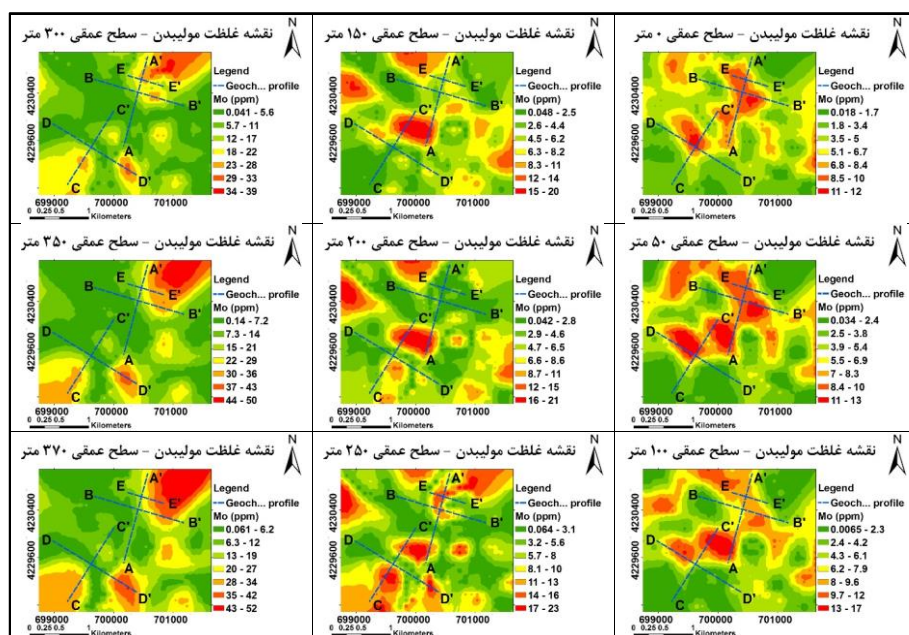


شکل ۲۱-۵: تغییرات غلظت سرب در سطوح مختلف عمقی

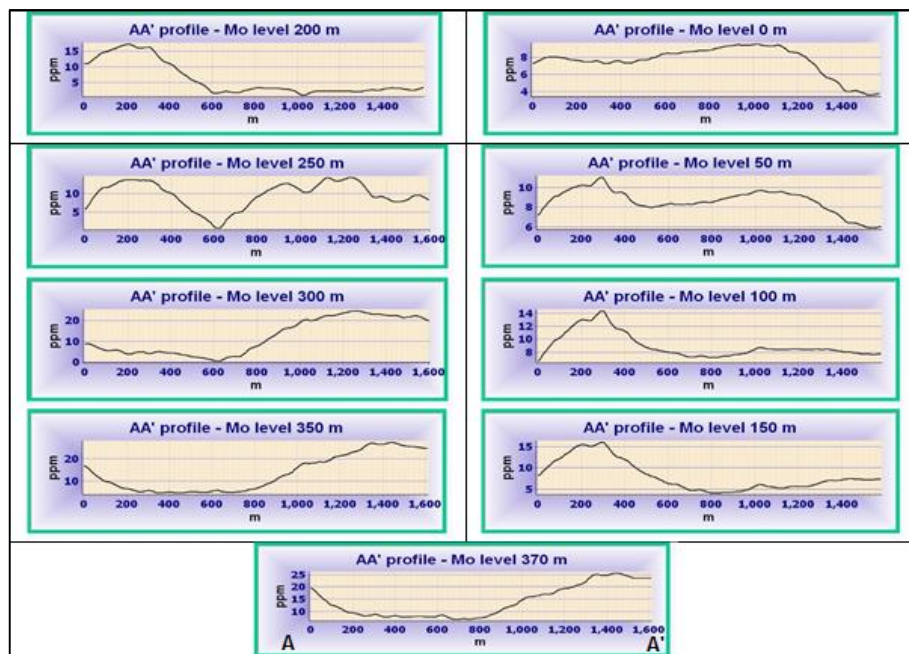


شکل ۲۲-۵: مقاطع عرضی تغییرات عنصر سرب در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی

مولیدن به عنوان یکی از عناصری است که می‌تواند به عنوان پاراژنز عنصر مس باشد. در نقشه تغییرات غلظت این عنصر (شکل ۵-۲۳ و شکل ۵-۲۴) مشاهده می‌گردد جاهایی که شدت مس کاهش یافته است مولیدن نیز کاهش یافته است و این نشان دهنده کف کانی‌سازی است.



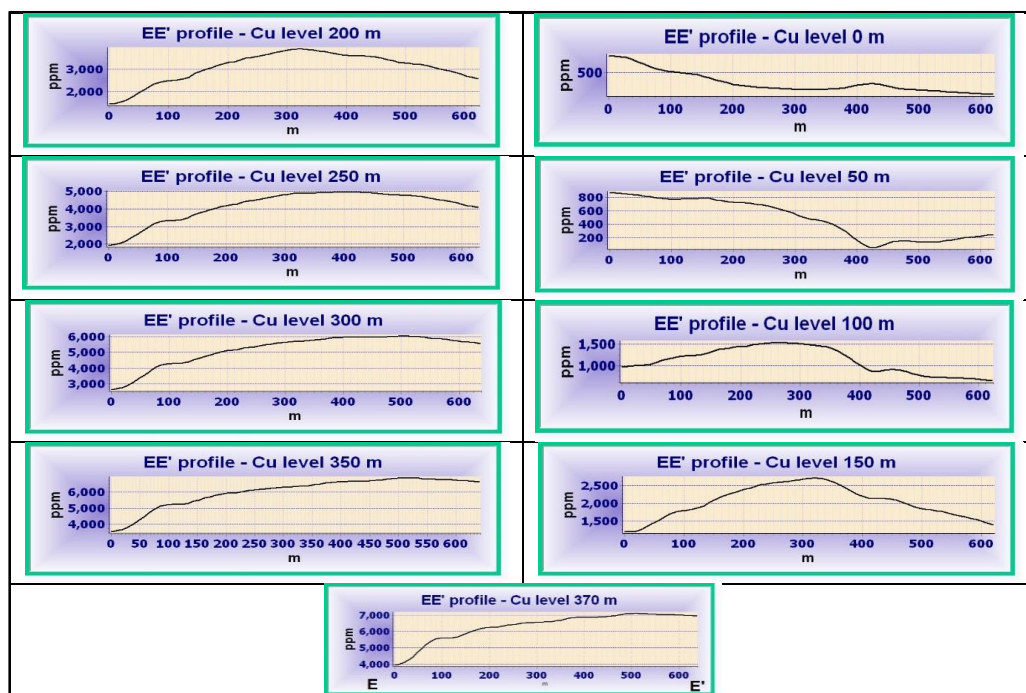
شکل ۵-۲۳: تغییرات غلظت مولیدن در سطوح مختلف عمقی



شکل ۵-۲۴: مقاطع عرضی تغییرات عنصر مولیدن در طول پروفیل ژئوشیمیایی AA' در سطوح مختلف عمقی

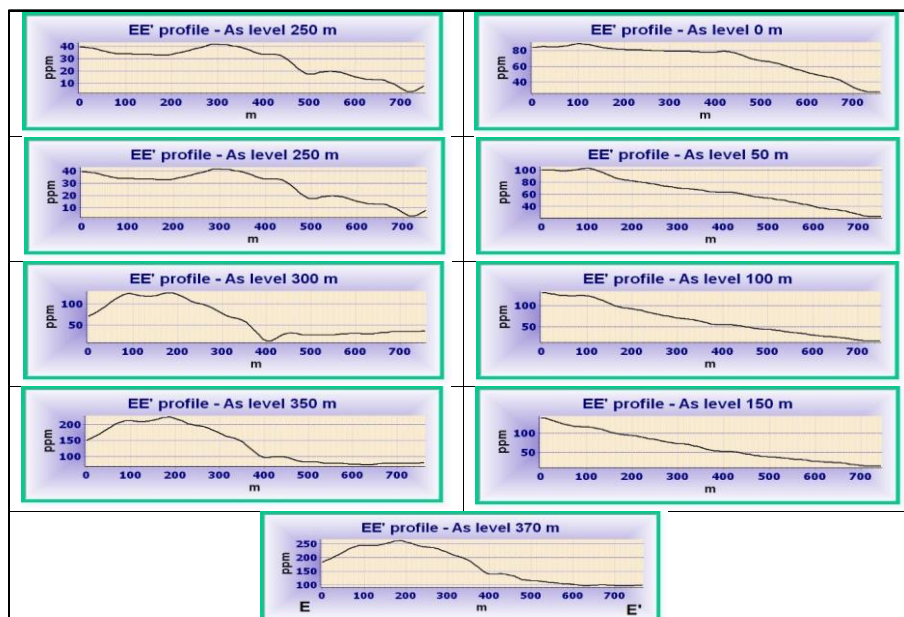
۵-۲-۲- پروفیل ژئوشیمیایی EE'

تغییرات غلظت مس در راستای پروفیل EE' (شکل ۵-۶ و شکل ۵-۲۵) نشان داد که کانی‌سازی مس از سطح با غلظت کم شروع می‌شود و افزایش غلظت مس از سطح ۱۰۰ متری با غلظت‌های بیش از ۰/۱ درصد شروع می‌شود. آنومالی شرقی از عمق ۱۵۰ متری با غلظت کم شروع شده و با افزایش عمق شدت مس زیاد شده تا ۰/۷ درصد می‌رسد. روند افزایش غلظت مس نشانگر آن است که کانی‌سازی در عمق نیز ادامه دارد.

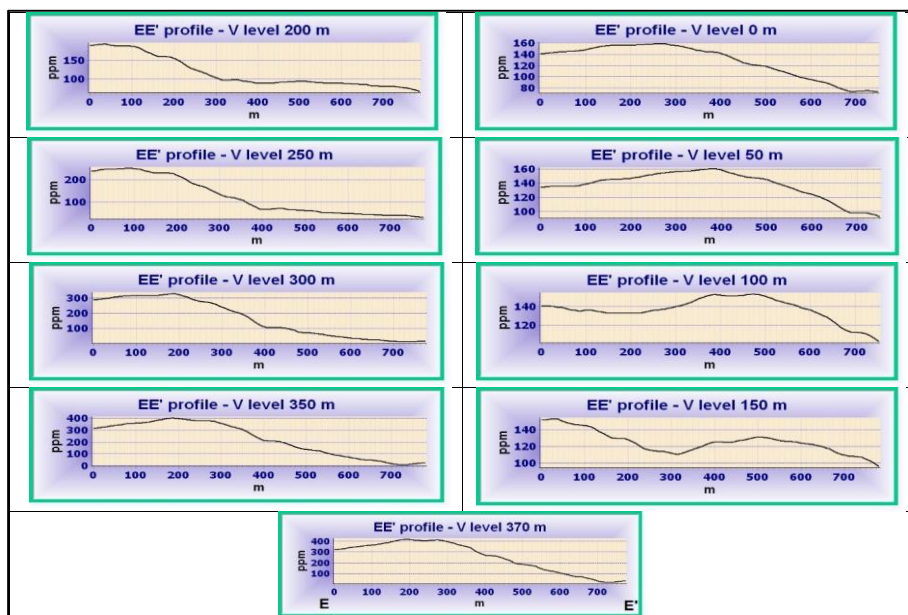


شکل ۵-۲۵: مقاطع عرضی تغییرات عنصر مس در طول پروفیل ژئوشیمیایی EE' در سطوح مختلف عمقی تغییرات غلظت آرسنیک (شکل ۵-۷ و شکل ۵-۲۶) نشان می‌دهد که در ابتدای پروفیل به سمت غرب منطقه، غلظت آرسنیک در سطح زیاد است و با توجه به اینکه آرسنیک نقش فوق‌کانساری در پروفیل AA'، این امر شروع کانی‌سازی در عمق بخش شمال - شمال شرقی منطقه را گواهی می‌کند (شکل ۵-۶). از عمق ۱۰۰ متری به بعد غلظت آرسنیک در مرکز پروفیل افزایش زیادی نشان می‌دهد و تا عمق ادامه داشته و به سمت غرب منطقه متمایل می‌شود و این احتمال را می‌دهد که در این بخش کانی‌سازی در

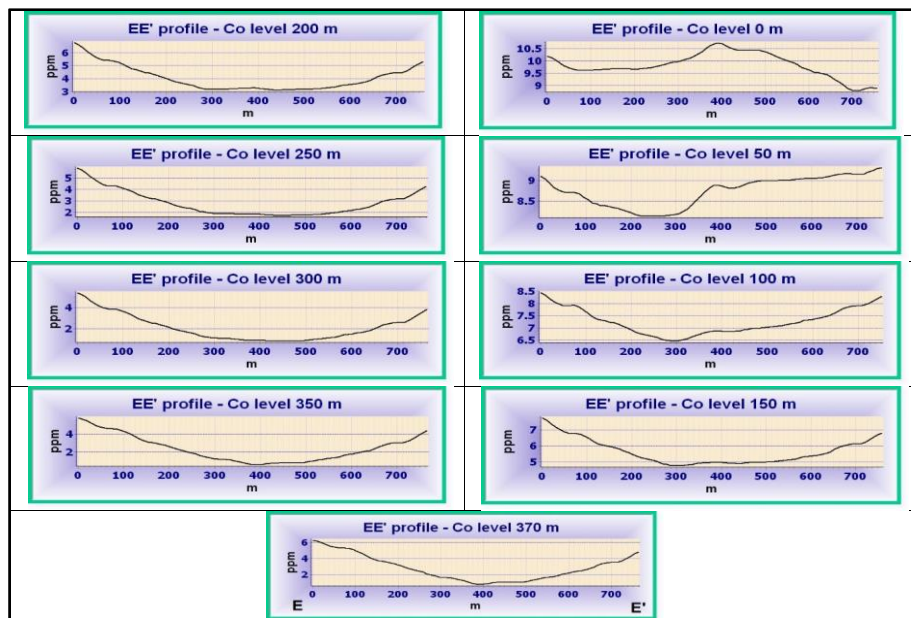
عمق واقع شده است. شاید این امر گواهی بر جهت شیب شمال غربی کانی سازی بخش شمال - شمال شرقی منطقه نیز باشد.



شکل ۵-۲۶: مقاطع عرضی تغییرات عنصر آرسنیک در طول پروفیل ژئوشیمیایی EE' در سطوح مختلف عمقی به خوبی افزایش غلظت عنصر وانادیوم با نزدیک شدن به عمق رفتار تحت کانساری آنرا تایید می کند (شکل ۵-۹ و شکل ۵-۲۷) تا عمق ۲۰۰ متری حداکثر آن به حدود ۲۰۰ ppm رسیده، در بازه عمقی ۱۰۰-۲۵۰ این غلظت به ۲۵۰ - ۱۵۰ ppm و از عمق ۳۷۰ - ۳۰۰ متر به بالاتر از ۴۰۰ ppm رسیده است.

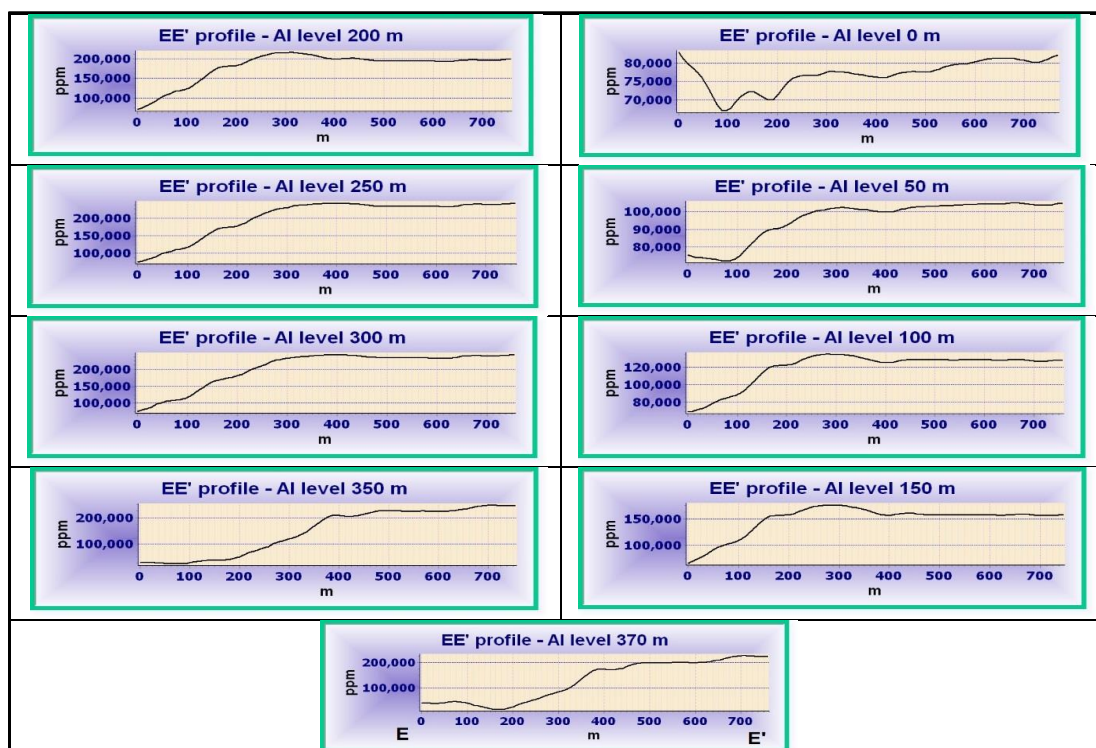


شکل ۵-۲۷: مقاطع عرضی تغییرات عنصر وانادیوم در طول پروفیل ژئوشیمیایی EE' در سطوح مختلف عمقی بررسی رفتار تغییرات کبالت (شکل ۵-۲۸) نشان می‌دهد که این عنصر رفتاری شبیه به عناصر تحت‌کانساری دارد بدین معنا در جاهایی که کانی‌سازی در عمق وجود دارد غلظت کبالت کم و در جاهایی که کانی‌سازی از بین رفته است غلظت کبالت افزایش یافته است.



شکل ۵-۲۸: مقاطع عرضی تغییرات عنصر کبالت در طول پروفیل ژئوشیمیایی EE' در سطوح مختلف عمقی

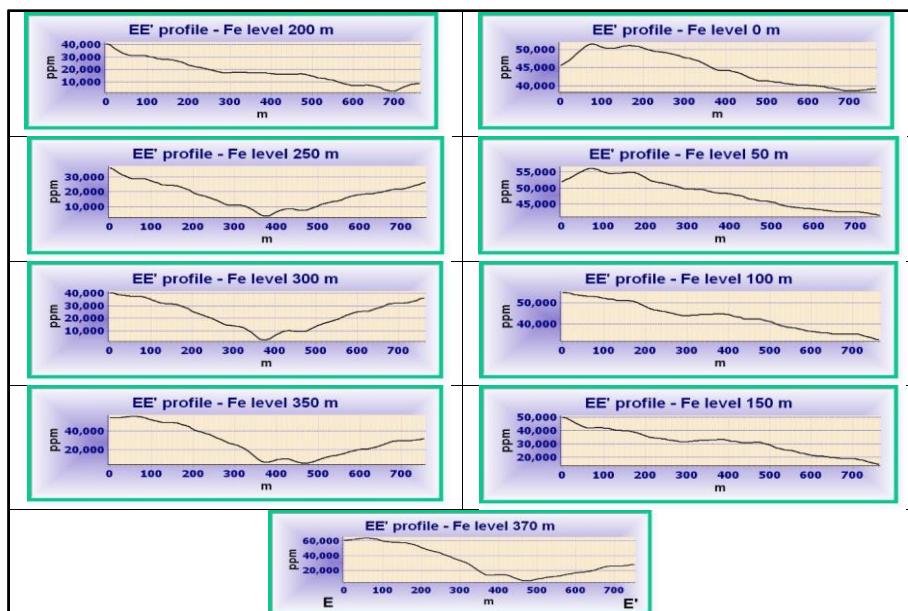
احاطه شدن محدوده کانی‌سازی عمقی توسط غلظت‌های بالای آلومینیوم که در پروفیل و نقشه‌های مقاطع عمقی پروفیل AA' مشاهده شد، در اینجا بر روی پروفیل EE' قابل مشاهده است. این امر به نوعی گسترش کانی‌های رسی در این مناطق می‌تواند باشد. در قسمت مرکزی پروفیل که کانی‌سازی در عمق وجود دارد آلومینیوم نسبت به بقیه قسمت‌ها غلظت بالاتری نشان داده و از مقیاس ۸ درصد به مقیاس ۲۰ درصد می‌رسد (شکل ۵-۲۹).



شکل ۵-۲۹: مقاطع عرضی تغییرات عنصر آلومینیوم در طول پروفیل ژئوشیمیایی EE' در سطوح مختلف عمقی

در بررسی‌های پروفیل AA' به این نتیجه رسیدیم که آهن و گوگرد به عنوان دو پارائنز کانی‌سازی هستند. رفتاری که تغییرات غلظت این دو عنصر (شکل ۵-۳۰ و شکل ۵-۳۱) در عمق نشان می‌دهند همخوانی مناسبی با تغییرات غلظت مس در عمق دارند. تفسیرهای پروفیل AA' اصلاح می‌گردد که کانی‌سازی بخش شمال - شمال شرقی منطقه در عمق حاوی غلظت‌های آنومال مس، آهن و گوگرد از

خود نشان می‌دهد. این امر گواهی حضور پیریت، کالکوپیریت و بورنیت در عمق می‌تواند باشد (۴ تا ۶ درصد آهن و سولفور). بالا بودن آهن در سطح می‌تواند معرف گوسن بالای ذخیره باشد.

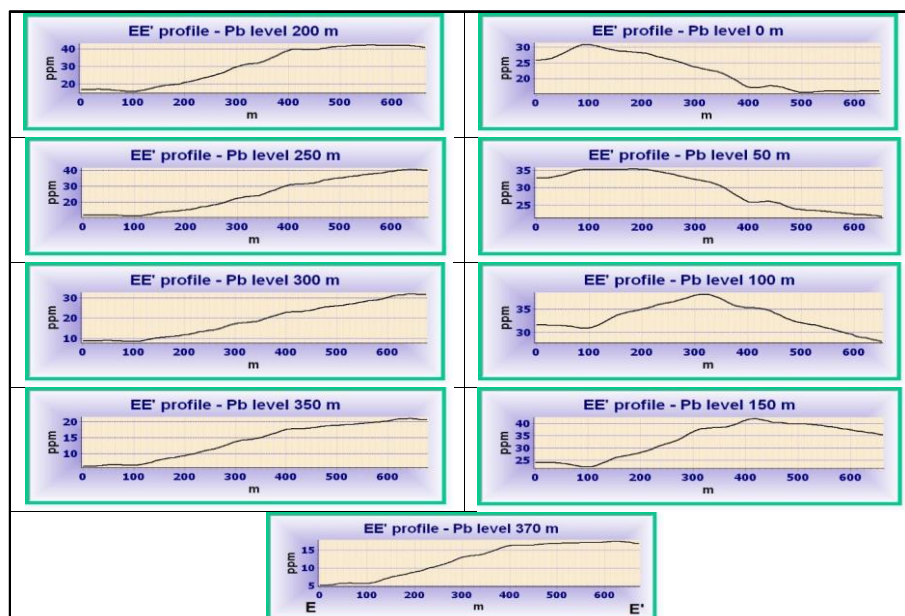


شکل ۳-۵: مقاطع عرضی تغییرات عنصر آهن در طول پروفیل ژئوشیمیایی EE' در سطوح مختلف عمقی

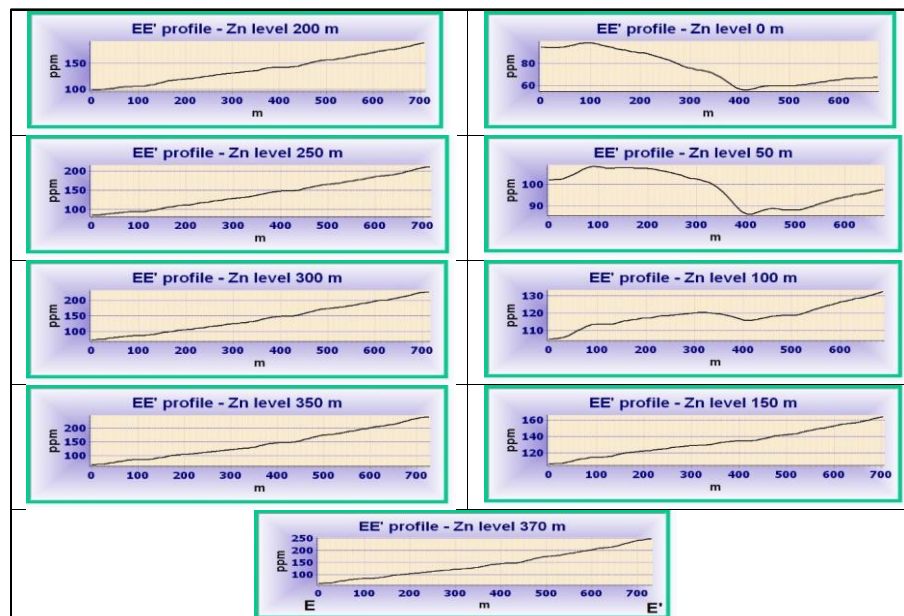


شکل ۳-۱۱: مقاطع عرضی تغییرات عنصر گوگرد در طول پروفیل ژئوشیمیایی EE' در سطوح مختلف عمقی

عناصر سرب و روی (شکل ۵-۳۲ و شکل ۵-۳۳) در سطح همراه با کانی‌سازی دیده می‌شوند و با افزایش عمق در حاشیه کانی‌سازی قرار می‌گیرند و می‌توانند به عنوان راهنمای اکتشافی برای پی بردن به محل کانی‌سازی استفاده شوند.



شکل ۵-۳۲: مقاطع عرضی تغییرات عنصر سرب در طول پروفیل ژئوشیمیایی EE' در سطوح مختلف عمقی

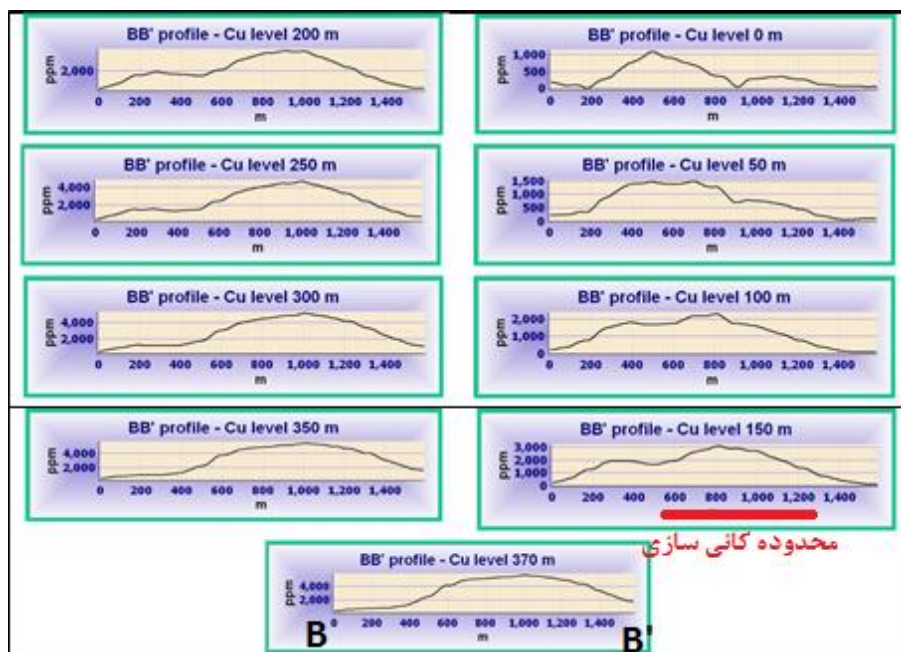


شکل ۵-۳۳: مقاطع عرضی تغییرات عنصر روی در طول پروفیل ژئوشیمیایی EE' در سطوح مختلف عمقی

توالی منطقه‌بندی $As - Pb - Zn - (Cu - S - Fe) - Co - V$ برای منطقه مورد مطالعه تایید می‌گردد، لذا بررسی تغییرات غلظت عناصر نشان داد که عنصر آرسنیک نقش عنصر فوق‌کانساری و وانادیوم و کبالت نقش عنصر تحت‌کانساری و عناصر $Pb - Zn$ حاشیه‌کانی‌سازی را نشان می‌دهند. در ادامه تغییرات غلظت عناصر $Cu - As - V$ در پروفیل‌های BB' ، CC' و DD' بررسی خواهد شد.

۳-۲-۵- پروفیل ژئوشیمیایی BB'

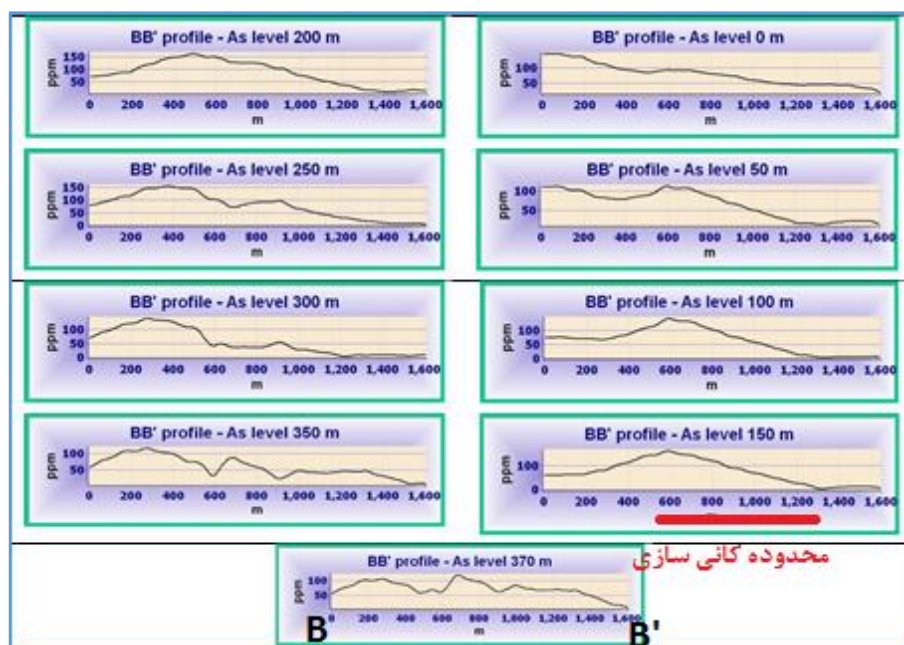
تغییرات غلظت مس در راستای پروفیل BB' (شکل ۳۴-۵) نشان داد که در قسمت غربی این پروفیل هاله‌های ژئوشیمیایی از سطح شروع شده و کانی‌سازی از عمق ۱۵۰ متر تا غلظت بالای ۰/۳ درصد مس شروع و تا عمق ادامه پیدا می‌کند. آنومالی شمال - شمال شرقی ناحیه به نظر می‌رسد شبیه به سمت شمال شرق منطقه داشته باشد.



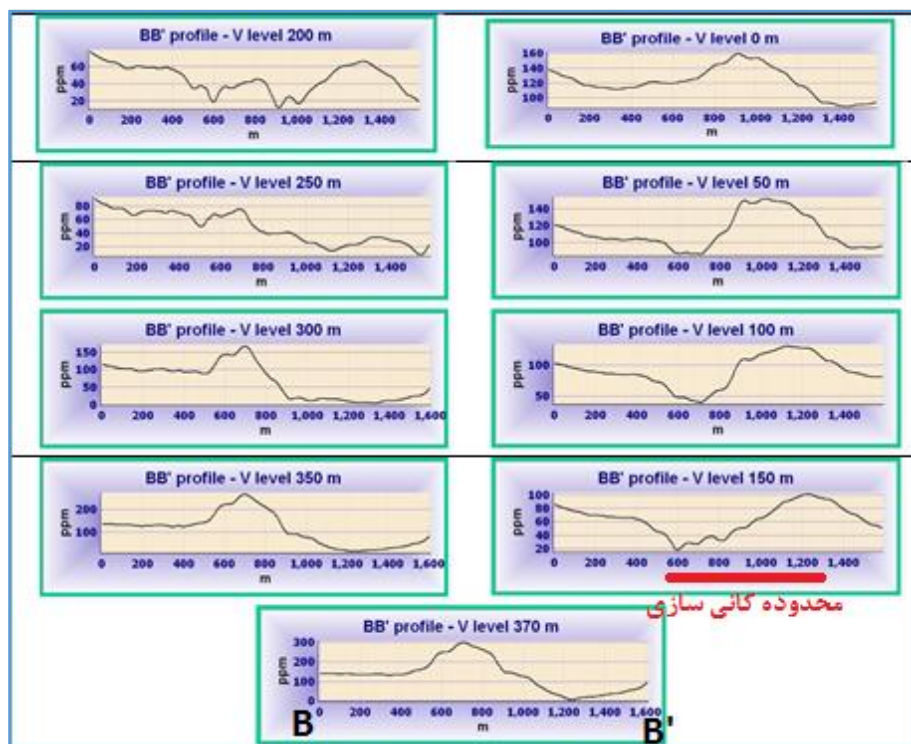
شکل ۳۴-۵: مقاطع عرضی تغییرات عنصر مس در طول پروفیل ژئوشیمیایی BB' در سطوح مختلف عمقی تغییرات غلظت آرسنیک (شکل ۳۵-۵) نشان دهنده آن است که این عنصر از سطح تا عمق روند تدریجی افزایش شدت غلظت و گسترش طول هاله را نشان می‌دهد و سپس از عمق ۲۵۰ متری شدت غلظت و

گسترش طول هاله آن روند کاهشی را شروع می‌کند. این ویژگی بنوعی رفتار فوق کانساری بودن آنرا تایید می‌کند.

غلظت عنصر وانادیوم از سطح به عمق در مرکز آنومالی مس با افت غلظت همراه است ولی از عمق ۱۵۰ متری در مرکزیت کانی‌سازی شروع به افزایش غلظت و وسعت هاله‌ها می‌کند که تا عمق ادامه دارد. این امر مجدداً رفت تحت کانساری این عنصر را تایید می‌کند (شکل ۵-۳۶).



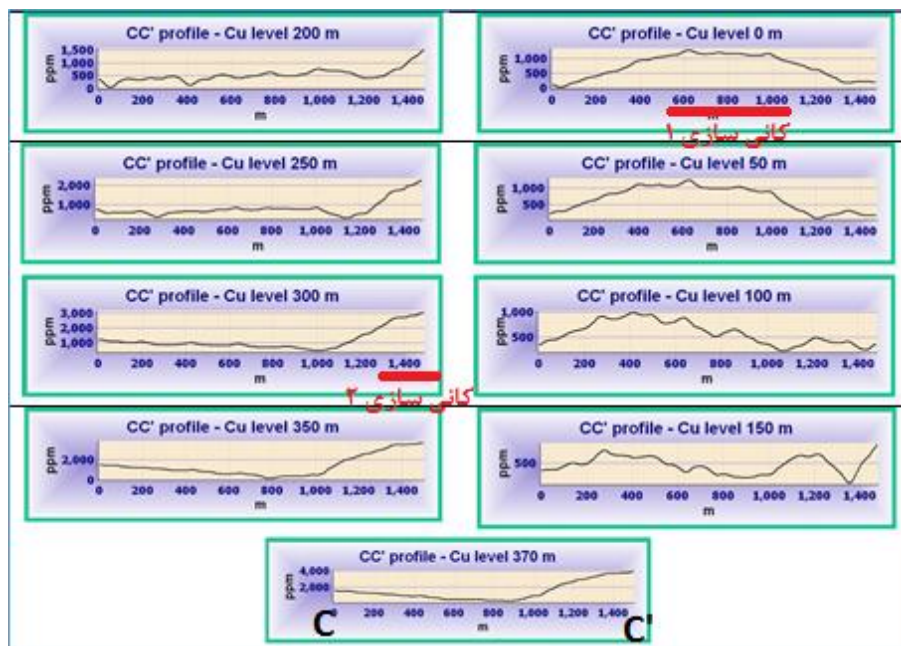
شکل ۵-۳۶: مقاطع عرضی تغییرات عنصر آرسنیک در طول پروفیل ژئوشیمیایی BB' در سطوح مختلف عمقی



شکل ۵-۳۶: مقاطع عرضی تغییرات عنصر وانادیوم در طول پروفیل ژئوشیمیایی BB' در سطوح مختلف عمقی

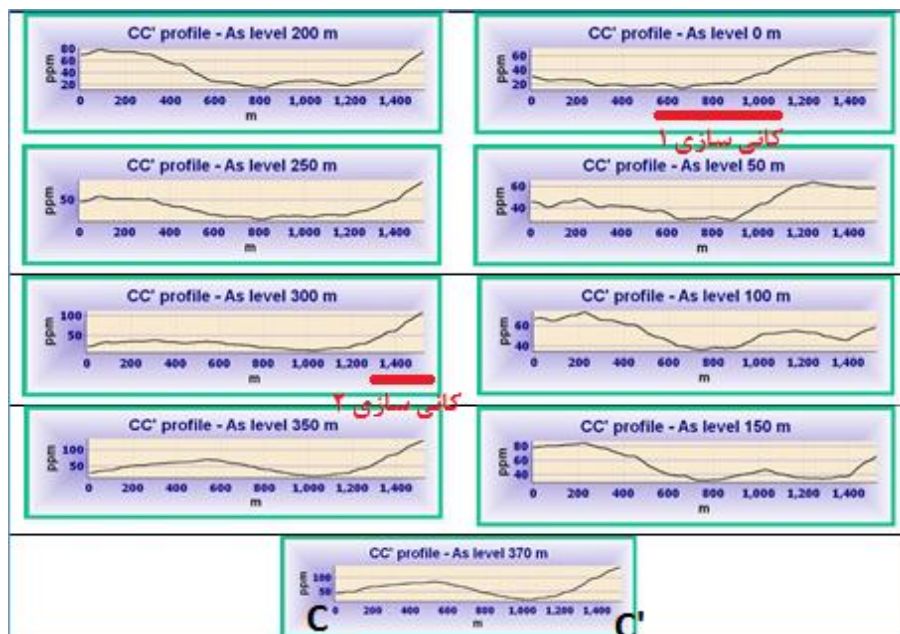
۵-۲-۴- پروفیل ژئوشیمیایی CC'

نقشه‌های تغییرات غلظت مس (شکل ۵-۶ و شکل ۵-۳۷) نشانگر آن است که در قسمت جنوب‌غربی محدوده مورد مطالعه، آنومالی مس از سطح تا عمق ۱۰۰ متری (محدوده کانی‌سازی ۱) وجود دارد. بررسی گمانه‌های اکتشافی (شکل ۵-۴) عدم کانی‌سازی در عمق را نشان می‌دهد. لذا تغییرات غلظت مس در قسمت جنوبی پروفیل CC'، گواه آن است که کانی‌سازی در این قسمت فرسایش یافته است و به انتهای کانی‌سازی رسیده‌ایم و با افزایش عمق اثری از کانی‌سازی مشاهده نمی‌گردد. در قسمت شمالی پروفیل از ۱۵۰ متری هاله ژئوشیمیایی مس مشاهده شده و از عمق ۲۵۰ متری شدت مس به حد اقتصادی (بالای ۰/۳ درصد - کانی‌سازی ۲) می‌رسد.

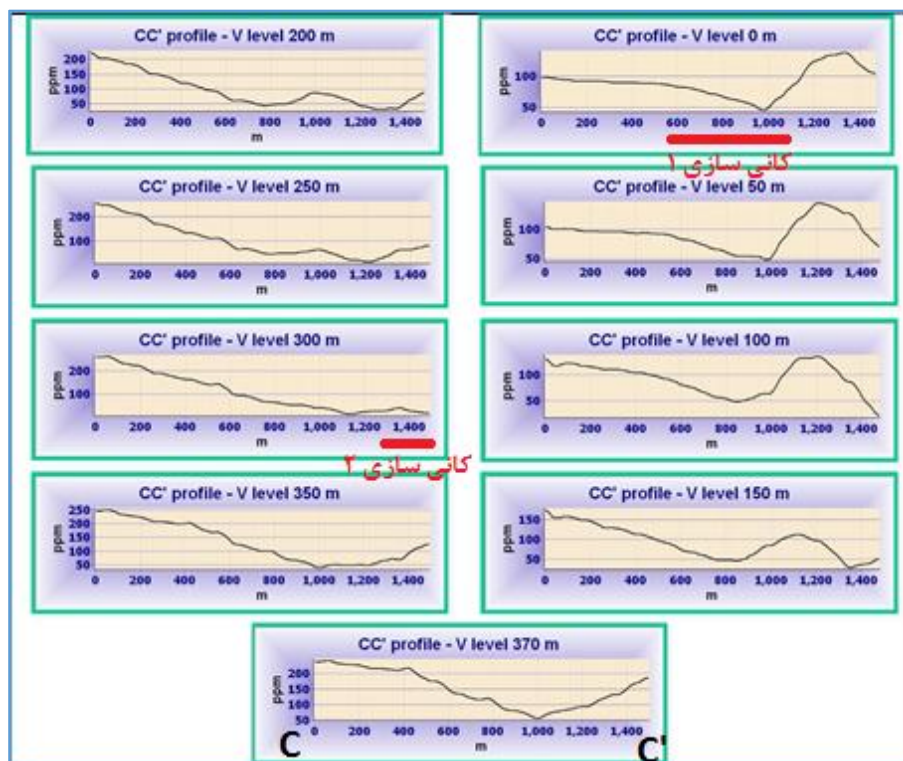


شکل ۵-۳۷: مقاطع عرضی تغییرات عنصر مس در طول پروفیل ژئوشیمیایی CC' در سطوح مختلف عمقی

تغییرات غلظت وانادیوم و آرسنیک (شکل ۵-۳۸ و شکل ۵-۳۹) به نحوی اتفاق افتاده است که هم‌خوانی مناسبی با تغییرات غلظت مس دارند و به خوبی نقش تحت (وانادیوم) و فوق کانساری (آرسنیک) را ایفا کرده‌اند. در قسمت جنوبی پروفیل با افزایش عمق غلظت وانادیوم افزایش می‌یابد و نشانگر اتمام کانی‌سازی در این بخش است و در قسمت شمالی پروفیل، با افزایش عمق غلظت آرسنیک زیاد شده است که نشان می‌دهد در عمق کانی‌سازی وجود دارد.



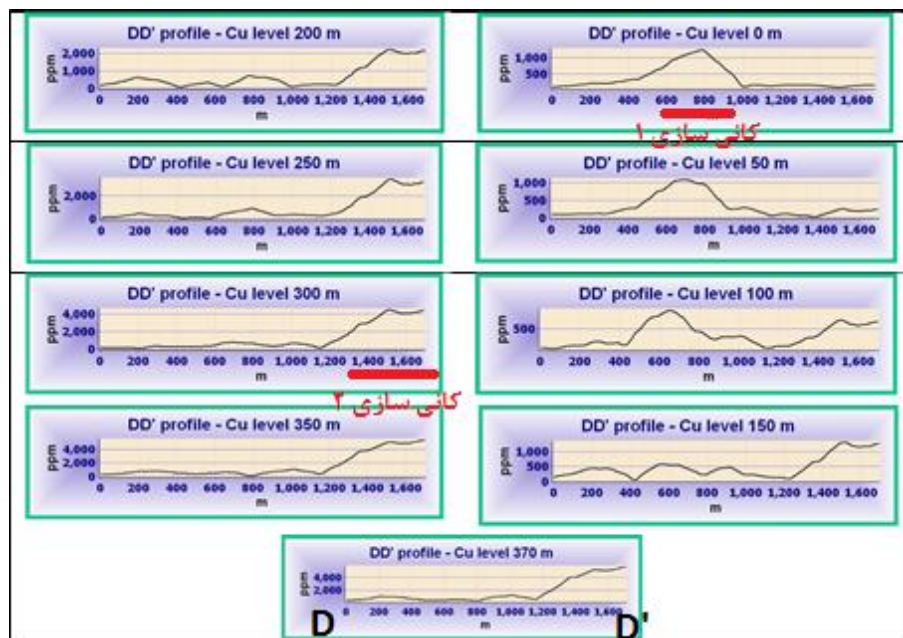
شکل ۵-۳۸: مقاطع عرضی تغییرات عنصر آرسنیک در طول پروفیل ژئوشیمیایی CC' در سطوح مختلف عمقی



شکل ۵-۳۹: مقاطع عرضی تغییرات عنصر وانادیوم در طول پروفیل ژئوشیمیایی CC' در سطوح مختلف عمقی

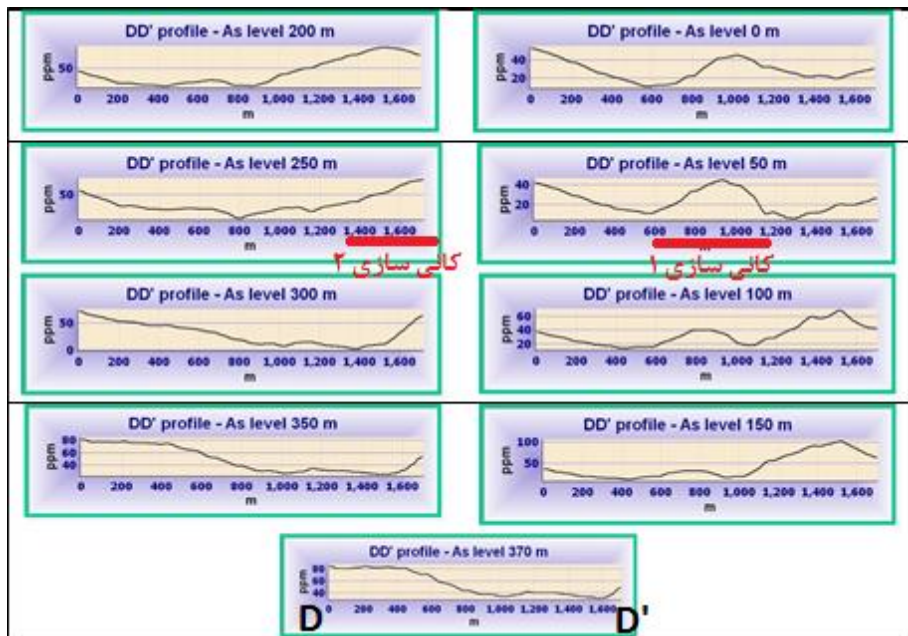
۵-۲-۵- پروفیل ژئوشیمیایی DD'

پروفیل DD' عمود بر پروفیل CC' و در راستای غربی - شرقی ترسیم شده است. بررسی بخش مرکزی پروفیل CC' نشان داد که در این قسمت فرسایش زیادی اتفاق افتاده و به ریشه کانی‌سازی رسیده‌ایم (کانی‌سازی ۱) (شکل ۵-۶ و پروفیل DD'، شکل ۵-۴). در انتهای پروفیل به سمت شرق منطقه، از عمق ۱۰۰ متری هاله‌های مس شروع به افزایش یافته و در عمق ۲۵۰ متر به بعد به حد اقتصادی می‌رسد (بیش از ۰/۴ درصد - کانی‌سازی ۲).

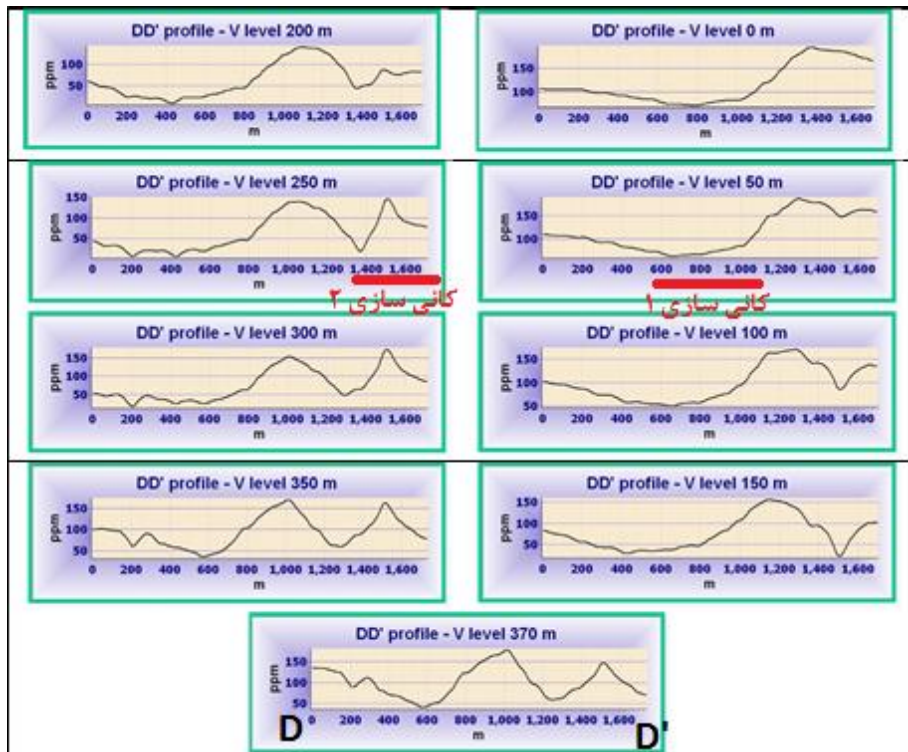


شکل ۵-۴: مقاطع عرضی تغییرات عنصر مس در طول پروفیل ژئوشیمیایی DD' در سطوح مختلف عمقی

تغییرات غلظت عناصر وانادیوم و آرسنیک (شکل ۵-۴۱ و شکل ۵-۴۲) تایید کننده نتایج حاصل از بررسی تغییرات غلظت مس هستند و شدت وانادیوم در مرکز پروفیل که کانی‌سازی در عمق ادامه ندارد با افزایش عمق، بیشتر شده است. میزان آرسنیک در انتهای پروفیل به سمت شرق منطقه، از عمق ۱۰۰ متری افزایش پیدا می‌کند و احتمال وجود کانی‌سازی پنهان را شدت می‌دهد.



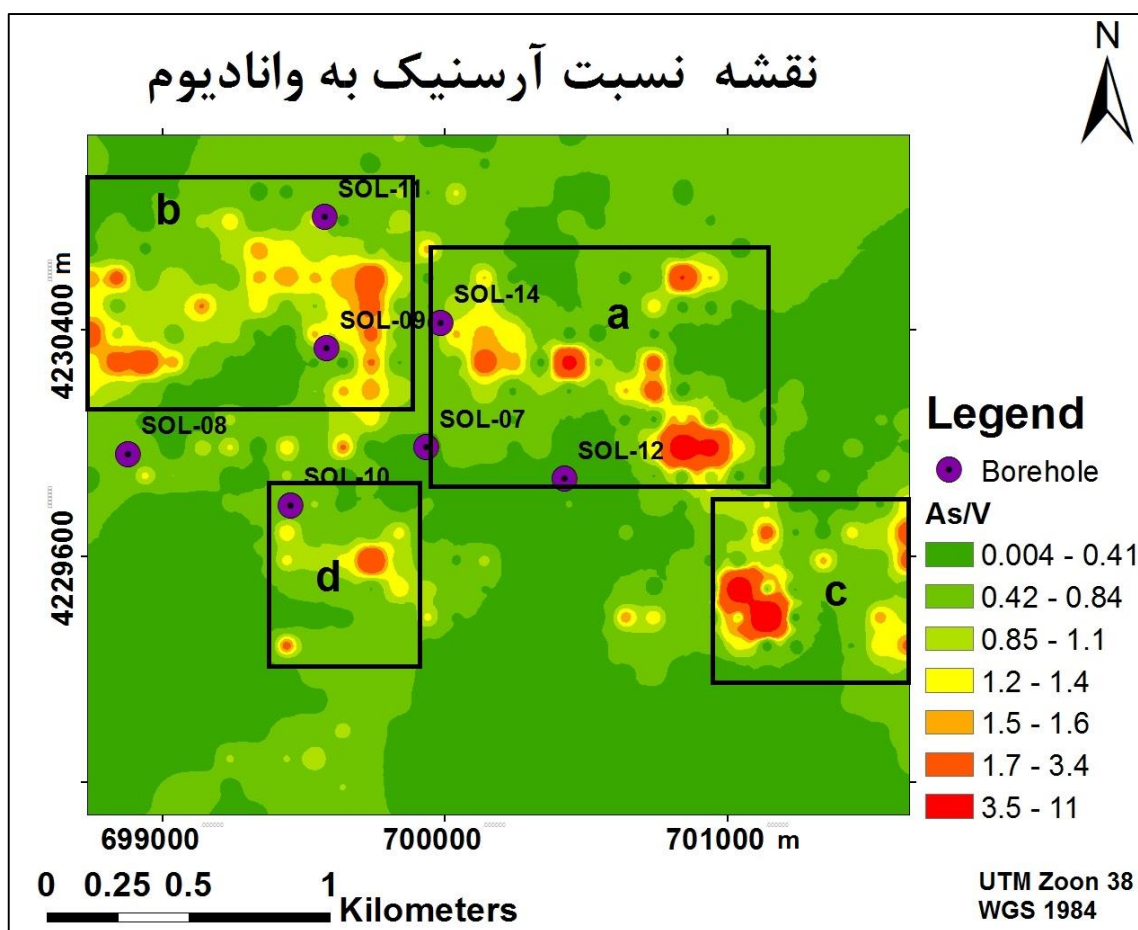
شکل ۵-۴۱: مقاطع عرضی تغییرات عنصر آرسنیک در طول پروفیل ژئوشیمیایی DD' در سطوح مختلف عمقی



شکل ۵-۴۲: مقاطع عرضی تغییرات عنصر وانادیوم در طول پروفیل ژئوشیمیایی DD' در سطوح مختلف عمقی

۵-۲-۶- نسبت عنصر فوق کانساری به تحت کانساری

نتایج بررسی تغییرات غلظت عناصر مختلف در سطوح مختلف عمقی نشان داد که عنصری مانند آرسنیک نقش فوق کانساری و وانادیوم نقش تحت کانساری ایفا می‌کند. جهت شناسایی مناطق پتانسیل‌دار در عمق و کانی‌سازی عمیق می‌توان از نسبت فوق به تحت کانساری استفاده نمود که نقشه نسبت این دو عنصر در (شکل ۵-۴۳) نشان داده شده است.



شکل ۵-۴۳: نقشه نسبت آرسنیک به وانادیوم

نقشه نسبت آرسنیک به وانادیوم مناطقی را با مقادیر بالایی از این نسبت نشان می‌دهد که شامل چهار منطقه a, b, c و d است که در بررسی نقشه‌های غلظت در سطوح مختلف عمقی نتیجه بر این شد که کانی‌سازی در محدوده a در عمق ادامه دارد و در محدوده d فرسایش انجام یافته و در انتهای کانی‌سازی

قرار گرفته‌ایم که این نتایج با نتایج حاصل از نسبت عنصر فوق‌کانساری به تحت‌کانساری مطابقت دارد. دو محدوده c و b نیز مقادیر بالایی از نسبت فوق به تحت‌کانساری را نشان می‌دهند. منطقه b بر طبق اطلاعات گمانه‌های SOL-08، SOL-09 و SOL-11 (شکل ۵-۴) فاقد کانی‌سازی در عمق هست. برای اعتبارسنجی محدوده c اطلاعاتی در دست نیست، برای حل این مشکل به جای استفاده از عناصر محدودی مانند عناصر فوق‌کانساری و تحت‌کانساری برای شناسایی مناطق کانی‌سازی، بهترین کار این است که از همه عناصر مهم مرتبط در کانی‌سازی مس برای مدل‌سازی تغییرات غلظت مس در سطوح مختلف عمقی استفاده گردد که این کار با استفاده از الگوریتم‌های طبقه‌بندی انجام خواهد گرفت.

۵-۳- طبقه‌بندی تغییرات غلظت مس در سطوح مختلف عمقی

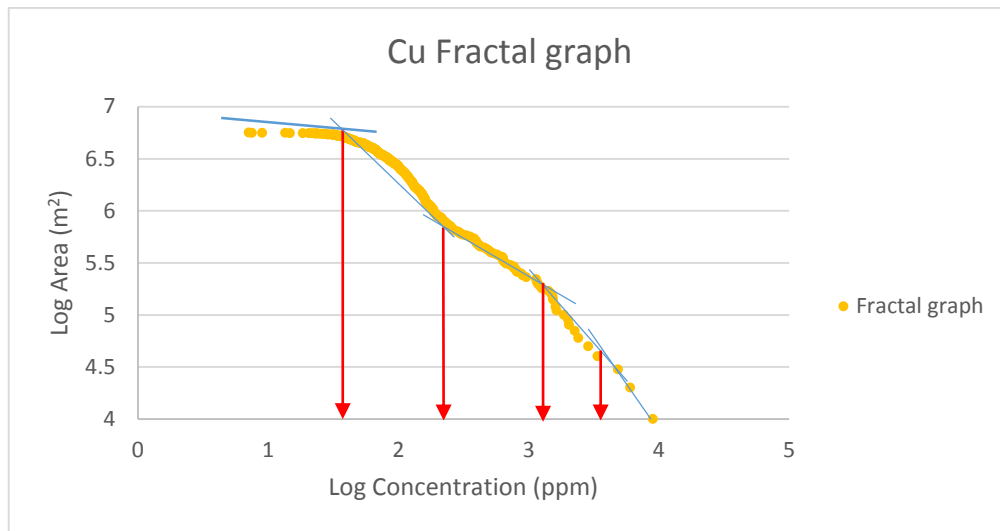
۵-۳-۱- نقشه فرکتالی تغییرات غلظت مس

به منظور تعیین کلاس غلظت عنصر مس در داده‌های آموزشی و آزمون از روش فرکتالی عیار - مساحت استفاده شده است. نمودار فرکتالی عنصر مس در شکل ۵-۴۴ نشان می‌دهد که ۵ کلاس مس در بین داده‌ها وجود دارد.

تغییرات غلظت مس در جدول ۵-۲ نشان داده شده است.

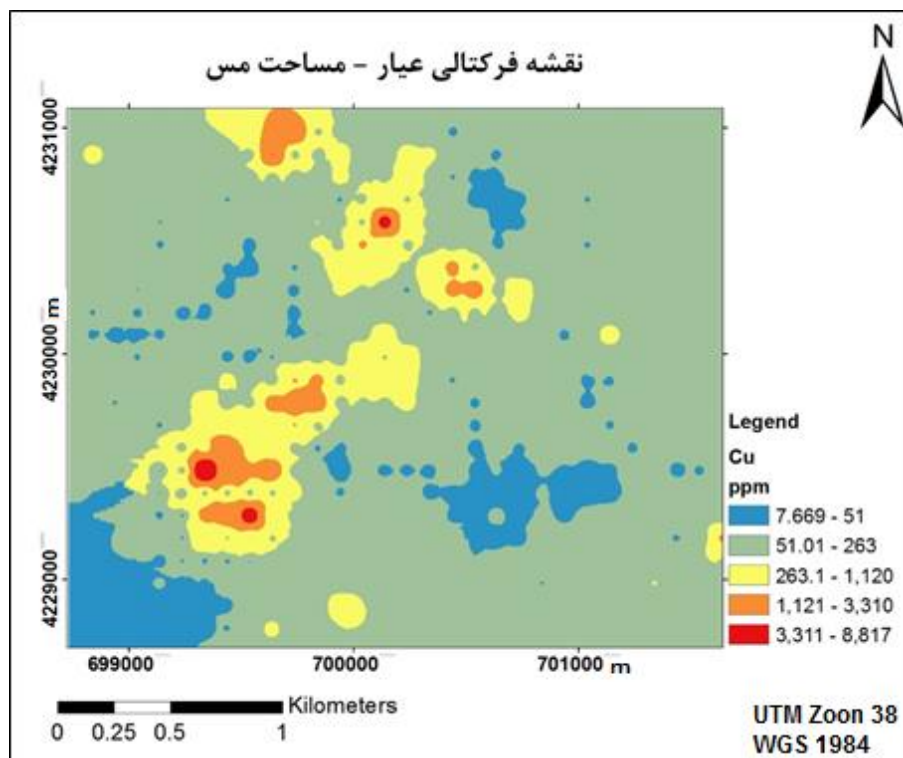
جدول ۵-۲: تغییرات فرکتالی غلظت مس

کلاس مس	کلاس ۱	کلاس ۲	کلاس ۳	کلاس ۴	کلاس ۵
عیار مس (ppm)	کمتر از ۵۱	۵۱ - ۲۶۳	۲۶۴ - ۱۱۲۰	۱۱۲۱ - ۳۳۱۰	بیشتر از ۳۳۱۰



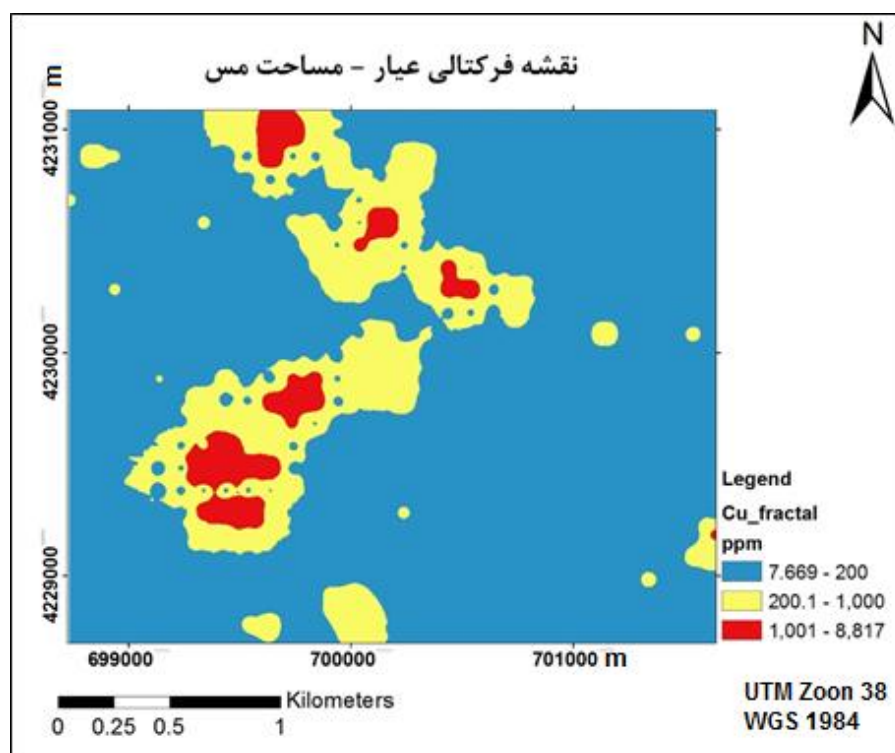
شکل ۴۴-۵: نمودار فرکتالی عنصر مس در منطقه

نقشه تغییرات غلظت مس بر اساس نتایج بدست آمده از روش فرکتالی عیار مس در شکل ۴۵-۵ رسم گردیده است.



شکل ۴۵-۵: نقشه فرکتالی تغییرات غلظت مس در سطح

در استفاده از روش‌های شناسایی الگو دقت بدست آمده از این روش‌ها ارتباط معکوسی با تعداد کلاس‌ها دارد. بنابراین باید تلاش شود تا جایی که امکان پذیر است تعداد کلاس‌ها را کاهش داد. با بررسی نقشه فرکتالی بدست آمده برای مس مشاهده می‌شود که درون داده‌ها دو جامعه زمينه (تا غلظت ۲۶۳ ppm)، یک جامعه حد گذر (از غلظت ۲۶۳ ppm تا ۱۱۲۰ ppm) و دو جامعه آنومالی (غلظت بالاتر از ۱۱۲۰ ppm) داریم که می‌توان این جوامع را باهم تلفیق و تعداد کلاس‌ها را کاهش داد. به طوریکه غلظت ۲۶۳ ppm را به عنوان مرز زمينه و از ۱۱۲۰ ppm بیشتر را به عنوان مرز آنومالی در نظر گرفته شده است. در انجام مراحل آموزش و آزمون داده‌ها با روش‌های طبقه‌بندی داده‌ها نتیجه‌گیری شد اگر مرز زمينه ۲۰۰ و مرز آنومالی ۱۰۰۰ در نظر گرفته شود بیشترین دقت کسب خواهد شد (شکل ۵-۴۶).



شکل ۵-۴۶: نقشه فرکتالی تغییرات غلظت مس در سطح

۵-۳-۲- تحلیل مولفه‌های اصلی

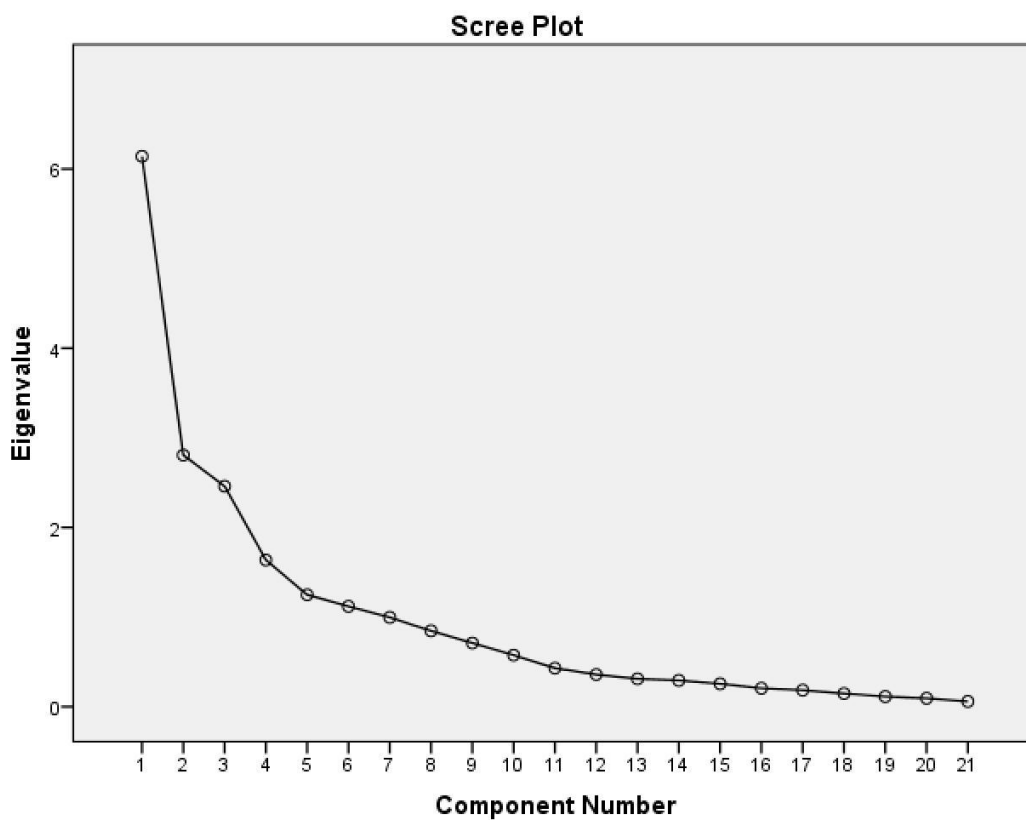
داده‌های مورد استفاده برای شناسایی کلاس‌های غلظت عنصر مس شامل ۲۱ مولفه است که این مولفه‌ها همان غلظت عناصر ژئوشیمیایی است که مورد استفاده قرار خواهند گرفت (جدول ۵-۳).

جدول ۵-۳: عناصر ژئوشیمیایی مورد استفاده برای طبقه‌بندی مس

Mn	Mg	La	K	Fe	Co	Ca	Be	Ba	As	Al
	Zn	V	Sr	Sc	S	Pb	P	Ni	Na	Mo

انتخاب بهترین ویژگی‌ها در تحلیل داده‌ها باعث بیشترین همبستگی داده‌های ورودی و خروجی و دقت بالاتر نتایج بدست آمده می‌شود. روش تحلیل مولفه‌های اصلی روشی است که با نداشت داده‌ها در یک فضای آماری جدید، سعی در شناسایی بهترین ویژگی‌ها به منظور تحلیل‌ها را دارد.

نتایج بدست آمده از تحلیل مولفه‌های اصلی داده‌ها نشان می‌دهد که از بین ۲۱ مولفه در فضای آماری جدید، ۱۰ مولفه بیشترین ضریب را در بیان تغییرات جامعه نشان می‌دهند و با کاهش ابعاد ویژگی‌ها از ۲۱ به ۱۰ تحلیل داده‌ها با دقت بیشتری انجام خواهد شد (شکل ۵-۴۷).

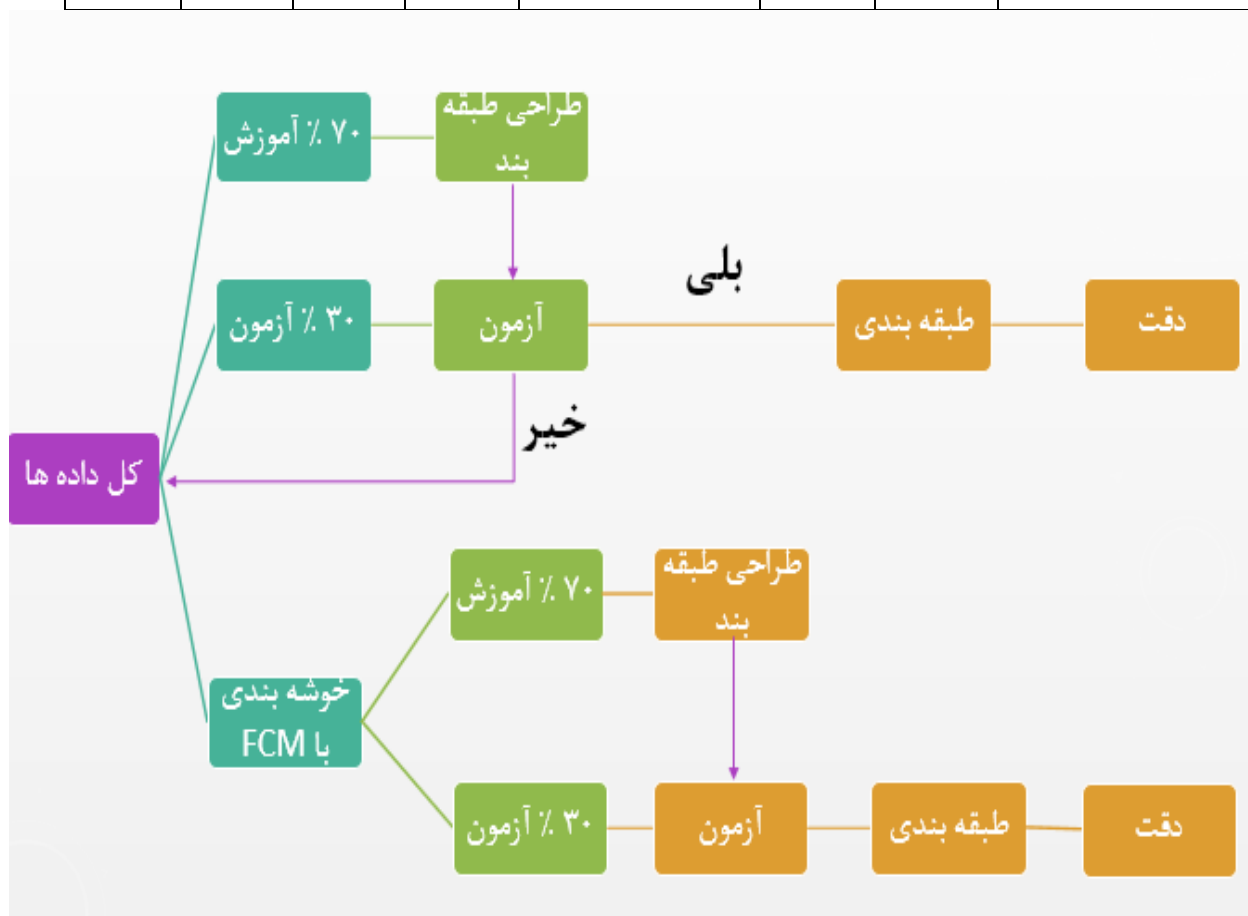


شکل ۵-۴۷: نمودار مقادیر ویژه مولفه‌های اصلی داده‌ها

در انجام روش‌های طبقه‌بندی ۷۰ درصد داده‌ها به عنوان داده‌های آموزشی و ۳۰ درصد داده‌ها به عنوان داده‌های آزمون مورد استفاده قرار می‌گیرند. برای انجام طبقه‌بندی داده‌ها ابتدا کل داده‌ها مورد استفاده قرار گرفته و سپس داده‌ها خوشه‌بندی شده به عنوان داده‌های ورودی استفاده شده که نشان می‌دهد استفاده از داده‌های خوشه‌بندی شده با روش FCM دقت طبقه‌بندی را افزایش داده است (شکل ۵-۴۸). در بررسی‌ها صورت گرفته مشخص گردید که استفاده از روش تحلیل مولفه‌های اصلی نیز دقت طبقه‌بندی را افزایش می‌دهد.

جدول ۴-۵: داده‌های مورد استفاده برای طبقه‌بندی

تعداد داده درون هر کلاس			عنصر هدف	تعداد مولفه PCA مورد استفاده	تعداد عنصر	تعداد داده‌ها	داده‌ها
کلاس سوم	کلاس دوم	کلاس اول					
۱۵۱	۳۰۵	۱۲۹۹	مس	۱۰	۲۱	۱۷۵۵	کل داده‌ها
۱۱۷	۱۴۴	۳۰۹	مس	۱۰	۲۱	۵۶۹	داده‌های خوشه‌بندی شده با FCM



شکل ۴-۵: استفاده از داده‌های خوشه‌بندی شده بوسیله الگوریتم FCM جهت طراحی طبقه‌بندی

۵-۳-۳- طبقه‌بندی داده‌ها با روش بیزین

۵-۳-۳-۱- استفاده از کل داده‌ها

ابتدا کل داده‌ها را که شامل ۱۷۵۵ داده با ۲۱ عنصر بود به عنوان داده‌های آموزشی و داده‌های آزمون مورد استفاده قرار داده و طبقه‌بند طراحی شده است و سپس طبقه‌بند طراحی شده با داده‌های آزمون مورد اعتبارسنجی قرار گرفته است. نتیجه به دست آمده نشان می‌دهد هرچند دقت کلی (۷۱/۴) دقت مناسبی به نظر می‌رسد اما با بررسی دقت هر کدام از کلاس‌ها متوجه می‌شویم که عملکرد طبقه‌بند در شناسایی کلاس‌های ۲ و ۳ دارای دقت‌های پایین ۱۷/۶ درصد و ۴۲/۵ درصد بوده که این نتایج مورد قبول نمی‌باشد (شکل ۵-۴۹).

کلاس واقعی	کلاس تصمیم			
	۱	۲	۳	
۱	۲۹۵	۳۲	۶	٪۸۸,۶
۲	۵۴	۱۵	۱۶	٪۱۷,۶
۳	۲۱	۲	۱۷	٪۴۲,۵
				٪۷۱,۴

شکل ۵-۴۹: ماتریس دقت طبقه‌بندی داده‌های آزمون با استفاده از روش بیزین و استفاده از کل داده‌ها

۵-۳-۲- استفاده از داده‌های خوشه‌بندی

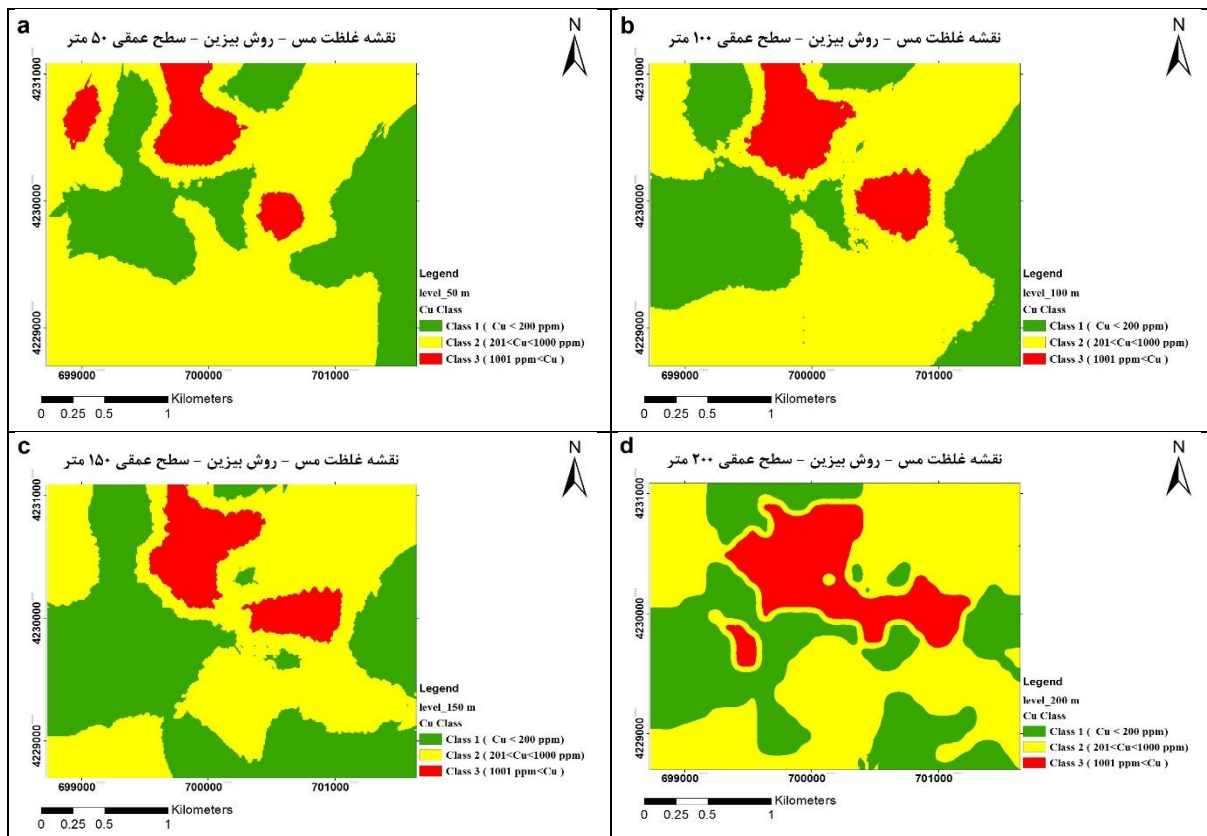
در این مرحله از داده‌های خوشه‌بندی شده با استفاده از روش FCM استفاده و بهترین مجموعه داده از بین خوشه‌ها برای استفاده به عنوان داده‌های آموزشی و آزمون انتخاب شده است. داده‌ها شامل ۵۶۹ داده و ۲۱ عنصر می‌باشد. پس از طراحی طبقه‌بند با استفاده از ۷۰ درصد داده‌ها (به عنوان داده‌های آموزشی)، طبقه‌بند طراحی شده را با ۳۰ درصد از داده‌ها که به صورت تصادفی انتخاب شده مورد اعتبارسنجی قرار گرفته است. نتایج بدست آمده نشان می‌دهد که دقت طبقه‌بند افزایش خوبی داشته است و از ۷۱/۴ درصد به ۸۴/۵ درصد افزایش پیدا کرده است. علاوه بر آن دقت کلاس‌های ۲ و ۳ در هنگام استفاده از کل داده‌ها که به ترتیب ۱۷/۶ و ۴۲/۵ درصد بود به مقادیر ۷۳/۵ و ۷۸/۳ درصد افزایش یافته که نشان می‌دهد عملکرد طبقه‌بند در استفاده از داده‌های خوشه‌بندی به عنوان داده‌های آموزشی و آزمون نسبت به حالت استفاده از کل داده‌ها بهتر شده است (شکل ۵-۵۰).

کلاس واقعی	کلاس تصمیم			
	۱	۲	۳	
۱	۸۲	۷	۲	٪۹۰,۱
۲	۶	۲۵	۳	٪۷۳,۵
۳	۲	۳	۱۸	٪۷۸,۳
				٪۸۴,۵

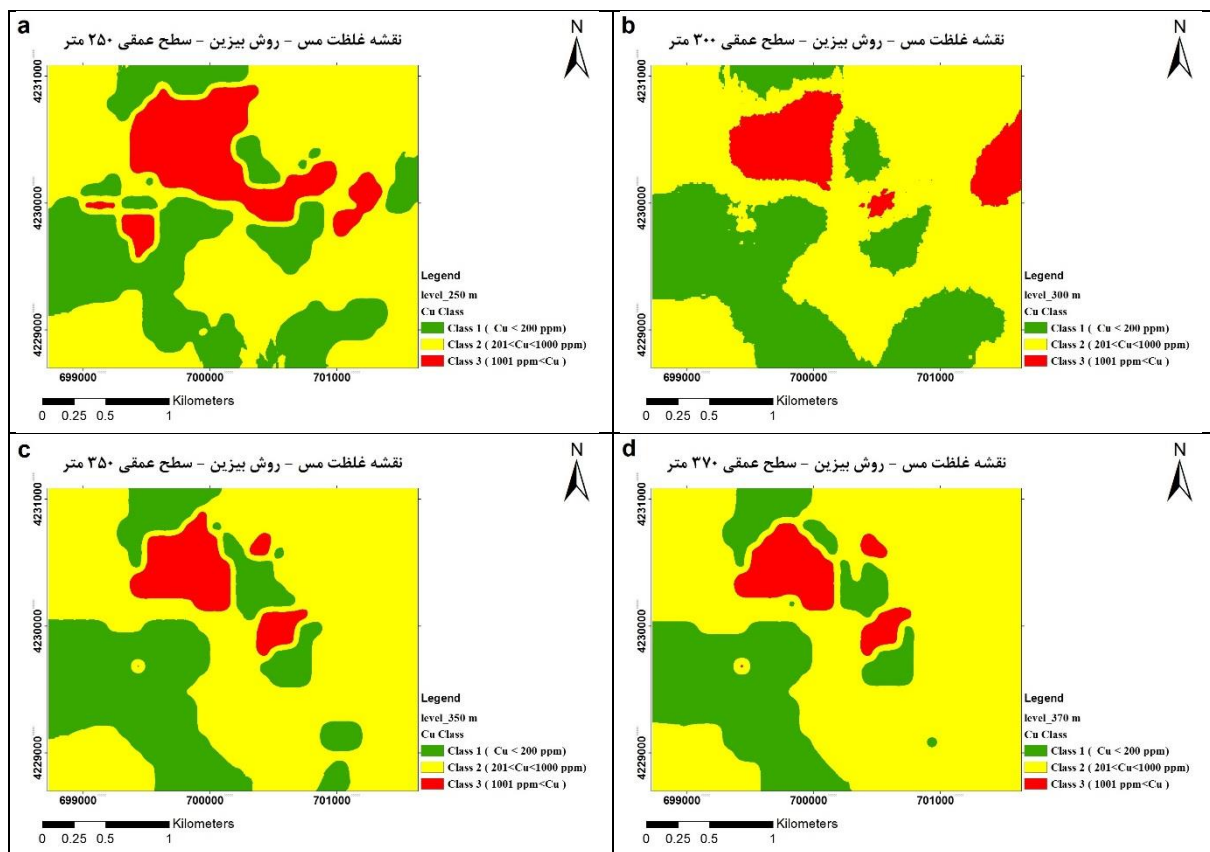
شکل ۵-۵۰: ماتریس دقت طبقه‌بندی داده‌های آزمون با استفاده از روش بیزین و استفاده از داده‌های خوشه‌بندی شده با

FCM

با استفاده از طبقه‌بند طراحی شده بوسیله روش بی‌زین و الگوریتم خوشه‌بندی FCM، اقدام به طبقه‌بندی غلظت مس در سطوح مختلف عمقی شده است (شکل ۵-۵۱ و شکل ۵-۵۲). این شکل گویای آن است که در عمق بین ۲۰۰ تا ۲۵۰ متری کانی‌سازی اصلی وجود دارد.



شکل ۵-۵۱: طبقه‌بندی غلظت مس در سطوح مختلف عمقی با استفاده از روش بی‌زین (a) سطح عمقی ۵۰ متر، (b) سطح عمقی ۱۰۰ متر، (c) سطح عمقی ۱۵۰ متر و (d) سطح عمقی ۲۰۰ متر.



شکل ۵-۵: طبقه‌بندی غلظت مس در سطوح مختلف عمقی با استفاده از روش بیزین (a) سطح عمقی ۲۵۰ متر، (b) سطح عمقی ۳۰۰ متر، (c) سطح عمقی ۳۵۰ متر و (d) سطح عمقی ۳۷۰ متر.

۵-۳-۴- طبقه‌بندی داده‌ها با روش SVM

۵-۳-۴-۱- استفاده از کل داده‌ها

برای طراحی طبقه‌بند با استفاده از روش ماشین بردار پشتیبان، ابتدا کل داده‌ها (شامل ۱۷۵۵ داده و ۲۱ عنصر) به عنوان داده‌های آموزشی و داده‌های آزمون مورد استفاده قرار گرفته و طبقه‌بند طراحی گردیده است. سپس طبقه‌بند طراحی شده با داده‌های آزمون مورد اعتبارسنجی قرار گرفته است. نتیجه به دست آمده نشان می‌دهد هرچند دقت کلی (۷۶/۶) دقت مناسبی به نظر می‌رسد اما با بررسی دقت هر کدام از کلاس‌ها متوجه می‌شویم که عملکرد طبقه‌بند در شناسایی کلاس‌های ۲ و ۳ دارای دقت‌های پایین ۲/۴ درصد و ۴۷/۵ درصد است و این نتایج مورد قبول نمی‌باشد به نحوی که طبقه‌بند طراحی شده قادر به

شناسایی کلاس ۲ مس نبوده و کلاس ۲ را به اشتباه در کلاس‌های ۱ و ۳ طبقه‌بندی کرده است (شکل ۵-۵۳).

کلاس واقعی	کلاس تصمیم			
	۱	۲	۳	
۱	۳۳۰	۱	۲	%۹۹,۱
۲	۷۱	۲	۱۲	%۲,۴
۳	۲۱	۰	۱۹	%۴۷,۵
				%۷۶,۶

شکل ۵-۵۳: ماتریس دقت طبقه‌بندی داده‌های آزمون با استفاده از روش SVM و استفاده از کل داده‌ها

۵-۳-۴-۲- استفاده از داده‌های خوشه‌بندی

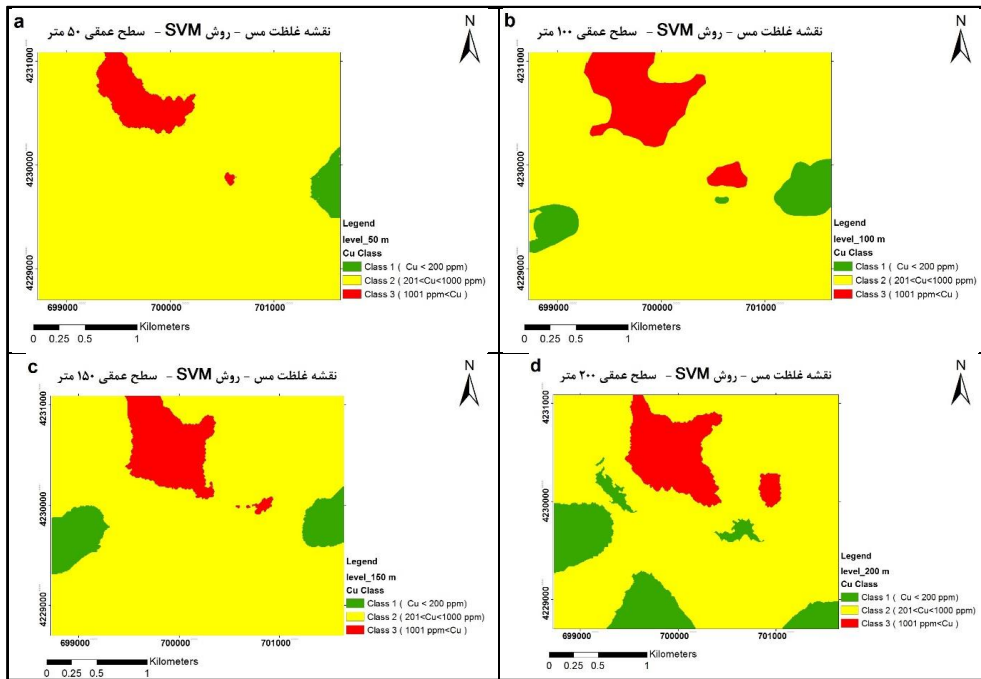
در این مرحله از داده‌های خوشه‌بندی شده با استفاده از روش FCM، استفاده و بهترین مجموعه داده از بین خوشه‌ها برای استفاده به عنوان داده‌های آموزشی و آزمون انتخاب شده است. تعداد داده‌ها شامل ۵۶۹ داده و ۲۱ عنصر ژئوشیمیایی است. پس از طراحی طبقه‌بند با استفاده از ۷۰ درصد داده‌ها به عنوان داده‌های آموزشی، طبقه‌بند طراحی شده با ۳۰ درصد از داده‌ها که به صورت تصادفی انتخاب شده مورد اعتبارسنجی قرار گرفته است. نتایج بدست آمده نشان می‌دهد که دقت طبقه‌بند افزایش خوبی داشته است و از ۷۱/۴ درصد به ۸۴/۵ درصد افزایش پیدا کرده است. علاوه بر آن دقت کلاس‌های ۲ و ۳ در هنگام استفاده از کل داده‌ها که به ترتیب ۲/۴ و ۴۷/۵ درصد بودند به مقادیر ۶۴/۷ و ۶۹/۶ درصد افزایش

پیدا کرده‌اند که نشان می‌دهد عملکرد طبقه‌بند در استفاده از داده‌های خوشه‌بندی به عنوان داده‌های آموزشی و آزمون نسبت به حالت استفاده از کل داده‌ها بهتر شده است (شکل ۵-۵۴).

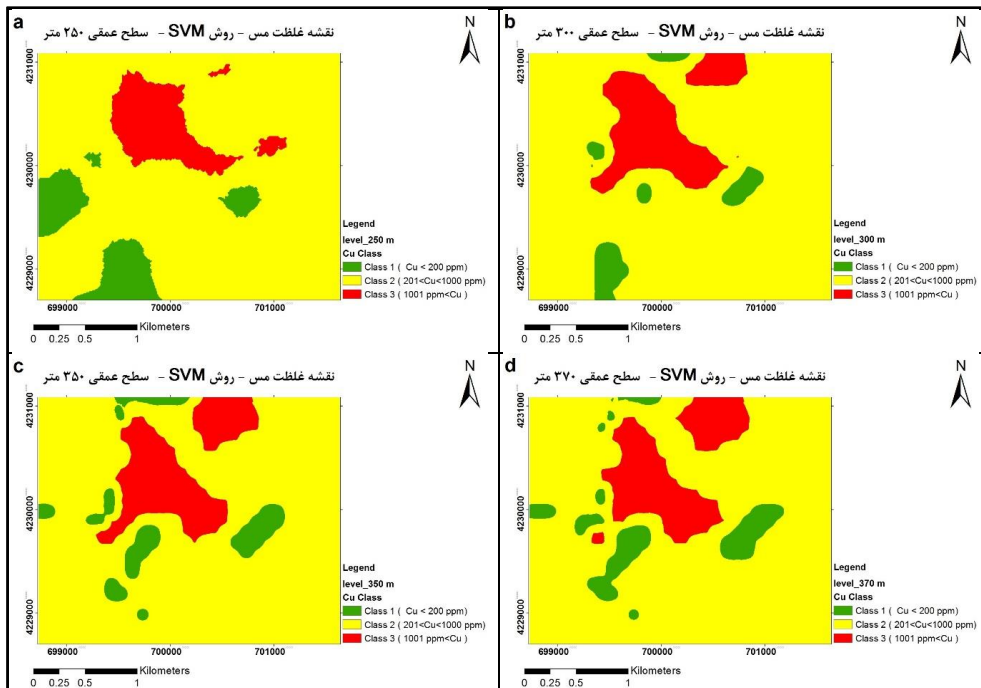
کلاس واقعی	کلاس تصمیم			
	۱	۲	۳	
۱	۸۷	۴	۰	%۹۵,۶
۲	۶	۲۲	۶	%۶۴,۷
۳	۰	۷	۱۶	%۶۹,۶
				%۸۴,۵

شکل ۵-۵۴: ماتریس دقت طبقه‌بندی داده‌های آزمون با استفاده از روش SVM و استفاده از داده‌های خوشه‌بندی شده با FCM

با استفاده از طبقه‌بند طراحی شده بوسیله روش ماشین بردار پشتیبان و الگوریتم خوشه‌بندی FCM اقدام به طبقه‌بندی غلظت مس در سطوح مختلف عمقی شده است (شکل ۵-۵۵ و شکل ۵-۵۶). این اشکال (بویژه شکل ۵-۵۶) گویای آن می‌باشند که از عمق ۳۰۰ متر به پایین با دو آنومالی متفاوت در بخش شمالی و مرکزی مواجه خواهیم بود.



شکل ۵-۵۵: طبقه‌بندی غلظت مس در سطوح مختلف عمقی با استفاده از روش SVM (a) سطح عمقی ۵۰ متر، (b) سطح عمقی ۱۰۰ متر، (c) سطح عمقی ۱۵۰ متر و (d) سطح عمقی ۲۰۰ متر.



شکل ۵-۵۶: طبقه‌بندی غلظت مس در سطوح مختلف عمقی با استفاده از روش SVM (a) سطح عمقی ۲۵۰ متر، (b) سطح عمقی ۳۰۰ متر، (c) سطح عمقی ۳۵۰ متر و (d) سطح عمقی ۳۷۰ متر.

۵-۳-۵- طبقه‌بندی داده‌ها با روش KNN

۵-۳-۵-۱- استفاده از کل داده‌ها

برای طراحی طبقه‌بند با استفاده از روش K نزدیکترین همسایگی، ابتداً کل داده‌ها (شامل ۱۷۵۵ داده و ۲۱ عنصر) به عنوان داده‌های آموزشی و داده‌های آزمون مورد استفاده و طبقه‌بند طراحی گردید. طبقه‌بند طراحی شده با داده‌های آزمون مورد اعتبارسنجی قرار داده شده است. نتیجه به دست آمده نشان می‌دهد هرچند دقت کلی ۸۱/۴ دقت مناسبی به نظر می‌رسد اما عملکرد طبقه‌بند در شناسایی کلاس ۲ دارای دقت‌های پایین ۲۹/۴ درصد بود که مورد قبول نمی‌باشد و باید تلاش شود تا دقت شناسایی کلاس ۲ و در نهایت دقت کلی افزایش داده شود (شکل ۵-۵۷).

کلاس واقعی	کلاس تصمیم			
	۱	۲	۳	
۱	۳۲۳	۸	۲	%۹۷,۰
۲	۵۳	۲۵	۷	%۲۹,۴
۳	۱۰	۵	۲۵	%۶۲,۵
				%۸۱,۴

شکل ۵-۵۷: ماتریس دقت طبقه‌بندی داده‌های آزمون با استفاده از روش KNN و استفاده از کل داده‌ها

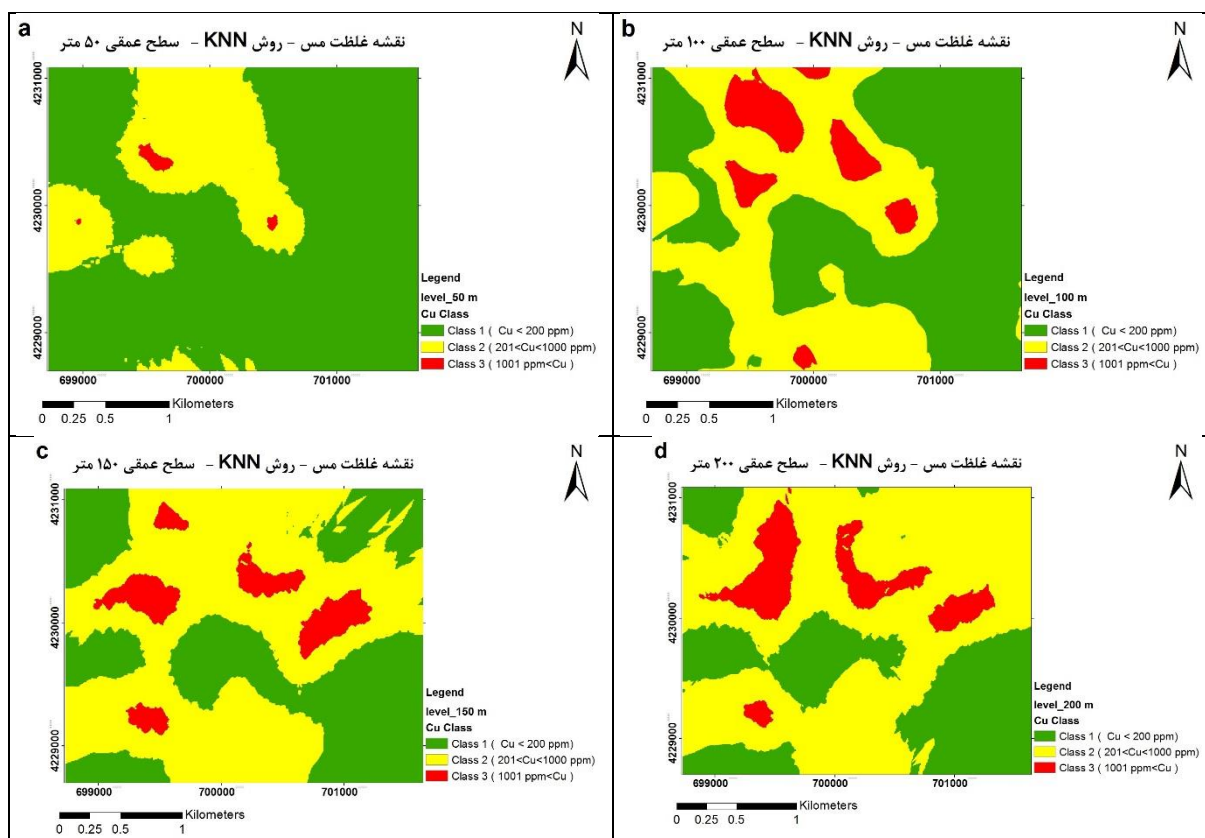
۵-۳-۲- استفاده از داده‌های خوشه‌بندی

داده‌های خوشه‌بندی شده با استفاده از روش FCM خوشه‌بندی شده و بهترین مجموعه داده از بین خوشه‌ها برای استفاده به عنوان داده‌های آموزشی و آزمون انتخاب شده است. تعداد داده‌های استفاده شده ۵۶۹ داده با ۲۱ عنصر ژئوشیمیایی می‌باشد. پس از طراحی طبقه‌بند با استفاده از ۷۰ درصد داده‌ها به عنوان داده‌های آموزشی، طبقه‌بند طراحی شده را با ۳۰ درصد از داده‌ها که به صورت تصادفی انتخاب شده بودند مورد اعتبارسنجی قرار گرفت. نتایج بدست آمده نشان می‌دهد که دقت طبقه‌بند افزایش خوبی داشته است و از ۸۱/۴ درصد به ۸۸/۵ درصد افزایش پیدا کرده است. علاوه بر آن دقت کلاس‌های ۲ و ۳ در هنگام استفاده از کل داده‌ها که به ترتیب ۲۹/۴ و ۶۲/۵ درصد بودند به مقادیر ۷۰/۶ و ۷۸/۳ درصد افزایش پیدا کرده‌اند که نشان می‌دهد عملکرد طبقه‌بند در استفاده از داده‌های خوشه‌بندی به عنوان داده‌های آموزشی و آزمون نسب به حالت استفاده از کل داده‌ها بهتر شده است (شکل ۵-۵۸).

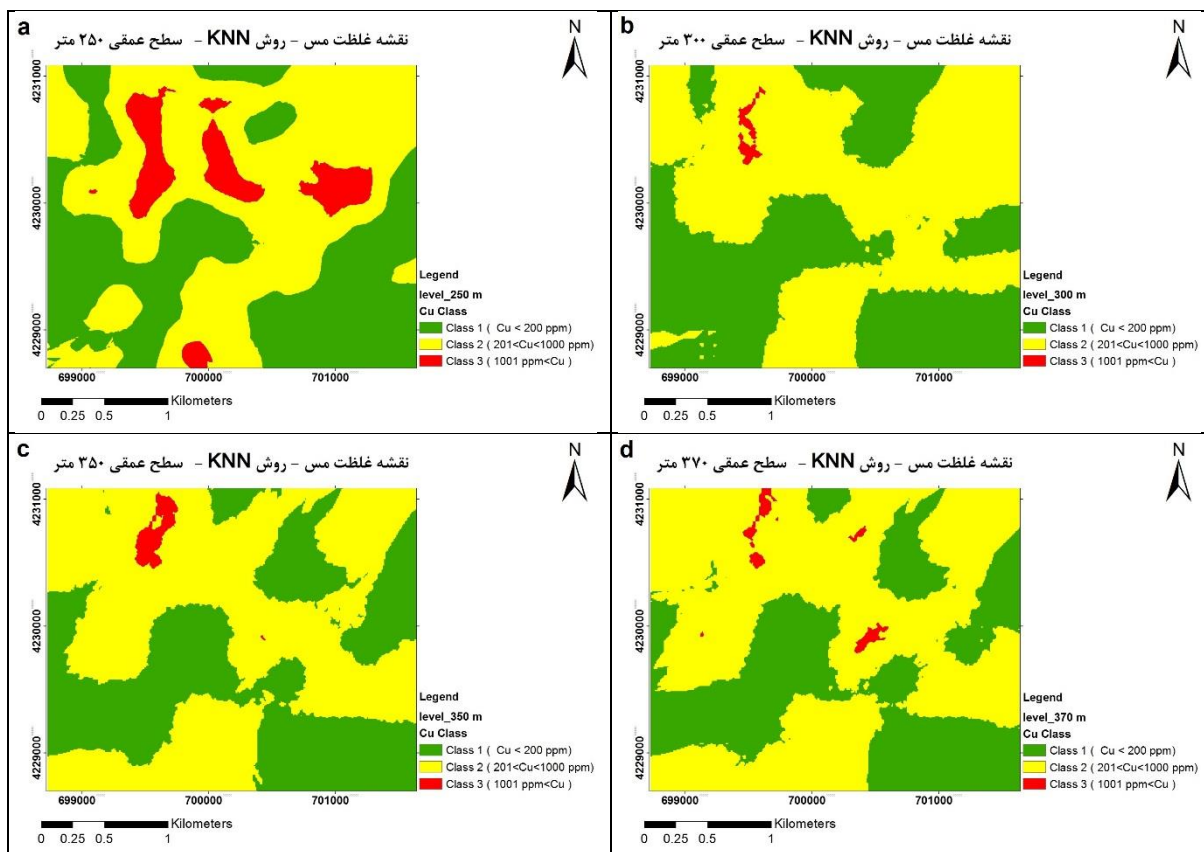
کلاس واقعی	کلاس تصمیم			
	۱	۲	۳	
۱	۸۹	۲	۰	٪۹۷,۸
۲	۶	۲۴	۴	٪۷۰,۶
۳	۰	۵	۱۸	٪۷۸,۳
				٪۸۸,۵

شکل ۵-۵۸: ماتریس دقت طبقه‌بندی داده‌های آزمون با استفاده از روش KNN و استفاده از داده‌های خوشه‌بندی شده با FCM

با استفاده از طبقه‌بند طراحی شده بوسیله روش K نزدیکترین همسایگی و الگوریتم خوشه‌بندی FCM اقدام به طبقه‌بندی غلظت مس در سطوح مختلف عمقی شده است (شکل ۵-۵۹ و شکل ۵-۶۰). بطور کلی این روش آنومالی‌ها را به خوبی تفکیک کرده است.



شکل ۵-۵۹: طبقه‌بندی غلظت مس در سطوح مختلف عمقی با استفاده از روش KNN (a) سطح عمقی ۵۰ متر، (b) سطح عمقی ۱۰۰ متر، (c) سطح عمقی ۱۵۰ متر و (d) سطح عمقی ۲۰۰ متر.

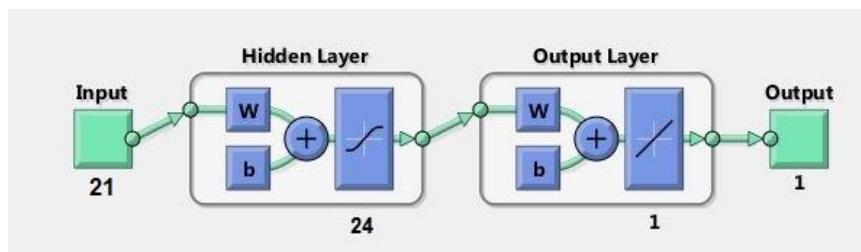


شکل ۵-۶: طبقه‌بندی غلظت مس در سطوح مختلف عمقی با استفاده از روش KNN (a) سطح عمقی ۲۵۰ متر، (b) سطح عمقی ۳۰۰ متر، (c) سطح عمقی ۳۵۰ متر و (d) سطح عمقی ۳۷۰ متر.

۵-۳-۶- طبقه‌بندی داده‌ها با روش ANN

۵-۳-۶-۱- استفاده از کل داده‌ها

برای طراحی طبقه‌بند با استفاده از روش شبکه عصبی چندلایه، ابتداً کل داده‌ها (۱۷۵۵ داده با ۲۱ عنصر) را به عنوان داده‌های آموزشی و آزمون مورد استفاده قرار گرفته است و طبقه‌بند طراحی (شکل ۵-۶۱) و با داده‌های آزمون مورد اعتبارسنجی قرار گرفته است. نتیجه به دست آمده نشان می‌دهد هرچند دقت کلی ۸۱/۷ دقت مناسبی است اما تلاش می‌شود تا با استفاده از داده‌های خوشه‌بندی شده دقت را افزایش داد (شکل ۵-۶۲).



شکل ۵-۶۱: طراحی طبقه‌بند با استفاده از شبکه عصبی MLP

کلاس تصمیم

	۱	۲	۳	
۱	۱۲۷۱	۱۹۶	۳۷	%۸۴,۵
۲	۱۹	۶۳	۱۵	%۶۴,۹
۳	۹	۴۶	۹۹	%۶۴,۳
				%۸۱,۷

کلاس واقعی

شکل ۵-۶۲: ماتریس دقت طبقه‌بندی داده‌های آموزش و آزمون با استفاده از روش ANN و استفاده از کل داده‌ها

۵-۳-۶-۲- استفاده از داده‌های خوشه‌بندی

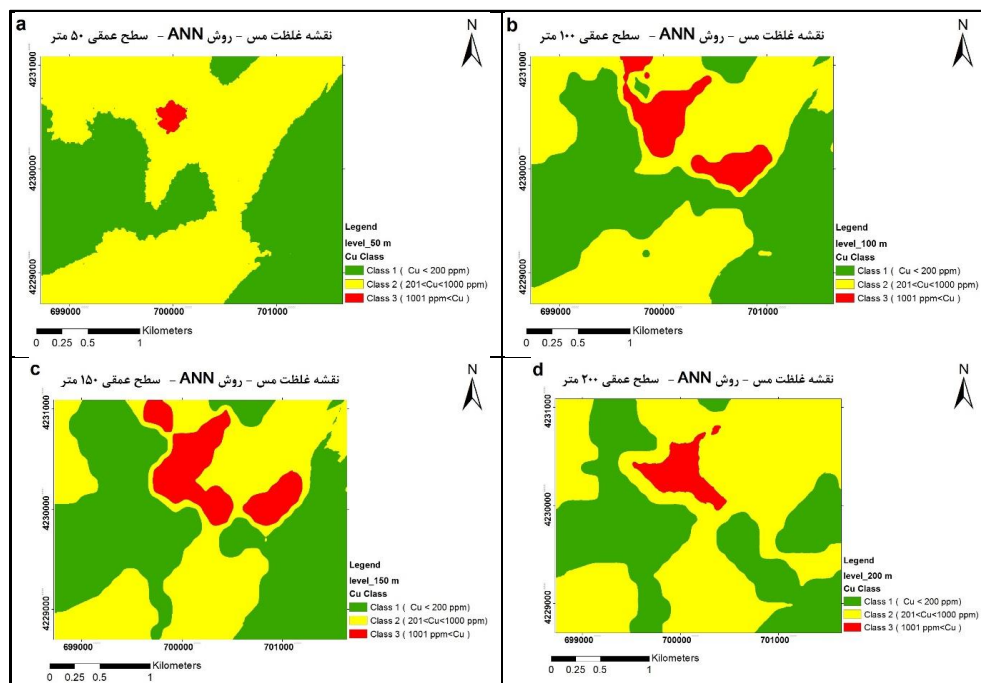
داده‌های خوشه‌بندی شده با استفاده از روش FCM استفاده و بهترین مجموعه داده از بین خوشه‌ها برای استفاده به عنوان داده‌های آموزشی و آزمون انتخاب شده است. تعداد داده برای ۲۱ عنصر استفاده شده است. پس از طراحی طبقه‌بند با استفاده از ۷۰ درصد داده‌ها به عنوان داده‌های آموزشی، طبقه‌بند طراحی شده را با ۳۰ درصد از داده‌ها که به صورت تصادفی انتخاب شده بودند مورد اعتبارسنجی قرار گرفت. نتایج بدست آمده نشان می‌دهد که دقت طبقه‌بند از ۸۱/۷ درصد به ۹۰/۶ درصد افزایش پیدا کرده

است. دقت بدست آمده نشان می‌دهد عملکرد طبقه‌بند در استفاده از داده‌های خوشه‌بندی به عنوان داده‌های آموزشی و آزمون نسبت به حالت استفاده از کل داده‌ها افزایش قابل توجه دقت داشته است (شکل ۵-۶۳).

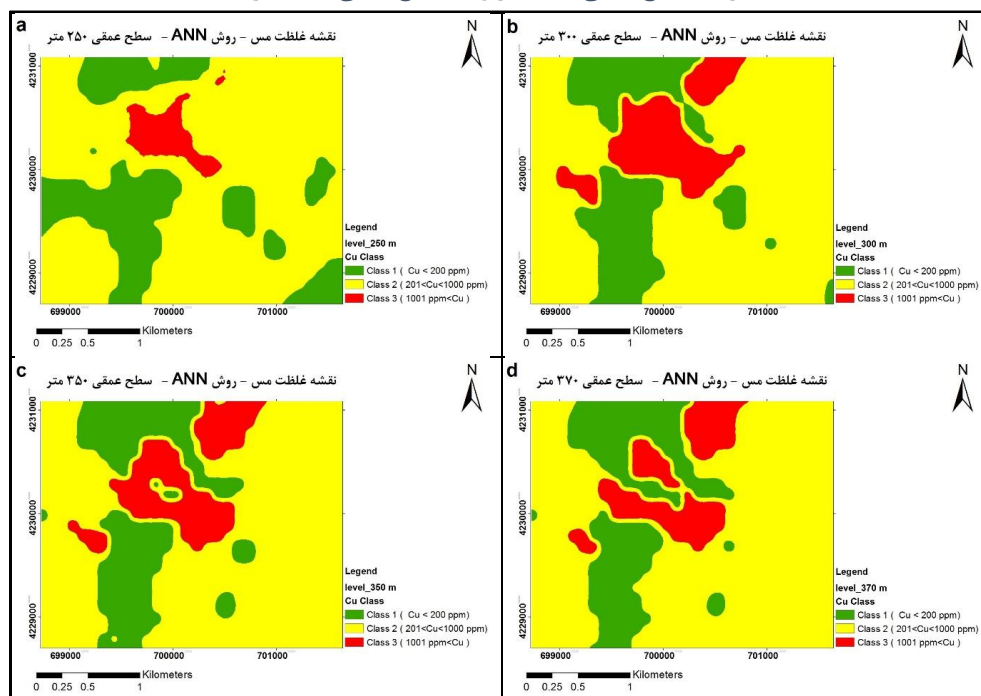
کلاس واقعی	کلاس تصمیم			
	۱	۲	۳	
۱	۴۷	۶	۰	%۸۸,۷
۲	۰	۱۳	۰	%۱۰۰
۳	۰	۲	۱۷	%۸۹,۵
				%۹۰,۶

شکل ۵-۶۳: ماتریس دقت طبقه‌بندی داده‌های آزمون با استفاده از روش ANN و استفاده از داده‌های خوشه‌بندی شده با FCM

با استفاده از طبقه‌بند طراحی شده بوسیله روش شبکه عصبی چندلایه و الگوریتم خوشه‌بندی FCM اقدام به طبقه‌بندی غلظت مس در سطوح مختلف عمقی شده است (شکل ۵-۶۴ و شکل ۵-۶۵). این نتایج نیز گویای حضور دو آنومالی تفکیک شده از عمق ۳۰۰ متری به بعد می‌باشند.



شکل ۵-۶۴: طبقه‌بندی غلظت مس در سطوح مختلف عمقی با استفاده از روش ANN (a) سطح عمقی ۵۰ متر، (b) سطح عمقی ۱۰۰ متر، (c) سطح عمقی ۱۵۰ متر و (d) سطح عمقی ۲۰۰ متر.



شکل ۵-۶۵: طبقه‌بندی غلظت مس در سطوح مختلف عمقی با استفاده از روش ANN (a) سطح عمقی ۲۵۰ متر، (b) سطح عمقی ۳۰۰ متر، (c) سطح عمقی ۳۵۰ متر و (d) سطح عمقی ۳۷۰ متر.

۵-۳-۷- انتخاب بهترین طبقه‌بند

برای انجام طبقه‌بندی غلظت عناصر در این پژوهش از ۴ روش بیزین، نزدیکترین K همسایگی، ماشین بردار پشتیبان و شبکه عصبی استفاده شده است. استفاده از روش‌های مختلف برای طبقه‌بندی کلاس عناصر این امکان را می‌دهد که با مقایسه نتایج بدست آمده از هر یک از این روش‌ها، بهترین روش را انتخاب و در ادامه تحلیل‌ها از این روش استفاده نماییم. دقت شناسایی کلاس‌های مس با استفاده از روش‌های مذکور در حالت استفاده از کل داده‌ها و در حالت استفاده از داده‌های خوشه‌بندی شده با استفاده از روش FCM در جدول ۵-۵ و جدول ۶-۵ آمده است.

جدول ۵-۵: نتایج طبقه‌بندها در حالت استفاده از کل داده‌ها

روش طبقه‌بندی	داده‌های آموزش و آزمون	دقت شناسایی کلاس اول	دقت شناسایی کلاس دوم	دقت شناسایی کلاس سوم	دقت کل
ANN	کل داده‌ها	۸۴/۵	۶۴/۹	۶۳/۴	۸۱/۷
KNN		۹۷	۲۹/۴	۶۲/۵	۸۱/۴
SVM		۹۹/۱	۲/۴	۴۷/۵	۷۶/۶
بیزین		۸۸/۶	۱۷/۶	۴۲/۵	۷۱/۴

جدول ۶-۵: نتایج طبقه‌بندها در حالت استفاده از داده‌های خوشه‌بندی شده با FCM

روش طبقه‌بندی	داده‌های آموزش و آزمون	دقت شناسایی کلاس اول	دقت شناسایی کلاس دوم	دقت شناسایی کلاس سوم	دقت کل
ANN	داده‌های خوشه‌بندی شده با FCM	۸۸/۷	۱۰۰	۸۹/۵	۹۰/۶
KNN		۹۷/۸	۷۰/۶	۷۸/۳	۸۸/۵
SVM		۹۵/۵	۶۴/۷	۶۹/۶	۸۴/۵
بیزین		۹۰/۱	۷۳/۵	۷۸/۳	۸۴/۵

مقایسه نتایج بدست آمده نشان می‌دهد که دقت شناسایی کلاس‌های مختلف غلظت مس در حالت استفاده از داده‌های خوشه‌بندی افزایش قابل توجهی داشته است به طوری که در روش شبکه عصبی دقت کلاس دوم از ۶۴/۹ به ۱۰۰ و دقت کلاس سوم از ۶۴/۳ به ۸۹/۵، در روش نزدیکترین K همسایگی دقت کلاس دوم از ۲۹/۹ به ۷۰/۶ و دقت کلاس سوم از ۶۲/۵ به ۷۸/۳، در روش ماشین بردار پشتیبان دقت کلاس دوم از ۲/۴ به ۶۴/۷ و دقت کلاس سوم از ۴۷/۵ به ۶۹/۶ و در روش بیزین دقت کلاس دوم از ۱۷/۶ به ۷۳/۵ و دقت کلاس سوم از ۴۲/۵ به ۷۸/۳ افزایش پیدا کرده‌اند. مقایسه دقت کل حاصل از شناسایی کلاس‌های مختلف با روش‌های مورد استفاده نمایان‌گر آن است که شبکه عصبی با دقتی برابر با ۹۰/۶ درصد بهترین عملکرد را داشته است.

فصل

۶- اعتبارسنجی، نتیجه‌گیری و پیشنهادات

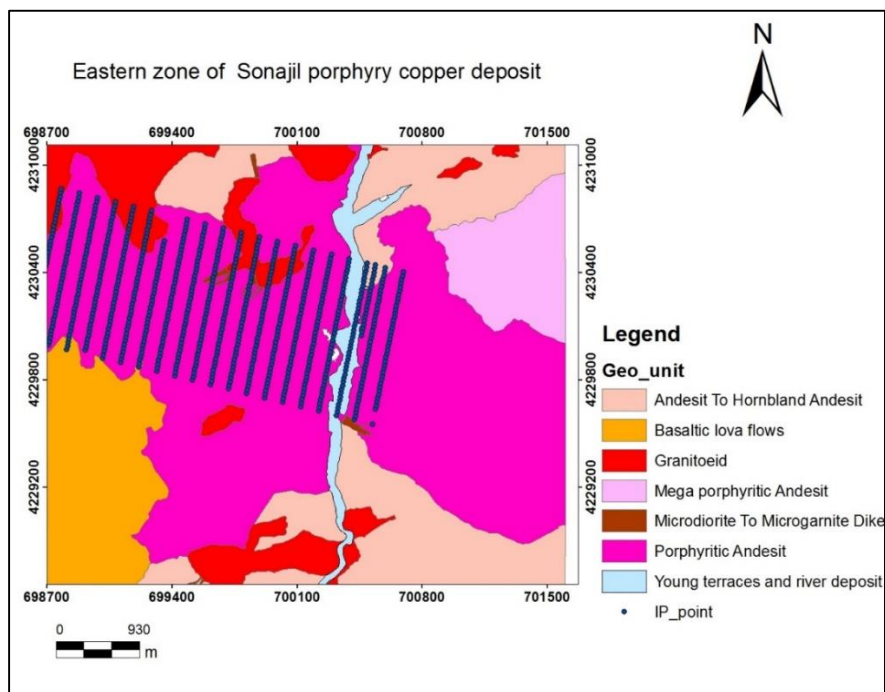
۶-۱- داده‌های اکتشافی منطقه و اعتبارسنجی

از داده‌های ژئوفیزیکی برای اعتبارسنجی نتایج بدست آمده از پردازش داده‌های ژئوشیمیایی استفاده خواهد گردید. از نتایج روش برداشت مستطیلی ژئوفیزیک برای شناسایی گسترش سطحی مناطق کانی‌سازی و از نتایج پروفیل‌های ژئوفیزیکی جهت شناسایی عمق شروع کانی‌سازی استفاده خواهد شد.

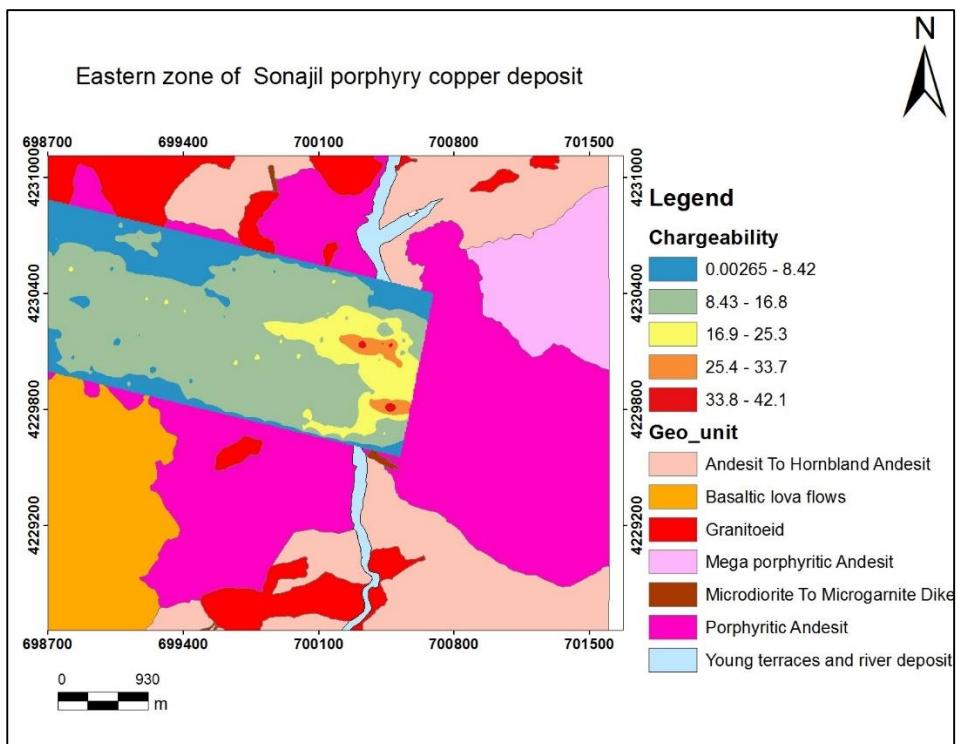
۶-۱-۱- روش پلاریزاسیون القایی

پلاریزاسیون القایی از روش‌های کارآمد در تشخیص کانی‌های سولفیدی است. اندازه‌گیری‌های IP می‌تواند در حوزه زمان یا فرکانس انجام شود. بارپذیری، معمولترین مقدار اندازه‌گیری زمان - حوزه‌ای محسوب می‌شود.

در این منطقه از آرایش مستطیلی (شکل ۶-۱) با فاصله خطوط جریان ۱۲۰۰ متر و الکترودهای پتانسیل ۲۰ متر برای بررسی تغییرات کل منطقه پرداخته شده است. در برداشت‌های IP هر ایستگاه علاوه بر بارپذیری، مقاومت‌ویژه ظاهری نیز به طور همزمان اندازه‌گیری می‌شود. در (شکل ۶-۲) نقشه بارپذیری در کل منطقه مشاهده می‌گردد. برای تعیین مناطق زمینه و آنومالی می‌توان نقشه تغییرات مقادیر بارپذیری را تهیه و محدوده‌های آنومال را مشخص کرد. بدین ترتیب روی نقشه بارپذیری فقط دو بی‌هنجاری موجود در بخش شرقی نقشه ارزش بررسی بیشتر را از نظر ژئوفیزیکی دارا می‌باشد (شکل ۶-۲).

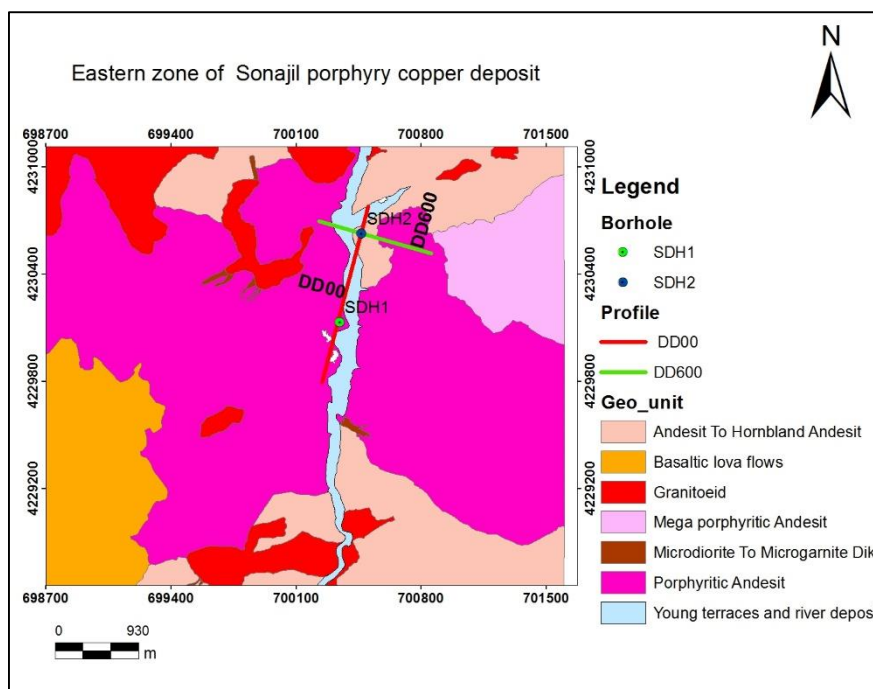


شکل ۶-۱: نمایش نقاط برداشت الکتریکی با آرایش مستطیلی بر روی نقشه زمین‌شناسی



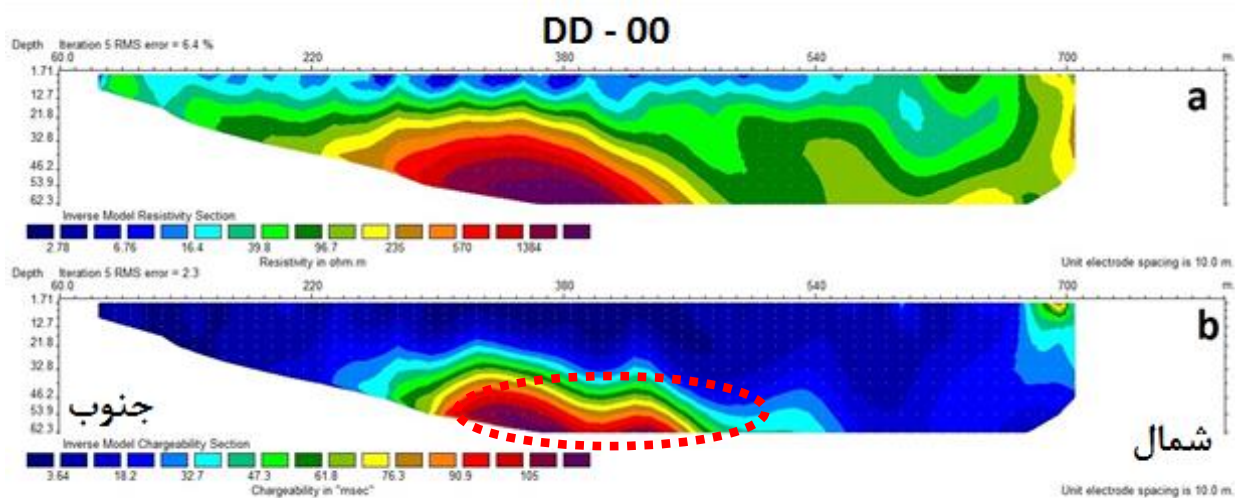
شکل ۶-۲: نمایش تغییرات کلی بارپذیری برداشت ژئوفیزیکی انجام شده بر روی نقشه زمین‌شناسی منطقه

از آرایش قطبی - دوقطبی و دوقطبی - دوقطبی با فاصله الکترودی ۲۰ متر و تعداد گام‌های از ۱ تا ۸ در دو پروفیل استفاده شده است. پروفیل DD600 که بصورت قطبی - دو قطبی برداشت شده است و راستای این پروفیل شرقی - غربی بوده و در قسمت شمالی محدوده قرار دارد و پروفیل DD00 که بصورت دوقطبی - دوقطبی برداشت شده، راستای شمال - جنوبی داشته و در قسمت مرکزی محدوده مطالعاتی قرار گرفته است. لازم به ذکر است که دو پروفیل همدیگر را قطع کرده‌اند (شکل ۳-۶).

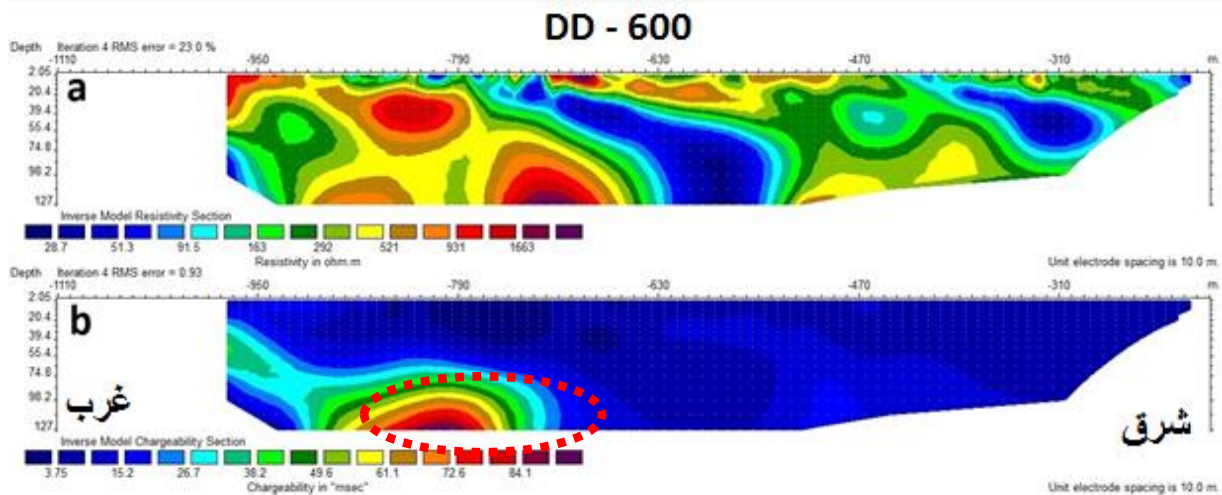


شکل ۳-۶: نمایش پروفیل‌های دوقطبی - دوقطبی DD600 و DD00 بر روی نقشه زمین‌شناسی (Hezarkhani, 2003). در مورد پروفیل‌های قطبی - دوقطبی DD00 و دوقطبی - دوقطبی DD600 نیز تفسیرهای لازم انجام گرفت. براساس نتایج بدست آمده از بارپذیری و مقاومت ویژه اندازه‌گیری شده در پروفیل شماره DD00 در (شکل ۴-۶) مشاهده می‌شود که یک توده کانی‌سازی در منطقه نمایان شده است. منطقه کانی‌سازی بر اساس افزایش شدت بارپذیری از عمق ۳۵ متری شروع شده و تا عمق ادامه یافته است. مدل‌سازی داده‌های مقاومت ویژه نشان داد که ادامه کانی‌سازی در عمق وجود دارد.

همچنین نتایج بدست آمده از تفسیرهای پروفیل دوقطبی- دوقطبی DD600 نیز نشانگر وجود کانی سازی در این منطقه می شود. با توجه به (شکل ۶-۵) مشاهده می شود که قسمتی از منطقه بارپذیری بالایی نشان می دهد و این می تواند به خاطر وجود کانی سازی پراکنده در منطقه باشد.



شکل ۴-۶: نتایج مدل سازی مقاومت ویژه (a)، بارپذیری (b) برای پروفیل قطبی- دوقطبی DD00



شکل ۵-۶: نتایج مدل سازی مقاومت ویژه (a)، بارپذیری (b) برای پروفیل دوقطبی- دوقطبی DD600

۲-۶- اعتبارسنجی

نتایج ژئوفیزیک نشان داد که کانی‌سازی در عمق ادامه دارد. در برداشت مستطیلی مشخص گردید که گسترشی سطحی از ماده معدنی در مرکز منطقه مطالعاتی وجود دارد (شکل ۲-۶) و پروفیل‌های ژئوفیزیکی نیز وجود توده کانی‌سازی در عمق را تایید می‌نمایند (شکل ۳-۶ و شکل ۴-۶). مدلسازی تغییرات غلظت مس بوسیله روش‌های طبقه‌بندی برای سطوح مختلف عمقی وجود یک توده کانی‌سازی در عمق را نشان می‌دهد به طوری که از عمق ۵۰ متری شروع و تا عمق ۳۷۰ متری ادامه دارد (شکل ۵-۶۴ و شکل ۵-۶۵) بررسی پروفیل‌های قائم گمانه‌ها برای عنصر مس (شکل ۵-۳)، از وجود کانی‌سازی در عمق حکایت دارد که این نتایج با سایر نتایج همخوانی دارد.

۳-۶- نتیجه‌گیری

۱. ترکیب الگوریتم‌های خوشه‌بندی با تخمینگر شبکه عصبی، دقت تخمین را به خوبی افزایش داد و باعث گردید نتایج بدست آمده از صحت و دقت بالایی برخوردار باشد.
۲. در بین الگوریتم‌های خوشه‌بندی، الگوریتم خوشه‌بندی FCM با افزایش دقت ۱۳ درصدی دقت میانگین تخمین عناصر را از ۷۵ درصد به ۸۸ درصد افزایش و بهترین عملکرد را نشان داد.
۳. بررسی غلظت عناصر ژئوشیمیایی نشان داد عنصر آرسنیک به عنوان عنصر فوق‌کانساری و وانادیوم به عنوان عنصر تحت‌کانساری کانسار مس پورفیری سوناجیل می‌باشند.
۴. الگوریتم طبقه‌بندی شبکه عصبی با دقت ۹۰/۶ درصد در بین سایر الگوریتم‌های استفاده شده در این پژوهش قادر به شناسایی کلاس‌ها مس در عمق شد.
۵. با استفاده از روش طبقه‌بندی شبکه عصبی و ویژگی سایر عناصر، تغییرات غلظت مس در سطوح مختلف عمقی و محل کانی‌سازی در عمق مشخص گردید.

۶. نتایج بدست آمده نشان داد که کانسار در قسمت جنوب غربی محدوده فرسایش یافته و در قسمت شمال - شمال شرق منطقه در عمق ادامه دارد.

۷. بررسی های تغییرات غلظت عنصر مس در سطح نشان داد که دو آنومالی در سطح وجود دارد که آنومالی قسمت شمالی در نزدیکی گمانه SOL-14 از عمق ۵۰ متری با عیار بیشتر از ۰/۲ درصد و در گمانه های SOL-07 و SOL-12 با عیار نزدیک به ۰/۲ درصد از عمق ۲۸۰ متری تا عمق ادامه دارد. نتایج بدست آمده برای تغییرات غلظت کلاس مس با روش شبکه عصبی برای سطوح مختلف عمقی محدوده مورد مطالعه، بوسیله اطلاعات ژئوفیزیکی و حفاری تایید گردید.

۴-۶- پیشنهادات

۱. الگوریتم خوشه بندی FCM با شکستن فضای ناهمگن کلی داده ها به فضا های همگن، همبستگی بین داده های ورودی و خروجی تخمینگر شبکه عصبی را افزایش می دهد بنابراین پیشنهاد می گردد برای بررسی سطح از فرسایش کانسار و شناسایی موقعیت کانی سازی در عمق، از ترکیب تخمینگر شبکه عصبی با الگوریتم خوشه بندی FCM استفاده گردد.

۲. در این مطالعه برای آموزش الگوریتم های شناسایی الگو از داده های سطحی و عمقی استفاده شد. بررسی ها نشان داد که داده های سطحی و عمقی ارتباط معناداری دارند و امکان کشف این ارتباط وجود دارد بنابراین پیشنهاد می گردد امکان شناسایی تغییرات عمقی غلظت عناصر ژئوشیمیایی با تحلیل داده های سطحی مورد بررسی قرار گیرد.

۳. برای بررسی دقیق تر کانی سازی قسمت شمالی محدوده مورد مطالعه، پیشنهاد می گردد نقشه زمین شناسی ۱:۱۰۰۰ با جزئیات بیشتر شامل دگرسانی ها، سنگ شناسی و ساختارها همراه با نقشه توپوگرافی تهیه گردد.

۴. پیشنهاد می‌گردد در قسمت شمالی کانسار که کانی‌سازی در عمق ادامه دارد مطالعات ژئوفیزیکی الکتریکی انجام گیرد تا مناطق دقیق کانی‌سازی و عمق شروع کانی‌سازی مشخص گردد.
۵. در محدوده مورد مطالعه تعداد گمانه محدودی حفر گردیده است و اغلب این گمانه‌ها با کانی-سازی برخورد نکرده‌اند، پیشنهاد می‌گردد در قسمت شمالی محدوده مورد مطالعه که کانی‌سازی در عمق ادامه داشته و گمانه اکتشافی نیز در این قسمت حفر نشده است چندین گمانه اکتشافی حفاری گردد.

٧- مراجع

۱. حسنی پاک، علی اصغر؛ ۱۳۸۲؛ مدیریت خطا و ریسک در اکتشاف؛ انتشارات دانشگاه تهران؛ چاپ اول.
۲. حسنی پاک، علی اصغر؛ ۱۳۸۴؛ تحلیل داده‌های اکتشافی؛ انتشارات دانشگاه تهران.
۳. صنایع مس ایران، ۱۳۸۰؛ مطالعات زمین‌شناسی محدوده اکتشافی سوناجیل.
۴. طهماسبی، پژمان؛ هزارخانی، اردشیر؛ ۱۳۹۰؛ استفاده از شبکه‌های عصبی-فازی-ژنتیکی به منظور تخمین عیار در کانسارمس پرفیری دره زار-کرمان؛ نشریه علمی-پژوهشی "مهندسی معدن" دوره ششم، شماره دوازدهم.
۵. قنبری، آرش؛ هداوندی، اسماعیل؛ عباسیان نقتنه، سلمان؛ ۱۳۸۹؛ ترکیب تکنیک خوشه بندی داده ها و سیستم های فازی تکاملی برای مدل‌سازی مسائل پیش بینی، هفتمین کنفرانس بین المللی مهندسی صنایع، اصفهان؛ انجمن مهندسی صنایع ایران، دانشگاه صنعتی اصفهان.
۶. قوامی ریایی، سید رضا؛ دارای گلستان، فرشاد؛ ۱۳۹۳؛ تحلیل و مدل سازی اکتشافی داده های ژئوشیمیایی ذخایر معدنی؛ انتشارات دانشگاه صنعتی شاهرود؛ چاپ اول.
۷. کیا، مصطفی، ۱۳۸۷، شبکه‌های عصبی در MATLAB، چاپ دوم، انتشارات کیان رایانه سبز.
۸. گرانیان، حمید؛ طباطبایی، سید حسن؛ اسدی هارونی، هوشنگ؛ ۱۳۹۲؛ کاربرد روش‌های طبقه‌بندی بر اساس نظریه بی‌زین در پتانسیل یابی طلا به کمک داده‌های ژئوشیمیایی خاک در کانسارم طلای اپی ترمال ساری گونئی؛ نشریه علمی-پژوهشی "ژئوشیمی" سال اول، شماره چهارم.
۹. منہاج، محمدباقر، ۱۳۸۷، مبانی شبکه های عصبی (هوش محاسباتی)، جلد اول، چاپ پنجم، مرکز نشر دانشگاه صنعتی امیرکبیر، تهران.

10. Abedi, M., Norouzi, G.H. and Bahroudi, A., 2012. Support vector machine for multi-classification of mineral prospectivity areas. *Computers & Geosciences*, 46, pp.272-283.
11. Abedi, M., Norouzi, Gh.H., Fathianpour, N. 2013. Fuzzy outranking approach: A knowledge-driven method for mineral prospectivity mapping, *International Journal of Applied Earth Observation and Geoinformation*, Volume 21, Pages 556-567, ISSN 0303-2434.
12. Abedi, M., Torabi, S.A., Norouzi, G.H., Hamzeh, M. 2012. ELECTRE III: A knowledge-driven method for integration of geophysical data with geological and geochemical data in mineral prospectivity mapping, *Journal of Applied Geophysics*, Volume 87, Pages 9-18, ISSN 0926-9851.
13. Al-Anazi, A. and Gates, I.D., 2010. On the capability of support vector machines to classify lithology from well logs. *Natural resources research*, 19(2), pp.125-139.
14. Allahkarami, E., Nuri, O. S., Abdollahzadeh, A., Rezai, B., and Maghsoudi, B. 2017. Improving estimation accuracy of metallurgical performance of industrial flotation process by using hybrid genetic algorithm-artificial neural network (ga-ann). *Physicochem. Probl. Miner. Process*, 53(1), 366-378.
15. Altrock, C.V., 1995. *Fuzzy logic & Neurofuzzy applications explained*, Prentice Hall, P.384.
16. Andrew, F., Anthony, W., AndrewWare, J.H. 2007. Two supervised neural networks for classification of sedimentary organic matter images from palynological preparations, *Math Geol*, 39: 657-671.
17. Asadi, H., Kianpouryan, S., Lu, L., McCuaig, T. 2014, Exploratory data analysis and C-A fractal model applied in mapping multi-element soil anomalies for drilling: A case study from the Sari Gunay epithermal gold deposit, NW Iran, *Journal of Geochemical Exploration*, Volume 145, Pages 233-241, ISSN 0375-6742.

18. Balasko, B., Abonyi, J. Feil, B. 2003. *Fuzzy Clustering and Data Analysis Toolbox*.
19. Ben-Jemaa, F. and Marino, M.A., 1990, April. Optimization of a groundwater well monitoring network. In *International Conference on Optimizing the Resources for Water Management, Forth worth, Texas, April* (pp. 17-21).
20. Bensaid, A.M., Hall, L.O., Bezdek, J.C., Clarke, L.P., Silbiger, M.L., Arrington, J.A., Murtagh, R.F. 1996. Validity-guided (Re) clustering with applications to image segmentation, *IEEE Transactions on Fuzzy Systems* 4, 112–123.
21. Beus, A.A and Grigorian, S.V. 1962. *Geochemical exploration methods for mineral deposits*, Trans. by R.T. Schneider, ed. by A.A. Levinson, Illinois: Applied publishing.
22. Bezdek, J.C. 1981. *Pattern recognition with fuzzy objective function algorithms*. Plenum Press.
23. Bishop, C.M. 2006. *Pattern recognition and machine learning*. Springer science. 738pp.
24. Breiman, L. 1996. Bagging predictors. *Mach. Learn.* 24, 123–140.
25. Breiman, L. 2001. Random forests. *Mach. Learn.* 45, 5–32.
26. Carranza, E.J.M. 2002. *Geologically-constrained mineral potential mapping*. Phd thesis, Delft University of technology, The Netherlands, 480 pp.
27. Carranza, E.J.M. 2009. Geochemical anomaly and mineral prospectivity mapping in GIS. *Elsevier Science*.
28. Carranza, E.J.M., Laborte, A.G. 2015a. Data-driven predictive mapping of gold prospectivity, Baguio district, Philippines: Application of Random Forests algorithm, *Ore Geology Reviews*, Volume 71, Pages 777-787, ISSN 0169-1368.
29. Carranza, E.J.M., Laborte, A.G. 2015b. Random forest predictive modeling of mineral prospectivity with small number of prospects and data with missing values in Abra (Philippines), *Computers & Geosciences*, Volume 74, Pages 60-70, ISSN 0098-3004.
30. Carranza, E.J.M., Sadeghi, M. 2010. Predictive mapping of prospectivity and quantitative estimation of undiscovered VMS deposits in Skellefte district (Sweden), *Ore Geology Reviews*, Volume 38, Issue 3, Pages 219-241, ISSN 0169-1368.
31. Carranza, E.J.M., Sadeghi, M., Billay, A. 2015. Predictive mapping of prospectivity for orogenic gold, Giyani greenstone belt (South Africa), *Ore Geology Reviews*, Volume 71, Pages 703-718, ISSN 0169-1368.
32. Chen, Y. 2015. Mineral potential mapping with a restricted Boltzmann machine, *Ore Geology Reviews*, Volume 71, Pages 749-760, ISSN 0169-1368.
33. Cheng, Q. 2004. Application of Weights of Evidence Method for Assessment of Flowing Wells in the Greater Toronto Area, Canada. *Natural Resources Research* Volume 13, Issue 2 , pp 77-86.
34. Cheng, Q. 2008. Special Issue of Mathematical Geosciences for the 33rd IGC, *Journal of China University of Geosciences*, Volume 19, Issue 4, Pages 307-308, ISSN 1002-0705.
35. Cheng, Q., Agterberg, F.P. and Ballantyne, S.B., 1994. The separation of geochemical anomalies from background by fractal methods. *Journal of Geochemical Exploration*, 51(2), pp.109-130.
36. Cheng, Q., et al., 1994, The separation of geochemical anomalies from background by fractal methods, *Journal of Geochemical Exploration*, vol. 5, p. 109–130.
37. Czarnota, K., Blewett, R.S., Goscombe, B. 2010. Predictive mineral discovery in the eastern Yilgarn Craton, Western Australia: An example of district scale targeting of an orogenic gold mineral system, *Precambrian Research*, Volume 183, Issue 2, Pages 356-377, ISSN 0301-9268.
38. Daya, A.A., 2015. Comparative study of C–A, C–P, and N–S fractal methods for separating geochemical anomalies from background: A case study of Kamoshgaran region, northwest of Iran. *Journal of Geochemical Exploration*, 150, pp.52-63.
39. Dietterich, T. 2000. An experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting, and randomization. *Mach. Learn.* 40, 139–157.

40. Dominy, S.C., Platten, I.M., Raine, M.D. 2003. Grade and geological continuity in highnugget effect gold-quartz reefs: *implications for resource estimation and reporting Applied Earth Science*. Trans. Inst. Min. Metall. B 112, 239–259.
41. Dowd, P.A and Sarac, C. 1994. Aneural Network Approach to Geostatistical Simulation, *Mathematical Geology*, Vol.26.No.4, p.491-503
42. Duda, R. A., Hart, P. E., Stork, D.G. 2002. *Pattern Classification*, Springer.
43. Duda, R. A., Hart, P. E., Stork, D.G. 2012. *Pattern Classification*. John Wiley & Sons.
44. Dunn, J. 1974. A fuzzy relative of the ISODATA process and its use in detecting compact well separated cluster, *Journal of Cybernetics* 3, 32–57.
45. El-Khoribi, R.A. 2008. Support Vector Machine Training of HMT Models for Land Cove Image Classification. ICGST-GVIP1687-398X8, 7–11.
46. Freund, Y and Schapire, R. 1996. Experiments with a new boosting algorithm. *International Conference on Machine Learning*, pp. 148–156.
47. Friedman, N., Geiger, D., Goldszmidt, M. 1997. Bayesian network classifiers. *Mach. Learn.* 29, 131–163.
48. Galiano, V.R., Castillo, M.S., Olmo, M.Ch., Rivas, M.Ch. 2015. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines, *Ore Geology Reviews*, Volume 71, Pages 804-818, ISSN 0169-1368.
49. Geranian, H., Mokhtari, A.R. and Cohen, D.R., 2013. A comparison of fractal methods and probability plots in identifying and mapping soil metal contamination near an active mining area, Iran. *Science of the Total Environment*, 463, pp.845-854.
50. Geranian, H., Tabatabaei, S.H., Asadi, H.H. and Carranza, E.J.M., 2016. Application of discriminant analysis and support vector machine in mapping gold potential areas for further drilling in the Sari-Gunay Gold Deposit, NW Iran. *Natural Resources Research*, 25(2), pp.145-159.
51. Geranian, H., Tabatabaei, S.H., Asadi, H.H. and Carranza, E.J.M., 2015. Multivariate regression analysis of lithogeochemical data to model subsurface mineralization: a case study from the Sari Gunay epithermal gold deposit, NW Iran. *Journal of Geochemical Exploration*, 148, pp.249-258.
52. Geranian, H., Tabatabaei, S.H., Asadi, H.H., Carranza, E.J.M. 2014. Multivariate regression analysis of lithogeochemical data to model subsurface mineralization: a case study from the Sari Gunay epithermal gold deposit, NW Iran, *Journal of Geochemical Exploration*, Volume 148, Pages 249-258.
53. Ghaffari, A., Abdollahi, H., Khoshayand, M, R., Soltani, I., Dadgar, A., Tehrani, M., 2006. Performance comparison of neural network training algorithms in modeling of bimodal drug delivery, *International Journal of Pharmaceutics* (327) pp.126–138.
54. Ghavami-Riabi, R., Seyedrahimi-Niarraq, M.M., Khalokakaie, R., Hazareh, M.R. 2010. U-spatial statistic data modeled on a probability diagram for investigation of mineralization phases and exploration of shear zone gold deposits, *Journal of Geochemical Exploration*, Volume 104, Issues 1–2, Pages 27-33, ISSN 0375-6742,
55. Goovaerts, P., 1997. *Geostatistics for natural resources evaluation*. Oxford University Press on Demand.
56. Guo, W. W. 2010. A novel application of neural networks for instant iron-ore grade estimation. *Expert Systems with Applications*, 37(12), 8729-8735.
57. Gustafson, E and Kessel, W. 1978. Fuzzy clustering with a fuzzy covariance matrix, in: *IEEE Conference on Decision and Control*, pp. 761–766.
58. Han, R.S., Chen, J., Wang, F., Wang, X.K. and Li, Y., 2015. Analysis of metal–element association halos within fault zones for the exploration of concealed ore-bodies—A case study of the Qilinchang Zn–Pb–(Ag–Ge) deposit in the Huize mine district, northeastern Yunnan, China. *Journal of Geochemical Exploration*, 159, pp.62-78.

59. Harraz, H.Z. and Hamdy, M.M., 2015. Zonation of primary haloes of Atud auriferous quartz vein deposit, Central Eastern Desert of Egypt: A potential exploration model targeting for hidden mesothermal gold deposits. *Journal of African Earth Sciences*, 101, pp.1-18.
60. Hechenbichler, K and Schliep, K.P. 2004. Weighted k-nearest-neighbor techniques and ordinal classification. Discussion Paper 399, SFB 386. *Ludwig-Maximilians University Munich*.
61. Hezarkhani, A. 2003. *Exploration of Sonajil copper deposit, Iranian company of copper*, northwestern report exploration. Retrieved from.
62. Hill, E. J., Oliver, N.H.S., Fisher, L., Cleverley, j. S., Nugus, M.j. 2014. Using geochemical proxies to model nuggety gold deposits: An example from Sunrise Dam, Western Australia, *Journal of Geochemical Exploration*, Volume 145, Pages 12-24, ISSN 0375-6742.
63. Hoppner, F., Klawonn, F., Kruse, R and Runkler, T. 1999. *Fuzzy Cluster Analysis*. Wiley, Chichester.
64. Hornik, K., Stinchcombe, M., White, H. 1989. Multilayer feedforward networks are universal approximators. *Neural Network* 2 (5), 359–366.
65. Huang, C., Davis, L.S., Townshend, J.R.G. 2002. An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing* 23, 725–749.
66. Jafrasteh, B., and Fathianpour, N. 2017. A hybrid simultaneous perturbation artificial bee colony and back-propagation algorithm for training a local linear radial basis neural network on ore grade estimation. *Neurocomputing*. doi:http://dx.doi.org/10.1016/j.neucom.2017.01.016
67. Jalloh, A. B., Kyuro, S., Jalloh, Y. & Barrie, A. K. (2016). *International Journal of Mining Science and Technology* 26, 581-585.
68. Journel, A.G and Huijbregts, C. 1978. *Mining Geostatistics*. Academic Press, London 600 pp.
69. Jozanikohan, G., Norouzi, G. H., Sahabi, F., Memarian, H. & Moshiri, B. 2015. *Journal of Natural Gas Science and Engineering* 22, 119-131.
70. Kashani, M.H., Ghorbani, M.A., Dinpashoh, Y. and Shahmorad, S., 2014. Comparison of volterra model and artificial neural networks for rainfall–runoff simulation. *Natural resources research*, 23(3), pp.341-354.
71. Kavzoglu, T and Colkesen, I. 2009. A kernel functions analysis for support vector machines for land cover classification. *International Journal of Applied Earth Observation and Geoinformation* 11, 352–359.
72. Keller, J.M., Gray, M.R and Givens, J.R. 1985. A Fuzzy k-nearest neighbor algorithm, *IEEE Transaction on Systems*, vol. 15, pp. 580-585.
73. Koike, K. and Matsuda, S., 2003. Characterizing content distributions of impurities in a limestone mine using a feedforward neural network. *Natural resources research*, 12(3), pp.209-222.
74. Lacassie, J.P., Roser, B.P., Ruiz-del-Solar, J., Herve, F. 2004. Visualization of geochemical datasets by using neural networks: a novel perspective for sedimentary provenance analysis: *Sedimentary Geol.*, 165, 1: 175–191.
75. Lacassie, J.P., Solar, J.R., Roser, B., Herve, F. 2006. Visualization of Volcanic Rock Geochemical Data and Classification with Artificial Neural Networks, *Mathematical Geology*, 38, 6.
76. Lesot, M.J and Kruse, R. 2008. Gustafson-Kessel-like clustering algorithm based on typicality degrees, Book title: Uncertainty and Intelligent Information Systems, Publisher: *World scientific Publishing Company*.117-130.
77. Li, C., Ma, T., Shi, J., 2003. Application of a fractal method relating concentrations and distances for separation of geochemical anomalies from background. *Journal of Geochemical Exploration* 77, 167–175.
78. Lisitsin, V. 2015. Spatial data analysis of mineral deposit point patterns: Applications to exploration targeting, *Ore Geology Reviews*, Volume 71, Pages 861-881, ISSN 0169-1368.
79. Lowell, J.D. and Guilbert, J.M., 1970. Lateral and vertical alteration-mineralization zoning in porphyry ore deposits. *Economic Geology*, 65(4), pp.373-408.

80. Mahmoudabadi, H., Izadi, M., Menhaj, M.B. 2009. A hybrid method for grade estimation using genetic algorithm and neural networks. *Computational Geosciences* 13, 91–101.
81. Mandelbrot, B.B. and Pignoni, R., 1983. *The fractal geometry of nature* (Vol. 173). New York: WH freeman.
82. Mars, J.C. and Rowan, L.C., Geosphere (2006) Illustrated deposit model of a porphyry copper deposit (modified from Lowell and Guilbert, 1970). *Geol Soc Am*, 2, pp.161-186.
83. Matheron, G. 1962. *Discusses applied geostatistics*, Volume I: Memories of the Office of Geological and Mining Research, no. 14, Editions Technip, Paris, 333 p.
84. McClenaghan, M.B., Thorleifson, L.H., DiLabio, R.N.W. 2000. Till geochemical and indicator mineral methods in mineral exploration, *Ore Geology Reviews* 16, 145–166
85. Mohammadi Gonbadi, A., Tabatabaei, S.H., Carranza, E. J. M. 2015. Supervised geochemical anomaly detection by pattern recognition, *Journal of Geochemical Exploration*, Volume 157, Pages 81-91, ISSN 0375-6742.
86. Najafi, A., Karimpour, M.H., Ghaderi, M. 2014. Application of fuzzy AHP method to IOCG prospectivity mapping: A case study in Taherabad prospecting area, eastern Iran, *International Journal of Applied Earth Observation and Geoinformation*, Volume 33, Pages 142-154, ISSN 0303-2434.
87. Nakhaei, F., and Irannajad, M. 2013. Comparison between neural networks and multiple regression methods in metallurgical performance modeling of flotation column. *Physicochemical Problems of Mineral Processing*, 49(1), 255--266.
88. Nejadi, S. and Leung, J.Y., 2015. Estimation of facies boundaries using categorical indicators with P-Field simulation and ensemble Kalman filter (EnKF). *Natural Resources Research*, 24(2), pp.121-138.
89. Padilla, G. R.A., Titley, S.R., Francisco, P. B. 2001. *Geology of the Escondida porphyry*.
90. Pazand, K., Hezarkhani, A., Ataei, M. and Ghanbari, Y., 2011. Application of multifractal modeling technique in systematic geochemical stream sediment survey to identify copper anomalies: a case study from Ahar, Azarbaijan, Northwest Iran. *Chemie der Erde-Geochemistry*, 71(4), pp.397-402.
91. Porwal, A., Carranza, E.J.M. and Hale, M., 2003. Artificial neural networks for mineral-potential mapping: a case study from Aravalli Province, Western India. *Natural resources research*, 12(3), pp.155-171.
92. Porwal, A., Carranza, E.J.M., Hale, M. 2006. Bayesian network classifiers for mineral potential mapping, *Computers & Geosciences*, Volume 32, Issue 1, Pages 1-16, ISSN 0098-3004.
93. Purwar, S., Jablonowski, C.J. and Nguyen, Q.P., 2011. Development optimization using reservoir response surfaces: Methods for integrating facility and operational options. *Natural resources research*, 20(1), pp.1-9.
94. Rendu, J.M., 1979, October. Kriging, logarithmic Kriging, and conditional expectation: comparison of theory with actual results. In *Proceedings, 16th APCOM Symposium* (pp. 199-212).
95. Richards, J.P., Boyce, A.J. and Pringle, M.S., 2001. Geologic evolution of the Escondida area, northern Chile: a model for spatial and temporal localization of porphyry Cu mineralization. *Economic Geology*, 96(2), pp.271-305.
96. Rizzo, D.M. and Dougherty, D.E., 1994. Characterization of aquifer properties using artificial neural networks: Neural kriging. *Water Resources Research*, 30(2), pp.483-497.
97. Rooki, R., Doulati Ardejani, F., Aryafar, A. & Bani Asadi, A. 2011. *Environmental Earth Sciences* 64, 1303-1316.
98. Roshani, P., Mokhtari, A.R., Tabatabaei, S.H. 2013. Objective based geochemical anomaly detection—Application of discriminant function analysis in anomaly delineation in the Kuh Panj porphyry Cu mineralization (Iran), *Journal of Geochemical Exploration*, Volume 130, Pages 65-73, ISSN 0375-6742.
99. Samanta, B., Bandopadhyay, S., Ganguli, R. 2002. Data segmentation and genetic algorithms for sparse data division in Nome placer gold grade estimation using neural network and geostatistics. *Mining Exploration Geology* 11 (1–4), 69–76.

- 100.Samanta, B., Bandopadhyay, S., Ganguli, R., and Dutta, S. 2005. A comparative study of the performance of single neural network vs. Adaboost algorithm based combination of multiple neural networks for mineral resource estimation. *Journal of South African institute of mining and metallurgy*, 105(4), 237-246.
- 101.Schölkopf, B., Smola, A.J., Williamson, R.C. and Bartlett, P.L., 2000. New support vector algorithms. *Neural computation*, 12(5), pp.1207-1245.
- 102.Serir, L., Ramasso, E., Zerhouni, N. 2012. Evidential evolving Gustafson–Kessel algorithm for online data streams partitioning using belief function theory, *International Journal of Approximate Reasoning*, Volume 53, Issue 5, Pages 747-768, ISSN 0888-613X.
- 103.Shahi, H., Ghavami, R. and Rouhani, A.K., 2016. Detection of deep and blind mineral deposits using new proposed frequency coefficients method in frequency domain of geochemical data. *Journal of Geochemical Exploration*, 162, pp.29-39.
- 104.Sim, D.G., Kwon, O.K. and Park, R.H., 1999. Object matching algorithms using robust Hausdorff distance measures. *IEEE Transactions on image processing*, 8(3), pp.425-429.
- 105.Singer, D.A. and Kouda, R., 1997. Classification of mineral deposits into types using mineralogy with a probabilistic neural network. *Natural Resources Research*, 6(1), pp.27-32.
- 106.Singer, D. 2006. Typing mineral deposits using their associated rocks and grades and tonnages in a probabilistic neural network. *Math Geol* 38(4):465–475.
- 107.Smirnoff, A., Boisvert, E., Paradis, S.J. 2008. Support vector machine for 3D modeling from sparse geological information of various origins. *Computers & Geosciences* 34, 127–143.
- 108.Solomatine, D.P. and Ostfeld, A., 2008. Data-driven modelling: some past experiences and new approaches. *Journal of hydroinformatics*, 10(1), pp.3-22.
- 109.Strebelle, S. 2002. Conditional simulation of complex geological structures using multiple-point geostatistics. *Mathematical Geology* 34 (1), 1–22.
- 110.Tahmasebi, P and Hezarkhani, A. 2012. A hybrid neural networks-fuzzy logic-genetic algorithm for grade estimation. *Computers and Geosciences*, Volume 42, May, Pages 18-27.
- 111.Tahmasebi, P., and Hezarkhani, A. 2010a. 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.
- 112.Tahmasebi, P., and Hezarkhani, A. 2010b. Comparison of optimized neural network with fuzzy logic for ore grade estimation. *Australian Journal of Basic and Applied Sciences*, 4(5), 764-772.
- 113.Tahmasebi, P., Hezarkhani, A. and Sahimi, M., 2012. Multiple-point geostatistical modeling based on the cross-correlation functions. *Computational Geosciences*, 16(3), pp.779-797.
- 114.Theodoridis, S and Kourtombas, K. 2009, "Pattern Recognition", fourth edition, Academic Press, P.967.
- 115.Theodoridis, S and Kourtombas, K. 2010. *An introduction of pattern recognition*, Academic Press.
- 116.Titley, S.R and Beane, R.E. 1981. Porphyry copper deposits. Part 1. Geologic setting, petrology, and tectogenesis: Part2. Hydrothermal alteration and mineralization. : *Economic Geology*, 75 th anniversary volume, p. 214-269.
- 117.Turcotte, D.L., 1986. A fractal approach to the relationship between ore grade and tonnage. *Economic Geology* 18, 1525–1532.
- 118.Vapnik, V., 1963. Pattern recognition using generalized portrait method. *Automation and remote control*, 24, pp.774-780.
- 119.Vapnik, V.N. 1995. *The Nature of Statistical Learning Theory*. Springer-Verlag, New York.
- 120.Verbiest, N., Cornelis, C., Jensen, R. 2012. Fuzzy rough Positive region based Nearest Neighbour Classificaton, WCCI IEEE World congress on Computational Intelligence.
- 121.Wang, H. and Zuo, R., 2015. A comparative study of trend surface analysis and spectrum–area multifractal model to identify geochemical anomalies. *Journal of Geochemical Exploration*, 155, pp.84-90.

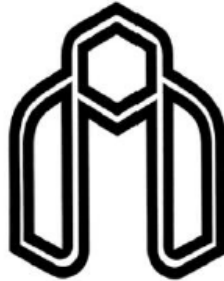
122. Wang, Zh., Cheng, Q., Xu, D., Dong, Y. 2008. Fractal Modeling of Sphalerite Banding in Jinding Pb-Zn Deposit, Yunnan, Southwestern China, *Journal of China University of Geosciences*, Volume 19, Issue 1, Pages 77-84, ISSN 1002-0705.
123. Watermeyer, G. A. 1919. Application of the theory of probability in the determination of ore reserves, *Journal of the Chemical Metallurgical and Mining Society of South Africa*, 19, 97–107.
124. Webb, A.R. 2002. *Statistical Pattern Recognition*, Second edition JohnWiley & Sons, Ltd., England, pp. 169–202.
125. Webb, A.R., 2003. *Statistical pattern recognition*. John Wiley & Sons.
126. Wei, S. and Pengda, Z., 2002. Theoretical study of statistical fractal model with applications to mineral resource prediction. *Computers & Geosciences*, 28(3), pp.369-376.
127. Weller, A.F., Corcoran, J., Harris, A.J. and Ware, J.A., 2005. The semi-automated classification of sedimentary organic matter in palynological preparations. *Computers & geosciences*, 31(10), pp.1213-1223.
128. Wu, X and Zhou, Y. 1993. Reserve estimation using neural network techniques: *Computers & geosciences.*, 19, 4:567–575.
129. Xiao-li, L., Yu-ling, X., Li-hong, L., and Qin-jin, G. 2009. A nonlinear grade estimation method based on Wavelet Neural Network. Paper presented at the Bio-Inspired Computing, 2009. BIC-TA'09. *Fourth International Conference on*.
130. Xie, X.L and Beni, G. 1991. A validity measure for fuzzy clustering, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 841–847.
131. Yang, Q., Li, X., Shi, X. 2008. Cellular automata for simulating land use changes based on support vector machines. *Computers & Geosciences* 34, 592–602.
132. Zarasvandi, A. 2004. *Magmatic and structural controls on localization of the Darreh-Zerreshk and Ali-Abad porphyry copper deposits*, Yazd Province, Central Iran, PhD thesis, Shiraz University, Shiraz, Iran, 280p
133. Zarasvandi, A., Liaghat, S., Zentilli, M. 2005. Geology of the Darreh-Zerreshk and Ali-Abad porphyry copper deposit, central Iran, *International Geology Reviews*, v. 47, no. 6, p. 620-646.
134. Zuo, R., Carranza, E.J.M. 2011. Support vector machine: A tool for mapping mineral prospectivity, *Computers & Geosciences*, Volume 37, Issue 12, Pages 1967-1975, ISSN 0098-3004.
135. Zuo, R., Carranza, E.J.M. and Wang, J., 2016. Spatial analysis and visualization of exploration geochemical data. *Earth-Science Reviews*, 158, pp.9-18.
136. Zuo, R., Xia, Q. and Zhang, D., 2013. A comparison study of the C–A and S–A models with singularity analysis to identify geochemical anomalies in covered areas. *Applied Geochemistry*, 33, pp.165-172.
137. Selim, S.Z. and Ismail, M.A., 1986. On the local optimality of the fuzzy isodata clustering algorithm. *IEEE transactions on pattern analysis and machine intelligence*, (2), pp.284-288.

Abstract

Intelligent methods and pattern recognition are the most common methods that have been used for data analysis in science and engineering fields. These methods have been used for anomaly of background separation in exploration of geochemistry. In the current research, clustering and classification algorithms were used together as a pattern recognition method using surface and depth information to estimate concentration of the elements in the missing points and different depth horizons. The data were clustered and the concentrations of the elements were estimated using the neural network and clustering methods. For this purpose, the number of optimum clusters was identified based on validation indices. The results showed that the combination of FCM and clustering algorithms increased the accuracy of the estimation from 75 up to 88 percents.

The geochemical maps at different depth levels were drawn. Front, tail and near ore haloes were identified. The behavior of elements showed that the arsenic as front halo and vanadium as tail halo reacted. Based on the classification of copper concentration changes by fractal-grade-area method and determining copper classes with classification algorithms, the model of the deposit was depicted in depth. The maps showed that the ore deposit was eroded in the southwestern part of the area, and there was another mineralized zone in depth at the north-central part of the region. Among the different used classification methods, the highest accuracy was related to the neural network algorithm with precision classification of 90.6%. The geophysical maps and sections together with drilling data were used for validation of the results. These data confirmed the geochemical results.

Keywords: Pattern recognition algorithms, accuracy of estimation, geochemical deep sections, deep modeling of deposit.



Shahrood University of Technology
Faculty of Mining, Petroleum and Geophysics
(Exploration Department)

PhD Dissertation in mining exploration

**Evaluation of deep geochemical dispersion map concentration
based on pattern recognition techniques to improve the grade
estimate and deposit model in depth**

By: Moharram Jahangiri

Supervisor

Dr. Seyed Reza Ghavami Riabi

Advisor

Dr. Behzad TokhmChi

January 2017