

تخمین عیار طلا در کانسار خونی با استفاده از بررسی رفتار عناصر طلا، آرسنیک و آنتیموان به روش خوشه‌بندی K-means

ندا ماهوش محمدی^۱، اردشیر هزارخانی^۲

۱- دانشجوی دکتری مهندسی معدن، دانشکده مهندسی معدن و متالورژی، دانشگاه صنعتی امیرکبیر

۲- استاد، دانشکده مهندسی معدن و متالورژی، دانشگاه صنعتی امیرکبیر

(دریافت: مرداد ۱۳۹۴، پذیرش: اسفند ۱۳۹۴)

چکیده

محدوده اکتشافی خونی در ۲۷۰ کیلومتری استان اصفهان در منطقه انارک و در زون ایران مرکزی واقع شده است. با توجه به وجود شواهدی حاکی از کانی‌سازی طلا در این ناحیه، شناسایی نواحی امید بخش معدنی در این منطقه ضروری است. بنابراین ضرورت یافتن اطلاعاتی در مورد ارتباط و چگونگی رفتار عناصر طلا، آرسنیک و آنتیموان نسبت به یکدیگر در این محدوده، برای پیدایش و تعیین وسعت هاله‌های ژئوشیمیایی، تعیین امیدبخش بودن منطقه و تخمین عیار اهمیت می‌یابد و باید مورد بررسی قرار گیرند. بدین منظور در تحقیق حاضر برای تعیین رفتار عناصر نام برده، با استفاده از روش معروف و کارآمد *K-means* انجام می‌شود که یکی از روش‌های خوشه‌بندی بوده و مبتنی بر کمینه نمودن مجموع فواصل اقلیدسی هر یک از نمونه‌ها از مرکز دسته‌ای که به آن تخصیص می‌یابد، است. در این پژوهش از تابع کیفیت خوشه‌بندی و میزان مطلوبیت نمونه در خوشه مورد نظر ($S(i)$)، برای تعیین تعداد خوشه بهینه استفاده شده است و در پایان با توجه به مراکز خوشه‌ها و نتایج حاصل، معادله‌ای جهت تخمین عیار عنصر طلا بر حسب چهار پارامتر عیار آرسنیک، عیار آنتیموان، طول و عرض نقاط نمونه برداری ارائه می‌شود.

واژگان کلیدی

طلا، کانسار خونی، روش خوشه بندی، تخمین عیار، *K-means*

ارجاع به این مقاله:

ماهوش محمدی، ن.، هزارخانی، الف.، (۱۳۹۴)، تخمین عیار طلا در کانسار خونی با استفاده از بررسی رفتار عناصر طلا، آرسنیک و آنتیموان به روش خوشه‌بندی *K-means*، روش‌های تحلیلی و عددی در مهندسی معدن، ۵(۱۰)، ۷۷-۹۲.

۱- مقدمه

کمینه شود [۶]. مواردی چند از کاربرد روش *K-means* که به منظور تحلیلی مناسب از رفتار داده‌ها و آنالیزهای موجود نسبت به یکدیگر داشته است که می‌توان به تقسیم بندی عوارض زمین‌شناسی [۷]، تقسیم بندی تأثیر پوشش گیاهی و بازیابی سلامتی آب در جنگل‌های ساحل مدیترانه [۸]، ارائه برنامه شناسایی الگوهای ژئوشیمیایی در مناطق معدنی [۹]، پیش‌گویی در مورد ارگانیک‌های کربن در سیستم‌های هوشمند [۱۰] و تعیین اثر پخش شونده‌گی گاز در محیط شهری اشاره نمود [۱۱].

در این مقاله با استفاده از روش *K-means* و نرم افزارهای *MATLAB* و *SPSS* رفتار عناصر طلا، آرسنیک و آنتیموان بر اساس داده‌های رسوبات آبراهه‌ای جمع‌آوری شده مورد ارزیابی قرار گرفته و سپس به کمک نتایج حاصل، عیار طلا تخمین زده شده است.

۲- منطقه مورد مطالعه

منطقه مورد مطالعه در فاصله ۲۷۰ کیلومتری شمال شرق اصفهان، ۷۶ کیلومتری شمال شرقی نایین، ۶۰ کیلومتری شمال شرق انارک و در فاصله ۴۰ کیلومتری شرق معدن سرب و روی نخلک قرار گرفته است. این منطقه در حاشیه جنوبی دشت کویر (کویر مرکزی)، بین طول‌های جغرافیایی ۵۵-۵۴ درجه شرقی و عرض‌های جغرافیایی ۳۳-۳۴ درجه شمالی واقع شده است. معدن خونی و کالکافی در گذشته فعال بوده و یکی از منابع مهم تأمین طلای کشور محسوب می‌شده است. معدن متروکه خونی در حاشیه شمالی توده کالکافی قرار دارد. تکتونیک پویای منطقه و گسل‌ها در جای‌گیری رگه‌ها و دگرسانی‌ها نقش اساسی ایجاد نموده‌اند. پیدایش کانسارها در انارک از یک سو متأثر از ماگماتیسزم ائوسن و از سوی دیگر محصول عملکرد تکتونیک و فازهای متعدد دگرگونی است [۱۲]. بر اساس مطالعات نظام‌پور (۱۳۸۴)، کانه‌زایی‌های منطقه‌ی خونی به ۶ بخش قابل تقسیم هستند. در میان این کانه‌زایی‌ها، سه بخش از نوع اکسیدی و مابقی سولفیدی هستند. در دامنه شمالی کوه خونی، زبانه‌هایی از سنگ‌های آذرین رخمون دارند که توانسته‌اند سنگ‌های مجاور خود را دگرگون نموده و آنها را به اسکارن تبدیل نماید [۱۳].

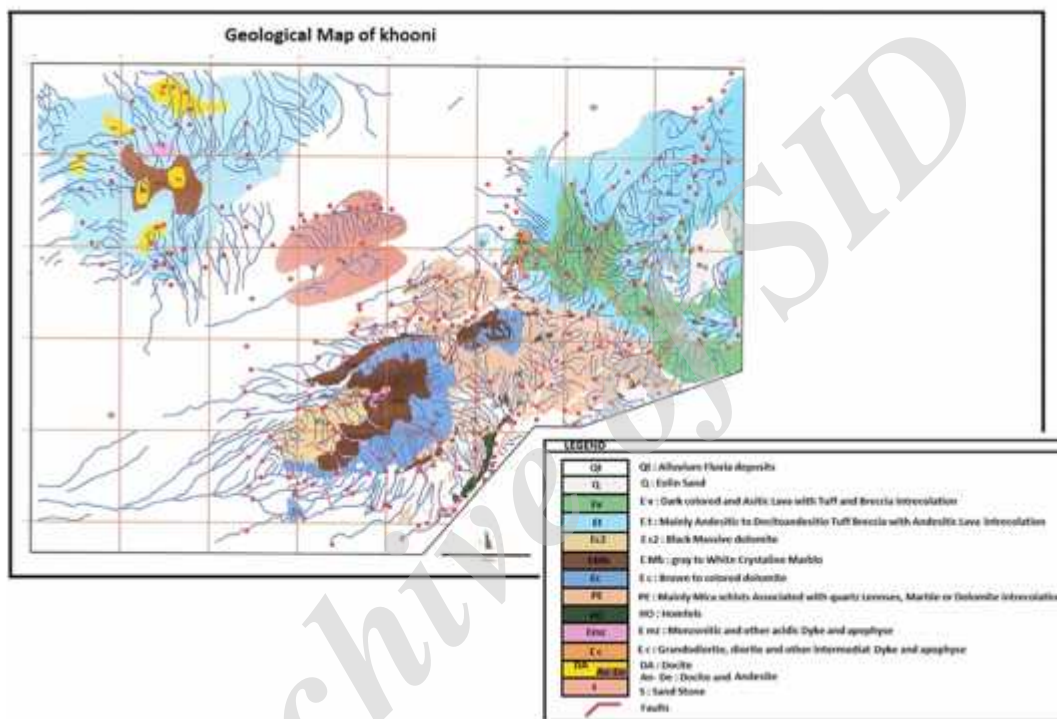
در سال‌های اخیر به علت وابستگی زیاد پروژه‌های معدنی به تعیین دقیق‌تر تناژ ماده‌ی معدنی، روش‌های گوناگونی برای تخمین عیار به وجود آمده‌اند که از آن جمله می‌توان به روش‌های هندسی، مبتنی بر فاصله و زمین آماری اشاره کرد. اما هر یک از روش‌ها دارای محدودیت‌ها و معایبی هم هستند که بر روی دقت تخمین تأثیر می‌گذارند [۱]. یکی از روش‌های نوین تخمین عیار به کمک روش خوشه‌بندی است. روش‌های تحلیل خوشه‌ای کاربرد وسیعی در علوم زمین دارند. گروه‌بندی خوشه‌ای روشی است که برای دسته‌بندی داده‌های ژئوشیمیایی به کار می‌رود. تحلیل خوشه‌ای، مشاهداتی که با هم شباهت زیاد دارند را به یکدیگر مرتبط می‌کند، سپس به طور متوالی مشاهداتی را که بیشترین شباهت نسبت به مشاهدات قبلی را دارند به آنها متصل می‌کند [۲]. به عبارتی در خوشه‌بندی سعی می‌شود تا داده‌ها به خوشه‌هایی تقسیم شوند که شباهت بین داده‌های درون هر خوشه حداکثر و شباهت بین داده‌های درون خوشه‌های متفاوت حداقل شود [۳]. در روش خوشه‌بندی هیچ دسته‌ای از قبل وجود ندارد و در واقع متغیرها به صورت مستقل و وابسته تقسیم نمی‌شوند، بلکه در این جا جستجو برای دستیابی به گروه‌هایی از داده‌ها انجام می‌شود که به هم شباهت دارند و با کشف این شباهت‌ها می‌توان رفتارها را بهتر شناسایی کرد و بر مبنای آنها، طوری عمل کرد که نتیجه بهتری حاصل شود [۴]. روش خوشه بندی یک روش غیر مستقیم است، بدین معنی که این روش را می‌توان حتی هنگامی که هیچ نوع اطلاعات قبلی از ساختار داخلی پایگاه داده‌ها وجود ندارد، استفاده نمود. از این روش می‌توان برای کشف الگوهای پنهان و بهبود عملکرد روش‌های مستقیم نیز استفاده نمود [۵].

روش *K-means* یکی از روش‌های خوشه‌بندی داده‌ها در داده‌کاوی است. این روش روشی انحصاری و مسطح محسوب می‌شود که به طور وسیعی توسط پژوهشگران مختلف مورد بررسی قرار گرفته است و در تلاش است که به خوشه‌بندی تعدادی نمونه با تعداد دسته مشخص (*K*) بپردازد، به نحوی که مجموع فواصل اقلیدسی هر یک از نمونه‌ها از مرکز دسته‌ای که به آن تخصیص یافته است،

۲-۱- زمین‌شناسی عمومی منطقه

کرتاسه که به طور دگرشیب بر روی واحدهای قدیمی‌تر قرار گرفته‌اند، رخنمون دارند. قسمت‌های کم ارتفاع و پست دشت نیز توسط تراس‌های آبرفتی قدیمی، رسوبات پهنه‌ای دشت‌ها و آبرفت‌های جوان و رودخانه‌ای پوشیده شده است. به طور کلی، سنگ‌شناسی محدوده مورد مطالعه از واحدهای مختلف لیتولوژیک شامل واحدهای شیست دگرگونه (متعلق به دگرگونی‌های انارک) تا رسوبات جوان کواترنری تشکیل شده است (شکل ۱) [۱۴، ۱۵].

چینه‌شناسی محدوده مورد بررسی، از پرکامبرین تا کواترنر را در برمی‌گیرد. رخنمون‌های سنگی بخش غربی، عمدتاً متشکل از واحدهای دگرگونه کامبرین و در بخش شرقی متشکل از واحدهای ولکانیک و پیروکلاستیک انوسن با ترکیب غالب حدواسط (آندزیت- تراکی آندزیت) است که توسط نفوذی‌ها و دایک‌های با ترکیب غالب مونزونیتی قطع شده‌اند. در منتهی الیه شمال غربی محدوده، آهک‌های



شکل ۱: نقشه زمین‌شناسی و موقعیت نمونه‌های آبراه‌های در محدوده خونی [۱۶]

اینکه اکثر داده‌های ژئوشیمیایی حالت بسته یا ترکیبی دارند، اگر تبدیل مناسب برای باز نمودن داده‌ها صورت نگیرد منجر به نتایج اشتباه می‌شود. بنابراین قبل از انجام هرگونه پردازش بر روی داده‌ها باید آنها را باز نمود. در این پژوهش از روش [۱۷] *Ilr* برای باز نمودن داده‌ها استفاده شده است.

۳- روش مورد مطالعه

از منطقه خونی تعداد ۲۵۶ نمونه رسوبات آبراه‌های جمع آوری شده و به روش *ICP-MS* تجزیه شدند که در این مقاله به منظور تخمین عیار طلا و بررسی رفتار عناصر آرسنیک، آنتیموان و طلا، تنها این سه عنصر، به دلیل پاراژنز بودن و همچنین با توجه به مقدار ضریب همبستگی‌شان، مورد بررسی و پردازش قرار گرفته‌اند. همانطور که در جدول ۱ مشاهده می‌شود مقادیر ضریب همبستگی این سه عنصر بالا بوده و نشان دهنده‌ی ارتباط نزدیک این سه عنصر با یکدیگر است. همچنین با توجه به

۳-۱- الگوریتم *K-means*

الگوریتم *K-means* با مقداری مشخص برای K (تعداد دسته‌ها) شروع کرده و سعی در تخمین موارد زیر دارند: بدست آوردن نقاطی به عنوان مراکز خوشه‌ها این نقاط در واقع همان میانگین نقاط متعلق به هر خوشه هستند.

۳. محاسبه تابع هدف ناشی از دسته‌بندی $\{C_1, C_2, \dots, C_k\}$ بر اساس رابطه ۲ که به محاسبه‌ی مجموع فواصل نمونه‌ها از مرکز دسته‌ها می‌پردازد.

$$f(C_1, C_2, \dots, C_k) = \sum_{j=1}^k \sum_{x \in C_j} |x - z_j|^2 \quad (2)$$

۴. کمینه نمودن تابع هدف رابطه ۲ و یافتن دسته‌بندی مناسب روی مجموعه M با تعداد دسته K [۲۰].

۴- نتایج و بحث

۴-۱- رفتارسنجی به کمک روش K -means

در مطالعات مختلفی نظیر رابطه‌ی دیوریت دگرسان شده با کانی مگنتیت در کمر بند آهنی شیلی [۲۱]، رابطه بین مس و مولیبدن کانسار مس پورفیری [۲۲] و رابطه بین عناصر گروه پلاتین کانسار مس پورفیری [۲۳] رفتار عناصر به روش‌های گوناگونی نسبت به یکدیگر سنجیده شده است. در مطالعه پیش رو با به کارگیری روش K -means برای خوشه‌بندی داده‌های رسوبات آبراه‌های منطقه خونی، با سه مقدار عیار عناصر طلا، آرسنیک و آنتیموان (با در نظر گرفتن مختصات نقاط نمونه‌برداری)، مقدار بهینه K در هر مورد محاسبه شده است زیرا که عناصر آرسنیک و آنتیموان از عناصر مهم در تعیین هاله‌های ژئوشیمیایی عنصر طلا به حساب می‌آیند.

در این پژوهش به منظور مشخص نمودن مقدار مناسب K برای تعیین تعداد خوشه‌ها، از دو معیار مناسب استفاده شده است. اولین معیار به کار گرفته شده معیار $S(i)$ است، که بر اساس آن، تعداد خوشه‌ها از ۳ الی ۱۰ تغییر داده می‌شوند و سپس به کمک معیار مذکور، نتایج حاصل شده مورد تحلیل قرار می‌گیرند تا K بهینه انتخاب شود [۲۴]. معیار مناسب جهت تعیین K بهینه مطابق با رابطه ۳ تعیین شده است که بر اساس آن، دسته‌بندی‌های انجام شده مورد سنجش قرار می‌گیرد.

$$S(i) = \frac{\text{Min}(\text{AVEG} - \text{BETWEEN}(i, k)) - \text{AVEG} - \text{WITHIN}(i)}{\text{Max}[\text{AVEG} - \text{WITHIN}(i), \text{Min}(\text{AVEG} - \text{BETWEEN}(i, k))]} \quad (3)$$

پارامتر $\text{AVEG_BETWEEN}(i, k)$ بیان‌کننده میانگین فاصله نمونه i ام از سایر نمونه‌هایی است که در دسته دیگری مانند K قرار دارند [۲۰، ۲۴].

نسبت دادن هر نمونه داده به یک خوشه که آن داده کمترین فاصله تا مرکز آن خوشه را دارا باشد [۶].

جدول ۱: ماتریس ضریب همبستگی پیرسون، عناصر طلا، آرسنیک و آنتیموان

Variables	Sb	As	Au
Sb	۱		
As	۰/۸۱۹	۱	
Au	۰/۷۳۵	۰/۷۷۹	۱

در نوع ساده‌ای از این روش ابتدا به تعداد خوشه‌های مورد نیاز، نقاطی به صورت تصادفی انتخاب می‌شود. سپس داده‌ها، با توجه به میزان نزدیکی (شبهات) به یکی از این خوشه‌ها نسبت داده می‌شوند و بدین ترتیب خوشه‌های جدیدی حاصل می‌شود. با تکرار همین مراحل می‌توان در هر تکرار با میانگین‌گیری از داده‌ها مراکز جدیدی برای آنها محاسبه کرد و مجدداً داده‌ها را به خوشه‌های جدید نسبت داد [۱۸]. مراحل مهمی که در این الگوریتم طی می‌شود به طور خلاصه به شرح ذیل است [۲۰، ۱۹]:

۱. ابتدا k عضو (که k تعداد خوشه‌ها است) به صورت تصادفی از میان n عضو به عنوان تعداد خوشه‌ها انتخاب می‌شود.

۲. محاسبه بردار Z_j بر اساس رابطه ۱ که نشان دهنده مرکز یا نماینده هر دسته C_j است.

$$Z_j = \frac{\sum_{x \in C_j} x}{\#C_j} \quad \text{for } j = 1, 2, \dots, K \quad (1)$$

در این رابطه، x نشان‌دهنده بردار یک نمونه‌ی عضو C_j و $\#C_j$ نشان‌دهنده تعداد نمونه‌های عضو کلاس C_j است. لازم به ذکر است که رابطه ۱ برای محاسبه مرکز هر دسته در خلال حل استفاده می‌شود و در شروع الگوریتم معمولاً به صورت تصادفی K نمونه انتخاب شده و به عنوان مرکز هر دسته در نظر گرفته می‌شوند [۲۰].

در رابطه ۳، $S(i)$ بیان‌کننده میزان مطلوبیت نمونه i ام در کلاس مورد نظر خودش، پارامتر $\text{AVEG_WITHIN}(i)$ نشان‌دهنده میانگین فاصله نمونه i ام از سایر نمونه‌های دیگر در آن کلاسی است که خود نمونه i ام وجود دارد و

خوشه C^n و همه عناصر خوشه O^n است، m میانگین شباهت بین مرکز خوشه C^m و همه عناصر خوشه O^m است و در نهایت nm به عنوان شباهت C^n و O^n تعریف می‌شود [۲۵].

۴-۲- رفتار سنجی عناصر طلا و آرسنیک و آنتیموان نسبت به یکدیگر

به منظور بررسی رفتار عناصر مربوطه نسبت به یکدیگر در ابتدا نیمرخ خوشه‌ها و میزان مطلوبیت هر نمونه برای دسته‌بندی‌های $K=3$ تا $K=10$ برای عناصر طلا و آرسنیک و آنتیموان به صورت دو به دو انجام شده است و نتایج میزان مطلوبیت دسته‌ها با هم مقایسه شده است و بهترین دسته بر اساس میزان مطلوبیت دسته‌ها مشخص شده و سپس مراکز خوشه‌های هر دسته با توجه به آن مشخص می‌شود.

همان‌طور که در شکل ۲ مشاهده می‌شود، مطابق با نمودارهای نیمرخ کلاس و با توجه به مقادیر مطلوبیت بهترین کلاس برای دو عنصر طلا و آرسنیک، کلاس $K=4$ ، به عنوان بهترین دسته انتخاب می‌شود، زیرا هر چقدر میزان مطلوبیت به مقدار ۱ نزدیک باشد، نمونه‌ها به درستی در کلاس قرار گرفته‌اند. با توجه به نمودار، مقادیر منفی تقریباً کمی در این دسته‌بندی نیز دیده می‌شود. مقدار متوسط مطلوبیت در این دسته‌بندی برابر با 0.8064 است که نسبت به مقدار متوسط مطلوبیت در دسته‌های دیگر این مقدار بیش‌تر بوده و به همین روال و با توجه به شکل‌های ۳ و ۴ برای دیگر عناصر، بهترین دسته‌بندی با توجه به نیمرخ خوشه‌ها و مقدار متوسط مطلوبیت انتخاب می‌شود که برای دو عنصر طلا و آنتیموان کلاس $K=3$ با مقدار متوسط مطلوبیت 0.9647 و برای دو عنصر آرسنیک و آنتیموان کلاس $K=3$ با مقدار متوسط مطلوبیت 0.8584 به عنوان بهترین دسته‌بندی انتخاب می‌شود. برای اطمینان بیشتر و با توجه به نوسانات موجود در مقادیر مطلوبیت، مقدار K تا عدد ۲۰ افزایش داده شده است، اما مقدار مطلوبیت از مقدار مطلوبیت بهترین دسته‌بندی هر گروه فراتر نرفته و به ازای $K=20$ به بعد، روند کاهشی داشته است.

با محاسبه مقدار مطلوبیت تحت عنوان میانگین مطلوبیت، به تحلیل نتایج پرداخته می‌شود. مقدار مطلوبیت بین (-1) و $(+1)$ تغییر می‌کند به صورتی که هر چقدر این مقدار به $(+1)$ نزدیک‌تر باشد، نمونه مورد نظر در دسته‌بندی مناسب و هر چقدر به (-1) نزدیک باشد نمونه مورد نظر در دسته بندی نامناسب قرار گرفته است و عدد صفر به این معنا است که حضور نمونه مورد نظر در دسته فعلی و یا دسته دیگر دارای اهمیت چندانی نیست [۲۰، ۲۴].

دومین معیار استفاده شده، تابع کیفیت است. بر اساس دانسته‌ها، بهترین خوشه‌بندی به صورتی است که مجموع تشابه بین مرکز خوشه و همه اعضای خوشه را حداکثر و مجموع تشابه بین مراکز خوشه‌ها را حداقل بشود. برای انتخاب بهترین خوشه، ابتدا یک محدوده پیشنهادی برای تعداد خوشه‌ها مشخص می‌شود. که در این پژوهش محدوده بین ۳ تا ۱۰ کلاس پیشنهاد شده است. سپس مقدار (k) ... برای هر یک از مقادیر k محاسبه می‌شود. مقداری از k که در آن (k) ... حداکثر شود، به عنوان تعداد بهینه خوشه‌ها انتخاب می‌شود. به این ترتیب می‌توان تعداد خوشه‌ای را انتخاب نمود که به ازای آن فاصله بین مراکز خوشه‌ها و شباهت مراکز خوشه با اعضای درون هر خوشه حداکثر است. کیفیت نتایج خوشه بندی با k خوشه به صورت زیر تعریف می‌شود [۲۵]:

$$O = \{c^n \mid n = 1, \dots, k\} \quad (4)$$

$$O^n = \{C_i \mid i = 1, \dots, \|T^c - O\|\} \quad (5)$$

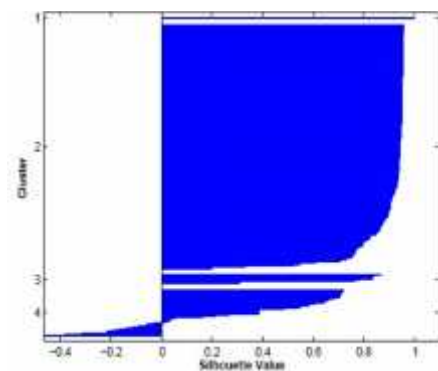
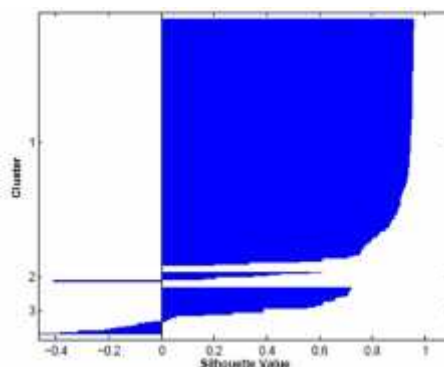
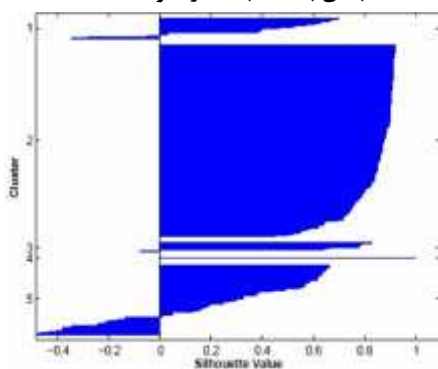
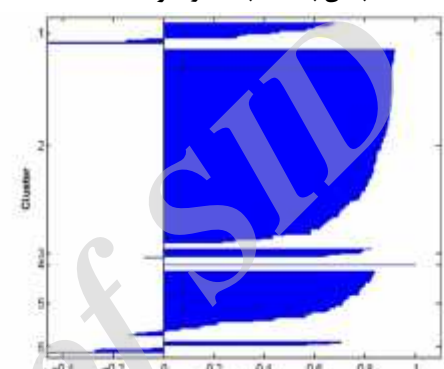
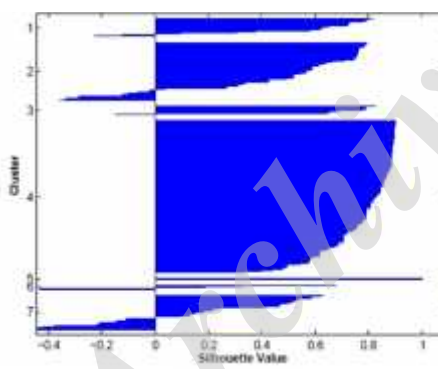
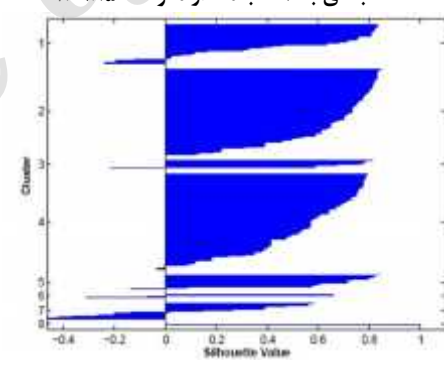
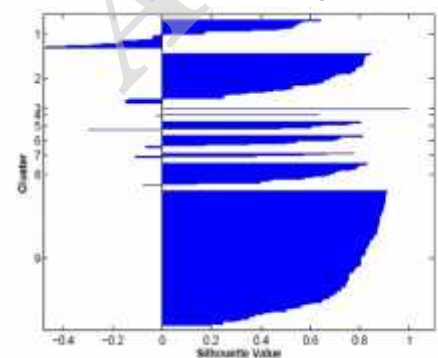
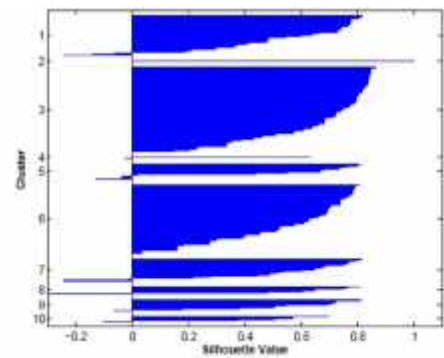
$$\dots(k) = \frac{1}{k} \sum_{n=1}^k (\min \{ \frac{y_n + y_m}{u_{nm}} \}) \quad (6)$$

$$y_n = \frac{1}{\|O^n\|} \sum_{c_j \in O^n} Sim(c_j, c^n) \quad (7)$$

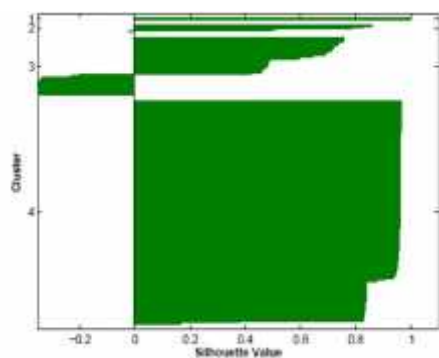
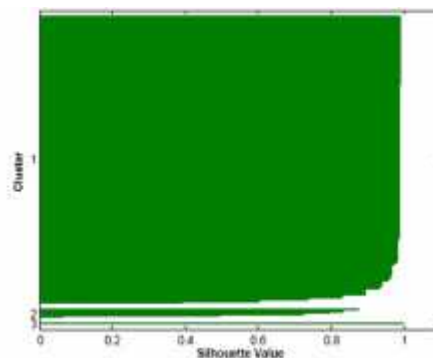
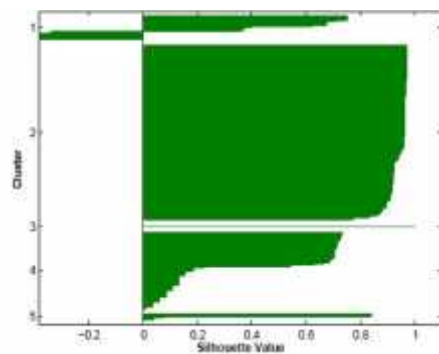
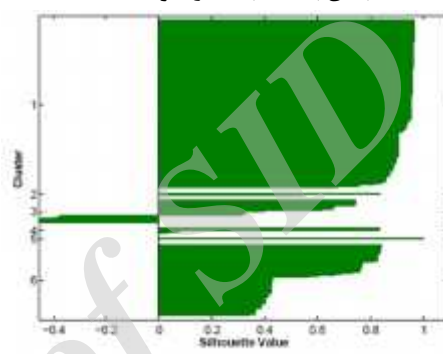
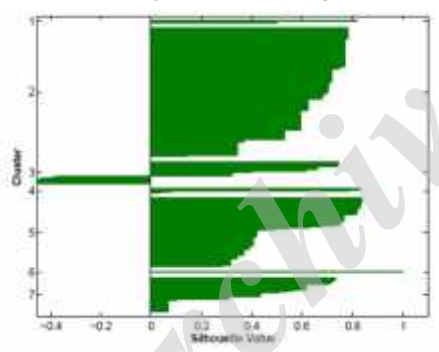
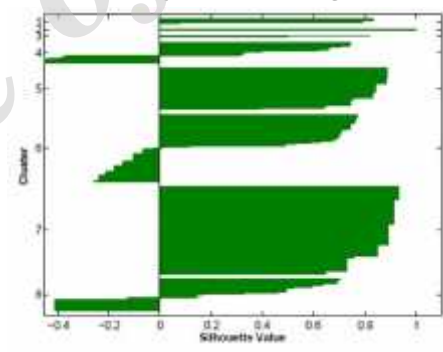
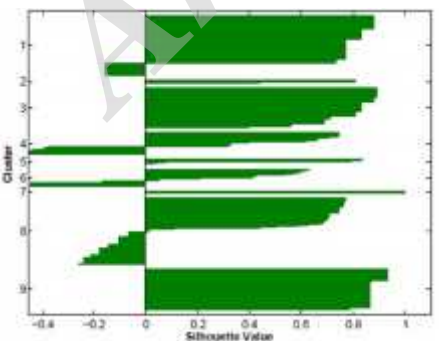
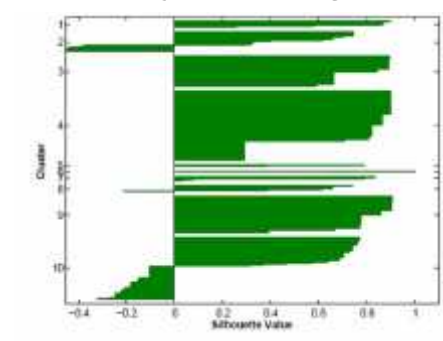
$$y_m = \frac{1}{\|O^m\|} \sum_{c_j \in O^m} Sim(c_j, c^m) \quad (8)$$

$$u_{nm} = Sim(c^n, c^m) \quad (9)$$

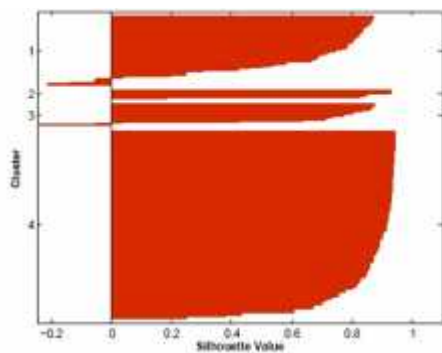
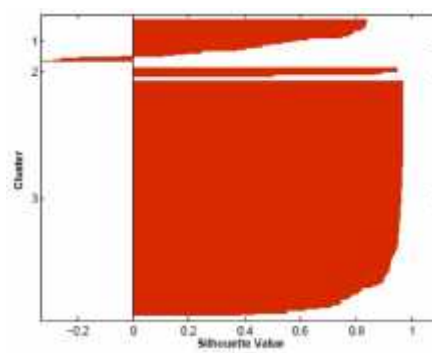
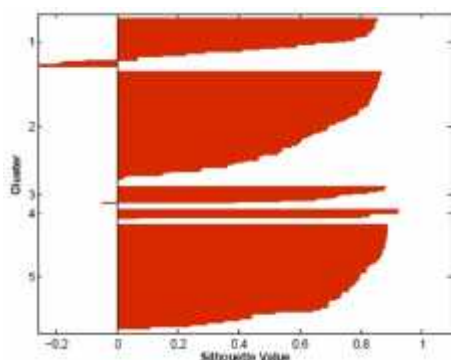
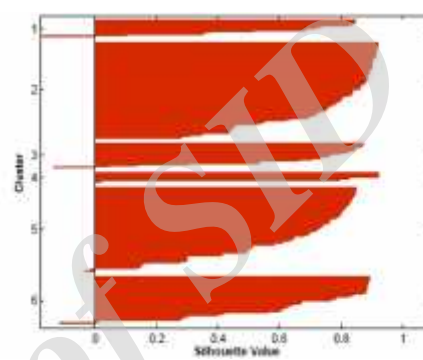
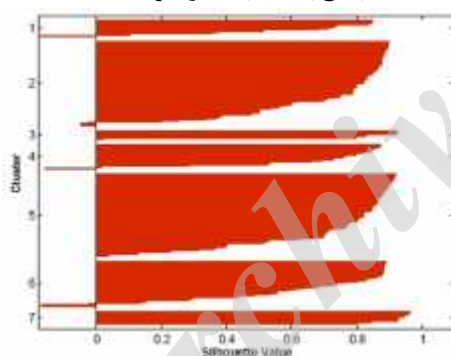
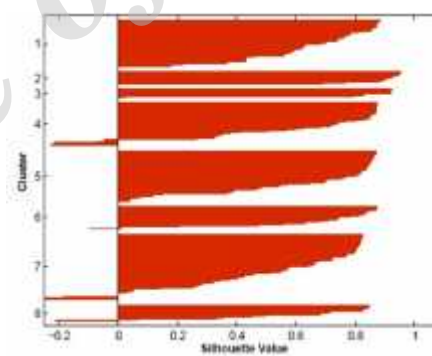
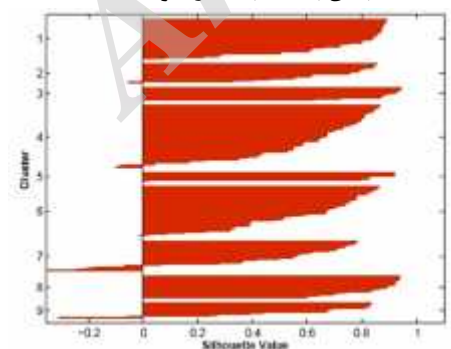
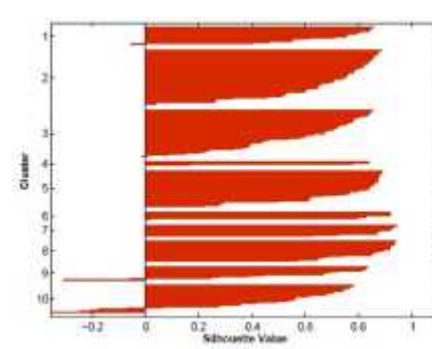
در این معادلات O مجموعه مراکز خوشه‌ها، c^n مراکز خوشه‌ها، O^n مجموعه عناصری که به عنوان مراکز خوشه انتخاب نشده‌اند، T^c مجموعه کلیه عناصری که خوشه‌بندی روی آنها صورت گرفته است، m میانگین شباهت بین مرکز

دسته‌بندی با $K=4$ با مقدار متوسط 0.8064 دسته‌بندی با $K=3$ با مقدار متوسط 0.7892 دسته‌بندی با $K=5$ با مقدار متوسط 0.6607 دسته‌بندی با $K=6$ با مقدار متوسط 0.6985 دسته‌بندی با $K=7$ با مقدار متوسط 0.605 دسته‌بندی با $K=8$ با مقدار متوسط 0.5437 دسته‌بندی با $K=9$ با مقدار متوسط 0.6164 دسته‌بندی با $K=10$ با مقدار متوسط 0.5864

شکل ۲: نیم‌رخ خوشه‌ها و مقادیر مطلوبیت با $K=3$ تا $K=10$ ، مربوط به دو عنصر طلا و آرسنیک.

دسته‌بندی با $K=4$ با مقدار متوسط 0.7886 دسته‌بندی با $K=3$ با مقدار متوسط 0.9647 دسته‌بندی با $K=5$ با مقدار متوسط 0.7239 دسته‌بندی با $K=6$ با مقدار متوسط 0.7743 دسته‌بندی با $K=7$ با مقدار متوسط 0.5543 دسته‌بندی با $K=8$ با مقدار متوسط 0.5472 دسته‌بندی با $K=9$ با مقدار متوسط 0.5258 دسته‌بندی با $K=10$ با مقدار متوسط 0.5542

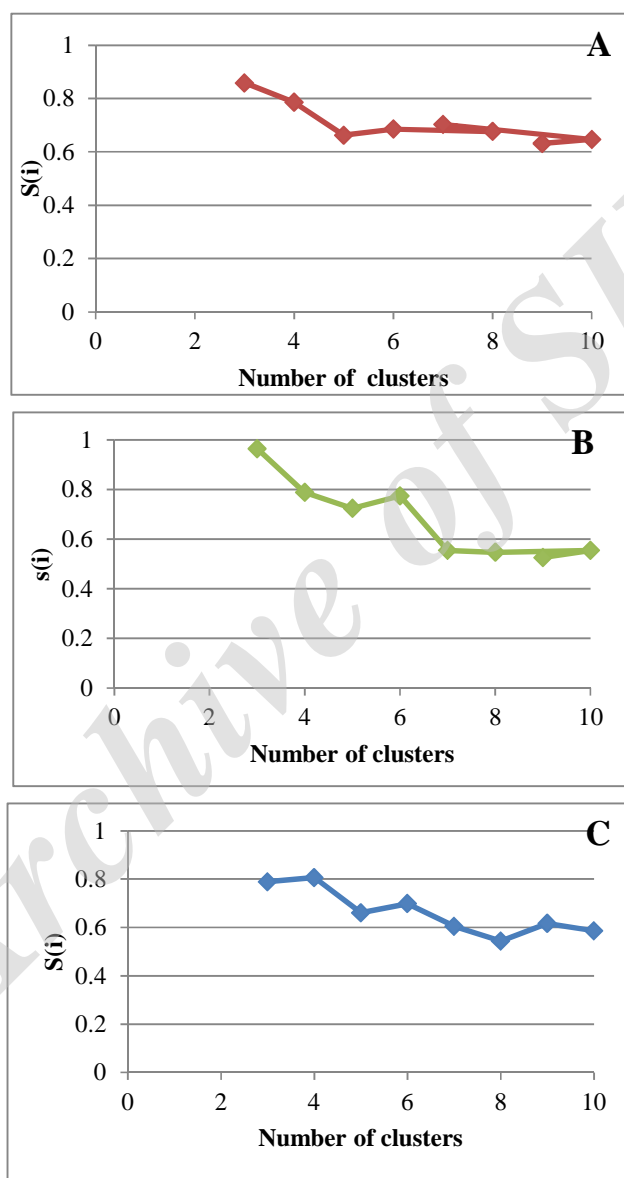
شکل ۳: نیمرخ خوشه‌ها و مقادیر مطلوبیت به ازای $K=3$ تا $K=10$ در مورد عناصر طلا و آنتیموان.

دسته‌بندی با $K=4$ با مقدار متوسط 0.7858 دسته‌بندی با $K=3$ با مقدار متوسط 0.8584 دسته‌بندی با $K=5$ با مقدار متوسط 0.662 دسته‌بندی با $K=6$ با مقدار متوسط 0.6854 دسته‌بندی با $K=7$ با مقدار متوسط 0.7036 دسته‌بندی با $K=8$ با مقدار متوسط 0.6762 دسته‌بندی با $K=9$ با مقدار متوسط 0.6318 دسته‌بندی با $K=10$ با مقدار متوسط 0.6462

شکل ۴: نیمرخ خوشه‌ها و مقادیر مطلوبیت برای دسته‌بندی با $K=3$ تا $K=10$. برای دو عنصر آرسنیک و آنتیموان.

آرسنیک و آنتیموان و شکل (B-5) برای دو عنصر طلا و آنتیموان- برای هر دو زوج نامبرده شده- تعداد خوشه ۳ دارای بیشترین مقدار و در شکل (C-5) تعداد خوشه ۴ برای دو عنصر طلا و آرسنیک، $S(i)$ دارای بیشترین مقدار است.

همچنین می‌توان برای انتخاب تعداد خوشه بهینه، نمودار تغییر ملاک اعتبارسنجی $S(i)$ را بر اساس تغییر تعداد خوشه‌ها به صورت شکل ۵ نشان داد که مقایسه آن راحت‌تر و ساده‌تر است. به عبارتی هر خوشه که بیشترین مقدار $S(i)$ را داشته باشد، به عنوان تعداد خوشه بهینه انتخاب می‌شود. شکل (A-5) مقدار $S(i)$ برای دو عنصر



شکل ۵: تغییر مقدار ملاک اعتبارسنجی $S(i)$ براساس تعداد خوشه‌ها (A: برای دو عنصر آرسنیک و آنتیموان، B: دو عنصر طلا و آنتیموان، C: دو عنصر طلا و آرسنیک)

مقادیر مختلف k محاسبه شده است. همانطور که بیان شد، مقدار حداکثر (k) ، نشان‌دهنده تعداد مناسب خوشه‌ها است. جدول ۲ مقادیر (k) به دست آمده متناظر با تعداد

همچنین با توجه به تابع کیفیت و با استفاده از مقدار (k) ... تعداد خوشه مناسب تعیین می‌شود. برای تعیین تعداد خوشه‌ها، مقدار (k) ... با استفاده از رابطه ۶ برای

۳ و برای عناصر طلا، آرسنیک و آنتیموان با توجه به موقعیت نمونه‌برداری آن، تعداد خوشه برابر با ۵، مناسب‌ترین تعداد خوشه است. همانطور که مشهود است، تعداد خوشه مناسب به دست آمده از تابع کیفیت با نتایج معیار $S(i)$ مطابقت دارد.

خوشه‌ها را نشان می‌دهد. در رفتارسنجی دو عنصر طلا و آرسنیک، $k=0/۸۵۵۶۲$... بیشترین مقدار است، بنابراین مناسب‌ترین تعداد خوشه برابر با چهار است و به همین ترتیب برای دو عنصر طلا و آنتیموان تعداد خوشه برابر با ۳ و برای دو عنصر آرسنیک و آنتیموان، تعداد خوشه برابر با

جدول ۲: مقدار k ... برای مقادیر مختلف k

عناصر	تعداد خوشه	k ...	عناصر	تعداد خوشه	k ...		
دو عنصر طلا و آرسنیک	۳	۰/۸۲۶۲۴	دو عنصر طلا و آنتیموان	۳	۰/۷۹۳۴۵		
	۴	۰/۸۵۵۶۲		۴	۰/۷۱۱۲۵		
	۵	۰/۸۱۲۳۵		۵	۰/۶۵۴۲۹		
	۶	۰/۷۶۲۳۵		۶	۰/۵۶۲۲۷		
	۷	۰/۷۲۵۶۹		۷	۰/۵۲۱۲۵		
	۸	۰/۷۱۲۵۶		۸	۰/۵۱۵۵۹		
	۹	۰/۶۹۸۲۳		۹	۰/۵۱۱۶۳		
	۱۰	۰/۶۹۱۲۵		۱۰	۰/۵۱۳۲۵		
	دو عنصر آرسنیک و آنتیموان	۳		۰/۹۱۳۳۹	سه عنصر طلا، آرسنیک و آنتیموان با توجه به موقعیت نمونه‌ها	۳	۰/۶۶۵۲۳
		۴		۰/۹۰۱۴۲		۴	۰/۶۵۵۲۸
۵		۰/۸۷۷۵۲	۵	۰/۶۹۸۸۴			
۶		۰/۸۷۴۴۰	۶	۰/۶۵۲۴۷			
۷		۰/۸۱۸۹۱	۷	۰/۶۱۱۲۵			
۸		۰/۸۱۱۲۵	۸	۰/۶۱۰۵۱			
۹		۰/۸۱۰۰۲	۹	۰/۶۰۳۷۸			
۱۰		۰/۸۰۸۸۶	۱۰	۰/۶۰۰۲۳			

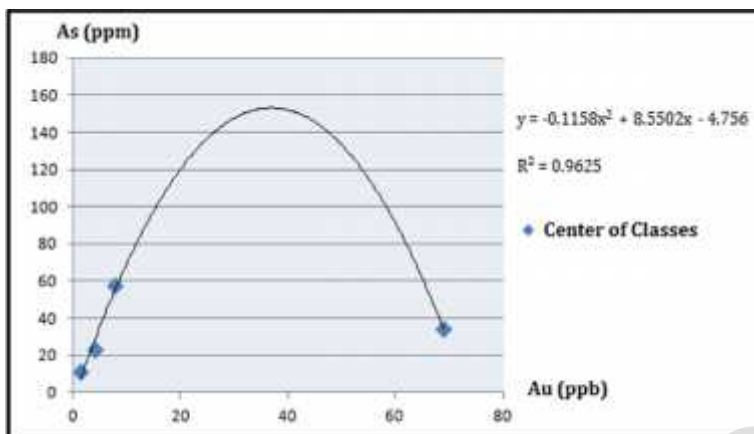
افزایش، بهترین منحنی برازش شده به آن، منحنی درجه دو با تقعر منفی است. معادله منحنی مذکور به صورت $y = -0.0012x^2 + 0.0928x + 1.134$ و ضریب همبستگی آن برابر با $R^2 = 1.00$ به دست آمده است.

با توجه به رفتار طلا نسبت به آرسنیک و آنتیموان، که با افزایش طلا، هم آرسنیک و هم آنتیموان افزایش پیدا می‌کنند به همین دلیل قابل انتظار است که این دو عنصر آرسنیک و آنتیموان دارای رابطه مستقیم باشند که بر اساس دسته‌بندی که در شکل ۸ مشخص می‌باشد، به ازای افزایش عیار عنصر آرسنیک، عیار عنصر آنتیموان نیز افزایش می‌یابد. معادله‌ای خط برازش شده $y = 0.0247x + 0.9713$ و ضریب همبستگی معادله

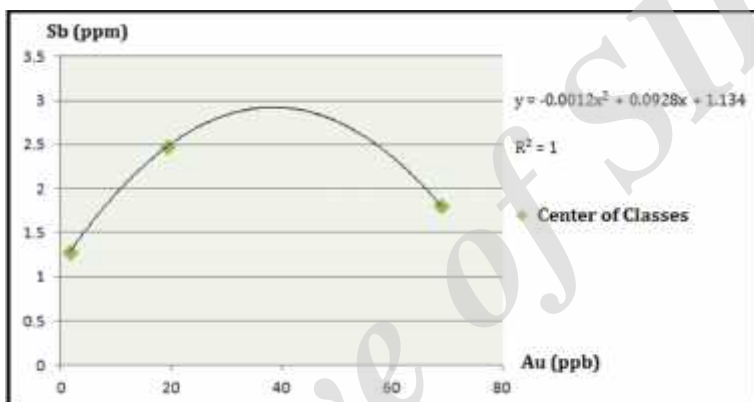
برازش شده به مرکز دسته‌ها نیز برابر با $R^2 = 0.9992$ است.

مراکز دسته‌های تعیین شده به ازای بهترین دسته‌بندی که برای دو عنصر طلا و آرسنیک به ازای $K=4$ در شکل ۶ و به ازای $K=3$ برای دو عنصر طلا و آنتیموان در شکل ۷ و به ازای $K=3$ برای دو عنصر آرسنیک و آنتیموان در شکل ۸ به صورت نمودار نشان داده می‌شود. بر اساس این دسته‌بندی که در شکل ۶ مشخص است، به ازای افزایش عیار عنصر طلا، عیار عنصر آرسنیک، ابتدا افزایش و سپس کاهش می‌یابد. با توجه به این افزایش و کاهش، بهترین منحنی درجه دو با تقعر منفی است. خط برازش شده به صورت $y = -0.1158x^2 + 8.5502x - 4.756$ و ضریب همبستگی آن نیز برابر با $R^2 = 0.9625$ به دست آمده است.

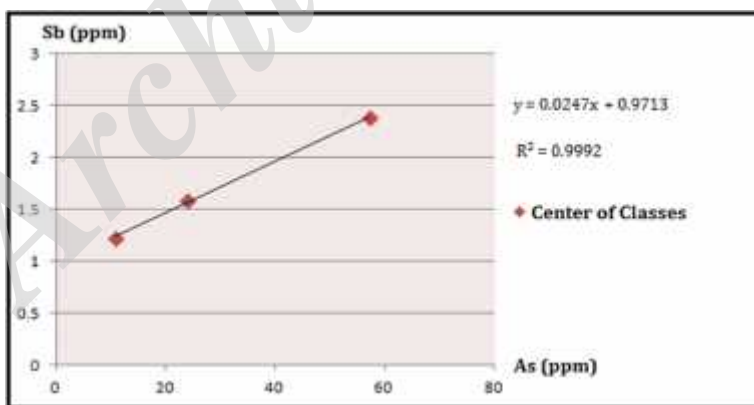
برای عناصر طلا و آنتیموان با توجه به شکل ۷، به ازای افزایش عیار عنصر طلا، عیار عنصر آنتیموان در ابتدا افزایش و سپس کاهش می‌یابد. بر اساس این کاهش و



شکل ۶: بهترین خط برازش شده به مراکز دسته‌ها به ازای $K=4$ برای عناصر طلا و آرسنیک



شکل ۷: بهترین خط برازش شده به مراکز دسته‌ها به ازای $K=3$ برای عناصر طلا و آنتیموان

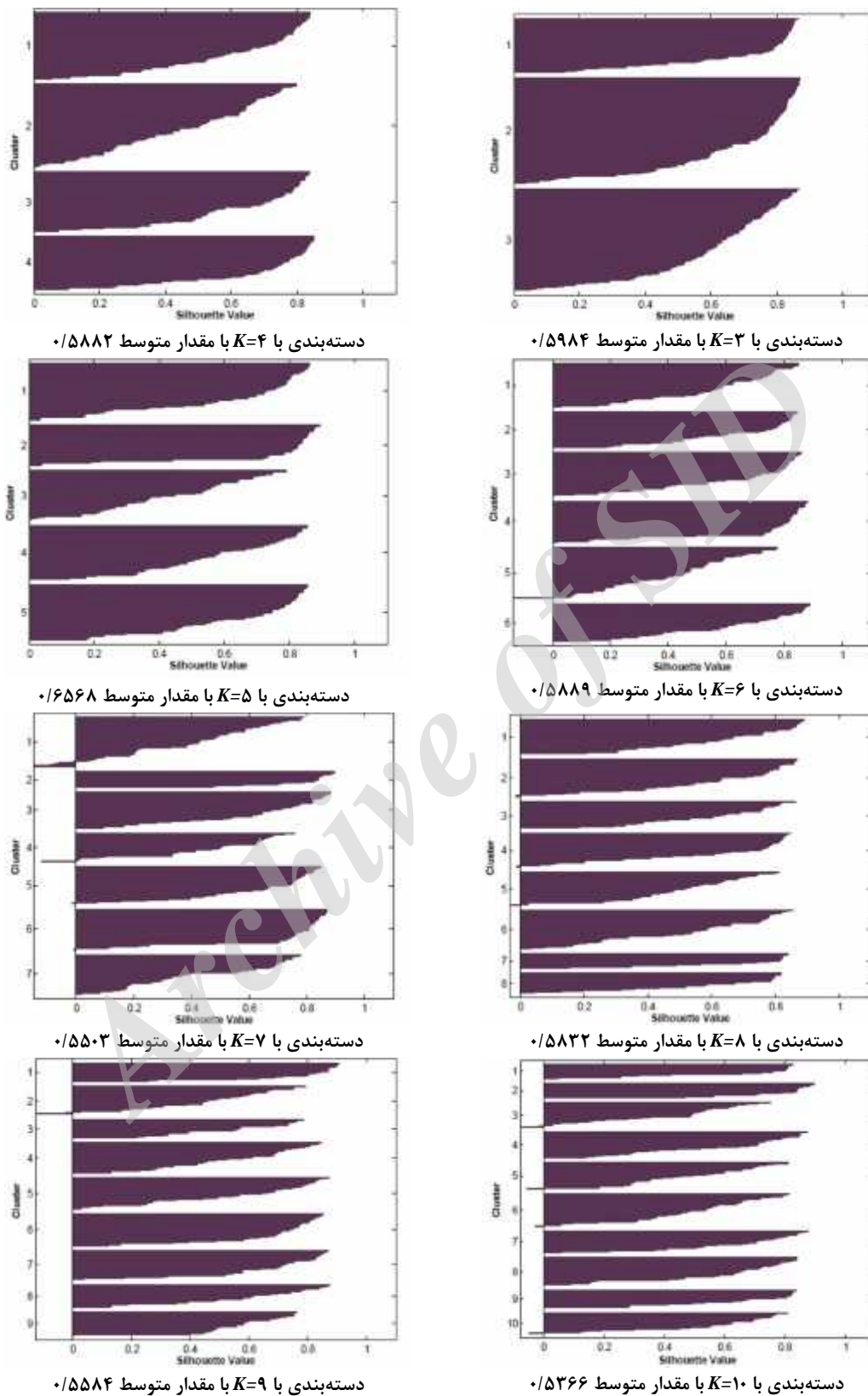


شکل ۸: بهترین خط برازش شده به مراکز دسته‌ها به ازای $K=3$ برای عناصر آرسنیک و آنتیموان

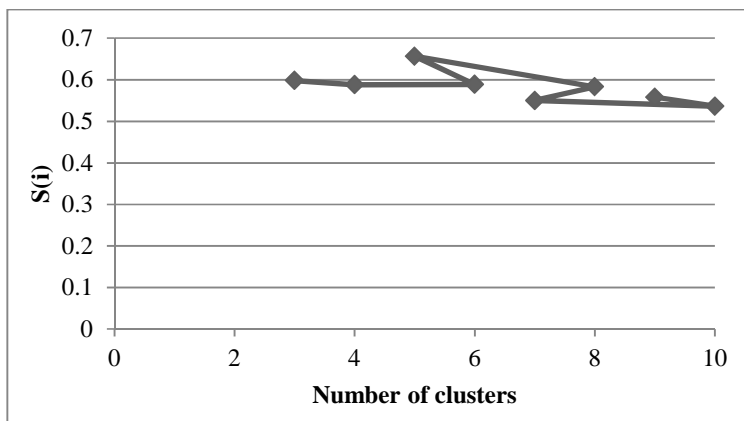
عنصر طلا، آرسنیک و آنتیموان (با در نظر گرفتن طول و عرض نقاط) در شکل ۹ آورده شده است.

۳-۴- بررسی رفتار طلا با توجه به عیار عناصر آرسنیک و آنتیموان

نیمرخ خوشه‌ها و میزان مطلوبیت دسته‌بندی‌های صورت گرفته با توجه به مقادیر $K=3$ تا $K=10$ در مورد سه



شکل ۹: نیمرخ خوشه‌ها و مقادیر مطلوبیت برای دسته‌بندی با $K=3$ تا $K=10$ برای سه عنصر طلا و آرسنیک و آنتیموان



شکل ۱۰: تغییر مقدار ملاک اعتبارسنجی $S(i)$ براساس تعداد خوشه‌ها (برای سه عنصر طلا، آرسنیک و آنتیموان)

توجه به اینکه مختصات به عنوان ویژگی ورودی در کنار عیار عناصر طلا و آرسنیک و آنتیموان قرار گرفته است و اینکه بازه تغییرات مختصات و مقادیر عیار متفاوت است، برای جلوگیری از ایجاد خطا در محاسبات و به دست آوردن مقدار تخمینی صحیح، باید همه مقادیر ورودی در یک بازه استاندارد قرار بگیرند. بدین منظور با استفاده از رابطه‌ی ۱۰ تمام ورودی‌ها در بازه‌ی $[0, 1]$ قرار گرفتند. مشخصات مراکز خوشه‌ها با $K=5$ در جدول ۳ آورده شده است.

$$X_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (10)$$

بر اساس نتایج حاصل شده به ازای مقادیر متفاوت برای K از ۳ الی ۱۰، همان‌طور که مشاهده شد، دسته‌بندی روی نمونه‌های تعریف شده با مشخصه‌های عیار عنصر طلا، آرسنیک و آنتیموان و طول و عرض نمونه‌های برداشت شده، به ازای $K=5$ دارای بهترین دسته‌بندی است. همچنین با توجه به شکل ۱۰، بیشترین مقدار $S(i)$ متعلق به $K=5$ است. برای اطمینان بیشتر و با توجه به نوسانات موجود در مقادیر مطلوبیت، مقدار K تا عدد ۲۰ افزایش داده شده است، اما مقدار مطلوبیت از ۰/۶۵۶۸ فراتر نرفته و به ازای $K=20$ به بعد، روند کاهشی داشته است. بنابراین برای به دست آوردن مشخصات مربوط به مرکز دسته‌ها، با

جدول ۳: مشخصات مرکز دسته‌ها (خوشه‌بندی با مقدار $K=5$)

شماره مرکز دسته	مقدار عنصر طلا (ppb)	مقدار عنصر آرسنیک (ppm)	مقدار عنصر آنتیموان (ppm)	طول (m)	عرض (m)
اول	۰/۰۶۴	۰/۰۲۸	۰/۰۳۳	۰/۸۸۹	۰/۶۸۸
دوم	۰/۰۱۷	۰/۱۵۷	۰/۰۶۱	۰/۷۰۸	۰/۳۲۱
سوم	۰/۰۱۰	۰/۰۳۸	۰/۰۳۶	۰/۵۵۲	۰/۵۶۲
چهارم	۰/۰۰۶	۰/۰۶۰	۰/۰۴۶	۰/۳۳۹	۰/۱۹۹
پنجم	۰/۰۰۶	۰/۰۴۹	۰/۰۲۸	۰/۱۶۷	۰/۷۲۷

در محدوده خونی، پرداخته می‌شود تا بتوان به کمک رابطه‌ی به دست آمده عیار عناصر را تخمین زد. مقادیر عنصر طلا به عنوان متغیر وابسته و مقادیر عنصر آرسنیک و آنتیموان، طول و عرض نقاط به عنوان متغیرهای مستقل به نرم‌افزار معرفی می‌گردند. سپس نتایج موجود در جدول ۴

۵- تخمین عیار طلا

در این قسمت با استفاده از رگرسیون چند متغیره در نرم‌افزار SPSS به تعیین رابطه در مورد طلا و آرسنیک و آنتیموان با توجه به طول و عرض نمونه‌های برداشت شده

به عنوان مشخصات و ضرایب رگرسیون چند متغیره رابطه (۱۱) $y = a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + b$ محاسبه و گزارش شده‌اند.

جدول ۴: مشخصات و ضرایب مربوط به خط رگرسیون چندگانه.

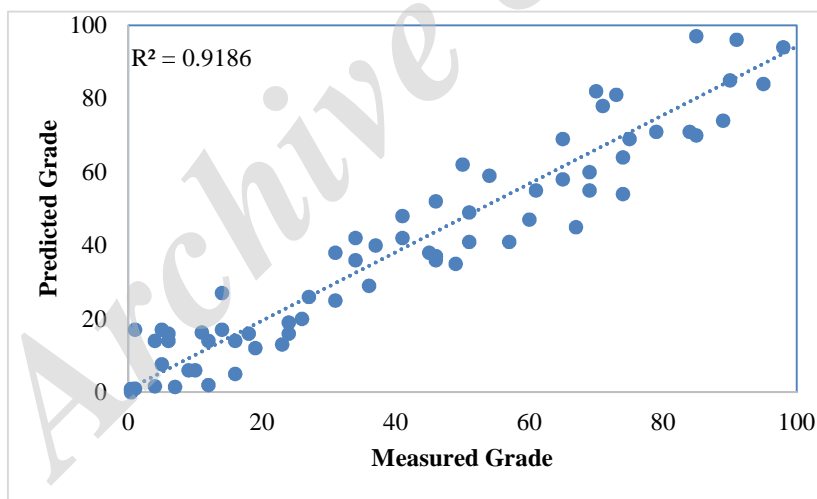
a_4	a_3	a_2	a_1	b	r^2	R
-۰/۳۷۵	۰/۲۲۱	-۱۴/۸۹۲	۱/۹۹۵	۰/۵۶۱	۰/۷۳۲۷	۰/۸۵۶

با مقادیر به دست آمده از معادله مقایسه گردد، تا دقت و صحت تخمین‌گر سنجیده شود. بنابراین ۳۰ درصد از نمونه‌ها به طور تصادفی انتخاب شد و براساس مقادیر عناصر آنتیموان و آرسنیک و طول و عرض نقاط نمونه‌برداری و قرار دادن آن در معادله ۱۲، مقدار عنصر طلا تخمین زده شده و با مقدار واقعی آن مقایسه می‌شود. نتایج به صورت یک نمودار پراکندگی در شکل ۱۱ نشان داده شده است. شکل ۱۱ با ضریب همبستگی ۹۱ درصد، به خوبی بهترین تطابق میان واقعیت و عبارهای تخمین زده شده را نشان می‌دهد، که این نشان‌دهنده صحت روش به کار گرفته شده است.

با توجه به ضرایب موجود در جدول فوق فرمول خط رگرسیون چندگانه تعیین می‌شود (رابطه ۱۲).

$$Au = 1.995As - 14.892Sb + 0.221X - 0.375Y + 0.561 \quad (12)$$

میزان r^2 حاصل شده نشان‌دهنده رابطه یک سهمی است که مدل رگرسیون توانسته تغییرات را نسبت به Y (یعنی Au) توضیح دهد. در این جا $r^2 = 0.7327$ شده است و بنابراین ۷۳٪ تغییرات میزان طلا (Y) به واسطه X ها یعنی (آرسنیک، آنتیموان، طول و عرض نقاط نمونه‌برداری) است. برای اعتبار سنجی تخمین عیار طلا براساس معادله‌ی به دست آمده (رابطه‌ی ۱۲) باید تعدادی از داده‌های واقعی



شکل ۱۱: نمودار پراکندگی مقادیر واقعی عیار طلا در برابر مقادیر تخمینی برای اعتبار سنجی

و آنتیموان در محدوده خونی نسبت به یکدیگر پرداخته شد و بین هر دو عنصر مورد بررسی، رابطه و معادله‌ای همراه با ضریب همبستگی ارائه گردید. سپس با استفاده از روش مذکور، رابطه عناصر با در نظر گرفتن طول و عرض جغرافیایی نمونه‌ها به منظور تخمین عیار و برآورد دقیق‌تر از پیدایش و وسعت هاله‌های ژئوشیمیایی در محدوده مورد

۶- نتیجه گیری

با توجه به وجود شواهد کانی‌سازی طلا در منطقه خونی، بررسی وسعت هاله‌های ژئوشیمیایی و رفتار عناصر پارائز طلا در منطقه اهمیت می‌یابد بدین منظور با استفاده از روش K -means، به بررسی رفتار عناصر طلا و آرسنیک

prospecting: A case study of the Sanandaj-Sirjan metallogenic zone, Iran. *Journal of Geochemical Exploration*, Elsevier, 108, 183-195

[10] Sfidari, E., Kadkhodaie-Ilkhchi, A., Najjari, S., (2012), Comparison of intelligent and statistical clustering approaches to predicting total organic carbon using intelligent systems. *Journal of Petroleum Science and Engineering*, Elsevier, 63, 190-205.

[11] Wegner, T., Hussein, T., Hämeri, K., Vesala, T., Kulmala, M., Weber, S., (2012). Properties of aerosol signature size distributions in the urban environment as derived by cluster analysis. *Atmospheric Environment*, Elsevier, 61, 350-360.

[12] Ahmadian, J., Michael, H., McDonald, I., Regelous, M., Ghorbani, M. R., and Murata, M. (2009). High magmatic flux during Alpine-Himalayan collision: Constraints from the Kal-e-Kafi complex, Central Iran. *Geological Society of America Bulletin*. 121, 857-868.

[13] Nezampoor, M.H., (2005). Determination of promise points in Kal-e-Kafi area with remote sensing technology, Anarak, central Iran, *Mineraleum deposita conference*, China.

[14] Kan Azin consulting engineers Co. (2005). Report of geology, mining and Lithochemical 1: 5000 to explore gold in the khooni district. (in Persian)

[15] Heydarian, N., Rassa, A., (2012). Characteristics and genesis of gold mineralization in Eocene volcanic units of khooni Cheshmeh, Anarak, the nature of mineralizing fluids and comparison with other types of gold deposits. *Journal of Geology of Iran* 17, 73-85. (in Persian)

[16] Pourjabar, A., (2005). Geochemical investigations on the polymetallic vein in Khouni (Esfahan Province), MSc Thesis, Amirkabir University of Technology, Tehran, 217. (in Persian)

[17] Egozcue, J.J., Pawlowsky-Glahn, V., Mateu-Figueras, G. (2003). Isometric Logratio Transformations for Compositional Data Analysis. *Mathematical Geology*, 35, 279-300.

[18] Jain, A. K. (2012). Data clustering: 50 years beyond K-Means, *Pattern Recognition Letters*, Elsevier, 31, 651-666.

[19] Saha, S., Bandy opadhyay, S. (2013). A generalized automatic clustering algorithm in a multi objective framework. *Applied Soft Computing*, Elsevier, 13, 89-108.

[20] Ghannad pour, S., Hezarkhani, A., Mahvash mohammadi, N. (2013). Investigation of Pb behavior circumstance than Zn in porphyry copper system of Parkam, Kerman using K-means method. *First*

مطالعه تعیین شد. با توجه به نتایج حاصل از پردازش عناصر نامبرده، معادله بدست آمده جهت تخمین عیار عنصر طلا بر حسب چهار پارامتر عیار آرسنیک، عیار آنتیموان، طول و عرض نقاط نمونه برداری، به صورت رابطه ۱۲ است و ضریب همبستگی اعتبارسنجی آن برابر با ۹۱ درصد گزارش شده است.

مراجع

[1] Majdifar, S., Kamali, GH. (2014). Iron Grade Estimation Using ANFIS Algorithm at Tappeghermez Anomaly of Sangan Mine. *Journal of Analytical and Numerical Methods in Mining Engineering (ANM)*. No 5. (in Persian)

[2] Abbas zadeh, S., Rahimi pour, GH., Najmodini, M. (2013). Recognition of Cu-Porphyry Mineralization Areas by Using One and multivariate integration methods on drainage geochemical data in Ghale Askar area, Kerman province. *Journal of Analytical and Numerical Methods in Mining Engineering (ANM)*. No. 6. (in Persian)

[3] ShayesteFar, M., Rasa, I. (2005). Multivariate Data Analysis of Qanat Marvan Lead-Zinc Deposit., Kerman. *Geosciences Scientific Quarterly Journal*. No, 57. (in Persian)

[4] Malyszko, D., Wierzchon, S. T. (2007). Standard and Genetic K-means Clustering Techniques in Image Segmentation. In 6th International Conference on Computer Information Systems and Industrial Management Applicatins.

[5] Abolhassani, B., Salt, J. E. (2005). A Simplex K-means Algorithm for Radio-Port Placement in Cellular Networks. In Canadaian Conference on Electrical and Computer Engineering.

[6] Chen, T. W., Chien, S. Y. (2009). Bandwidth Adaptive Hardware Architecture of K-Means Clustering for Video Analysis. *IEEE Transactions on VLSI Systems*.

[7] Yang, J., Zhuang, Y., Wu, F. (2012). ESVC-based extraction and segmentation of texture features. *Computers & Geosciences*, Elsevier, 49, 238-247.

[8] Mora, J.L., Armas-Herrera, C.M., Guerra, J.A., Rodríguez-Rodríguez, A., Arbelo, C.D., (2012), Factors affecting vegetation and soil recovery in the Mediterranean woodland of the Canary Islands (Spain). *Journal of Arid Environments*, Elsevier, 87, 58-66.

[9] Meshkani, S. A., Mehrabi, B., Yaghubpur, A., Alghalandis, Y. F., (2011). The application of geochemical pattern recognition to regional

Geological Conference of Iranian plateau. (in Persian)

[21] Menard, J. J. (1995). Relationship between altered pyroxene diorite and the magnetite mineralization in the Chilean Iron Belt, with emphasis on the El Algarrobo iron deposits (Atacama region, Chile)., *Mineral Deposita*, Springer-Verlag, 30, 268-274.

[22] Leiluo, X., Xianwu, B., Ruizhong, H., Xingchun, Z. (2012). Relationships between porphyry Cu-Mo mineralization in the Jinshajiang-Red River metallogenic belt and tectonic activity: Constraints from zircon U-Pb and molybdenite Re-Os geochronology., *Ore Geology Reviews*, Elsevier, 48, 460-473.

[23] Tarkian, M., Stribny, B. (1999). Platinum-group elements in porphyry copper deposits: a reconnaissance study. *Mineralogy and Petrology*, Springer-Verlag, 65, 161-183.

[24] Yaghini, M., Ghannad pour, F., Khadmatlu, S., (2008). Offering an clustering innovative method in data mining by using genetic algorithms to solve a real case study in the rail transportation industry, Iran Data Mining Conference. Amirkabir university. Tehran. (in Persian)

[25] Shina, H.W. and Sohb, S.Y. (2004). Segmentation of Stock Trading Customers According to Potential Value, *Expert Systems with Applications*, 27, 27-33.

Estimation of Grade Gold in Khooni Deposit Using the Behavior of Gold, Arsenic and Antimony Elements by Clustering K-Means Method

N. Mahvash Mohammadi¹, A. Hezarkhani^{2*}

1- PhD. Student in Mining, Dept. of Mining and Metallurgy, Amirkabir University of Technology, Iran
2- Professor, Dept. of Mining and Metallurgy, Amirkabir University of Technology, Iran

* Corresponding Author: ardehez@aut.ac.ir

(Received: July 2015, Accepted: March 2016)

<i>Keywords</i>	<i>English Extended Abstract</i>
Gold Khooni Deposits Clustering Estimated Grade K-Means	Summary Khooni exploratory area is located at 270 kilometers of Esfahan, and belongs to Central Iran geological zone. According to some instances expressing gold mineral forming in this area, make reconnaissance of favorable area as an essential fact. A well-known algorithm of clustering is K-Means by which the data are divided into K clusters on the basis of distance. In this analysis, using the K-Means method to classify the sampling of khooni district for Gold, Arsenic and Antimony elements. The optimal K value was determined and then the data were clustered and the behavioral characteristics were analyzed, and at the end relationships and equations with correlation coefficients were identified and the grade of gold were estimated based on four parameters Arsenic, Antimony content; length and width of the sampling points.

Introduction

There are many methods to estimate the grade because of mining projects have high dependency to determine the tonnage accurately. One of the new methods of estimating the grade is clustering. Thus, it is necessary to find information about the relationships and behavior elements of Gold, Arsenic and Antimony to each other. in this area to determine the extent of the geochemical halo and to identify favorable area as well as estimation.

Methodology and Approaches

In this study for determination the behavior of the mentioned elements, the famous and efficient method of K-Means, which is one of the clustering methods, are used, that is based on minimizing total of Euclidean distance from the center each class. For this purpose, the quality function and silhouette criterion have been used to determine the optimal number of clusters. All data ranged the number of classes from k=3 to k=10 and afterwards the optimum number of clusters were selected by mentioned methods. At the end, according to the Center of clusters, the equation for estimating the grade of gold is provided. In this study, MATLAB and SPSS software are used to achieve results.

Results and Conclusions

The relationship of gold, arsenic and antimony according to length and width of the sampling points were determined for estimation grade of gold, Results showed that clustering with k=5 in case of Au, As and Sb were better than other classes number in each case. Obtained equation for estimating grade of gold element based on four parameters Arsenic, Antimony content; length and width of the sampling points, is $Au = 1.995 As - 14.892 Sb + 0.221 X - 0.375 Y + 0.561$ and the correlation coefficient 73% is reported. Also to evaluate the accuracy of the estimator, the results were validated. Estimated gold element values almost match with actual values of gold element that represents the accuracy of the used method.

Cite This Paper:

Mahvash Mohammadi, N., Hezarkhani, A., (2015). "Estimation of Grade Gold in Khooni Deposit Using the Behavior of Gold, Arsenic and Antimony Elements by Clustering K-Means Method" Journal of Analytical and Numerical Methods in Mining Engineering 5(10): 77-92.