



خوشه‌بندی بدون‌ناظر برای استخراج ساختار دگرسانی پروپیلتیک در داده‌های سنجش‌ازدور

هومص | HOOMAS

۳	پس‌زمینه و ضرورت پروژه
۳	چالش‌های فعلی و نیاز به نوآوری
۴	هدف و چشم‌انداز پروژه
۴	رویکرد و مسیر حل مسئله
۵	ساختار و جریان پروژه
۵	ارزش و تأثیر پروژه
۶	معماری راه‌حل و رویکرد فناورانه
۶	قیود اجرایی و چالش‌های عملی
۷	مراحل اجرایی پروژه
۷	شاخص موفقیت و معیارهای ارزیابی
۸	نتایج و تفسیر مدل
۹	چشم‌انداز و توسعه آینده

پس‌زمینه و ضرورت پروژه

در فرایند اکتشاف معدنی، هر تصمیم اشتباه می‌تواند هزینه‌ای معادل ماه‌ها کار و ده‌ها میلیارد تومان سرمایه را هدر دهد؛ به‌ویژه در پروژه‌هایی که داده‌های اندک اما اثرگذار از سنجش‌ازدور به‌عنوان نخستین ابزار شناخت منطقه به‌کار گرفته می‌شود. نرم‌افزارهایی مانند ENVI سال‌هاست به زمین‌شناسان کمک می‌کنند تا از طریق نسبت‌بندی و پردازش‌های کلاسیک، نقشه‌هایی از مناطق دگرسان شده تولید کنند؛ اما این نقشه‌ها در بهترین حالت تنها «بازتابی از مقدار طیفی» هستند و رفتار منطقه را توصیف نمی‌کنند.

در مقابل، واقعیت زمین‌شناسی هرگز صرفاً یک تصویر نیست؛ دگرسانی پروپلیتیک یک مفهوم طیفی-مکانی است؛ از رگه‌ها منشأ می‌گیرد، در امتداد خطواره‌ها رشد می‌کند، در نقاطی افزایش شدت دارد و در نقاطی محو می‌شود. این رفتار را نمی‌توان صرفاً با یک مقدار خاکستری یا یک آستانه ثابت مشاهده کرد.

در این پروژه تلاش شد به جای تحلیل ایستا، از یک نگاه «یادگیرنده» استفاده شود تا داده‌های سنجش‌ازدور نه تنها از نظر مقدار، بلکه از نظر بافت، ساختار و ارتباط فضایی نیز تفسیر شوند. هدف ما جایگزینی ENVI نبود؛ بلکه ایجاد یک لایه تحلیلی فراتر برای استخراج دانشی بود که در روش‌های کلاسیک دیده نمی‌شود. تمام این تحلیل‌ها برداشت ماست از داده موجود و ادعای قطعیت ندارد؛ تنها راهی است برای کشف الگوهای پنهان و هدایت بهتر تصمیم‌گیری در مراحل بعدی اکتشاف.

چالش‌های فعلی و نیاز به نوآوری

داده‌های سنجش‌ازدور در نگاه اول ساده‌اند: یک تصویر یا یک نقشه نسبت‌بندی. اما در واقع، این داده‌ها یکی از پیچیده‌ترین انواع داده‌های اکتشافی‌اند.

تفاوت‌های کوچک در مقدار طیفی، وجود نویز، پیکسل‌های پرت، تأثیر سایه‌روشن و اثرات لیتولوژیک باعث می‌شود تفسیر دستی این نقشه‌ها بسیار دشوار باشد. در نرم‌افزارهایی مانند ENVI، نتیجه‌های ما به آستانه‌ها وابسته است؛ اگر آستانه کمی بالا یا پایین رود، کل نقشه تغییر می‌کند و ساختاری که واقعاً وجود دارد محو یا اغراق می‌شود.

از طرفی، نسبت‌بندها تنها یک بعد از رفتار طیفی را نشان می‌دهند؛ در حالی که دگرسانی یک پدیده چندپه‌بعدی است. یک پیکسل ممکن است مقدار طیفی مشابهی با دیگری داشته باشد، اما از نظر مکانی در نقطه‌ای کلیدی قرار گرفته باشد. این ارتباط مکانی در ENVI دیده

نمی‌شود. در نتیجه، تصمیم‌گیری درباره‌ی شدت، اهمیت و شکل واقعی دگرسانی معمولاً به تجربه فردی واگذار می‌شود؛ تجربه‌ای که هم ارزشمند است و هم می‌تواند خطاپذیر باشد. به همین دلیل به یک ابزار جدید نیاز بود؛ ابزاری که بتواند بدون تکیه بر برچسب، از دل داده‌ها الگوها را بیاموزد، نه این‌که صرفاً یک نقشه خاکستری را رنگ‌آمیزی کند.

هدف و چشم‌انداز پروژه

هدف این پروژه آن بود که داده‌ی خام نسبت‌باندی منطقه تبدیل به یک مدل قابل‌تفسیر شود؛ مدلی که بتواند زیرزون‌های متفاوت دگرسانی پروپیلیتیک را شناسایی کند و نشان دهد کدام نقاط واقعاً شبیه یکدیگر رفتار می‌کنند.

این چشم‌انداز مبتنی بر سه ایده بود:

نخست، اینکه رفتار طیفی، هرچند کم‌دامنه، می‌تواند حامل نشانه‌های دگرسانی باشد. دوم، اینکه مختصات مکانی پیکسل‌ها نباید نادیده گرفته شوند؛ چرا که دگرسانی‌ها ذاتاً پیوسته، خطی و ساختارمند هستند.

سوم، اینکه مدل بدون نظارت می‌تواند درک خودکار از این الگوها بسازد؛ درکی که در ذهن ما وجود دارد اما ابزارهای کلاسیک توان ثبت آن را ندارند.

چشم‌انداز نهایی این بود که محدوده دگرسانی نه فقط از نظر مقدار طیفی، بلکه از نظر رفتار فضایی نیز تحلیل شود و این تحلیل به ابزاری راهبردی برای انتخاب مناطق مناسب برای برداشت ژئوشیمیایی، نمونه‌برداری یا حفاری تبدیل شود.

رویکرد و مسیر حل مسئله

نقطه شروع، داده خامی بود که از نرم‌افزار سنجش‌ازدور استخراج شده بود؛ مختصات، مقدار نسبت‌بندی و موقعیت جغرافیایی. تمام مقادیر صفر کنار گذاشته شد تا تنها روی نواحی هدف تمرکز شود. پس از آن، داده‌ها نرمال شدند تا اختلاف واحد، ابعاد و دامنه‌ها رفتار مدل را منحرف نکنند.

در مرحله نخست، چند الگوریتم بدون‌ناظر بررسی شد. KMeans به عنوان مدلی ساده اما قدرتمند اولین انتخاب بود. این مدل توانست سه خوشه اصلی را استخراج کند که از نظر طیفی-مکانی کاملاً قابل تفسیر بودند. خوشه‌ای که بالاترین میانگین مقدار طیفی را داشت

به عنوان هسته دگرسانی شناخته شد و دو خوشه دیگر بیانگر زون های مرزی و کم قدرت تر بودند.

پس از آن، Gaussian Mixture Model برای تحلیل دقیق تر طیف استفاده شد. اگرچه GMM تفاوت های طیفی را بهتر نشان می داد و به نوعی ساختار احتمالاتی نقشه را ترسیم می کرد، اما از نظر پیوستگی مکانی رفتار نسبتاً پیچیده ای نشان داد.

در مرحله بعد، DBSCAN به عنوان مدل چگالی محور بررسی شد. این روش، برخلاف انتظار، تفکیک زیادی ارائه نداد؛ دلیل آن یکنواختی نسبی چگالی طیفی در منطقه و همپوشانی گسترده نقاط بود. بنابراین DBSCAN بیشتر به عنوان ابزار ارزیابی و کنترل کیفیت استفاده شد نه روش اصلی تفکیک.

در نهایت، با مقایسه نتایج، KMeans به عنوان مدل مناسب برای این پروژه انتخاب شد؛ زیرا هم ساختار طیفی را به خوبی ضبط کرده بود و هم پیوستگی فضایی را حفظ کرده بود.

ساختار و جریان پروژه

فرآیند پروژه به صورت گام به گام طراحی شد تا هر مرحله بسترساز مرحله بعد باشد. ابتدا پاک سازی و استاندارد سازی داده انجام شد. سپس تحلیل آماری اولیه برای شناخت دامنه و پراکندگی نسبت باند صورت گرفت. پس از آن مدل های بدون نظارت به کار گرفته شدند و خروجی ها از چند جهت چون رفتار طیفی، ساختار مکانی و قابلیت تفسیر زمین شناسی ارزیابی شدند.

در پایان، نتایج مدل منتخب دوباره روی نقشه اصلی نمایش داده شد تا شکل واقعی دگرسانی و ارتباط آن با ساختارهای خطی منطقه مشخص شود. این چرخه به گونه ای طراحی شده بود که تحلیل قابلیت بازتولید داشته باشد و بتوان در مراحل بعدی با داده های جدید گسترش یابد.

ارزش و تأثیر پروژه

ارزش اصلی این پروژه در تبدیل یک نقشه خاکستری ساده به یک مدل رفتاری است. این مدل نشان داد که دگرسانی پروپیلیتیک، هرچند از نظر طیفی کم دامنه است، اما الگوی مشخصی در فضا دارد. چنین تحلیلی فراتر از توان نرم افزارهای کلاسیک سنجش از دور است؛ چرا که این ابزارها تنها مقدار را نشان می دهند نه ساختار را.

از منظر اقتصادی، چنین تحلیلی می‌تواند مسیر نمونه‌برداری و حتی حفاری را به شکل قابل توجهی اصلاح کند. انتخاب محل مناسب برای یک گمانه می‌تواند میلیون‌ها تومان صرفه‌جویی ایجاد کند و احتمال موفقیت عملیات را افزایش دهد. این مدل توانایی دارد مناطق با پتانسیل کمتر را حذف کند و مناطق با تمرکز طیفی مشابه را برجسته کند.

از نظر علمی نیز این پروژه یک مرحله فراتر از نسبت‌بندی ساده است. این تحلیل نشان می‌دهد که حتی داده‌ای که در ظاهر تنها شامل یک مقدار طیفی است، می‌تواند حامل الگوهای عمیق باشد، به شرط آنکه ابزار تحلیل مناسب استفاده شود.

اوج ارزش این روش زمانی است که با داده‌های ژئوشیمیایی ترکیب شود. نقاطی که هم از نظر طیفی (سنجش‌ازدور) و هم از نظر شیمیایی (ژئوشیمی) همگرایی داشته باشند، بهترین نقاط برای عملیات میدان و حفاری هستند. این هم‌پوشانی می‌تواند ریسک پروژه را کاهش داده و ارزش نهایی آن را افزایش دهد.

معماری راه‌حل و رویکرد فناورانه

معماری این پروژه بر پایه سه اصل طراحی شد: سادگی در ورودی، صراحت در پردازش، و قابلیت تفسیر در خروجی. مدل به‌گونه‌ای طراحی شد که بدون نیاز به برچسب یا نمونه مرجع، بتواند تنها با اتکا به رفتار داده در فضا و طیف، خوشه‌بندی مؤثری ایجاد کند.

تمام پردازش‌ها در یک چارچوب بازتولیدپذیر انجام شد. داده ورودی استاندارد شد، مدل روی بخش هدف داده آموزش دید و خروجی به صورت نقشه طیفی-مکانی ارائه شد. مدل هیچ فرضی درباره اینکه کدام نقطه دگرسان است نداشت و تمام تفکیک حاصل از خود داده استخراج شد.

این ساختار باعث می‌شود پروژه نه وابسته به شخص باشد و نه به یک نرم‌افزار خاص؛ بلکه قابل تکرار، توسعه و ترکیب با سایر داده‌های آینده باشد.

قیود اجرایی و چالش‌های عملی

پروژه با چالش‌هایی نیز همراه بود. عدم وجود برچسب واقعی موجب می‌شود ارزیابی مدل کاملاً داده‌محور باشد و نیاز به تفسیر زمین‌شناسی وجود داشته باشد. دامنه محدود مقدار نسبت‌بندی نیز باعث شد برخی مدل‌های پیشرفته نتوانند تفکیک مناسبی ارائه دهند. نویزهای طیفی و تفاوت رفتار پیکسل‌ها در مرزها نیز از چالش‌های اصلی بودند.

این تحلیل جایگزین کار صحرایی یا داده ژئوشیمی نیست؛ بلکه مکمل آن است. خروجی مدل نباید به تنهایی مبنای تصمیم نهایی باشد و باید در کنار دانش منطقه، برداشت صحرایی و سایر شواهد قرار گیرد.

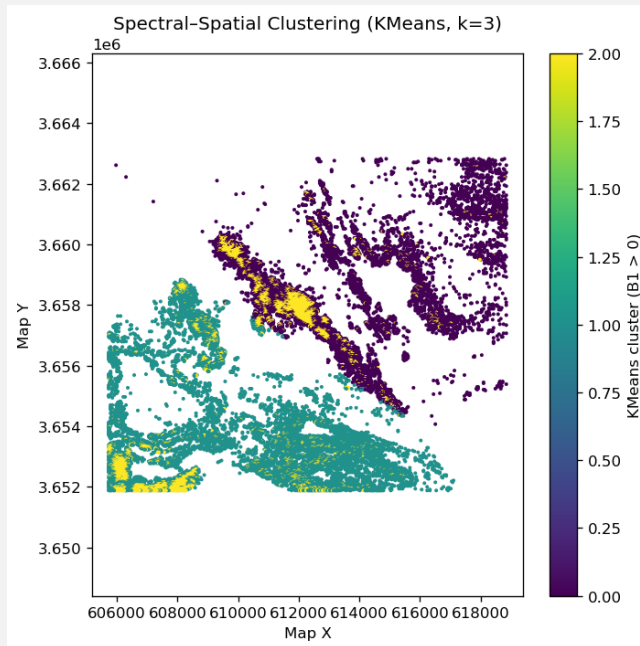
مراحل اجرایی پروژه

مراحل اجرایی شامل پنج فاز اصلی بود.

در فاز اول، داده‌های خام استخراج و پاک‌سازی شد. در فاز دوم، تحلیل اولیه داده و شناخت رفتار طیفی انجام شد. در فاز سوم، چندین مدل بدون‌ناظر اجرا و نتایج مقایسه شد. در فاز چهارم، تفسیر زمین‌شناسی و تحلیل ساختاری نتایج انجام گرفت. در نهایت در فاز پنجم، خروجی نهایی تهیه، مستندسازی و برای استفاده در فازهای بعدی پروژه آماده شد.

شاخص موفقیت و معیارهای ارزیابی

از آنجا که پروژه بدون برچسب انجام شد، شاخص‌های ارزیابی مبتنی بر رفتار داده و قابلیت تفسیر زمین‌شناسی بودند. کیفیت تفکیک خوشه‌ها، همپوشانی آن‌ها با ساختارهای خطی، شکل‌گیری هسته دگرسانی و پایداری مدل نسبت به داده‌های ورودی از مهم‌ترین معیارها بودند. این شاخص‌ها نشان دادند که مدل توانسته رفتار طیفی-مکانی دگرسانی را با دقت مناسبی استخراج کند.



خروجی نهایی مدل KMeans پس از اعمال آن بر داده‌های $B1 > 0$ ، سه خوشه اصلی را استخراج کرد که هر کدام رفتار طیفی-مکانی متفاوتی را نشان می‌دهند. بررسی آماری این سه خوشه نشان داد که میانگین مقدار طیفی در هر خوشه، اگرچه در نگاه اول اختلاف اندکی دارد، اما همین اختلاف‌های کوچک حامل معنای زمین‌شناسی قابل توجهی هستند. خوشه اول با ۷۲۳۹ پیکسل میانگین طیفی

حدود ۱۰۴۵۴ را نشان می‌دهد. خوشه دوم که بزرگ‌ترین خوشه منطقه است، با ۱۰۴۳۰ پیکسل میانگین طیفی ۱۰۴۵۰ دارد. این دو خوشه از نظر شدت طیفی بسیار به هم نزدیک‌اند و عمدتاً بخش‌های گسترده و کم‌قدرت‌تر دگرسانی را پوشش می‌دهند.

در مقابل، خوشه سوم که تنها شامل ۴۴۰۰ پیکسل است میانگین طیفی ۱۰۷۴۶ دارد و حداکثر مقدار طیفی آن به بیش از ۱.۱۴ می‌رسد. این بخش از داده نمایانگر شدیدترین زون دگرسانی پروپلیتیک در منطقه است و رفتار آن در تمام چارک‌ها - از کمترین مقدار تا بیشترین مقدار - نشان می‌دهد که این خوشه به‌طور واضح از دو خوشه دیگر جداست. درصد نسبی این خوشه در کل داده‌های هدف حدود ۱۹ درصد است، اما همین بخش کوچک‌تر از نظر زمین‌شناسی اهمیت بیشتری دارد، زیرا شدت دگرسانی و احتمال همراهی با کانی‌سازی در آن بالاتر است.

از آنجا که این مدل تنها از یک نسبت طیفی و مختصات استفاده کرده و هیچ داده ساختاری، لیتولوژیکی یا ژئوشیمیایی دریافت نکرده است، شکل‌گیری چنین خوشه‌ای که هم از نظر طیفی و هم از نظر مکانی معنادار است، نشان‌دهنده قابلیت الگوریتم در تشخیص الگوهای پنهان است. خوشه سوم توسط مدل به‌عنوان شدیدترین منطقه دگرسانی شناخته شد و این نتیجه با بررسی بصری نیز کاملاً هماهنگ است.

تصویر خروجی مدل KMeans این موضوع را به وضوح نشان می‌دهد. نقاط زردرنگ که نمایش‌دهنده خوشه سوم هستند، دقیقاً بر روی رگه‌ها، ساختارهای کششی و مرزهای سنگی قرار می‌گیرند؛ نقاطی که در تحلیل‌های کلاسیک سنجش‌ازدور نیز معمولاً به عنوان مناطق با پتانسیل بالای دگرسانی شناخته می‌شوند. این پراکنش خطی و متصل، رفتاری تصادفی نیست و کاملاً با الگوی شناخته‌شده دگرسانی پروپلیتیک در سیستم‌های پورفیری سازگار است.

از سوی دیگر، نقاط سبز و بنفش - یعنی خوشه‌های ۰ و ۱ - دامنه‌های وسیع‌تر و کم‌قدرت‌تر دگرسانی را تشکیل می‌دهند. این دو خوشه رفتار طیفی به‌مراتب یکنواخت‌تری دارند و بیشتر بخش‌های حاشیه‌ای و مرزی منطقه را پوشش می‌دهند. این ترکیب سه‌گانه باعث می‌شود نقشه خروجی نه تنها شدت را نشان دهد، بلکه «ساختار» را نیز آشکار کند؛ موضوعی که نرم‌افزارهای کلاسیک سنجش‌ازدور قادر به ثبت آن نیستند و معمولاً تنها مقدار طیفی را به‌صورت خام نمایش می‌دهند.

تحلیل بصری تصویر نشان می‌دهد که خوشه سوم بدون اینکه مدل هیچ اطلاعاتی از زون‌های ساختاری داشته باشد، دقیقاً بر مسیرهایی قرار گرفته که در برداشت میدانی معمولاً به‌عنوان نقاط کلیدی دگرسانی شناخته می‌شوند. این موضوع یکی از مهم‌ترین دستاوردهای پروژه است؛ مدل نه فقط مقدار طیفی، بلکه «معنا»ی مکانی آن را نیز درک کرده است.

این الگو در عمل می‌تواند راهنمای ارزشمندی باشد برای برنامه‌ریزی نمونه‌برداری ژئوشیمیایی، تعیین نقاط مناسب برای برداشت صحرایی و حتی انتخاب موقعیت‌های پیشنهادی برای حفاری اکتشافی. اگر در آینده داده‌های ژئوشیمی یا اطلاعات زمین‌شناسی با این نتایج ترکیب شود، انتظار می‌رود نقشه‌های پتانسیلی دقیق‌تری ساخته شود که توان هدایت عملیات اکتشاف را بیشتر کند.

چشم‌انداز و توسعه آینده

این پروژه تنها نقطه آغاز است. با افزودن داده‌های ژئوشیمی، اطلاعات ساختاری، DEM و سایر لایه‌های اکتشافی، می‌توان یک مدل چندمنبعی ایجاد کرد که قادر است مناطق با پتانسیل بالای دگرسانی را با دقت بیشتری مشخص کند. همچنین ترکیب این مدل با تکنیک‌های نیمه‌نظارتی می‌تواند امکان ایجاد یک مدل جامع پتانسیل‌یابی را فراهم کند.

در نهایت، هدف این پروژه تبدیل داده به دانش و ایجاد یک رویکرد هوشمند در اکتشاف است؛ رویکردی که نه جایگزین روش‌های کلاسیک است و نه در تقابل با آن‌ها، بلکه ابزاری مکمل برای تصمیم‌سازی دقیق‌تر و کاهش ریسک مالی و عملیاتی در پروژه‌های معدنی است.