

Análisis de muestreos geoquímicos de suelos superficiales usando métodos no supervisados de aprendizaje automático como estrategia para la prospección de yacimientos minerales en Perú.

Integrantes: [Diana Urbano](#), [Edinson Fernandez](#), [Daren Rodríguez](#)

1. Resumen:

Entre el año 2000 al 2019 se realizaron diversos muestreos geoquímicos por el Instituto Geológico, Minero y Metalúrgico de Perú, de los cuales se tomó una base de datos con 536 muestras de suelos superficiales para la prospección de yacimientos minerales.

En este análisis se usaron técnicas de aprendizaje automático no supervisado hallando 9 componentes principales que explican el 80% de la varianza y usando el clúster aglomerativo con mejores agrupaciones que el DBSCAN, hallando 13 clústeres de los cuales 4 tienen correlaciones significativas con elementos indicadores de yacimientos minerales, con el clúster 6 indicando la presencia de un yacimiento, asociado a Au, Ag, Hg, Bi mientras que los otros 2 clústeres (8 y 10) se asocian con elementos de interés como: Au, Ag, Cu, Zn, Pb, In entre otros, y en menor medida el clúster 9 se asocia con Mo, V, Ni, Tl indicando yacimientos posiblemente hidrotermales o se sulfuros masivos. Estas agrupaciones delimitan nuevas zonas de prospección minera de 50 km entre la zona de Huancayo y Jauja siguiendo la rivera del río Mantaro y otra zona puntual cerca a Cochamarca. Para estas dos zonas se recomienda realizar un mapeo geológico más detallado y decidir si se realizan estudios especializados.

2. Resultados y Discusión

Se encuentran 9 componentes principales, con lo cual se reducen las 69 variables a 9 conservando el 80% de la varianza de los datos.

Con los 9 componentes principales se genera un dendograma usando el método Ward y una distancia de 30, cortando 13 clústeres Figura 1.

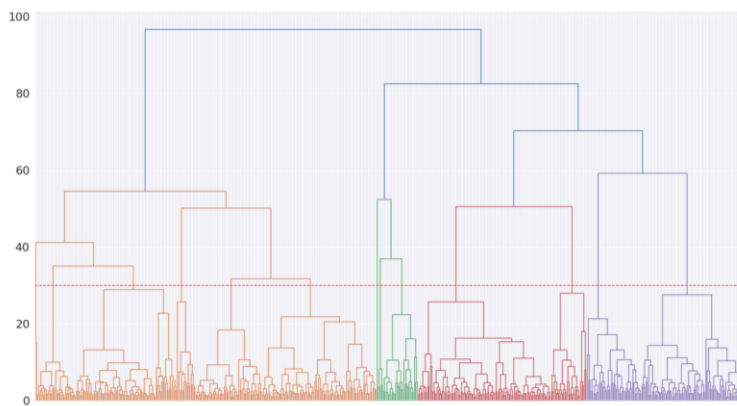


Figura 1. Dendograma usando las 9 componentes principales y una distancia de 30.

Usando el linkage Ward, los 9 componentes principales y 13 clústeres se desarrolla el algoritmo de clúster aglomerativo, Figura 2.



Figura 2. Agrupación de 13 clústeres usando el método aglomerativo.

Para el algoritmo DBSCAN se usa un eps de 4.48 usando la librería kneed para hallar los clústeres. Las gráficas muestran muy poca separación en los datos y tienden a agrupar todos los valores en un solo clúster por lo cual se descarta este método ya que el objetivo es ver diferencias entre los diferentes clústeres y su relación con yacimientos de interés, solo se logra una división en 3 clusters usando min samples =2 Figura 3.

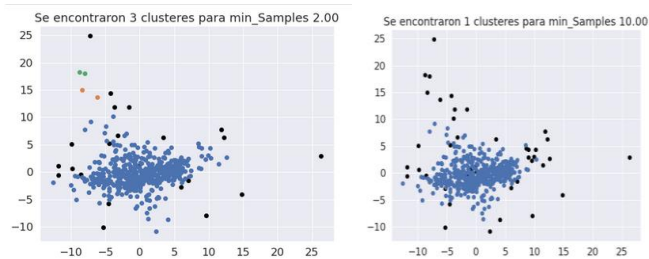


Figura 3. 3 clústeres usando el método DBSCAN para min samples =2.

De los 13 clústeres que se usan con el método aglomerativo, 4 tienen resultados prometedores debido a su asociación con minerales indicadores. El clúster 6 con un punto tiene una correlación moderada a muy alta con los elementos Au, Ag, Hg y Bi por lo que esta zona entre el sector de Cochamarca y la reserva de Junín puede ser un buen prospecto para la búsqueda de yacimientos de oro hidrotermales o de sulfuros masivos en la zona (Punto rojo) Figura 4. Los otros 2 clústeres 8 y 10 se asocian a sulfuros masivos e hidrotermales con elementos de interés como: Oro, Plata, Cobre, Zinc, Plomo, indio entre otros. Estas agrupaciones delimitan nuevas zonas de prospección minera o nuevos “Target” de 50 km entre la zona de Huancayo y Jauja, siguiendo la rivera del rio Mantaro (puntos amarillos y magenta) Figura 4. El clúster 9 con dos puntos presenta baja a moderada correlación con Mo, V, Ni, Tl que se asocian con yacimientos de oro y de sulfuros masivos en esta misma franja al sur de Huancayo (puntos aguamarina)Figura 4.

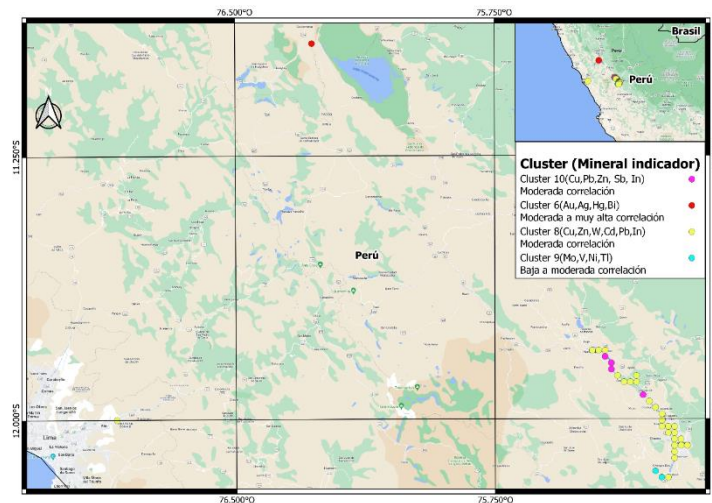


Figura 4. Los 4 clústeres que se pueden usar como indicadores de yacimientos minerales.

Para estas dos zonas se recomienda realizar un mapeo geológico más detallado preliminar y posteriormente decidir si ampliar el estudio con métodos geofísicos, sondajes exploratorios, así como campañas de muestreo más densas para delimitar aún más los posibles yacimientos minerales.

3. Conclusión

Se realizó el análisis de 536 muestras geoquímicas de suelo superficial en Perú con fines de prospección de yacimientos minerales, para este análisis se usaron técnicas de aprendizaje automático no supervisado hallando 9 componentes principales que explican el 80% de la varianza y usando el clúster aglomerativo con mejores agrupaciones que el DBSCAN, usando 13 clústeres de los cuales 4 tienen correlaciones significativas con elementos indicadores asociados a yacimientos minerales, de estos el clúster 6 puede indicar la presencia de un yacimiento, asociado a Au, Ag, Hg, Bi, mientras que los otros 2 clústeres (8 y 10) se asocian con elementos de interés como: Au, Ag, Cu, Zn, Pb, In entre otros, y en menor medida el clúster 9 se asocia a Mo, V, Ni, Ti indicando todos estos clúster yacimientos posiblemente hidrotermales o de sulfuros masivos. Estas agrupaciones delimitan una nueva zona de prospección minera de 50 km entre la zona de Huancayo y Jauja siguiendo la riva del río Mantaro y otra zona puntual cerca a Cochamarca. Para estas dos zonas se recomienda realizar un mapeo geológico más detallado preliminar y posteriormente decidir si ampliar el estudio con métodos geofísicos, sondajes exploratorios, así como campañas de muestreo más densas para delimitar aún más los posibles yacimientos minerales.

4. Bibliografía

Amat, J. (2020). Clustering con Python. <https://www.cienciadedatos.net/documentos/py20-clustering-con-python.html>

Balaram, V.; Sawant, S.S. (2021). Indicator Minerals, Pathfinder Elements, and Portable Analytical Instruments in Mineral Exploration Studies. *Minerals* 2022, 12, 394.

Fernández Jauregui, A. (s.f.). *DBSCAN en Python: aprende cómo funciona*. Recuperado el 9 de 2022, de <https://anderfernandez.com/blog/dbscan-py>

Instituto Geológico, Minero y Metalúrgico de Perú. (2021). GEOCATMIN - Prospección Geoquímica del Perú, Serie B. Obtenido de <http://metadatos.ingemmet.gob.pe:8080/geonetwork/srv/spa/catalog.search#/metadata/b1cc5e47-88c8-4c5e-832f-f6dcdbd20211b>

Levitán, D.M.; Zipper, C.E.; Donovan, P.; Schreiber, M.E.; Seal, R.R.; Engle, M.A.; Chermak, J.A.; Bodnar, R.J.; Johnson, D.K.; Aylor, J.G.(2015). Statistical analysis of soil geochemical data to identify pathfinders associated with mineral deposits: An example from the Coles Hill uranium deposit, Virginia, USA. *Journal of Geochemical Exploration*, 154, 238–251.

Nude, P. M., Asigri, J. M., Yidana, S. M., Arhin, E., Foli, G., & Kutu, J. M. (2012). Identifying pathfinder elements for gold in multi-element soil geochemical data from the Wa-Lawra belt, northwest Ghana: A multivariate statistical approach. *International Journal of Geosciences*, 3, 62-70.

Patel, A. (2019). *Hands-On unsupervised learning using Python: How to build applied Machine Learning Solutions from unlabeled data*. O'Reilly Media.

Román, V. (12 de junio de 2019). Aprendizaje no supervisado en Machine Learning: agrupación. <https://medium.com/datos-y-ciencia/aprendizaje-no-supervisado-en-machine-learning-agrupaci%C3%B3n-bb8f25813edc>.