

Objeto de estudio

 El objetivo es utilizar técnicas de clustering para descubrir agrupaciones naturales en los datos geoquímicos y clasificar las ubicaciones en clústeres que compartan características comunes. Esto permite una comprensión más profunda de la variabilidad espacial y la distribución de los elementos de ocurrencias de trazas en la cordillera occidental del Ecuador.



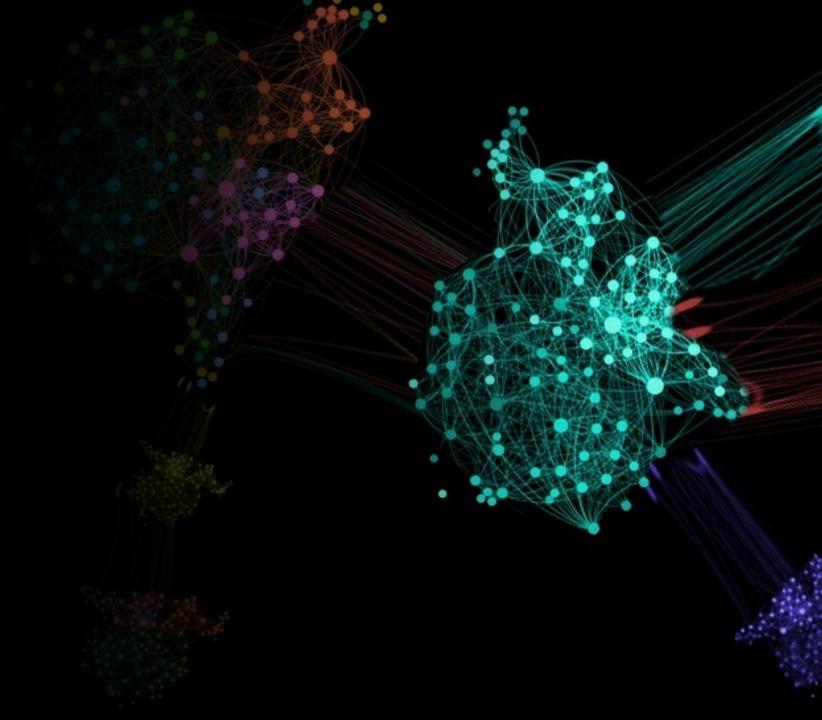
Plantamiento del problema

 Desarrollar un modelo de machine learning que pueda predecir y comprender las características y el comportamiento de los elementos geoquímicos de las ocurrencias de trazas en la Cordillera Occidentental del Ecudor.



Obejtivo Principal

Aplicar clustering en datos geoquímicos regionales para encontrar patrones y agrupaciones que ayuden a comprender mejor la variabilidad espacial y la distribución de los elementos de ocurrencias de trazas en la cordillera occidental. Esto proporciona información valiosa para la toma de decisiones y el desarrollo de estrategias en diferentes campos, incluyendo la geología, la minería y la conservación ambiental.



Objetivos específicos

- Realizar el procesamiento y limpieza de los datos geoquímicos, incluyendo la eliminación de valores faltantes o atípicos, y normalización de las variables si es necesario.
- Identificar las características geoquímicas más relevantes y significativas para el análisis de clustering. Esto podría incluir técnicas de selección de características, como análisis de componentes principales (PCA) o coeficientes de correlación.
- Utilizar algoritmos de clustering, como K-means, para agrupar las ubicaciones en clústeres basados en sus características geoquímicas. Estos algoritmos buscan encontrar grupos coherentes y compactos en función de la similitud de las observaciones.
- Evaluar la calidad y coherencia de los clústeres obtenidos utilizando métricas de evaluación de clustering, como el índice de Silhouette o el coeficiente de Calinski-Harabasz. Esto ayuda a determinar la calidad y robustez de los grupos identificados.

Revisión de Literatura

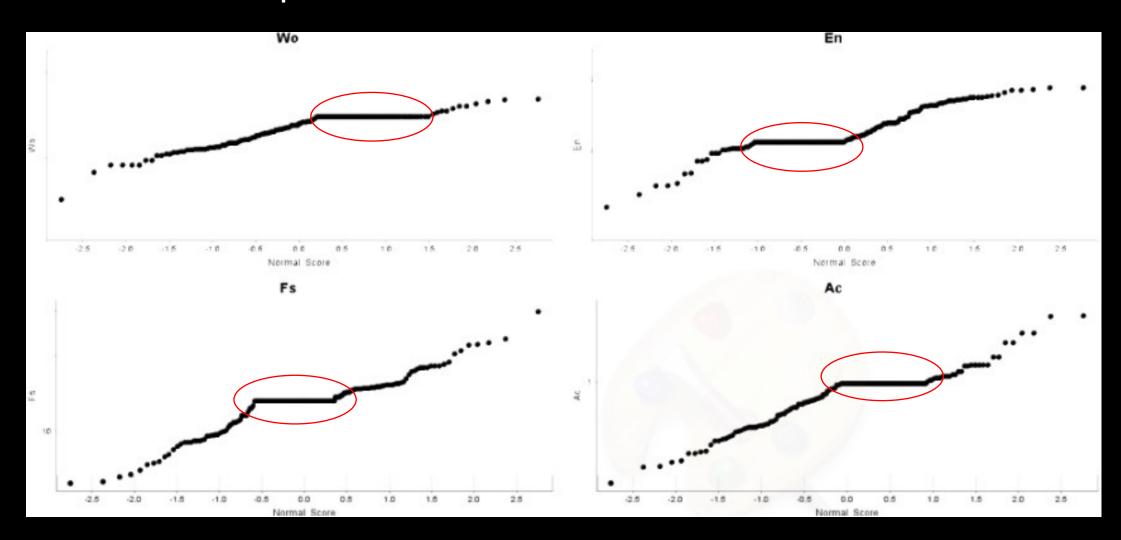
- Automatización del análisis exploratorio de datos y procesamiento geoquímico univariado empleando Python. (Castillo, 2023).
- A machine learning approach for regional geochemical data: Platinum-group element geochemistry vs geodynamic settings of the North Atlantic Igneous Province. (Lindsay, 2020).
- A study of the lake sediment geochemistry of the Melville Peninsula using multivariate methods: Applications for predictive geological mapping. (Grunsky C., 2014).
- Application of principal component analysis and cluster analysis to mineral exploration and mine geology. (Gazley, 2015).

Variables

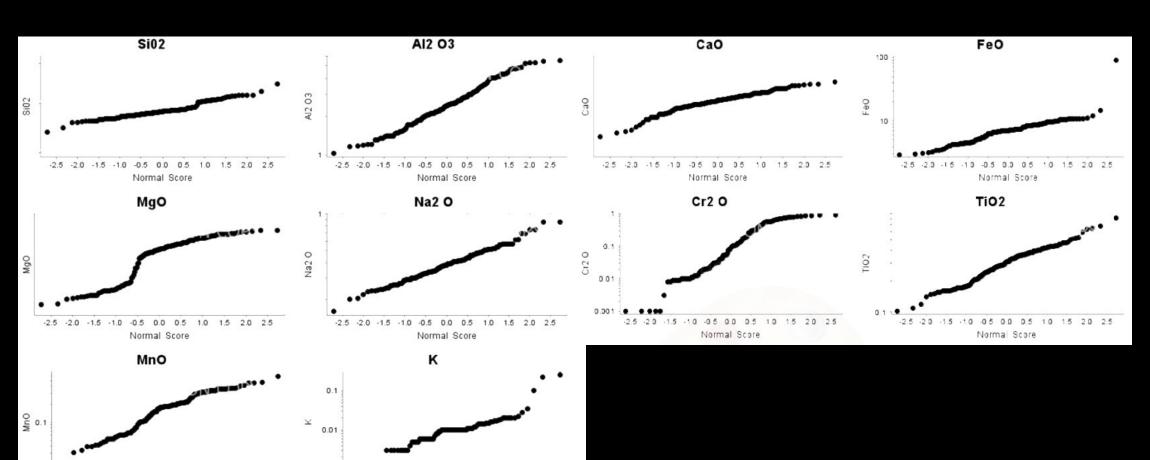
- Se tomaron los resultados geoquímicos de 10 muestras en un área de 27km2.
- Total de 26 variables, entre clinopiroxenos y oxidos.

Análisis exploratorio de Datos

Clinopiroxenos



Oxidos

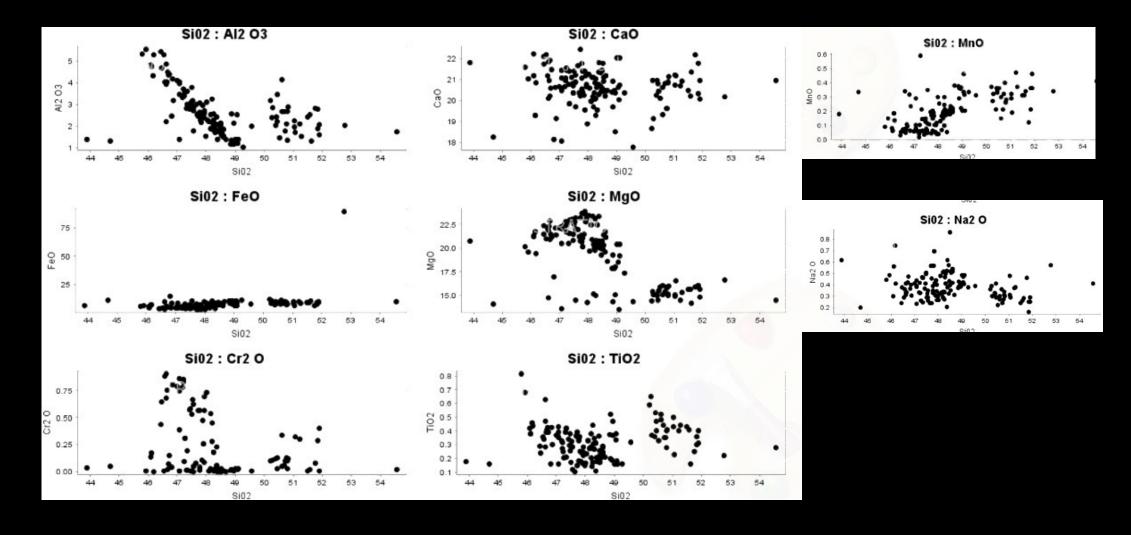


-2.5 -2.0 -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0 2.5

Normal Score

-2.5 -2.0 -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0 2.5 Normal Score

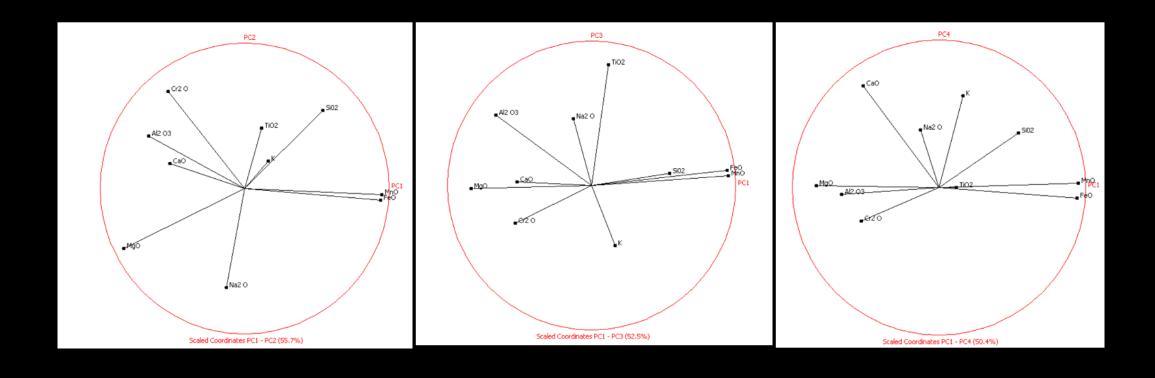
Gráficos de dispersión



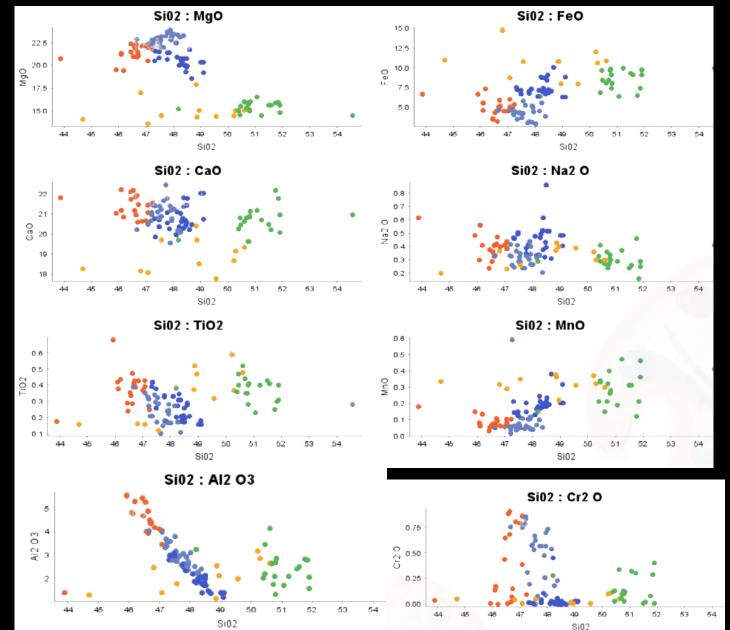
PCA

Scaled Coordinates	PC1	PC2	PC3	PC4
Si02	0.5423	0.5376	0.08229	0.3715
Al2 03	-0.6624	0.3627	0.4882	-0.04774
CaO	-0.5169	0.1736	0.02543	0.6908
K	0.164	0.1874	-0.4165	0.6214
Fe0	0.9418	-0.08167	0.1058	-0.07025
Mg0	-0.8372	-0.4111	-0.02018	0.01364
Na2 0	-0.1252	-0.681	0.464	0.3904
Cr2 O	-0.5298	0.6695	-0.2611	-0.2264
TiO2	0.1167	0.4139	0.8364	-7.1847E-4
MnO	0.9504	-0.04045	0.06932	0.03024

Relación de PCA con respecto a PCA1



Clusters con respecto a SiO2



Discusión

- El analisis llevado a cabo muestra resultados concluyentes que permiten caracterizar la composición de las rocas en cuestión como de origen volcánico intrusivo
- Es crucial destacar que este tipo de rocas no es el ambiente propicio para la formación de minerales económicamente rentables. Las características químicas y mineralógicas de las rocas influyen en la posibilidad de que se desarrollen depósitos minerales valiosos.
- Los procesos geológicos que dieron origen a estas rocas no implicaron condiciones adecuadas para la concentración y acumulación de minerales en niveles rentables.

Conclusiones y Recomendaciones

- El análisis realizado muestra resultados imporantes que facilitan la caracterización de la composición de las rocas estudiadas, originadas por procesos volcánicos intrusivos. Estos hallazgos están en línea con la clasificación de basaltos y sugieren que las rocas analizadas tienen características únicas pertenecientes a la formación geológica.
- La combinación de propiedades químicas y mineralógicas presentes en los basaltos andesíticos no coincide con las condiciones ideales para la acumulación y concentración de minerales de valor económico. Aunque estas rocas pueden tener relevancia en contextos geológicos y académicos, no son adecuadas como huéspedes para la formación de depósitos minerales explotables económicamente en la zona estudiada.

Referencias

Castillo, B. (2023). Automatización del análisis exploratorio de datos y procesamiento geoquímico univariado empleando Python.

Lindsay, J. (2020). A machine learning approach for regional geochemical data: Platinum-group element geochemistry vs geodynamic settings of the North Atlantic Igneous Province.

Grunsky, C. (2014). A study of the lake sediment geochemistry of the Melville Peninsula using multivariate methods: Applications for predictive geological mapping.

Gazley, C. (2015). Application of principal component analysis and cluster analysis to mineral exploration and mine geology.

Maimon, O. a. (2010). Data Mining and Knowledge Discovery Handbook. Israel: Oded Maimon and Lior Rokach Tel-Aviv University.

Pérez, L. (2022). Análisis de clústers K-means.

Grunsky, E. (2012). The interpretation of geochemical survery data. *Geochemistry: Exploration, Environment, Analysis 2012*, 27-74.

Davies, D. (1979). A Cluster Separation Measure(Article). En IEEE Transactions on Pattern Analysis and Machine Intelligence (Vols. PAMI-1, págs. 224-227).

Balamurali, M. (2016). t-SNE Based Visualisation and Clustering of Geological Domain. Neural Information Processing, 565-572.

Farnham, I. (2002). Treatment of nondetects in multivariate analysis of groundwater geochemistry data. Chemometrics and Intelligent Laboratory Systems, 265-281.

Ghannadpour, S. (2013). An Investigation of Pb Geochemical Behavior Respect to Those of Fe and Zn Based on k-Means Clustering Method. En Quartery Journal of Tethys (Vol. 1, págs. 291-2013)

Cahoon, J. (2021). Generalized inferential models for censored data. En I. J. Reasoning.

Hastie, T., & Tibshirani, R. (2009). The elements of statistical learning: data mining, inference, and prediction. Springer.

Mahmut, M. (2023). Basalt.