



DESARROLLO DE MODELOS DE QUANTUM MACHINE LEARNING NO SUPERVISADO PARA CLUSTERING DE DATOS GEOQUÍMICOS

Matías Orellana¹; Elizabeth Lam¹; Brian Keith¹

¹Universidad Católica del Norte, Av. Angamos 0610, Antofagasta, Chile.

brian.keith@ucn.cl

Propósito: Esta investigación desarrolla modelos de *Quantum Machine Learning* (QML) no supervisado para analizar datos geoquímicos de suelos en la región de Antofagasta, Chile. El objetivo principal es superar las limitaciones de los métodos tradicionales de *Machine Learning* en el análisis de datos escasos y ruidosos, aprovechando las ventajas computacionales de los sistemas cuánticos. Se busca identificar patrones de contaminación por elementos como As, Pb, Cu y otros metales pesados, facilitando procesos de monitoreo ambiental y toma de decisiones en áreas afectadas por actividad minera.

Material y Métodos: Se utilizaron 643 muestras geoquímicas provenientes de WSP-EMGRISA, CICITEM y CENMA, conteniendo concentraciones de 16 elementos químicos. Se implementaron tres algoritmos cuánticos principales: Quantum K-Means (con codificación angular y de amplitud), Quantum Boltzmann Machines (QBM) y Quantum Approximate Optimization Algorithm (QAOA). Los datos se procesaron mediante la metodología CRISP-DM, aplicando técnicas de normalización y división por bloques para optimizar el procesamiento cuántico. Se evaluaron los modelos mediante métricas estándar de clustering.

Resultados: Los modelos QML mostraron capacidad para identificar entre 3-6 clusters óptimos según el algoritmo. El QML K-Means con codificación angular por columnas alcanzó el mejor Silhouette Score (0.648), superando al ML tradicional (0.385). Los tiempos de ejecución variaron significativamente: ML tradicional (0.43s), QBM (12-46s), hasta QAOA (467-1128s). La división en bloques redujo considerablemente los tiempos computacionales manteniendo la calidad del clustering. Los valores de inercia fueron menores en modelos cuánticos (1.12-7.70) comparados con el tradicional (41.88-72.10), indicando clusters más cohesionados. La geovisualización reveló patrones espaciales coherentes en las comunas de la región de Antofagasta.

Conclusiones: Los modelos QML demostraron ventajas en la calidad del clustering sobre métodos tradicionales, especialmente en métricas de cohesión y separación. Sin embargo, requieren mayor tiempo computacional y conocimientos especializados en física cuántica y matemáticas avanzadas. La estrategia de división por bloques resultó efectiva para escalar los algoritmos cuánticos. Aunque prometedores, los modelos QML aún enfrentan desafíos de implementación práctica, requiriendo Python 3.10 y configuraciones específicas del entorno. El proyecto aporta documentación valiosa sobre errores comunes y recomendaciones para futuros desarrollos en QML aplicado a geociencias.

Palabras clave: quantum machine learning, clustering, datos geoquímicos, contaminación minera

Línea temática: Propiedades y Procesos del Suelo — Química del suelo.