Quantum-Enhanced Clustering of LA-ICP-MS Geochemical Data Using QCNN

Dr. Siddhartha Agarwal
Indian Institute of Technology (ISM) Dhanbad
Dhanbad, Jharkhand - 826004
sagarwal@iitism.ac.in

Gowtham Raju
Indian Institute of Technology, Ropar
Rupnagar, Punjab - 140001
2022mmb1378@iitrpr.ac.in

Abhishek Kumar
Indian Institute of Technology, Ropar
Rupnagar, Punjab - 140001
2022mmb1369@iitrpr.ac.in

Dr. Upama Dutta
Indian Institute of Technology (ISM) Dhanbad
Dhanbad, Jharkhand - 826004
upama@iitism.ac.in

Key Words: Critical Minerals, *Mineral Informatics, Quantum machine learning (QML), Quantum Convolutional Neural Network (QCNN), PennyLane*

1. Problem Description and Relevance

The world recognizes that critical minerals are the building blocks of future digital and energy-secure economies. An accurate mineral reserve estimation is necessary to initiate mining of critical minerals in a region based on economic feasibility. LA-ICP-MS (Laser Ablation–Inductively Coupled Plasma–Mass Spectrometry) [1] is an instrument that plays a crucial role in estimating and characterizing critical mineral reserves from drillhole samples, offering a unique combination of precision, spatial resolution, and multi-element detection, all of which are essential for modern mineral exploration and resource evaluation.

The datasets generated by LA-ICP-MS have high dimensionality, usually ranging from 30 to 60 elemental variables per sample[5]. This involves multicollinearity and element interdependency, which is a major hurdle in building predictive models. Another significant issue is that elemental concentrations are relative and always sum up to a constant total (100% on a ppm basis). This creates spurious correlations unless data transformation techniques are applied. Finally, identifying zones or clusters within the data is a tedious process. It requires domain expertise in handling geological and geochemical data, as this may involve non-linearities and anomalies, such as the presence of critical minerals at very low concentrations, which are often mixed with high background noise from major elements [2]. Mineral informatics is a new field that utilizes machine learning and artificial intelligence algorithms to identify clusters or patterns of mineralized zones within samples. Traditional methods of unsupervised clustering often fail to recognize the complex patterns that exist in mineral occurrences. The novelty of this research lies in the fact that:

1. This is the first attempt to classify minerals and determine the percentage proportions using quantum unsupervised clustering algorithms [3] from data generated by LA-ICP-MS [4]
2. Given a comparison of the results with classical clustering techniques
3. Automating the process and uncovering new strategies for mineral classification[12]

2. Methodology

The dataset used in the study comes from samples taken at the Yerington district, which is a significant host of Porphyry Cu-Mo-Au deposits that have critical Minerals in trace amounts [5]. The data consists of approximately 66,000 samples with x and y locations and 11 elements, as shown in Table 1. Figure 1 illustrates the process applied to the raw elemental data from raster maps, which are transformed using centered log-ratio (CLR) normalization and reduced via PCA before quantum encoding. A variational quantum circuit (VQC) was built with PennyLane (an open-source quantum simulator and machine learning framework) that processes the data, and clustering was done using hybrid quantum-classical convolutional neural networks (HQCNN)[3]. The performance is evaluated using silhouette-optimized embeddings, loss curves, and cluster maps.

Table 1: List of elements present in all samples (concentration in ppm)

Sample ID	x	y	Al_	Mn_	Fe_	As_	Sr_	Sb_	Ce_	Pb_	Bi_	Th_	U_	Total
YE16A A096	476.3842	24	1083 97	56.25 126	6646 6.2	2.936 372	2308. 007	1.074 556	1.617 326	0.729 648	0.361 722	0.219 928	1.372 077	1772 35.7
YE16A A096	478.3943	24	1083 97	88.64 346	6066 5.38	2.936 372	2285. 809	1.074 556	4.482 281	0.729 648	0.361 722	0.219 928	1.372 077	1714 48
YE16A A096	474.3742	18	1138 33.5	194.7 764	6139 0.38	2.936 372	2363. 76	1.074 556	4.482 281	2.472 21	0.361 722	0.219 928	1.849 427	1777 95.8

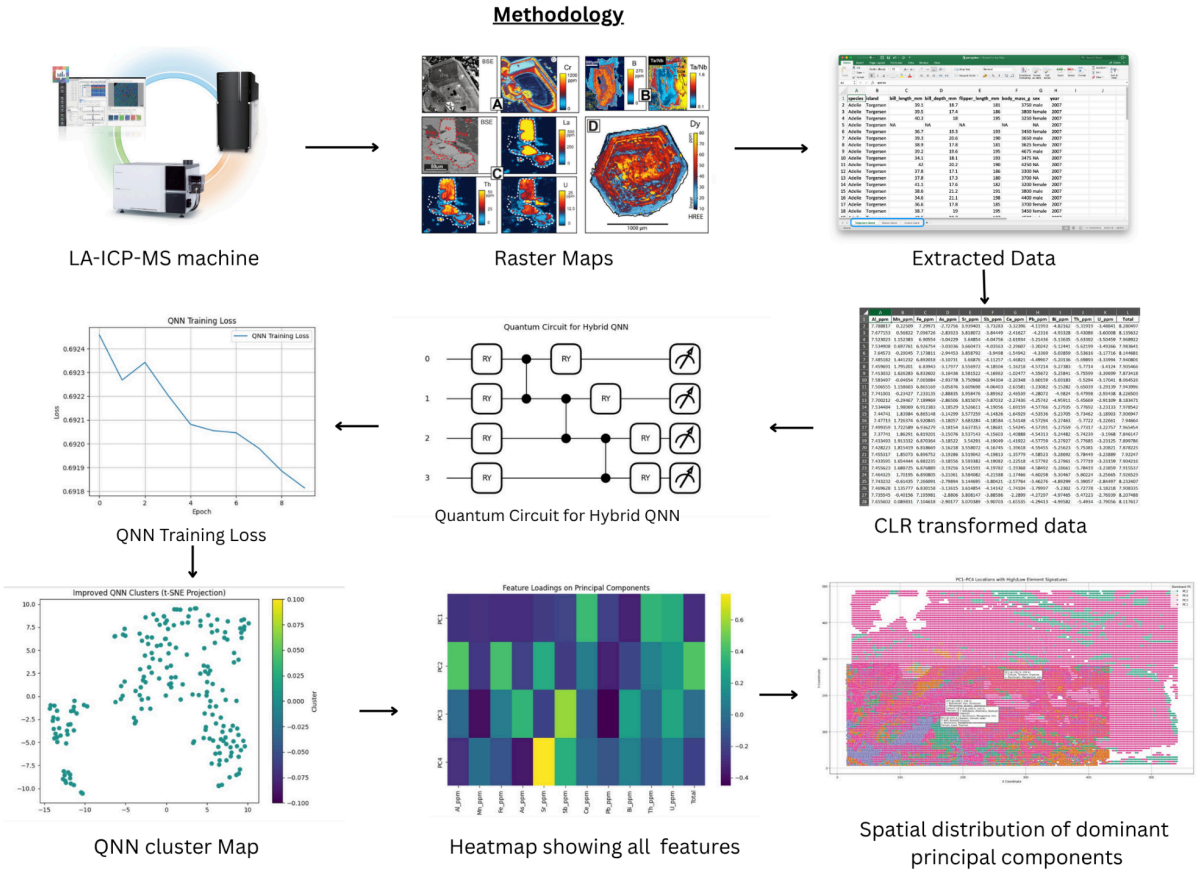

Figure 1. Overview of the proposed HQCNN pipeline for clustering LA-ICP-MS-derived geochemical data.

The pipeline begins by preprocessing raw trace element concentration data using the CLR transformation to manage the closed-sum constraint typical of compositional datasets. The data

is then standardized and reduced using PCA to 4 components for quantum encoding. A quantum circuit (as shown in Figure 2) is designed with four qubits, each initialized with RY rotations based on PCA-transformed features[5]. An entanglement layer using a controlled-Z (CZ) chain is applied, followed by trainable RY rotations to form a VQC. The quantum node outputs the expectation values of the Pauli-Z operator for each qubit, yielding a 4-dimensional vector representing a nonlinearly projected quantum feature space.

This quantum output is passed through a second classical layer to predict cluster labels. We compare this prediction against K-Means-derived clusters and compute Silhouette scores to evaluate performance. We intentionally avoid quantum preprocessing (e.g., QPCA), thereby preserving a hybrid structure that supports both classical preprocessing and quantum-enhanced inference.

3. Practical Demonstration

To validate our method, we implemented the entire quantum-classical pipeline using the PennyLane *default.qubit* simulator, which realistically simulates quantum behavior for small quantum circuits (up to ~20 qubits). We trained the HQCNN model on a real-world dataset of 66157 geochemical samples [5].

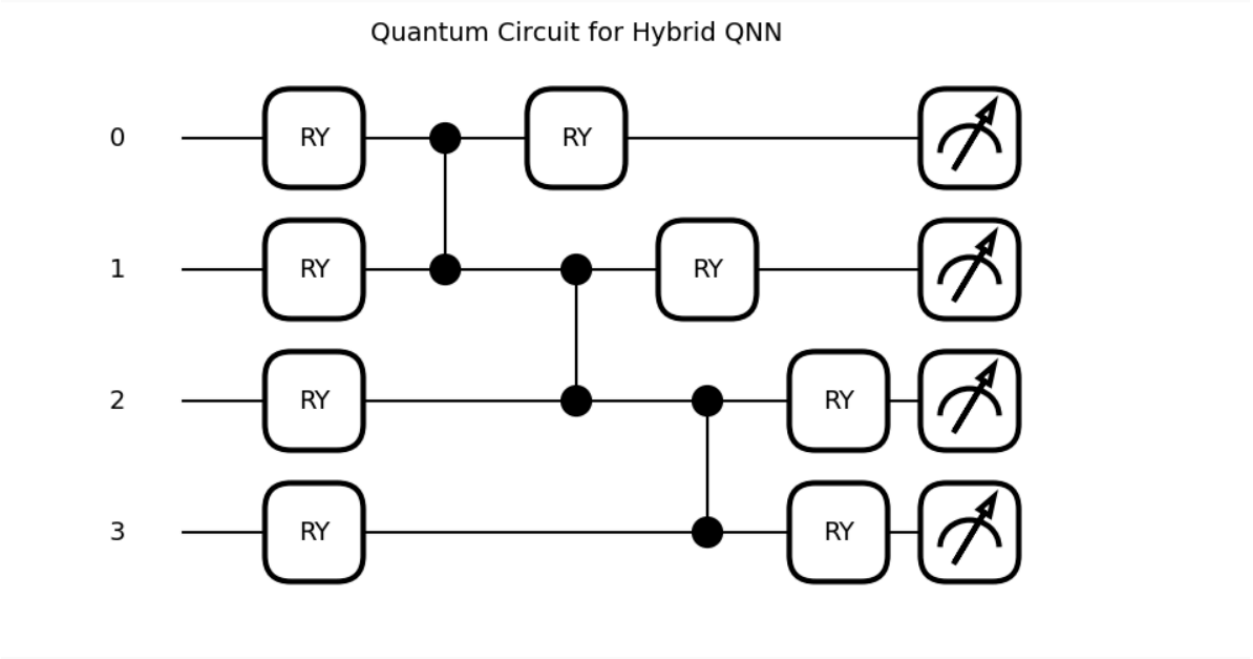

Figure 2. Quantum circuit architecture for the Hybrid Quantum Neural Network (HQNN

4. Application Potential

The quantum circuit is visualized using *draw_mpl()* and saved as a figure, demonstrating the actual structure used for feature projection. Quantum outputs from this circuit vary nonlinearly with input PCA vectors, demonstrating their ability to differentiate between subtle geochemical classes better than K-Means in our test case [3], [6], [7]. We compare QNN-derived cluster labels against those from classical K-Means using 2D and 3D PCA scatter plots, as well as t-SNE projections, and visualize them with Matplotlib and Seaborn [8]. We also include a loss curve during training to demonstrate model convergence and a feature loading heatmap for PCA

components, which facilitates the interpretation of classical transformations. Although we haven't deployed on actual quantum hardware yet, our approach is fully compatible with IBM Q or Xanadu's photonic devices. It can be ported with minimal changes using the same PennyLane framework.

Table 2: Clusters of elements predicted in all samples (concentration in ppm)

Cluster 1	**High Conc.**	Cerium	Thorium	Uranium					
	Low Conc.	Aluminum	Manganese	Iron	Arsenic	Strontium	Antimony	Lead	Bismuth
Cluster 2	**High Conc.**	Aluminum	Iron	Strontium					
	Low Conc.	Manganese	Arsenic	Antimony	Cerium	Lead	Bismuth	Thorium	Uranium
Cluster 3	**High Conc.**	Iron	Arsenic	Antimony	Bismuth				
	Low Conc.	Aluminium	Manganese	Strontium	Cerium	Lead	Thorium		
Cluster 4	**High Conc.**	Strontium	Antimony	Bismuth	Uranium				
	Low Conc.	Aluminum	Manganese	Iron	Arsenic	Cerium	Lead	Thorium	

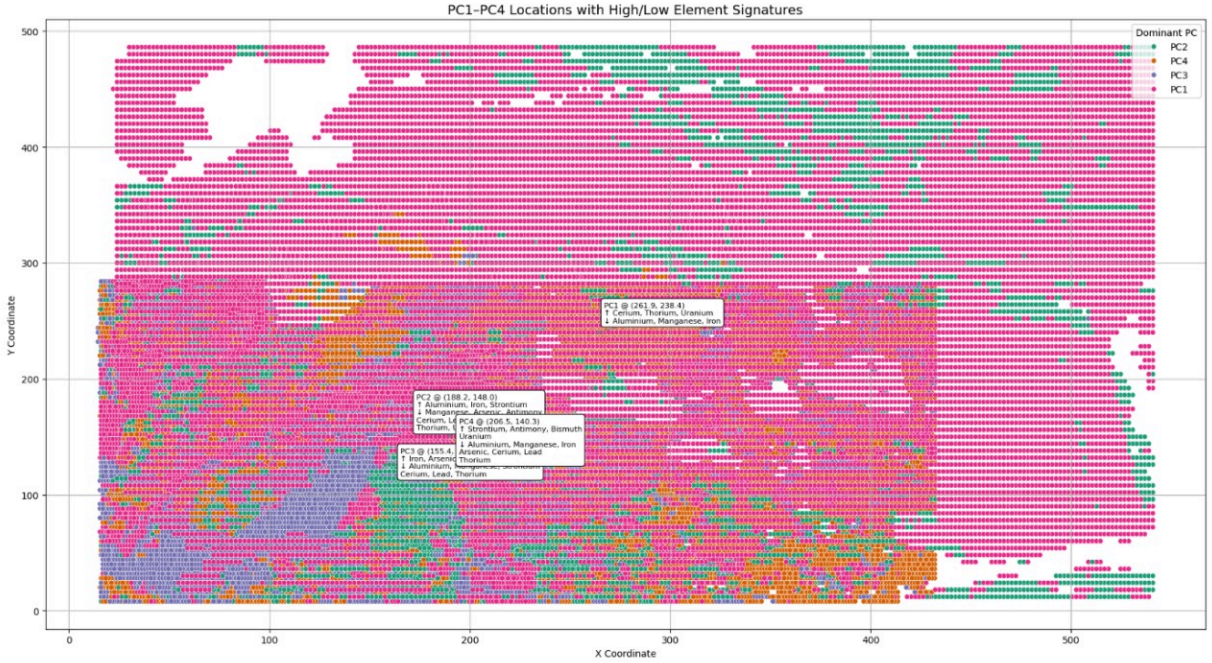

Figure 3. The four clusters are represented spatially on the x, y plane, and the colors indicate the concentration, showing which element is in high proportion and which is low.[5]

The other objective is to establish a technology, testing, and certification to fill the gap of advanced technologies required for characterizing the raw materials/ores for amenable extraction of REEs, which is currently lacking [1], [4]. Moreover, the global focus shifts toward space exploration and sampling extraterrestrial materials from planets, asteroids, and other celestial bodies; this method can play a pivotal role in preparing the world for a new frontier. By developing expertise in advanced material characterization, geochemical analysis, and mineralogical studies, researchers will be equipped to contribute to international space missions and planetary exploration efforts [12].

Presentation Preference:

☑ Oral Presentation
☑ Poster Presentation

REFERENCES

[1] Mo, X., Gong, C., Shang, Y., Wu, J., Wu, J., Qi, R., Wang, X., Guan, Q., & Kong, X. (2025). Zircon LA-ICP-MS Dating and Geochemical Characteristics of Rhyolites from the Qushi Area, Tengchong Terrane, Yunnan Province. Minerals, 15(3), 315.

[2] Jahangiri, M., Ghavami Riabi, S.R., & Tokhmechi, B. (2018). Estimation of geochemical elements using a hybrid neural network–Gustafson-Kessel algorithm. Journal of Mining & Environment, 9(2), 499-511.

[3] Haug, T., et al. (2024). qCLUE: a quantum clustering algorithm for multi-dimensional datasets. Frontiers in Quantum Science and Technology, 10.3389/frqst.2024.14620043

[4] Machado, N. (1997). A new technique for the synthesis of geochemical reference samples for laser ablation-ICP-MS analysis of zircons. Chemical Geology, 138, 1-7.4

[5] Unsupervised clustering of LA-ICP-MS raster map data for geological interpretation: A case study using epidote from the Yerington district, Nevada. (2020). .

[6] Horn, D. & Gottlieb, A. (2001). Algorithm for Data Clustering in Pattern Recognition Problems Based on Quantum Mechanics. Physical Review Letters.

[7] Kerenidis, I., Landman, J., Luongo, A., & Prakash, A. (2018). q-means: A quantum algorithm for unsupervised machine learning. arXiv, Dec 10, 2018.

[8] Sarma, A., Chatterjee, R., Gili, K., & Yu, T. (2019). Quantum Unsupervised and Supervised Learning on Superconducting Processors. arXiv, Sep 2019.

[9] Li, Q., Huang, Y., Jin, S., Hou, X., & Wang, X. (2022). Quantum spectral clustering algorithm for unsupervised learning. arXiv, Mar 2022.

[10] Frontiers in Quantum Science and Technology (2024). qCLUE: a quantum clustering algorithm for multi-dimensional datasets.

[11] LA-ICP-MS imaging in geosciences and its applications to geochronology. David Chew, Kerstin Drost, Jeffrey H. Marsh , Joseph A. Petrus

[12] Prabhu, A., Morrison, S. M. M., Fox, P., Ma, X., Wong, M. L. L., Williams, J. R. R., McGuinness, K. N. N., Krivovichev, S. V. V., Lehnert, K., Ralph, J., Lafuente, B., Downs, R. T. T., Walter, M. J. J., & Hazen, R. M. M. (2023). What is mineral informatics?