EGU23-10228
EGU General Assembly 2023

Interpreting Machine Learning Models for Geochemistry Data Classification using Decision Boundary Maps

Yu Wang[1,2], Kunfeng Qiu[1], Alexandru Telea[2], Zhaoliang Hou[3], and Haocheng Yu[1]
[1]School of Earth Sciences and Resources, China University of Geosciences, Beijing, China
[2]Department of Information and Computing Sciences, Utrecht University, Utrecht, Netherlands
[3]Department of Geology, University of Vienna, Vienna, Austria

Machine learning has been shown to be a highly effective method for classifying geochemistry data, such as mineral forming environments and rock tectonics. However, it can be difficult to understand the decision-making processes of these models. To address this issue, we propose the use of Decision Boundary Maps (DBMs) as a visualization tool for interpreting machine learning models. These maps project high-dimensional geochemistry data onto a 2D plane and depict the decision boundaries in the projected space, providing a visual representation of the algorithm's decision-making processes. In addition, DBMs can reveal trends, correlations, and outliers in the data, helping to interpret the results obtained from machine learning-based geochemistry data classification. Seeing the positions of data points, rather than just class labels, is especially valuable because samples in geological categories often follow a sequence, such as a magmatic to hydrothermal transition. Observing the positions of data points allows for the identification of trends from one class to an adjacent class.