Goldschmidt 2022 Abstract
https://doi.org/10.46427/gold2022.11779

Explainable Machine Learning in Uncovering H Diffusion Mechanism in Clinopyroxene

AN-ZHOU LI, SEN-SEN WU, HUAN CHEN, ZHEN-HONG DU AND PROF. QUN-KE XIA, PHD

Zhejiang University

Presenting Author: anzhouli@zju.edu.cn

Estimating the water content of mantle-derived magma through clinopyroxene (cpx) phenocrysts provides an effective way for constraining the water content in deep Earth. However, traditional geochemical approaches present uncertainties in determining whether primary water concentration has been preserved in cpx. Although previous study has applied machine learning (SVM) to build a classifier judging the occurance of H diffusion, it cannot show the basis of classification since most advanced machine learning models are hard-to-debug black boxes. Shapley Value in coalition game theory proposes a model-agnostic perspective solving feature contributions in a linear additive manner. We conduct XGBoost, a tree-based regularized gradient boosting framework, on a global dataset of cpx phenocrysts in basaltic melts to accurately model the process of H diffusion. Upon the Shapley value and the explainability of trees model, our study provides sample-level feature importance quantitatively. In aggregation up local attribution values, we show that Na takes the first place among major elements of cpx, indicating that the H associated with Na is easy to diffuse out of cpx. More importantly, through the interaction effects within pair-wise elements of Shapley value, we find that some high Mg content samples with high water content show a signal of H diffusion. This implies that during magma evolution, H probably diffuses into the early crystalized cpx, which leads to an overestimation of deep Earth water content. Our study first-time couples explainable machine learning models and geochemical principles to explain Hydrogen diffusion in minerals, suggesting broad implications in understanding the water content in deep Earth.