American Mineralogist, Volume 109, pages 1394–1405, 2024

Machine learning applied to apatite compositions for determining mineralization potential

Yu-yu Zheng[1,2,†], Bo Xu[1,2,*], David R. Lentz[3], Xiao-yan Yu[2], Zeng-qian Hou[4], and Tao Wang[4]

[1]State Key Laboratory of Geological Processes and Mineral Resources and Frontiers Science Center for Deep-time Digital Earth, China University of Geosciences, Beijing 100083, China
[2]School of Gemology, China University of Geosciences, Beijing 100083, China
[3]Department of Earth Sciences, University of New Brunswick, P.O. Box 4400, Fredericton, New Brunswick E3B 5A3, Canada
[4]Beijing SHRIMP Center, Institute of Geology, Chinese Academy of Geological Sciences, Beijing 100037, China

Abstract

Apatite major and trace element chemistry is a widely used tracer of mineralization as it sensitively records the characteristics of the magmatic-hydrothermal system at the time of its crystallization. Previous studies have proposed useful indicators and binary discrimination diagrams to distinguish between apatites from mineralized and unmineralized rocks; however, their efficiency has been found to be somewhat limited in other systems and larger-scale data sets. This work applied a machine learning (ML) method to classify the chemical compositions of apatites from both fertile and barren rocks, aiming to help determine the mineralization potential of an unknown system. Approximately 13 328 apatite compositional analyses were compiled and labeled from 241 locations in 27 countries worldwide, and three apatite geochemical data sets were established for XGBoost ML model training. The classification results suggest that the developed models (accuracy: 0.851–0.992; F1 score: 0.839–0.993) are much more accurate and efficient than conventional methods (accuracy: 0.242–0.553). Feature importance analysis of the models demonstrates that Cl, F, S, V, Sr/Y, V/Y, Eu*, $(La/Yb)_N$, and La/Sm are important variables in apatite that discriminate fertile and barren host rocks and indicates that V/Y and Cl/F ratios and the S content, in particular, are crucial parameters to discriminating metal enrichment and mineralization potential. This study suggests that ML is a robust tool for processing high-dimensional geochemical data and presents a novel approach that can be applied to mineral exploration.

Keywords: Apatite, major and trace element, machine learning, mineralization potential, XGBoost

Introduction

Apatite ($Ca_5[PO_4]_3[F,Cl,OH]$) is a ubiquitous accessory mineral in most igneous and metamorphic rocks and derived clastic sediments and is relatively resistant to weathering (O'Sullivan et al. 2020). It is considered to be an ideal indicator mineral, given its chemical composition sensitivity to the crystallization environment (Bruand et al. 2017; Mao et al. 2016). Trace elements, volatile chemistry, and isotopic signatures of apatites can characterize diverse crystallization environments, including magmatic systems (Cao et al. 2022; Gao et al. 2020; Li et al. 2021; Long et al. 2023; Palma et al. 2019; Qu et al. 2021; Tang et al. 2021; Xu et al. 2023; Zhang et al. 2021), low-grade metamorphic systems (Bea and Montero 1999; EL Korh et al. 2009; Henrichs et al. 2018; Nutman 2007), and sedimentary environments (Joosu et al. 2016). Accordingly, the trace element chemistry of apatite is widely used to characterize the lithology of source rocks (Belousova et al. 2002), including tracing detrital provenance (Bruand et al. 2017; Dill 1994; O'Sullivan et al. 2018, 2020), and used to constrain petrogenetic processes (Chu et al. 2009; La Cruz et al. 2020; Sun et al. 2022; Tollari et al. 2008; Zafar et al. 2019), especially for revealing the origin and evolution of magma (Gao et al. 2020; Meng et al. 2021; O'Reilly and Griffin 2000).

The major and trace element chemistry of apatite is applied to mineral exploration (Belousova et al. 2002; Cao et al. 2012; Mao et al. 2016; Sha and Chappell 1999; Xu et al. 2015). A series of indicators, including Sr/Y, Mn, Eu/Eu*, Th/U, La/Sm, and $(Ce/Yb)_N$ (Belousova et al. 2002), and several binary classification diagrams, such as Sr vs. F [Mn, Y, $(La/Yb)_N$, Eu/Eu*], F/Cl vs. F (Azadbakht et al. 2018; Cao et al. 2012; Zhong et al. 2018), Cl vs. Eu/Eu* (Mao et al. 2016), V/Y vs. REE+Y, Cl vs. SO_3, and $^{87}Sr/^{86}Sr$ vs. Cl/F, are commonly used to diagnose the metallogenic fertility of magmatic rocks. (Xu et al. 2021). Unfortunately as interest in apatite has recently increased and numerous major and trace element data have been reported (Adlakha et al. 2018; Bruand et al. 2019; Cao et al. 2022; Chakhmouradian et al. 2017; Chen and Zhang 2018; Chen et al. 2020; Gao et al. 2020; Glorie et al. 2019; Henrichs et al. 2018; Hoshino et al. 2017; La Cruz et al. 2020; Li et al. 2021; Liu et al. 2021; Long et al. 2023; Lupulescu et al. 2017; Meng et al. 2021; Mercer et al. 2020; Palma et al. 2019; Qu et al. 2021; Sun et al. 2022; Tang et al. 2021; Xie et al. 2018; Xu et al. 2023; Yang et al. 2018; Zafar et al. 2019; Zhang et al. 2021), it is challenging to validate these individual indicators and binary discrimination techniques due to a large overlap of compositional spots, suggesting that those traditional low-dimensional classifiers that seemed to work well in specific

* Corresponding author E-mail: xubo@outlook.com.cn
† Orcid https://orcid.org/0000-0003-1945-3440

0003-004X/24/0008–1394$05.00/DOI: https://doi.org/10.2138/am-2023-9115