This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

Revision 1

Apatite trace element composition as an indicator of ore deposit types: a machine learning approach

Kun-Feng Qiu[1], Tong Zhou[1], David Chew[2], Zhao-Liang Hou[3], Axel Müller[4,5], Hao-Cheng Yu[1], Robert G. Lee[6], Huan Chen[7], Jun Deng[1*]

[1]*State Key Laboratory of Geological Processes and Mineral Resources, School of Earth Science and Resources, China University of Geosciences, Beijing, Chin*

[2]*Department of Geology, School of Natural Sciences, Trinity College Dublin, Dublin, Ireland*

[3]*Department of Geology, University of Vienna, Vienna, Austria*

[4]*Natural History Museum, University of Oslo, Oslo, Norway*

[5]*Natural History Museum, London, UK*

[6]*Mineral Deposit Research Unit (MRDU), The University of British Columbia, Main Mall, Vancouver, British Columbia, Canada*

[7]*Institute of Marine Geology, College of Oceanography, Hohai University, Nanjing, China*

*Corresponding author

Jun Deng djun@cugb.edu.cn

Professor, China University of Geosciences, Beijing

No. 29 Xueyuan Road, Haidian District, Beijing, 100083, P.R. China

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

Abstract

The diverse suite of trace elements incorporated into apatite in ore-forming systems has important applications in petrogenesis studies of mineral deposits. Trace element variations in apatite can be used to distinguish between fertile and barren environments, and thus have potential as mineral exploration tools. Such classification approaches commonly employ two-variable scatterplots of apatite trace element compositional data. While such diagrams offer accessible visualization of compositional trends, they often struggle to effectively distinguish ore deposit types because they do not employ all the high-dimensional (i.e. multi-element) information accessible from high-quality apatite trace element analysis. To address this issue, we use a supervised machine learning-based approach (eXtreme Gradient Boosting, XGBoost) to correlate apatite compositions with ore deposit type, utilizing such high-dimensional information. We evaluated 8629 apatite trace element data from five ore deposit types (porphyry, skarn, orogenic Au, iron oxide copper gold, and iron oxide-apatite) along with unmineralized magmatic and metamorphic apatite to identify discriminating parameters for the individual deposit types as well as for mineralized systems. According to feature selection, eight elements (Th, U, Sr, Eu, Dy, Y, Nd and La) improve the model performance. We could show that the XGBoost classifier efficiently and accurately classifies high-dimensional apatite trace element data according to the ore deposit type (overall accuracy: 94% and F1 score: 89%). Interpretation of the model using the SHAPley Additive exPlanations (SHAP) tool shows that Th, U, Eu and Nd are the most indicative elements for classifying deposit types using apatite trace element chemistry. Our approach has broad implications for the better understanding of the sources, chemistry and evolution of melts and hydrothermal fluids resulting in ore deposit formation.

Keywords: Machine learning; apatite; trace elements; ore deposit fertility; XGBoost; LA-ICP-MS

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

52

53 ## Introduction

54 To develop a quantitative, process-based model for ore-forming systems, a
55 characterization of melt and hydrothermal fluid source, composition and evolution is
56 required (e.g., Andersson et al., 2019). Various minerals in ore-forming systems can
57 constrain the conditions of mineralization based on variations in their mineral chemistry,
58 thus recording the evolution of melts and hydrothermal fluids and yielding constraints
59 on the metallogenic processes (Clark & Williams-Jones, 2004; Pisiak et al., 2017;
60 Chapman et al., 2021; Qiu et al., 2021). As a common accessory mineral in igneous,
61 metamorphic and clastic sedimentary rocks, apatite has a broad range of applications in
62 the geosciences, including thermochronology studies to investigate tectonic unroofing
63 (Fitzgerald et al., 1991), fault slip rates (Brichau et al., 2006), landscape evolution
64 (Braun, 2006), petroleum system maturation (Burtner et al., 1994) and record of volatile
65 budgets and volcanic eruption triggering (Stock et al., 2016). The structure of apatite
66 also facilitates the substitution of more than half the stable members of the periodic
67 table as trace-elements (Hughes, 2015), including the rare earth elements and Sr, Y, Th,
68 and U (Sha & Chappell, 1999; Chew et al., 2011; Zhou et al., 2022a). Apatite trace
69 element chemistry thus has important applications in igneous and metamorphic
70 petrogenesis studies to improve the understanding of ore deposit formation (Chu et al.,
71 2009; O'Sullivan et al., 2020; Yu et al., 2021, 2022).

72 Previous studies that have employed apatite trace element chemistry to classify
73 protolith rock type or fertility have typically employed binary or ternary discrimination
74 diagrams with the variables being apatite trace element abundances or elemental ratios.
75 Belousova et al. (2002) analyzed trace elements in apatite from a variety of common
76 rock types and employed plots of Sr versus Y and Mn, $(Ce/Yb)_{cn}$ versus the sum of the
77 REE, and Y versus Eu/Eu^{*} to identify fields of apatite compositions from different rock
78 types. Bouzari et al., (2016) used cathodoluminescence combined with trace element
79 compositions to discriminate trace element variations due to alteration linked to the

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

ingress of hydrothermal fluids. Mao et al. (2016) evaluated trace element compositions in apatite from multiple deposit types and suggested several discrimination diagrams for the division of deposit types based on apatite trace element chemistry. O'Sullivan et al. (2020) applied compositional statistics, classification and a machine learning classifier to apatite trace element compositional data, and generated binary plots that discriminated between several types of igneous and metamorphic rocks. Zhou et al. (2022b) used a big data approach to investigate variations in apatite trace element chemistry and showed that an Eu/Y *vs* Ce diagram best discriminates apatite crystallized from different host rock types. However, while two-variable scatterplots or three-variable ternary diagrams offer easy and convenient visualization of discrimination trends, they can often fail to rigorously trace the sources, chemistry, and evolution of melts and hydrothermal fluids based on variations in apatite trace element chemistry (Li et al., 2015; Wang et al., 2021; Zhong et al., 2021). The first reason is that apatite has a complex chemistry with high partition coefficients for many trace elements, and trace element partition coefficients in apatite also differ significantly with varying temperature, pressure and melt compositions (Prowatke and Klemme, 2006). The range of possible substitutions in both anion and cation sites and significant tolerance to structural distortion and chemical substitution leads to highly diverse trace element and minor compositions. Another reason is the inherent difficulty of discrimination diagrams resulting in low classification accuracy. Although discrimination diagrams can have a robust geochemical basis, the discrimination fields themselves are defined based on statistics (Pearce, 1996). While the geochemical underpinnings of discrimination diagrams may be well understood, they are typically not sufficiently well constrained to accurately predict absolute elemental abundances for chemically complex systems (Snow, 2006). In addition, while an individual apatite trace element analysis can yield the abundances of tens of trace elements, discrimination diagrams typically only use the information from two or three variables (element contents and element ratios). Diagnostic geochemical signatures from apatite trace element data may not be

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

effectively extracted from these limited numbers of variables, potentially leading to different types of apatite not being discriminated between or, even worse, misclassified.

High-dimensional analysis methods using machine learning can overcome these challenges. As a rapidly growing approach to analyzing high-throughput experimental data in novel ways, machine learning focuses on the underlying relationships between features (measurable properties) and research targets (Jordan & Mitchell, 2015). In recent years, it has been successfully applied to a diverse suite of classification challenges on high-dimensional datasets in the geosciences (Petrelli & Perugini, 2016; Schönig et al., 2021; Zhong et al., 2021; Wang et al., 2022). These include estimating pre-eruptive temperatures and pressures using clinopyroxene-melt (Petrelli et al., 2020), evaluating the occurrence of H diffusion in the clinopyroxene phenocrysts of basaltic magma (Chen et al., 2021), proposing and improving thermobarometry for different magma types (biotite-bearing magma: Li and Zhang, 2022, amphibole -bearing magma: Higgins et al., 2022, clinopyroxene-bearing magma: Jorgenson et al., 2022), and distinguishing S-, I-, and A-type granites (Gion et al., 2022).

In this study, we have compiled a trace element dataset comprising 8629 apatite analyses from known mineralization types and ore-barren magmatic rocks from published literature to train and test the classification model. After comparing four commonly employed machine learning algorithms, we chose a scalable end-to-end tree boosting system called XGBoost as the optimal algorithm to tune and yield the final classifiers. XGBoost is an open-source machine-learning algorithm that combines 'weak classifiers' to form 'strong classifiers' based on a decision tree with gradient boosting (Chen & Guestrin, 2016). It provides a rapid and highly accurate approach to classifying high-dimensional data, such as distinguishing between ore-fertile and ore-barren provenance and classifying ore-fertile environments in this study. To address the black box problem commonly attributed to machine learning algorithms resulting from their potential opacity, we employed the SHAP (SHAPley Additive exPlanations) (Lundberg and Lee, 2017) visualization tool that makes a machine learning model more

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

explainable by visualizing its output. SHAP is a game theoretic method and applying it herein reveals the most diagnostic trace elements in apatite for classifying ore deposit types, while also revealing the variable geochemical behavior of different elements in ore deposit types. Our results demonstrate strong correlations between high-dimensional apatite trace-element geochemical data and ore deposit type thus furthering our knowledge of ore-forming systems, and have broad implications for understanding the sources, chemistry and evolution of melts and hydrothermal fluids.

Database

For the compilation of the apatite trace element dataset, 8629 analyses from 1685 rock samples were retrieved from 245 publications using the global petrological open-access database GEOROC (http://georoc.mpch-mainz.gwdg.de/georoc/). Apatite trace element compositions from these studies include data from five common ore deposit types located worldwide, including porphyry, skarn, orogenic Au, iron-oxide copper gold (IOCG), and iron-oxide apatite (IOA or Kiruna type) (Figure 1). Apatite trace element compositions were collected from various unmineralized (barren) magmatic and metamorphic rocks to identify any systematic differences between apatite from fertile and barren systems. Unmineralized samples in the database comprise both wall rocks from the respective mineral deposits but also include non-mineralized regions. As an example, three different types of quartz monzonite porphyry from Jia et al. (2020) were incorporated in our database. Two samples (PD02 and BR04) are ore-fertile samples containing sulfide veins, while sample PD01 is an ore-barren quartz monzonite porphyry containing minimal sulfide. Detailed information on the apatite analyses incorporated in the database is provided in Appendix Table 1.

Different experimental LA-ICP-MS procedures and protocols employed in the 245 publications result in a diverse suite of trace elements in the compiled dataset. The 14 most commonly analyzed trace elements. La, Ce, Pr, Nd, Sm, Eu, Gd, Dy, Yb, Lu, Sr, Y, Th and U were used to provide a consistent and optimized dataset. The data set

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

includes values below the detection limit (bdl) or values that were not reported. To improve the quality of the dataset, bdl analyses were replaced by a value of half of the detection limit (Zhong et al., 2021). Ultimately the dataset was reduced to 4085 analyses from 249 individual samples (unmineralized magmatic apatite: 148; porphyry: 29; skarn: 35; orogenic Au: 15; IOCG: 13 and IOA: 9) for further investigation by the different machine learning methods (Table 1). Figure 2 provides a compilation of the apatite trace element data based on deposit type and individual deposits. Apatite from IOA deposits has the highest La and Th contents, while IOCG apatite has the lowest Sr (Figure 2a, b). These diagrams show that the variation in concentration of some individual elements can distinguish apatite from different deposit types to a certain extent. However, most trace element ranges still overlap and are thus not entirely diagnostic. Therefore, although deposit type is unlikely to be identified using binary or ternary diagrams, the partial separation observed in some of the apatite compositional data implies that machine learning approaches in high-dimensional space have the potential to distinguish apatite derived from different ore deposit types.

Model development and performance

Machine learning is used to teach algorithms to construct self-learning systems which can handle large datasets more efficiently (Jordan & Mitchell, 2015; Mahesh, 2020). Machine learning is classified into two broad categories - supervised learning and unsupervised learning (Soofi & Awan, 2017). In this study, we used supervised machine learning (use of labeled datasets to train algorithms to classify data) to link apatite trace element composition to their source ore-deposit type. We tested four different established algorithms: k-nearest neighbors (KNN) (Bentley, 1975), random forest (RF) (Breiman, 2001), support vector machine (SVM) (Vapnik, 1995), and eXtreme Gradient Boosting (XGBoost) (Chen & He, 2015), before selecting the best classification model after hyperparameter optimization and comparison. Figure 3 outlines the detailed workflow of our approach.

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

Data pre-processing

Pre-processing of the data involves standardization and balance processing. A suitable standardization procedure is critical in applying machine learning algorithms, to avoid attributes in greater numeric ranges dominating those in smaller numeric fields, while also helping to eliminate potential numerical difficulties during the calculations in many machine learning approaches (Hsu et al., 2003). We first transformed the dataset in this study by applying a log-ratio transformation to obtain a Gaussian distribution which was then normalized using the Standardscaler () function in the Scikit-learn machine learning library for Python (more detail is provided in section 3.5 on the libraries employed in this study). This function centers data by setting the mean to zero for each feature, then scaling it by dividing non-constant features by their standard deviation to produce a standard normal distribution with the mean of observed values = 0 and a standard deviation = 1.

Dealing with imbalanced data is essential prior to building a machine learning model. Many algorithms may be biased towards classes with large sample sizes if the training set is imbalanced. For example, in our data set, 2300 analyses are from unmineralized magmatic apatite, while only 78 analyses are from IOCG deposits. Therefore, we applied the synthetic minority oversampling technique (SMOTE) using the imbalanced-Learn Library in Python to minimize the possible effects resulting from variations in sample size. SMOTE (Chawla et al., 2002) is an improved scheme based on a random oversampling algorithm, which artificially synthesizes new data to add to the dataset. Compared with most sampling methods, SMOTE has stronger robustness and achieved the real sense of combining the over-sampling minority class and under-sampling majority class.

The selected dataset is randomly divided into a training dataset (80%) and a testing dataset (20%) using the hold-out method while maintaining the exact proportions of each class. The training set was oversampled using the SMOTE algorithm, which was then used to train the classifier, while the testing set was utilized to evaluate the classifier.

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

Algorithm comparison

K-nearest neighbors (KNN), random forest (RF), support vector machine (SVM), and eXtreme Gradient Boosting (XGBoost) are widely used machine learning methods that can be applied to the classification of high-dimensional data, and have been commonly used in a variety of fields in the geosciences (Carranza & Laborte, 2015; Petrelli et al., 2017; Liu & Beaudoin, 2021; Shen et al., 2022). We compared these four supervised machine learning algorithms to select the optimal approach to train the machine learning model for determining ore-deposit type from apatite trace element data.

KNN is one of the simplest classification methods in that it calculates the similarity (proximity) between new and available data. It puts the new data case into the category most similar to the available categories. While this simple classification method has no explicit training step, it is not well suited for large datasets with high dimensionality due to the difficulties in calculating proximities for each data point in high dimensions and does not work well on imbalanced data or datasets with outliers (Bently, 1975; Alfeilat et al., 2018, Nathwani et al., 2022). RF employs an ensemble of decision tree classifiers on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting (Breiman, 2001). RF does not require significant tuning of parameters, tends not to overfit the data and can handle non-linear numeric and categorical predictors. Nevertheless, prediction accuracy on complex problems is generally inferior to that of gradient-boosted trees. RF classification is also more difficult to interpret than a single decision tree (which may be easily visualized as a sequence of decisions and outcomes). The objective of the SVM algorithm is to find a hyperplane in N-dimensional space (where N is the number of features, in this case elements) that distinctly classifies the data points. It is the most commonly used machine learning method in geosciences (Noble, 2006; Soofi & Awan, 2017), tends not to overfit data nor be overly influenced by outliers, and is most effective in high-dimensional spaces when there is a clear margin of separation between classes. SVM does not

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

perform very well when the dataset is noisy (i.e. target classes are overlapping) or on large datasets due to the training time involved. The final SVM model is not probabilistic and can be challenging to interpret and also requires selection of an appropriate kernel function and hyperparameters. XGBoost is a scalable machine learning system that combines 'weak classifiers' to form 'strong classifiers' based on a decision tree with gradient boosting (Chen & Guestrin, 2016). It typically outperforms all other algorithms in machine learning community competitions, can handle large datasets and is not prone to overfitting or the influence of outliers when properly tuned (Nielsen, 2016; Abou Omar, 2018; Ogunleye and Wang, 2020; Wang et al., 2020). It also does not require significant feature processing (i.e. no need for scaling or normalizing data, and it can also handle missing values well); thus feature importance can be ascertained, allowing for feature selection. It does not work well on sparse and unstructured data and can be difficult to tune due to the many hyperparameters involved. Similar to many of the other algorithms described above, interpretation of the final model can be difficult.

Grid search and cross-validation were introduced to optimize hyperparameters as appropriate hyperparameter selection can significantly improve the performance of the machine learning model. Grid search is the traditional approach to hyperparameter optimization, which finds the optimal hyperparameters by conducting a complete search over a given subset of hyperparameters space of the training algorithm (Liashchynskyi & Liashchynskyi, 2019). However, a single grid search is insufficient and therefore, we used k-fold cross-validation to undertake multiple grid searches using the GridSearchCV () function in Python's Scikit-learn machine learning library. The training set is divided into k groups, and one subset of data is selected randomly as a validation set and the remainder (k-1) of the subsets as training datasets. This step is repeated for k times to obtain k models, and the average classification accuracy of the final validation set of these k models is used as the performance indicator of the machine learning model.

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

We performed a grid search with 10-fold cross-validation to tune hyperparameters and used the testing set to evaluate the F1 score (which conveys the balance between the precision and the recall) of the four machine-learning algorithms. We set the random seed while splitting the training and testing sets. This ensures that the data is divided the same way every time the code is run and is also required because algorithms such as RF and XGBoost are non-deterministic (for a given input, the output is not always the same) and thus require a random seed argument for reproducible results and algorithm comparison. After tuning of the hyperparameters, the algorithms yielded the following performance: KNN algorithm (F1 score: 88.6%), random forest algorithm (F1 score: 89.8%), SVM algorithm (F1 score: 89.7%) and XGBoost algorithm (F1 score: 90.8%). Table 2 provides detailed information on the hyperparameters and test scores and Figure 4 shows the detailed classification information of the four algorithms on a confusion matrix. We chose XGBoost as the optimal supervised machine-learning algorithm as it produced the highest test score and the best and most balanced performance across the five ore deposit categories (Fig. 4).

Feature selection

To effectively apply machine learning methods, feature selection is a key step that helps understand the data, reduces computation and the curse of dimensionality (the explosive nature of increasing data dimensions and its resulting exponential increase in computational efforts) and improves learning performance (Kalousis et al., 2007; Chandrashekar & Sahin, 2014; Kumar & Minz, 2014; Li e al., 2017). The SHAP tool was employed to compute each trace element's contribution (SHAP value) in apatite in the initial dataset for a particular prediction. We list the SHAP values in descending order in Figure 5 and sequentially added more elements to the XGBoost algorithm in descending SHAP order to show the change (cross-validation and test score) in model performance. As shown in Figure 5, for n = 1 (Th) the cross-validation score is ~59% and the test score was only ~37%. Increasing the number of elements (n = 5; Th, U, Sr, Eu, Dy), the cross-validation score increased dramatically to ~98% with the test score

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

304 increasing to ~86% (n=5). When n=8, the cross-validation score and test score have
305 stabilized at ~99% and ~90%. The model could hence be built from these eight elements
306 (Th, U, Sr, Eu, Dy, Y, Nd, La) as there is minimal improvement when n>8, which is
307 geologically realistic as the remaining six elements (n = 9 to 14) are all REEs which
308 exhibit coupled geochemical behavior. Therefore, to improve the learning performance
309 and the application of the model, we built a filtered dataset using the XGBoost method
310 with eight elements (Th, U, Sr, Eu, Dy, Y, Nd, La).

311 **Retraining and testing the classifier**

312 The filtered dataset was again randomly split into a training set (80%) and a testing
313 set (20%) and the training set was then oversampled using the SMOTE algorithm, and
314 retrained to produce the final XGBoost classifier. Grid search and 10-fold cross-
315 validation were used to choose the optimal hyperparameters (gamma and max_depth,
316 Figure 6). The classifier was evaluated on the testing set. Randomly splitting the training
317 set and testing set will change the predicted results of the XGBoost model each time,
318 thus the test scores (mean score ± standard deviation) were calculated from 50 iterations.
319 The optimal XGBoost classification was determined for hyperparameters of
320 n_estimators=148, gamma=0, max_depth=9. (Table 3), with a precision of 0.89 ± 0.02,
321 recall of 0.90 ± 0.02, F1 score of 0.89 ± 0.02 and accuracy of 0.94 ± 0.01. Figure 6
322 shows the F1 score of different hyperparameter combinations. A summary of the
323 precision, recall, and F1 score for each class are provided in Table 3. The dataset and
324 code are available on the Zenodo website (http://doi.org/10.5281/zenodo.7094836).

325 **Libraries**

326 All operations on the reference dataset from pre-processing through to model
327 application were undertaken using the Python programming language. The following
328 libraries were used to complete the code: pandas (Snider and Swedo, 2004), numpy
329 (Oliphant, 2006) and imlearn (Ma and He, 2013) for data analysis; matplotlib (Barrette
330 et al., 2005) and seaborn (Waskom, 2021) for plotting the diagrams; scikit-learning
331 (Kramer, 2016) and xgboost (Chen & He, 2015) for machine learning; shap (Lundberg

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

and Lee, 2017) for feature selection and machine learning interpretation.

Discussion

Limitations of 2D classification diagrams employing two variables

The potential limitations of employing discrimination diagrams (e.g. 2D scatterplots with two variables) were initially discussed in the introduction and are explored further here. In this study, we first calculated the ratio of two random elements from the dataset and added them into the dataset as new features. A total of 5460 discrimination diagrams were constructed using any two features in the dataset with the best discrimination combination represented by a plot of Th/Pr *vs* U/Pr ratio (Figure 7a), with the silhouette coefficient used to investigate the separation distance between the resulting clusters. We also investigated the six elements (Th, U, Sr, Eu, Dy, and Y) with the highest SHAP values (Figure 5) to draw 2D scatterplots (Figure 7b, c, d).

As shown in Figure 7, these four discrimination diagrams cannot effectively distinguish between ore-fertile and ore-barren provenance. Apatite data from different ore-fertile environments overlap as well. This is the principal limitation of two-variable scatterplots – they only employ a small amount of information from the high-dimensional data, unlike the high dimensional machine learning approach undertaken in this study. Even though the apatite trace element data from the different ore deposit types overlap, the apatite data from individual deposit types still cluster together on the four discrimination diagrams (Figure 7). Unsurprisingly given the extremely broad variation in apatite trace element abundances in igneous rocks (O'Sullivan et al., 2020), the unmineralized magmatic apatite field is by far the largest, encompassing nearly all the ore deposit fields. The unmineralized magmatic apatite field exhibits bimodal Sr (Figure 7b) and U abundances (Figure 7c). This corroborates the findings of O'Sullivan et al. (2020), with U abundances low in ultramafic igneous and low-grade metamorphic apatite and higher in igneous and high-grade metamorphic apatite, and Sr low in all metamorphic rocks and I- and S-type igneous rocks, and higher in alkaline and

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

ultramafic igneous rocks (Figure 6 in O'Sullivan et al., 2020).

Apatite from IOA deposits define relatively restricted fields on all discrimination plots (Figure 7), while those from orogenic Au deposits show higher concentrations of Y and the geochemically-similar element Dy (Figure 7b, d). The kernel density curves of Sr contents in apatite from orogenic Au deposits also have two distinct peaks (Figure 7b). The kernel density curves of Eu and U abundances show that apatite from skarn deposits have lower concentrations of Eu and higher abundances of U compared with apatite from porphyry deposits (Figure 7c, d). These observations show that the trace element abundances of apatite from different ore deposits exhibit systematic trace element variations and thus have potential to be discriminated effectively using the high-dimensional data space through the machine-learning approach adopted in this study.

Classification in high-dimensional space

The classifier can effectively distinguish between ore-fertile and ore-barren environments (recall ratio > 95% for barren samples), and apatite from the different deposit types can be also successfully distinguished with F1 test scores of >88% for all four algorithms (Figure 4). This suggests that classifying deposit types using machine learning applied to apatite compositional data is a viable approach. The exception is IOCG apatite, for which 16% of analyses were predicted to belong to different classes (Figure 8), probably due to the small sample amount of this deposit type, even though SMOTE oversampled the training set. The predictions for porphyry and skarn deposits are better. However, both are less than 90% (porphyry deposits: 89%, skarn deposits: 88%), which is attributed mainly to the complexity of porphyry and skarn mineralization processes. Porphyry mineralization takes place across a very broad temperature range from 250 to 1000°C, and apatite forming during different porphyry crystallization stages may have very different trace element signatures (Sillitoe, 2010). Skarn mineralization also occurs across a wide range of formation temperatures, while additionally the diverse nature of host rock types in skarn systems may impart additional

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

trace element variability (Jia et al., 2020). Future work could include sub-division of apatite classes to incorporate differing crystallization stages and host rock chemistries in porphyry and skarn systems although this is likely to be a substantial undertaking. Nevertheless, the XGBoost classifier performs well on the classification of fertility and all deposit types in this dataset with an overall accuracy >94% and F1 score > 89%, with both high precision and recall ratios, especially for the IOA and orogenic Au deposits, from which almost all apatite data is predicted correctly (Figure 8).

Low-grade metamorphic apatite is very similar in terms of its trace element geochemistry to hydrothermal apatite (O'Sullivan et al., 2020). Therefore, an effective machine learning model must distinguish low-grade metamorphic apatite from the five mineralized classes. We selected 215 apatite analyses from 31 samples from the database of O'Sullivan et al. (2020) with different metamorphic grades (high-grade metamorphic apatite: 112; low- and medium-grade metamorphic apatite: 103) as a new testing set. Based on the XGBoost classifier, our predicted results show that most of the analyses accurately classified unmineralized apatite (181 out of 215, Appendix Table 2). Fourteen high-grade metamorphic apatite analyses were misclassified as IOCG apatite, while 98 high-grade metamorphic apatite analyses were correctly predicted as unmineralized apatite. For low- and medium-grade metamorphic apatite, 20 apatite were misclassified as a mineralized class (15 apatite predicted as orogenic Au, three apatite predicted as porphyry, one apatite predicted as skarn and one apatite predicted as IOA). In contrast, the remaining 83 apatite were predicted correctly. The performance (overall accuracy >84%) on this group of metamorphic samples shows that our XGBoost classifier can effectively distinguish low-grade metamorphic apatite from fertile classes and provides a rapid and highly accurate approach to predicting ore deposit type based on apatite trace element data.

Interpreting machine learning models

Machine learning methods have been widely used in geosciences and various algorithms have been proven to be useful tools for interpreting high-dimensional

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

geochemical data (Petrelli & Perugini, 2016; Chen et al., 2021; Wang et al., 2021). Despite their widespread application in the classification of big data sets, machine learning approaches are often referred to as a black box, where the dataset undergoes a series of calculations immediately followed by the output of results, without providing a transparent working process between the input and output data (Lancet Respiratory Medicine, 2018). Some studies have employed feature importance to select machine learning training parameters (Nathwani et al., 2022). However, such an approach does not help show the relationship between a given feature and the working target – feature importance is based on the decrease in model performance and contains no information beyond this. To improve the transparency and interpretation of our XGBoost classifier, a SHAP summary plot is presented in Figure 9. This summary plot combines feature importance with the magnitude of feature attributes, and features are ordered according to their importance. Each point on the summary plot is a SHAP value for a feature and an instance. The feature importance determines the position on the y-axis and on the x-axis by the SHAP value, while the color represents the value of the feature from low to high.

Strontium and Eu are the two most diagnostic elements for classifying IOCG deposits. For example, high concentrations of Sr (red colors) negatively influence the classification while low concentrations have a positive influence; the relationship is the opposite for Eu (Figure 9a). For IOA deposits, high Th contents, low U abundances and low Sr favor prediction as an IOA deposit (Figure 9b). Porphyry deposit apatite classification is favored by low Th and low Nd (Figure 9c) while low U and Eu abundances help to distinguish skarn deposits. The lowest U concentrations may be partly affected by values below the limit of detection. A larger dataset should confirm the relationship between apatite U contents and skarn deposits (Figure 9d). High concentrations of Dy and Sr help classify orogenic Au deposits (Figure 9e). Although there is wide variation in apatite trace element abundances in different types of igneous and metamorphic rocks (O'Sullivan et al., 2020) and the unmineralized magmatic

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

444 apatite dataset is very large and diverse, moderate Th and, in particular, high Nd are
445 indicative for unmineralized apatite (Figure 9f).

446 In summary, Th, U, Eu and Nd are the most effective elements for classifying ore
447 deposit types, especially Th for IOA (Figure 9b), Nd for porphyry and unmineralized
448 apatite (Figure 9c, f), U for skarn (Figure 9d) and Dy for orogenic Au deposits (Figure
449 9e). Other elements, like Sr, also improves the classification of some deposit types
450 (Figure 9a, e).

451

452 **Implications**

453 Traditional methods to discriminate (e.g. using two-variable scatterplots) only
454 result in partial separation of ore deposit classes because of the complexity of apatite
455 chemistry. Machine learning-based approach (XGBoost) fully exploit the high
456 dimensionality of apatite trace element data to produce a novel geochemical
457 classification system to link apatite trace element chemistry with ore deposit type. Based
458 on the classifier, apatite has strong potential as a fertility indicator to distinguish fertile and
459 barren environments effectively. To circumvent the 'black box' problem commonly
460 associated with machine learning models, SHAP (SHapley Additive exPlanations) tool
461 was introduced to explain individual predictions. Based on the selected elements (Th, U,
462 Sr, Eu, Dy, Y, Nd and La), the XGBoost algorithm accurately and efficiently classifies
463 apatite with ore deposit type (overall accuracy > 94%) and yields the optimal elements
464 (Th, U, Eu and Nd) to discriminate apatite from different ore deposit types. With the
465 increasing amount of high-throughput apatite trace element data produced by modern
466 analytical techniques, our XGBoost approach offers the potential to make more data-
467 driven decisions such as sub-division of porphyry and skarn mineralization stages.
468 Moreover, the novel SHAP-based analysis approach aids understanding of the sources,
469 chemistry, and evolution of mineralizing melts and fluids in ore deposit studies.

470

471 **Acknowledgments**

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

We gratefully acknowledge the constructive comments of the Editor Rachel Russell and Don Baker, Associate Editor Maurizio Petrelli, Oliver Higgins and anonymous journal reviewers. This work was financially supported by the National Natural Science Foundation of China (42130801, 42261134535, and 42072087), National Key Research Program (2019YFA0708603), the Beijing Nova Program (Z201100006820097), and the 111 Project (BP0719021). David Chew acknowledges past and present support from Science Foundation Ireland (SFI) under Grant Numbers 12/IP/1663 and 13/RC/2092_P2 (iCRAG, the SFI Research Centre in Applied Geosciences).

References

Abou Omar, K.B. (2018). XGBoost and LGBM for Porto Seguro's Kaggle challenge: A comparison. Preprint Semester Project.

Abu Alfeilat, H.A., Hassanat, A.B., Lasassmeh, O., Tarawneh, A.S., Alhasanat, M.B., Eyal Salman, H.S., and Prasath, V.S. (2019). Effects of distance measure choice on k-nearest neighbor classifier performance: a review. Big data, 7(4), 221-248.

Acosta-Vigil, A., Buick, I., Hermann, J., Cesare, B., Rubatto, D., London, D., and Morgan, G.B. (2010). Mechanisms of crustal anatexis: a geochemical study of partially melted metapelitic enclaves and host dacite, SE Spain. Journal of Petrology, 51(4), 785-821.

Adlakha, E., Hanley, J., Falck, H., and Boucher, B. (2018). The origin of mineralizing hydrothermal fluids recorded in apatite chemistry at the Cantung W–Cu skarn deposit, NWT, Canada. European Journal of Mineralogy, 30(6), 1095-1113.

Andersson, S.S., Wagner, T., Jonsson, E., Fusswinkel, T., and Whitehouse, M.J. (2019). Apatite as a tracer of the source, chemistry and evolution of ore-forming fluids: The case of the Olserum-Djupedal REE-phosphate mineralisation, SE Sweden. Geochimica et Cosmochimica Acta, 255, 163-187.

Barrett, P., Hunter, J., Miller, J.T., Hsu, J.C., and Greenfield, P. (2005). December.

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

matplotlib--A Portable Python Plotting Package. In Astronomical data analysis software and systems XIV (Vol. 347, p. 91).

Belousova, E.A., Griffin, W.L., O'Reilly, S.Y., and Fisher, N.I. (2002). Apatite as an indicator mineral for mineral exploration: trace-element compositions and their relationship to host rock type. Journal of Geochemical Exploration, 76(1), 45-69.

Bentley, J.L. (1975). Multidimensional binary search trees used for associative searching. Communications of the ACM, 18(9), 509-517.

Bouzari, F., Hart, C.J., Bissig, T., and Barker, S. (2016). Hydrothermal alteration revealed by apatite luminescence and chemistry: A potential indicator mineral for exploring covered porphyry copper deposits. Economic Geology, 111(6), 1397-1410.

Braun, J., Beek, P.v.d., and Batt, G. (2006). Quantitative thermochronology: numerical methods for the interpretation of thermochronological data. Cambridge University Press. Breiman, L., 2001. Random forests. Machine learning, 45(1), 5-32.

Brichau, S., Ring, U., Ketcham, R.A., Carter, A., Stockli, D., and Brunel, M. (2006). Constraining the long-term evolution of the slip rate for a major extensional fault system in the central Aegean, Greece, using thermochronology. Earth and Planetary Science Letters, 241(1-2), 293-306.

Burtner, R.L., Nigrini, A., and Donelick, R.A. (1994). Thermochronology of Lower Cretaceous source rocks in the Idaho-Wyoming thrust belt. AAPG bulletin, 78(10), 1613-1636.

Cao, M., Li, G., Qin, K., Seitmuratova, E.Y., and Liu, Y. (2012). Major and trace element characteristics of apatites in granitoids from Central Kazakhstan: implications for petrogenesis and mineralization. Resource Geology, 62(1), 63-83.

Carranza, E.J.M., and Laborte, A.G. (2015). Random forest predictive modeling of mineral prospectivity with small number of prospects and data with missing values in Abra (Philippines). Computers & Geosciences, 74, 60-70.

Chandrashekar, G., and Sahin, F. (2014). A survey on feature selection methods.

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

528 Computers & Electrical Engineering, 40(1), 16-28.
529 Chapman, R.J., Moles, N.R., Bluemel, B., and Walshaw, R.D. (2021). Detrital gold as
530 an indicator mineral. Geological Society, London, Special Publications, 516.
531 Chawla, N.V., Bowyer, K.W., Hall, L.O., and Kegelmeyer, W.P. (2002). SMOTE:
532 synthetic minority over-sampling technique. Journal of artificial intelligence
533 research, 16, 321-357.
534 Chen, H., Su, C., Tang, Y.Q., Li, A.Z., Wu, S.S., Xia, Q.K., and ZhangZhou, J. (2021).
535 Machine learning for identification of primary water concentrations in mantle
536 pyroxene. Geophysical Research Letters, 48(18), e2021GL095191.
537 Chen, T., and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In
538 Proceedings of the 22nd ACM SIGKDD International Conference On Knowledge
539 Discovery And Data Mining (pp. 785-794).
540 Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., and Chen, K. (2015).
541 Xgboost: extreme gradient boosting. R package version 0.4-2, 1(4), 1-4.
542 Chew, D.M., Sylvester, P.J., and Tubrett, M.N. (2011). U–Pb and Th–Pb dating of
543 apatite by LA-ICPMS. Chemical Geology, 280(1-2), 200-216.
544 Chu, M.F., Wang, K.L., Griffin, W.L., Chung, S.L., O'Reilly, S.Y., Pearson, N.J., and
545 Iizuka, Y. (2009). Apatite composition: tracing petrogenetic processes in
546 Transhimalayan granitoids. Journal of Petrology, 50(10), 1829-1855.
547 Clark, J.R., and Williams-Jones, A.E. (2004). Rutile as a potential indicator mineral for
548 metamorphosed metallic ore deposits. Rapport Final de DIVEX, Sous-projet SC2,
549 Montréal, Canada, 17.
550 Fitzgerald, P., Fryxell, J., and Wernicke, B. (1991). Miocene crustal extension and
551 uplift in southeastern Nevada: Constraints from fission track analysis. Geology,
552 19(10), 1013-1016.
553 Gion, A.M., Piccoli, P.M., and Candela, P.A. (2022). Characterization of biotite and
554 amphibole compositions in granites. Contributions to Mineralogy and Petrology,
555 177(4), 1-15.

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

556 Hazarika, P., Mishra, B., and Pruseth, K.L. (2016). Scheelite, apatite, calcite and
557 tourmaline compositions from the late Archean Hutti orogenic gold deposit:
558 Implications for analogous two stage ore fluids. Ore Geology Reviews, 72, 989-
559 1003.
560 Higgins, O., Sheldrake, T., and Caricchi, L. (2022). Machine learning thermobarometry
561 and chemometry using amphibole and clinopyroxene: a window into the roots of
562 an arc volcano (Mount Liamuiga, Saint Kitts). Contributions to Mineralogy and
563 Petrology, 177(1), pp.1-22.
564 Hsu, C.W., Chang, C.C., and Lin, C.J. (2003). A practical guide to support vector
565 classification. Hughes, J.M., 2015. The many facets of apatite. American
566 Mineralogist, 100(5–6), 1033–1039.
567 Jia, F., Zhang, C., Liu, H., Meng, X., and Kong, Z. (2020). In situ major and trace
568 element compositions of apatite from the Yangla skarn Cu deposit, southwest
569 China: Implications for petrogenesis and mineralization. Ore Geology Reviews,
570 127, 103360.
571 Jordan, M.I., and Mitchell, T.M. (2015). Machine learning: Trends, perspectives, and
572 prospects. Science, 349(6245), 255-260.
573 Jorgenson, C., Higgins, O., Petrelli, M., Bégué, F., and Caricchi, L. (2022). A Machine
574 Learning-Based Approach to Clinopyroxene Thermobarometry: Model
575 Optimization and Distribution for Use in Earth Sciences. Journal of Geophysical
576 Research: Solid Earth, 127(4), p.e2021JB022904.
577 Kalousis, A., Prados, J., and Hilario, M. (2007). Stability of feature selection algorithms:
578 a study on high-dimensional spaces. Knowledge and information systems, 12(1),
579 95-116.
580 Kramer, O. (2016). Scikit-learn. In Machine learning for evolution strategies (pp. 45-
581 53). Springer, Cham.
582 Krneta, S., Cook, N.J., Ciobanu, C.L., Ehrig, K., and Kontonikas-Charos, A. (2017).
583 The Wirrda Well and Acropolis prospects, Gawler Craton, South Australia:

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

584 Insights into evolving fluid conditions through apatite chemistry. Journal of
585 Geochemical Exploration, 181, 276-291.
586 Kumar, V., and Minz, S. (2014). Feature selection: a literature review. SmartCR, 4(3),
587 211-229.
588 Lancet Respiratory Medicine. (2018). Opening the black box of machine learning.
589 Lancet Respir Med, 6(11), 801.
590 Laurent, O., Zeh, A., Gerdes, A., Villaros, A., Gros, K., and Słaby, E. (2017). How do
591 granitoid magmas mix with each other? Insights from textures, trace element and
592 Sr–Nd isotopic composition of apatite and titanite from the Matok pluton (South
593 Africa). Contributions to Mineralogy and Petrology, 172(9), 1-22.
594 Li, C., Arndt, N.T., Tang, Q., and Ripley, E.M. (2015). Trace element indiscrimination
595 diagrams. Lithos, 232, 76-83.
596 Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R.P., Tang, J., and Liu, H. (2017).
597 Feature selection: A data perspective. ACM computing surveys (CSUR), 50(6), 1-
598 45.
599 Li, X., and Zhang, C. (2022). Machine Learning Thermobarometry for Biotite-Bearing
600 Magmas. Journal of Geophysical Research: Solid Earth, 127(9), p.e2022JB024137.
601 Liashchynskyi, P., and Liashchynskyi, P. (2019). Grid search, random search, genetic
602 algorithm: A big comparison for NAS. arXiv preprint arXiv:1912.06059.
603 Liu, H., and Beaudoin, G. (2021). Geochemical signatures in native gold derived from
604 Au-bearing ore deposits. Ore Geology Reviews, 132, 104066.
605 Lundberg, S.M., and Lee, S.I. (2017). A unified approach to interpreting model
606 predictions. Advances in neural information processing systems, 30.
607 Ma, Y., and He, H. eds. (2013). Imbalanced learning: foundations, algorithms, and
608 applications.
609 Mahesh, B. (2020). Machine learning algorithms-a review. International Journal of
610 Science and Research (IJSR). [Internet], 9, 381-386.
611 Mao, M., Rukhlov, A.S., Rowins, S.M., Spence, J., and Coogan, L.A. (2016). Apatite

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

612 trace element compositions: a robust new tool for mineral exploration. Economic
613 Geology, 111(5), 1187-1222.
614 Matusiak-Małek, M., Puziewicz, J., Ntaflos, T., Woodland, A., Uenver-Thiele, L.,
615 Büchner, J., Grégoire , M., and Aulbach, S. (2021). Variable origin of
616 clinopyroxene megacrysts carried by Cenozoic volcanic rocks from the eastern
617 limb of Central European Volcanic Province (SE Germany and SW Poland).
618 Lithos, 382, 105936.
619 Minissale, S., Zanetti, A., Tedesco, D., Morra, V., and Melluso, L. (2019). The
620 petrology and geochemistry of Nyiragongo lavas of 2002, 2016, 1977 and 2017
621 AD, and the trace element partitioning between melilitite glass and melilite,
622 nepheline, leucite, clinopyroxene, apatite, olivine and Fe-Ti oxides: a unique
623 scenario. Lithos, 332, 296-311.
624 Mukherjee, R., Venkatesh, A.S., and Fareeduddin. (2017). Chemistry of magnetite-
625 apatite from albitite and carbonate-hosted Bhukia Gold Deposit, Rajasthan,
626 western India–An IOCG-IOA analogue from Paleoproterozoic Aravalli
627 Supergroup: Evidence from petrographic, LA-ICP-MS and EPMA studies. Ore
628 Geology Reviews, 91, 509-529.
629 Nathwani, C.L., Wilkinson, J.J., Fry, G., Armstrong, R.N., Smith, D.J., and Ihlenfeld,
630 C. (2022). Machine learning for geochemical exploration: classifying
631 metallogenic fertility in arc magmas and insights into porphyry copper deposit
632 formation. Mineralium Deposita, pp.1-24.
633 Nielsen, D. (2016). Tree boosting with xgboost-why does xgboost win" every" machine
634 learning competition? (Master's thesis, NTNU).
635 Noble, W.S. (2006). What is a support vector machine?. Nature biotechnology, 24(12),
636 1565-1567.
637 Ogunleye, A., and Wang, Q.G. (2019). XGBoost model for chronic kidney disease
638 diagnosis. IEEE/ACM transactions on computational biology and bioinformatics,
639 17(6), pp.2131-2140.

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

640 Oliphant, T.E. (2006). A guide to NumPy (Vol. 1, p. 85). USA: Trelgol Publishing.

641 O'Sullivan, G., Chew, D., Kenny, G., Henrichs, I., and Mulligan, D. (2020). The trace
642 element composition of apatite and its application to detrital provenance studies.
643 Earth-Science Reviews, 201, 103044.

644 Pan, L.C., Hu, R.Z., Wang, X.S., Bi, X.W., Zhu, J.J., and Li, C. (2016). Apatite trace
645 element and halogen compositions as petrogenetic-metallogenic indicators:
646 Examples from four granite plutons in the Sanjiang region, SW China. Lithos, 254,
647 118-130.

648 Pearce, J.A. (1996). A user's guide to basalt discrimination diagrams. Trace element
649 geochemistry of volcanic rocks: applications for massive sulphide exploration.
650 Geological Association of Canada, Short Course Notes, 12(79), 113.

651 Petrelli, M., and Perugini, D. (2016). Solving petrological problems through machine
652 learning: the study case of tectonic discrimination using geochemical and isotopic
653 data. Contributions to Mineralogy and Petrology, 171(10), 1-15.

654 Petrelli, M., Bizzarri, R., Morgavi, D., Baldanza, A., and Perugini, D. (2017).
655 Combining machine learning techniques, microanalyses and large geochemical
656 datasets for tephrochronological studies in complex volcanic areas: New age
657 constraints for the Pleistocene magmatism of central Italy. Quaternary
658 Geochronology, 40, 33-44.

659 Petrelli, M., Caricchi, L., and Perugini, D. (2020). Machine learning thermo-barometry:
660 Application to clinopyroxene-bearing magmas. Journal of Geophysical Research:
661 Solid Earth, 125(9), e2020JB020130.

662 Pisiak, L.K., Canil, D., Lacourse, T., Plouffe, A., and Ferbey, T. (2017). Magnetite as
663 an indicator mineral in the exploration of porphyry deposits: A case study in till
664 near the Mount Polley Cu-Au deposit, British Columbia, Canada. Economic
665 Geology, 112(4), 919-940.

666 Prowatke, S., and Klemme, S. (2006). Trace element partitioning between apatite and
667 silicate melts. Geochimica et Cosmochimica Acta, 70(17), pp.4513-4527.

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

668 Qiu, K.F., Yu, H.C., Hetherington, C., Huang, Y.Q., Yang, T., and Deng, J. (2021).
669 Tourmaline composition and boron isotope signature as a tracer of magmatic-
670 hydrothermal processes. American Mineralogist: Journal of Earth and Planetary
671 Materials, 106(7), 1033-1044.
672 Schönig, J., von Eynatten, H., Tolosana-Delgado, R., and Meinhold, G. (2021). Garnet
673 major-element composition as an indicator of host-rock type: a machine learning
674 approach using the random forest classifier. Contributions to Mineralogy and
675 Petrology, 176(12), 1-21.
676 Sha, L.K., and Chappell, B.W. (1999). Apatite chemical composition, determined by
677 electron microprobe and laser-ablation inductively coupled plasma mass
678 spectrometry, as a probe into granite petrogenesis. Geochimica et Cosmochimica
679 Acta, 63(22): 3861-3881.
680 Shen, Y., Ruijsch, J., Lu, M., Sutanudjaja, E.H., and Karssenberg, D. (2022). Random
681 forests-based error-correction of streamflow from a large-scale hydrological model:
682 Using model state variables to estimate error terms. Computers & Geosciences,
683 159, 105019.
684 Sillitoe, R.H. (2010). Porphyry copper systems. Economic geology, 105(1), 3-41.
685 Snider, L.A., and Swedo, S.E. (2004). PANDAS: current status and directions for
686 research. Molecular psychiatry, 9(10), pp.900-907.
687 Snow, C.A. (2006). A reevaluation of tectonic discrimination diagrams and a new
688 probabilistic approach using large geochemical databases: Moving beyond binary
689 and ternary plots. Journal of Geophysical Research: Solid Earth, 111(B6).
690 Soofi, A.A., and Awan, A. (2017). Classification techniques in machine learning:
691 applications and issues. Journal of Basic and Applied Sciences, 13, 459-465.
692 Stock, M.J., Humphreys, M., Smith, V.C., Isaia, R., and Pyle, D.M. (2016). Late-stage
693 volatile saturation as a potential trigger for explosive volcanic eruptions. Nature
694 Geoscience, 9(3), pp.249-254.
695 Sun, J.F., Yang, J.H., Zhang, J.H., Yang, Y.H., and Zhu, Y.S. (2021). Apatite

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

696 geochemical and SrNd isotopic insights into granitoid petrogenesis. Chemical
697 Geology, 566, 120104.
698 Vapnik VN. (1995). The nature of statistical learning theory.
699 Wang, C., Deng, C., and Wang, S. (2020). Imbalance-XGBoost: leveraging weighted
700 and focal losses for binary label-imbalanced classification with XGBoost. Pattern
701 Recognition Letters, 136, pp.190-197.
702 Wang, Y., Qiu, K.F., Hou, Z.L., and Yu, H.C. (2022). Quartz Ti/Ge-P discrimination
703 diagram: A machine learning based approach for deposit classification. Acta
704 Petrologica Sinica, 38(1), 281-290.
705 Wang, Y., Qiu, K.F., Müller, A., Hou, Z. L., Zhu, Z.H., and Yu, H.C. (2021). Machine
706 Learning Prediction of Quartz Forming-Environments. Journal of Geophysical
707 Research: Solid Earth, 126(8), e2021JB021925.
708 Waskom, M.L. (2021). Seaborn: statistical data visualization. Journal of Open Source
709 Software, 6(60), p.3021.
710 Xing, K., Shu, Q., and Lentz, D.R. (2021). Constraints on the formation of the giant
711 Daheishan porphyry Mo deposit (NE China) from whole-rock and accessory
712 mineral geochemistry. Journal of Petrology, 62(4), egab018.
713 Yang, J.H., Kang, L.F., Peng, J.T., Zhong, H., Gao, J.F., and Liu, L. (2018). In-situ
714 elemental and isotopic compositions of apatite and zircon from the Shuikoushan
715 and Xihuashan granitic plutons: Implication for Jurassic granitoid-related Cu-Pb-
716 Zn and 646 W mineralization in the Nanling Range, South China. Ore Geology
717 Reviews, 93, 647 382-403.
718 Yu, H.C., Qiu, K.F., Chew, D., Yu, C., Ding, Z.J., Zhou, T., Li, S., and Sun, K.F. (2022).
719 Buried Triassic rocks and vertical distribution of ores in the giant Jiaodong gold
720 650province (China) revealed by apatite xenocrysts in hydrothermal quartz veins.
721 Ore 651Geology Reviews, 140, 104612.
722 Yu, H.C., Qiu, K.F., Hetherington, C.J., Chew, D., Huang, Y.Q., He, D.Y., Geng, J.Z.,
723 and Xian, H.Y. (2021). Apatite as an alternative petrochronometer to trace the

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

724 evolution of magmatic systems containing metamict zircon. Contributions to
725 Mineralogy and Petrology, 176(9), 1-19.
726 Zhang, G.L., Wang, S., Zhang, J., Zhan, M.J., and Zhao, Z.H. (2020). Evidence for the
727 essential role of CO2 in the volcanism of the waning Caroline mantle plume.
728 Geochimica et Cosmochimica Acta, 290, 391-407.
729 Zhong, R., Deng, Y., Li, W., Danyushevsky, L.V., Cracknell, M.J., Belousov, I., Chen,
730 Y.J., and Li, L.M. (2021). Revealing the multi-stage ore-forming history of a
731 mineral deposit using pyrite geochemistry and machine learning-based data
732 interpretation. Ore Geology Reviews, 133, 104079.
733 Zhou, R.J., Wen, G., Li, J.W., Jiang, S.Y., Hu, H., Deng, X. D., Zhao, X.F., Yan, D.R.,
734 Wei, K.T., Cai, H.A., Shang, S.C., Li, B.C., and Dai, X.K. (2022a). Apatite
735 chemistry as a petrogenetic–metallogenic indicator for skarn ore-related granitoids:
736 an example from the Daye Fe–Cu–(Au–Mo–W) district, Eastern China.
737 Contributions to Mineralogy and Petrology, 177(2), 1-21.
738 Zhou, T., Qiu, K.F., Wang, Y., Yu, H.C., and Hou, Z. L. (2022b). Apatite Eu/Y-Ce
739 discrimination diagram: A big data based approach for provenance classification.
740 Acta Petrologica Sinica, 38(1), 291-299.
741

742 **Figure Captions**

743 **Figure 1.** Locations of apatite samples investigated in this study. (a) The 245
744 publications with apatite compositional data cover 49 countries on six continents.
745 Countries are colored according to the number of apatite trace element data (orange
746 high, green low). (b) Pie chart of continent distribution. (c) Pie chart of deposit type
747 distribution. IOCG - iron oxide copper gold deposits, IOA - iron oxide-apatite deposits.
748

749 **Figure 2.** Box plots and line plots showing the abundances and dispersion of the
750 selected 14 trace elements in apatite. (a, b) The box plots of data categorized according
751 to deposit types. The height of the colored bars represents the interquartile range (25^{th}-

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

752 75th percentile). The horizontal lines within the colored bars are the median. Whiskers
753 show the 5th-95th percentile. The rhombuses (diamond shapes) represent outliers of
754 more than 1.5σ. Unknown denotes the deposit type is known but the locality is not
755 specified.
756
757 **Figure 3.** Workflow employed to develop the machine learning model. (a) Creating the
758 initial dataset after data pre-processing; (b) Using the training dataset to train four
759 different algorithms and then using the testing dataset to evaluate and compare their
760 performance to select the optimal one; (c) Calculating the SHAP value of each feature
761 (i.e. element) in the initial dataset and constructing the filtered dataset with the most
762 important (i.e. source-diagnostic) elements; (d) Retraining and testing the chosen
763 algorithm based on the filtered dataset to yield the final classifier; (e) Determining the
764 probable deposit type based on the trace element data.
765
766 **Figure 4.** Confusion matrix of the testing set used to evaluate the accuracy of the four
767 algorithms. (a) KNN; (b) Random Forest; (c) SVM; (d) XGBoost. The algorithm
768 method and its respective F1 score are presented above each panel while the numbers
769 at the top and bottom of each square represent the proportion of predicted deposit types
770 and the number of predicted deposit types respectively.
771
772 **Figure 5.** The mean SHAP value of each element and test F1 and cross-validation
773 scores of the XGBoost model. The bar plot shows the mean SHAP value of each element,
774 which reflects its contribution to the model prediction. The lines reflect the change in
775 algorithm performance with increasing number of elements (red = cross-validation
776 score; orange = test F1 score).
777
778 **Figure 6.** The cross-validation F1-score across the gamma and max_depth grid search.
779 The optimal combination is gamma=0 and max_depth=9.

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8805. http://www.minsocam.org/

780

781 **Figure 7.** Scatterplots and kernel density curves for different apatite trace element or
782 element ratio combinations. (a) Th/Pr *vs* U/Pr; (b) Sr *vs* Y; (c) Th *vs* U; (d) Eu *vs* Dy.

783

784 **Figure 8.** Confusion matrix of the testing set to evaluate the accuracy of the XGBoost
785 classifier. The numbers in the top and bottom of each square represent the proportion
786 of predicted deposit types and the number of predicted deposit types respectively. Note
787 the score in this confusion matrix and the evaluation report (Table 3) differ slightly from
788 the scores presented in the confusion matrix in Figure 4. In this figure and Table 3, the
789 XGBoost model was optimized further to use three hyperparameters (n_estimators,
790 gamma, and max_depth) and the splitting of the training set and testing set was iterated
791 50 times, both of which improved the classifier accuracy.

792

793 **Figure 9.** SHAP summary plots of apatite trace element data various deposit types. (a)
794 IOCG; (b) IOA; (c) Porphyry; (d) Skarn; (e) Orogenic Au; (f) Unmineralized rocks.
795 Each line represents one element from the dataset in decreasing order of importance,
796 and the abscissa is the SHAP value. When the SHAP value exceeds 0, the feature has a
797 positive impact and vice versa. A small circle (dot) represents an individual analysis
798 and the color represents the concentration of the respective element (red = high, blue =
799 low).

800

801 **Table**

802 **Table 1** Apatite trace element data description

Deposit type	Apatite type	Location	Country	Selected reference
IOCG	Magmatic/ Hydrothermal	Wernecke, Bhukia, Wirrda Well prospect, Acropolis prospect	USA, Australia, India	Mao et al., 2016; Mukherjee et al., 2017; Krneta et al., 2017
IOA	Magmatic/	Durango, Aoshan,	Mexico, Canada, China	Mao et al., 2016

Always consult and cite the final, published document. See http:/www.minsocam.org or GeoscienceWorld

	Hydrothermal	Great Bear		
Orogenic Au	Hydrothermal	Congress (Lou), Kirkland Lake, Dentonia, Seabee, Laodou, Xindigou, Hutti	Canada, China, USA, India	Mao et al., 2016; Hazarika et al., 2016; Zhang et al., 2020
Porphyry	Magmatic (/Hydrothermal)	Boss Mountain, Mount Polley, Shiko, Kemess South, Highmont, Highland Valley, Gibraltar, Brenda, Endako, Cassiar Moly, Dobbin, Lornex, Willa, Daheishan	Canada, China, USA, German, South Africa, Kazakhstan	Cao et al., 2012; Mao et al., 2016; Pan et al., 2016; Xing et al., 2021
Skarn	Hydrothermal	Racine, Minyari, Little Billie, Gold Canyon, O'Callagham's, Molly, Yangla, Shuikoushan, Cantung	Canada, China, USA, Kazakhstan	Cao et al., 2012; Mao et al., 2016; Adlakha et al., 2018; Yang et al., 2018; Jia et al., 2020
Unmineralized	Magmatic	Hawaiian Islands, European orogenic belt, Jan mayen, North Atlantic igneous province, Mexican volcanic belts, Sulawesi Arc	Canada, China, USA, German, South Africa, British, France, Brazil, Chile, Cabo Verde, Russia, Bolivia, Congo, Morocco, Czech, Finland, Greek, Hungary, Italy, Kenya, Norway, Spain, Tanzania, Turkey, Peru	Acosta et al., 2010; Laurent et al., 2017; Henrichs et al., 2018; Minissale et al., 2019; Matusiak et al., 2021; Sun et al., 2021

Table 2 Optimal hyperparameters and test scores of the four applied algorithms

Algorithms	**Best hyperparameters**	**Hyperparameter cross-validation score**	**Test score**
KNN	n_neighbors=2; p=5	99.0%	88.6%

RF	n_estimators=130	98.8%	89.8%
SVM	C=64; gamma=0.5	99.2%	89.7%
XGBoost	n_estimators=148	98.8%	90.8%

Table 3 Evaluation of 50 iterations of the final XGBoost classifier

	precision	**recall**	**F1 score**	**support**
IOCG	0.70±0.12	0.74±0. 12	0.71±0.09	15.80±3.63
IOA	0.99±0.01	0.98±0.02	0.98±0.01	52.08±5.26
Orogenic Au	0.91±0.03	0.90±0.04	0.90±0.03	49.33±6.40
Porphyry	0.86±0.04	0.87±0.04	0.87±0.03	84.57±7.71
Skarn	0.93±0.03	0.92±0.03	0.92±0.02	108.61±11.11
Unmineralized	0.96±0.01	0.96±0 .01	0.96±0.01	506.61±13.78
Accuracy			0.94±0.01	817.00
Macro avg.	0.89±0.02	0.90±0.02	0.89±0.02	817.00
Weighted avg.	0.94±0.01	0.94±0.01	0.94±0.01	817.00

Fig 1

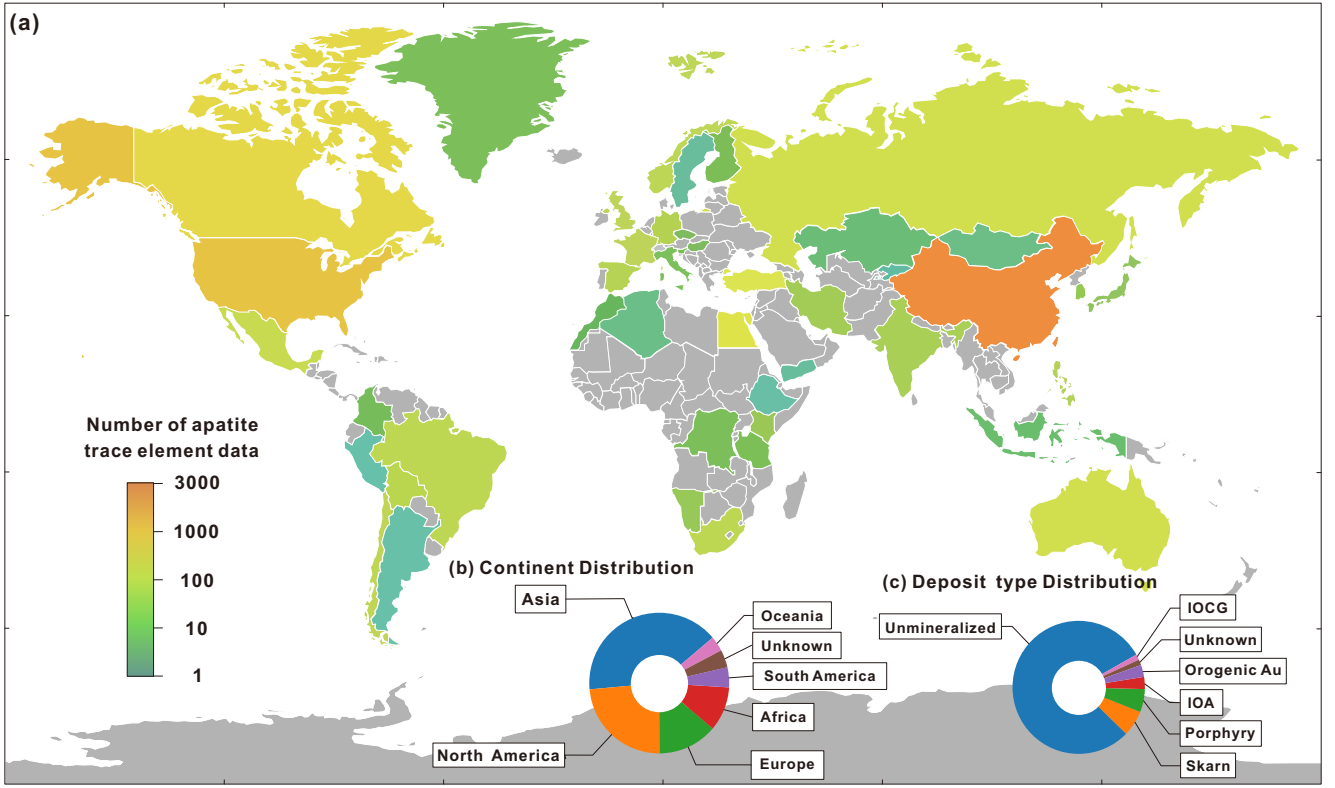

(a)
Number of apatite trace element data
3000
1000
100
10
1
(b) Continent Distribution
Asia
Oceania
Unknown
South America
Africa
Europe
North America
(c) Deposit type Distribution
Unmineralized
IOCG
Unknown
Orogenic Au
IOA
Porphyry
Skarn

Fig 2

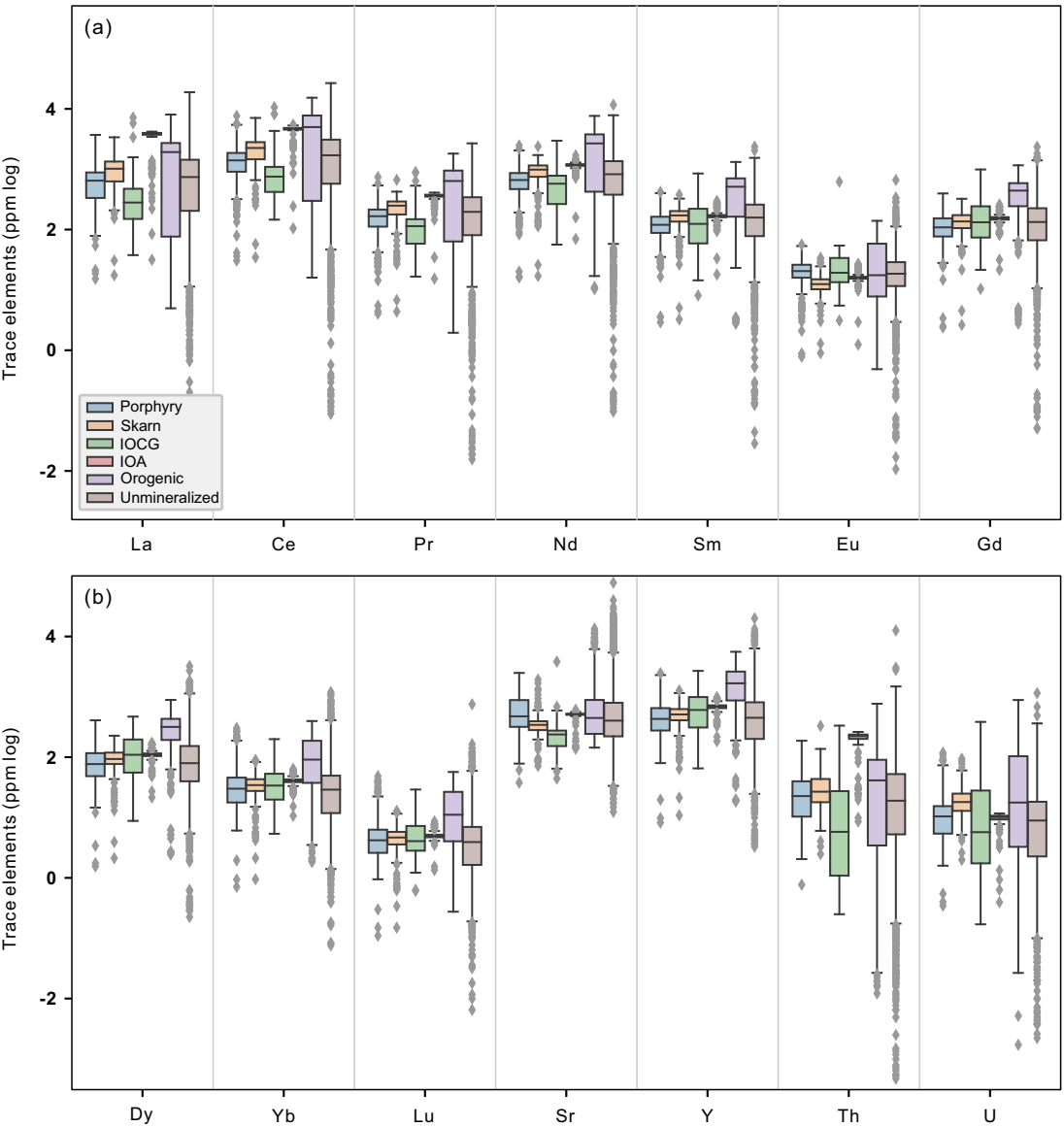

(a)
Trace elements (ppm log)
4
2
0
-2
Porphyry
Skarn
IOCG
IOA
Orogenic
Unmineralized
La
Ce
Pr
Nd
Sm
Eu
Gd
(b)
Trace elements (ppm log)
4
2
0
-2
Dy
Yb
Lu
Sr
Y
Th
U

Fig 3

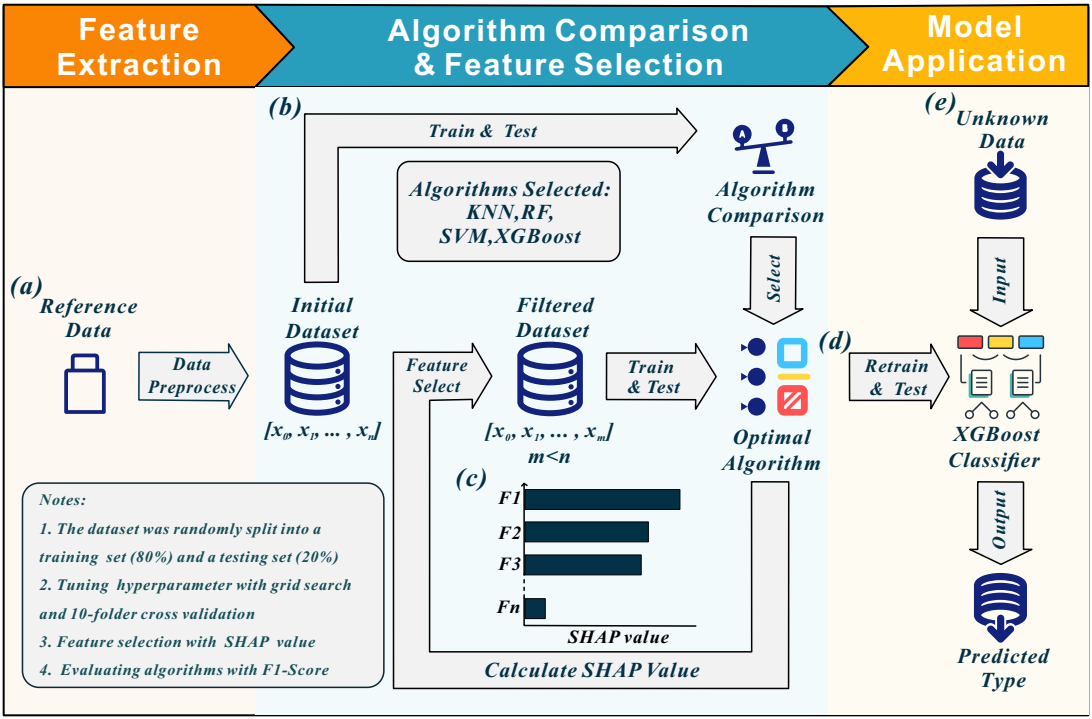

Feature Extraction
Algorithm Comparison & Feature Selection
Model Application
(a)
Reference Data
Data Preprocess
Initial Dataset
[x0, x1, ... , xn]
(b)
Train & Test
Algorithms Selected: KNN,RF, SVM,XGBoost
Algorithm Comparison
Select
Feature Select
Filtered Dataset
[x0, x1, ... , xm]
m<n
Train & Test
Optimal Algorithm
(c)
F1
F2
F3
Fn
SHAP value
Calculate SHAP Value
(d)
Retrain & Test
(e)
Unknown Data
Input
XGBoost Classifier
Output
Predicted Type
Notes:
1. The dataset was randomly split into a training set (80%) and a testing set (20%)
2. Tuning hyperparameter with grid search and 10-folder cross validation
3. Feature selection with SHAP value
4. Evaluating algorithms with F1-Score

Fig 4

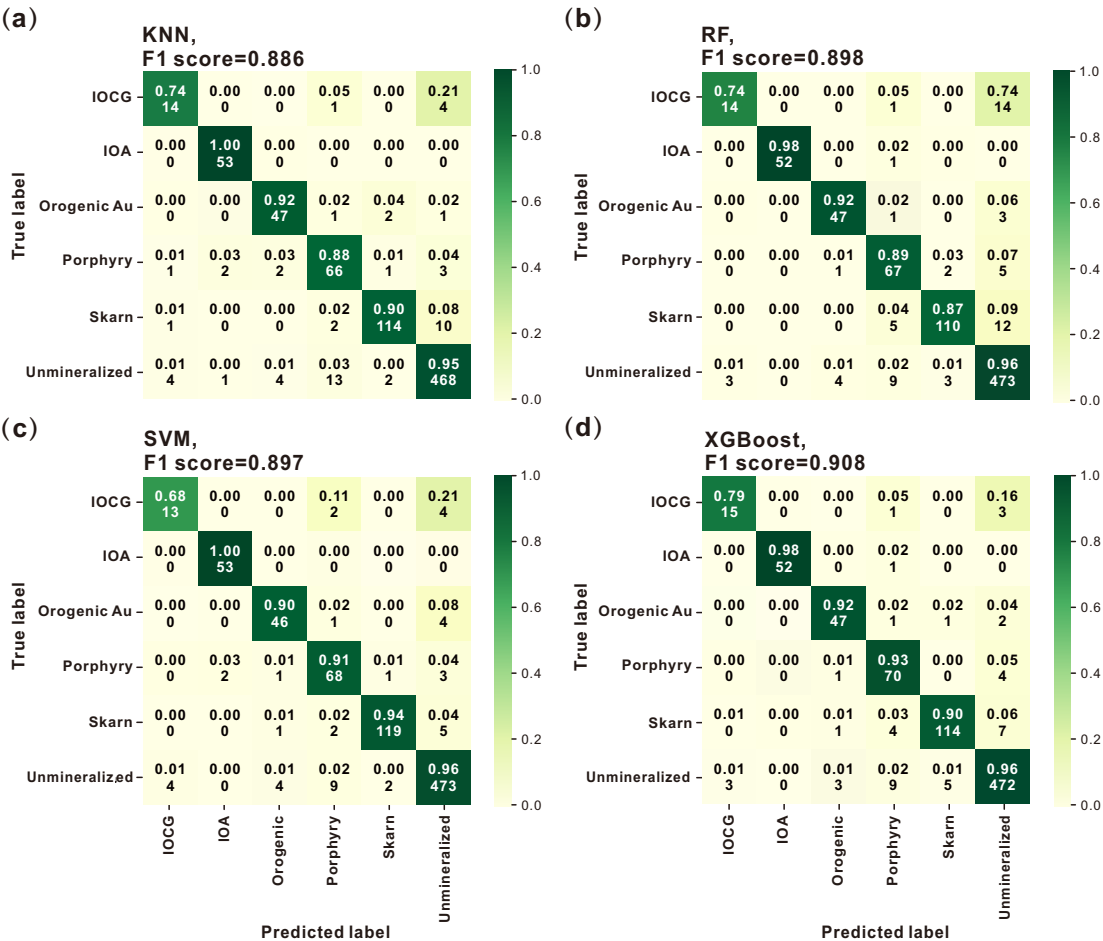

(a)
KNN,
F1 score=0.886
(b)
RF,
F1 score=0.898
(c)
SVM,
F1 score=0.897
(d)
XGBoost,
F1 score=0.908
True label
Predicted label
IOCG
IOA
Orogenic Au
Porphyry
Skarn
Unmineralized
Orogenic
1.0
0.8
0.6
0.4
0.2
0.0

Fig 5

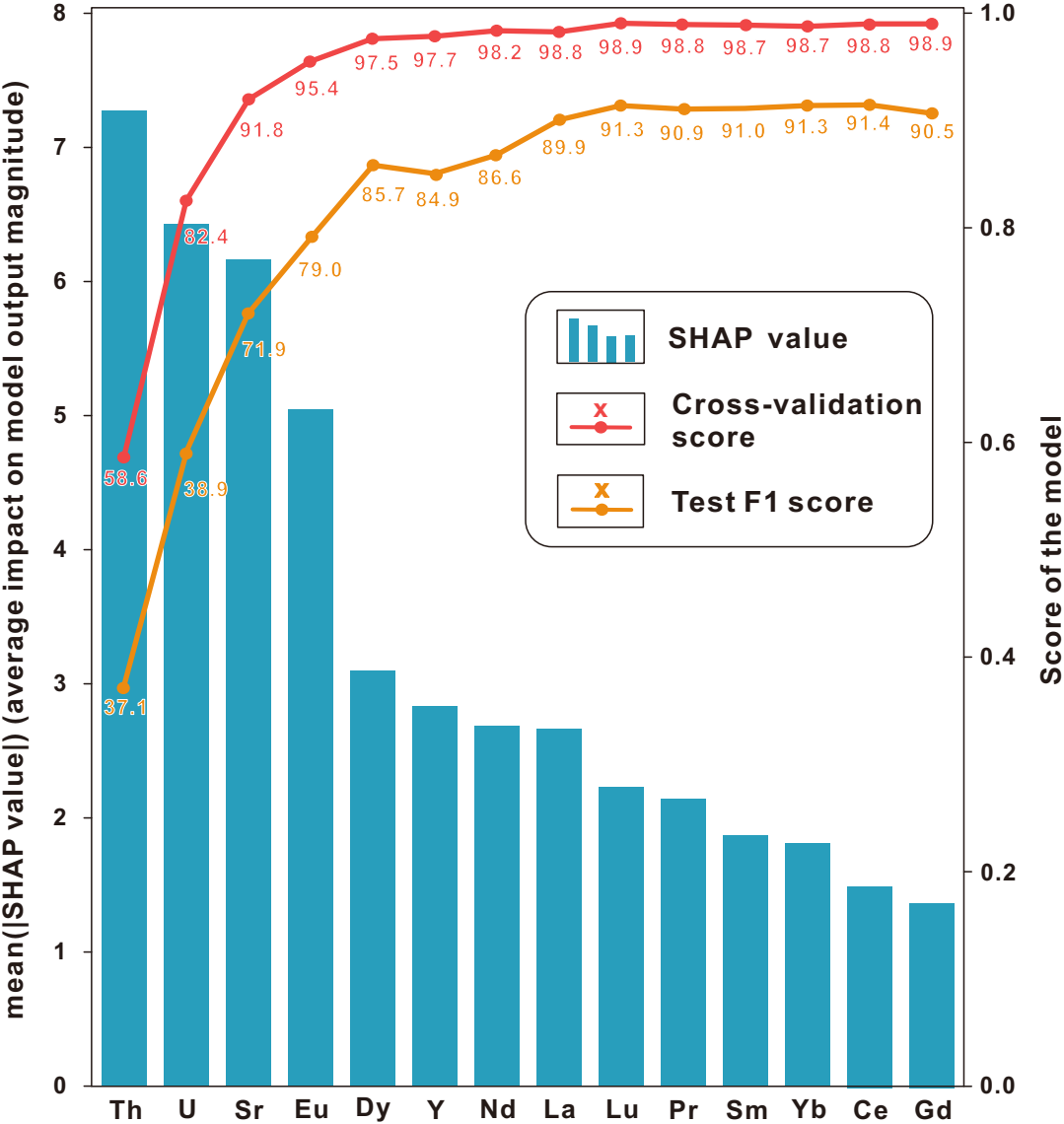

mean(|SHAP value|) (average impact on model output magnitude)
Score of the model
SHAP value
Cross-validation score
Test F1 score
58.6
91.8
95.4
97.5
97.7
98.2
98.8
98.9
98.8
98.7
98.7
98.8
98.9
37.1
71.9
79.0
85.7
84.9
86.6
89.9
91.3
90.9
91.0
91.3
91.4
90.5
Th
U
Sr
Eu
Dy
Y
Nd
La
Lu
Pr
Sm
Yb
Ce
Gd

Fig 6

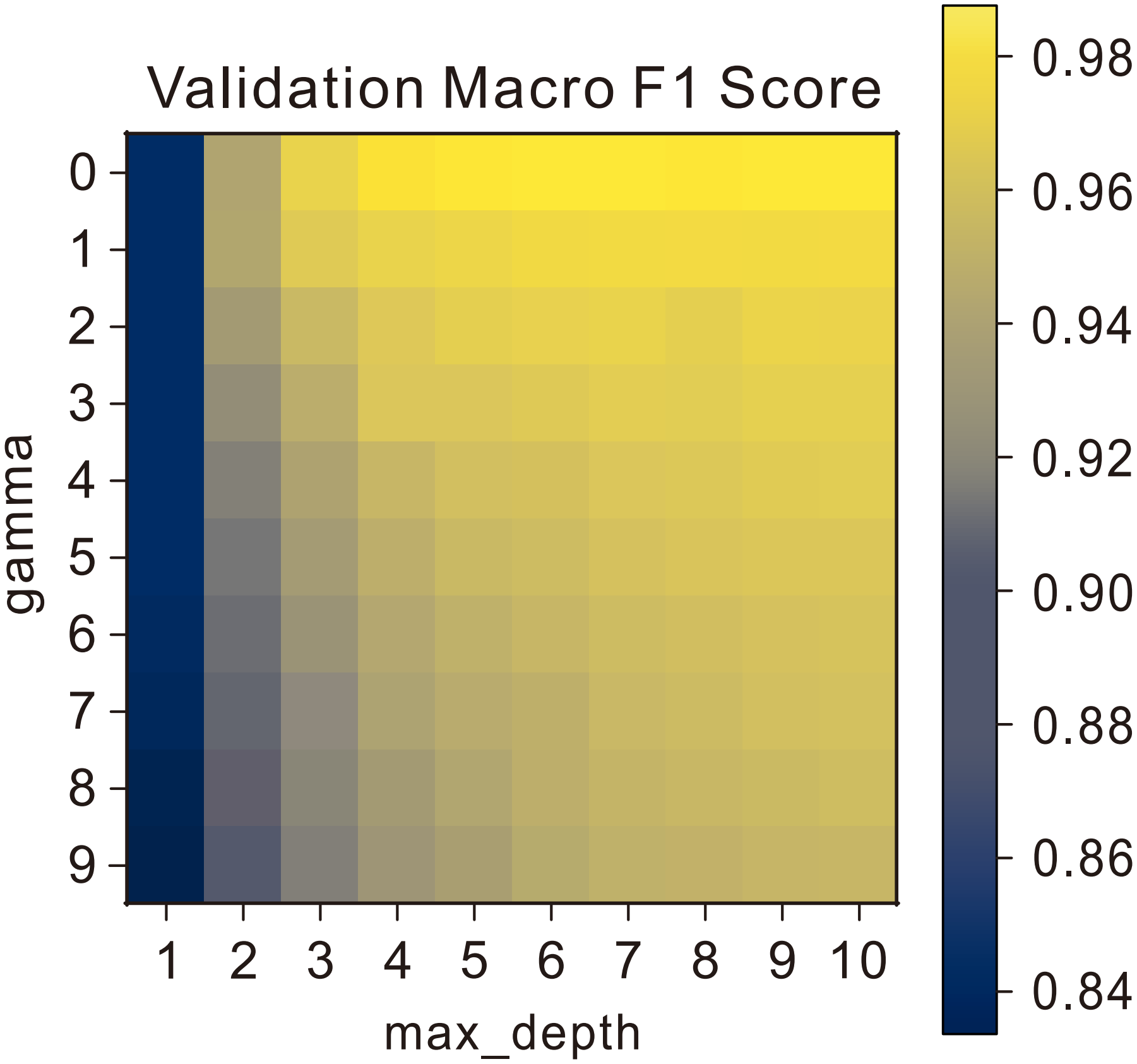

Validation Macro F1 Score
gamma
0
1
2
3
4
5
6
7
8
9
1
2
3
4
5
6
7
8
9
10
max_depth
0.98
0.96
0.94
0.92
0.90
0.88
0.86
0.84

Fig 7

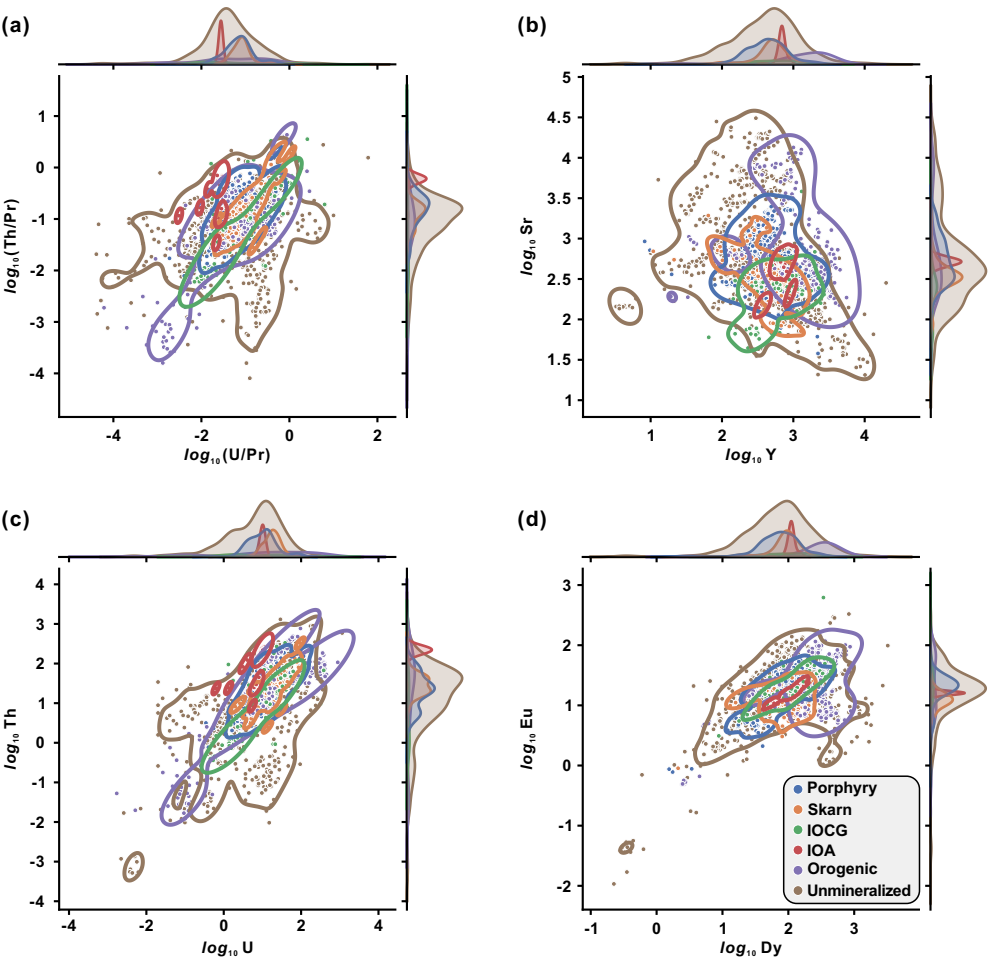

(a)
log10(Th/Pr)
log10(U/Pr)
(b)
log10 Sr
log10 Y
(c)
log10 Th
log10 U
(d)
log10 Eu
log10 Dy
Porphyry
Skarn
IOCG
IOA
Orogenic
Unmineralized

Fig 8

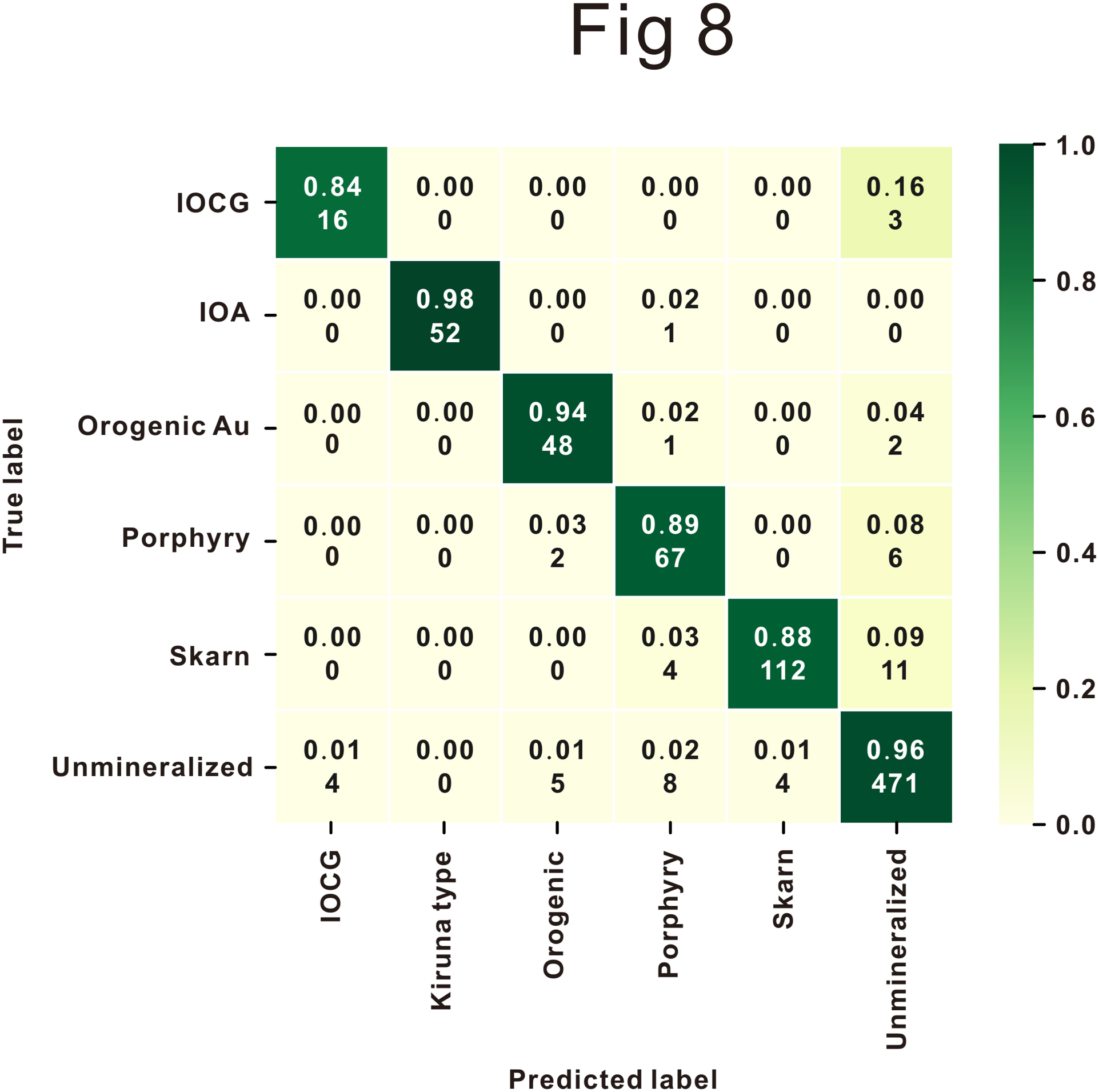
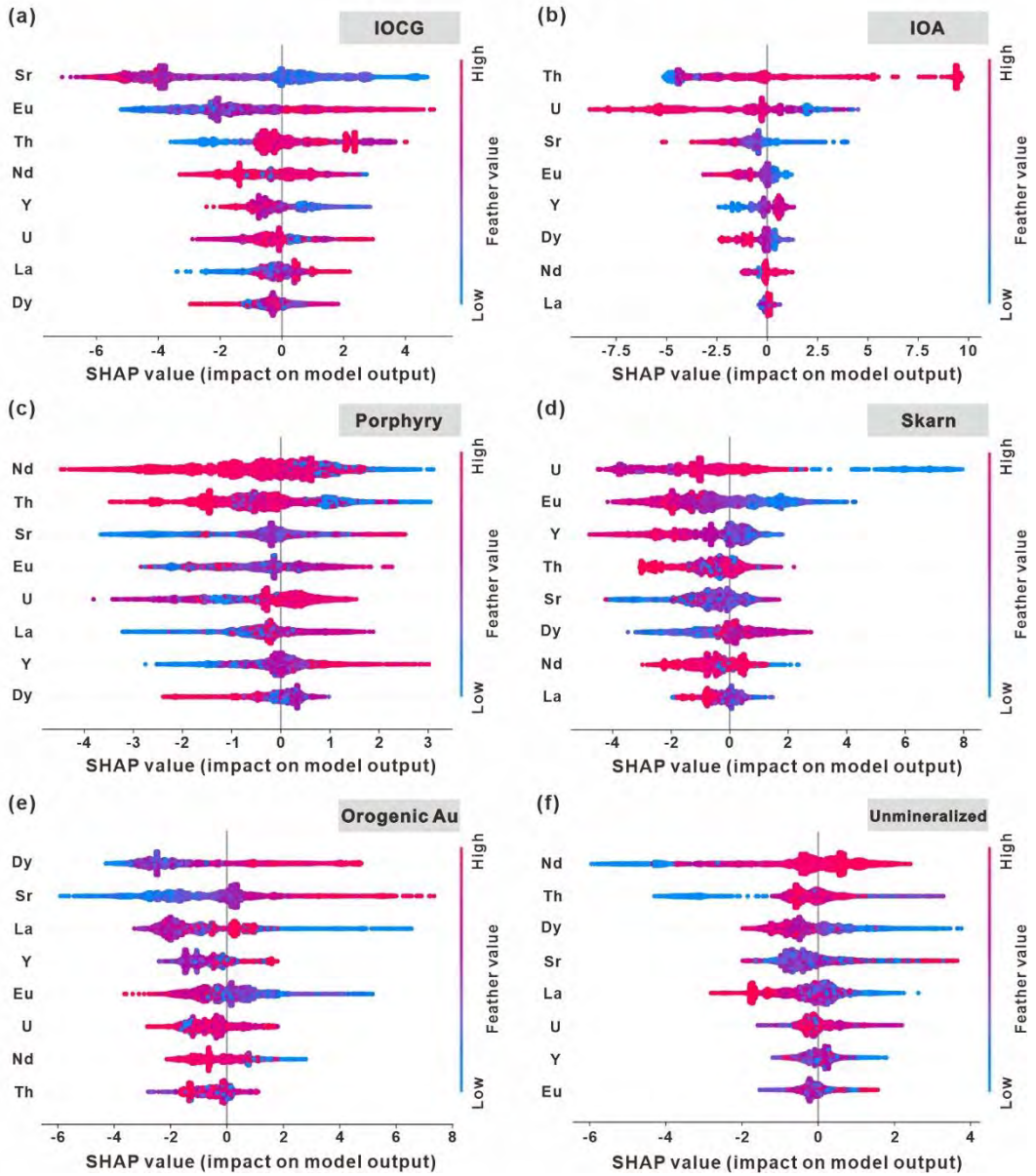

True label
IOCG
IOA
Orogenic Au
Porphyry
Skarn
Unmineralized
0.84 16
0.00 0
0.00 0
0.00 0
0.00 0
0.16 3
0.00 0
0.98 52
0.00 0
0.02 1
0.00 0
0.00 0
0.00 0
0.00 0
0.94 48
0.02 1
0.00 0
0.04 2
0.00 0
0.00 0
0.03 2
0.89 67
0.00 0
0.08 6
0.00 0
0.00 0
0.00 0
0.03 4
0.88 112
0.09 11
0.01 4
0.00 0
0.01 5
0.02 8
0.01 4
0.96 471
IOCG
Kiruna type
Orogenic
Porphyry
Skarn
Unmineralized
Predicted label
1.0
0.8
0.6
0.4
0.2
0.0

Fig 9


(a)
IOCG
Sr
Eu
Th
Nd
Y
U
La
Dy
-6
-4
-2
0
2
4
SHAP value (impact on model output)
High
Feather value
Low
(b)
IOA
Th
U
Sr
Eu
Y
Dy
Nd
La
-7.5
-5
-2.5
0
2.5
5
7.5
10
SHAP value (impact on model output)
High
Feather value
Low
(c)
Porphyry
Nd
Th
Sr
Eu
U
La
Y
Dy
-4
-3
-2
-1
0
1
2
3
SHAP value (impact on model output)
High
Feather value
Low
(d)
Skarn
U
Eu
Y
Th
Sr
Dy
Nd
La
-4
-2
0
2
4
6
8
SHAP value (impact on model output)
High
Feather value
Low
(e)
Orogenic Au
Dy
Sr
La
Y
Eu
U
Nd
Th
-6
-4
-2
0
2
4
6
8
SHAP value (impact on model output)
High
Feather value
Low
(f)
Unmineralized
Nd
Th
Dy
Sr
La
U
Y
Eu
-6
-4
-2
0
2
4
SHAP value (impact on model output)
High
Feather value
Low