

Classifying metallogenic fertility in magmatic arcs using supervised machine learning techniques and implications for porphyry copper deposit formation

CHE TAN NATHWANI^{1,2}, JAMIE J WILKINSON^{2,3} AND
ROBIN ARMSTRONG¹

¹Natural History Museum

²Imperial College London

³The Natural History Museum

Presenting Author: cln14@ic.ac.uk

Supervised machine learning techniques provide a powerful new approach in the challenge to discover ore deposits unexposed at the surface. An area of current exploration focus is the discovery of magmatic districts predisposed to host porphyry copper deposits, using large geochemical datasets. In this study, we train and test four common supervised machine learning techniques: logistic regression, support vector machines, artificial neural networks and Random ForestTM to classify metallogenic fertility in arc rocks using a global data compilation. Several properties of geochemical datasets, such as class imbalance, high multicollinearity, sparsity, missing values and compositional data effects preclude the application of traditional statistics and many machine learning techniques. Application of machine learning techniques on geochemical data therefore requires critical pre-processing steps. We evaluate the classification accuracy of each supervised machine learning technique, showing high accuracy for each model. A further validation test dataset, comprising three deposits of varying size, metal endowment and tectonic setting, shows similar classification performance in all cases. We demonstrate the value of our approach by comparing classification metrics to those from bivariate classification schemes. This highlights the value in utilising the high dimension space available in geochemical data for classification problems. Feature analysis of our models reveals that the most important parameters for classification reflect early and abundant fractionation of amphibole and/or garnet and suppression of plagioclase during lower crustal, hydrous magma evolution. We cannot exclude that some signatures may additionally reflect relatively subtle hydrothermal alteration effects. Empirically, our models underscore the potential of using supervised machine learning approaches in mineral exploration, and that the ‘black box’ nature of these techniques can be overcome.