Title:	Min-Py: a python-based tool for automated identification and classification of common-rock forming minerals
Authors:	Priyanshu
Keywords:	Algorithms Machine learning Electron Probe Model Evaluation
Issue Date:	May-2024
Publisher:	IISER Mohali
Abstract:	Minerals, most basic units of any lithology, are naturally occurring inorganic solids with constrained chemical compositions and atomic structures. Mineral chemistry and its changes are essential for lithological characterization and deciphering the evolutionary history of lithologies, making the estimation of mineral formula an essential step in any petrological workflow. The chemistry is often determined using non-destructive Electron Probe Micro- Analysis (EPMA) and reported in terms of oxides wt.%. A typical workflow involves manual identification of minerals from oxide wt.% data, oxygen assignment, cation calculation, site-assignment and end-member calculation. But, the high-dimensionality and noisy (resulting from external factors including sample nature, location of analysis point and instrument calibration) nature of the EPMA data makes the manual dataset processing and subsequent mineral identification, a time-consuming process with significant risk of unsystematic errors. This is especially true in the cases where EPMA datasets are significantly large, in order of 1000 or higher, and operator do not have access to samples or for the samples are cryptocrystalline in nature. In this context, this study explored the feasibility of automated program for mineral identification and formula calculation from natural mineral compositions. For this purpose, I developed and tested a Python-based program with a supervised machine-learning algorithm (i.e., Support Vector Machine) for identification of mineral, on an EPMA dataset comprising 3800 mineral compositions for 10 mineral classes. The minerals include amphibole, biotite, chlorite, feldspar, garnet, illite, kaolinite, muscovite, pyroxene, and vermiculite. Complete dataset was divided in to 80% and 20% training and validation data respectively. The overall classification accuracy of 99.5%, on validation data, indicates feasibility of such programs for mineral identification from EPMA datasets. The slight fall in accuracy for amphibole (98.2%) can be attributed to compositional similarity of the former with pyroxene. After mineral identification, the program performs cation calculation, site-assignment, Fe3+ estimation (for applicable minerals) and end-member calculations following methods proposed in published literatures. Compared to manual sorting and mineral identification, automation of mineral identification and cation calculation tasks using such programs will significantly reduce the operational time and risks of unsystematic errors. This will allow geoscientists to focus more on the interpretation of their data rather than its processing.
Description:	Under Embargo Period
URI:	http://hdl.handle.net/123456789/5659
Appears in Collections:	MS-19

Files in This Item:

File	Description	Size	Format	
Yet to obtain consent.pdf		144.56 kB	Adobe PDF	View/Open

Show full item record