Automating Lithology Prediction Using Python and ioGAS: A Streamlined Workflow for Geochemical Analysis

Ingrid Flemons[1], Adam Eskdale[1], Michael Whitbread[1]
[1]Rio Tinto Exploration, Australia

Geochemical analysis has traditionally relied upon manual methods, such as exploratory data analysis workflows, to interpret and define relationships between chemical data and the feature of interest, be it rock type, alteration, mineralisation or other parameters. These methods have evolved within mineral exploration to incorporate advanced similarity techniques (e.g., UMAP), enhancing the ability for geoscientists to handle large multi-element datasets.

Handling complex and diverse geochemical datasets still brings complications, such as missing or poorly recorded values and complex spatial relationships between variables. Greater controls and standardized workflows are essential to facilitate geoscientific efforts to leverage these datasets successfully.

This project aims to streamline the workflow for predicting lithology by leveraging similarity algorithms within ioGAS software and novel data science techniques powered by Python. Our approach maintains user control within the familiar ioGAS environment, which includes many published geochemical relationships and robust visualization capabilities that effectively communicate data within the geology and exploration communities. This retains the excellent capabilities within ioGAS to link plots and attribute data intuitively, keeping it accessible to geoscientists. Python packages can add ever-expanding algorithms and code bases (both public and in-house) to this environment.

By utilizing freely available Python packages alongside ioGAS tools, we developed a proof-of-concept workflow that integrates data science into geology and geochemistry processes within a large organization with much reduced intermediate product file handling. Key outcomes from this include: (i) a more efficient workflow for lithological prediction; (ii) reduced time-on-task for the responsible geoscientist, increasing capacity for stronger interpretation; (iii) a demonstration of potential future advancements in this field; (iv) evidence for the wider applicability of combined supervised-unsupervised workflows to facilitate successful discovery; (v) making data science techniques in Python accessible to non-Python proficient geoscientists.