Machine learning integration of hyperspectral and geophysical data for improved exploration targeting and OBK

R.A. Dutch[1], *T. Ostersen*[2], *B.P. Voutharoj*[3] *and M. Paknezhad*[4]

1.Head of Applied Science, Datarock Pty Ltd, Melbourne Victoria 3000. Email: riandutch@datarock.com.au
2.Senior Data Scientist, Datarock Pty Ltd, Melbourne Victoria 3000. Email: thomasostersen@datarock.com.au
3. ML Engineer, Datarock Pty Ltd, Melbourne Victoria 3000. Email: bhanuprakashvoutharoja@datarock.com.au
4.Snr ML Engineer, Datarock Pty Ltd, Melbourne Victoria 3000. Email: mahsapaknezhad@datarock.com.au

Keywords: Data integration, hyperspectral, geophysics, exploration targeting, ore body knowledge

ABSTRACT

With the proliferation of new sensor technologies, acquiring multiple datasets over the same ground is becoming cheaper and easier than ever. This new higher resolution multi-variate data provides a significant resource for exploration and resource geologists but comes with the added complexity of effectively integrating the various datasets in useful and meaningful ways to elucidate new geological understanding.

Both geophysical datasets (either down hole or regional surveys) and hyperspectral data (either on core datasets, airborne or satellite-based sensor systems) are extensively used for exploration targeting, providing different information at different resolutions and crustal scales. However, one of the biggest challenges comes from trying to effectively integrate datasets that record very different physical properties, across the different scales these data are captured at, in a way which can allow for a data-driven analysis of these combined datasets.

Machine learning techniques can be particularly useful as a means of extracting meaningful information from large datasets and reducing complex multi-variate datasets into simpler vector representations of the important features. Using a custom masked autoencoder (MAE) model with spatial and spectral attention, we can generate spectral-spatial feature embeddings at any scale from high-resolution hyperspectral data. Then, using a multi-channel convolutional neural network (CNN), we can extract meaningful textural feature embeddings from various geophysical datasets. By capturing these feature embeddings at the same spatial scale, we can effectively integrate these datasets together in a single unified latent space. This new approach allows us to bring multiple, disparate data types together to allow for more robust data-driven prospectivity analysis or target identification based on similarity to other known targets or prospects. Similarly, this methodology can effectively integrate these types of datasets across an ore body, allowing for a more rigorous assessment of the relationships between properties, and the potential to better understand and model a deposit.