Machine learning assisted lithology classification of drill hole data

V Salamakha[1]

1. Graduate Mining Engineer, Alkane Resources, Sydney NSW 2000.
   Email: val.salamakha@outlook.com

ABSTRACT

In mineral exploration, accurate lithology classification of drill hole data is paramount for assessing geological features and enhancing the profitability of mining operations. Traditional methods rely heavily on a geologists' visual inspection and subjective interpretation of core samples, which can be time-consuming and inconsistent. Machine learning algorithms applied in this classification context have the capability to assist in these challenges by producing rapid first-pass interpretations even when the core samples are missing or damaged (Horrocks, Holden and Wedge, 2015).

The main objective of this study is to apply a variety of machine learning methods to conduct supervised classification of lithology types using multivariate geophysical data that is recorded during drilling operations. A total of 11 drill holes and four distinct lithology types were utilised in the application of each method. The data was submitted to training, cross-validation, and testing in three different scenarios using Decision Tree (DT), Random Forest (RF), Multi-layer Perceptron (MLP), Support Vector Classifier (SVC), XGBoost and LightGBM algorithms. A novel approach to cross-validation was employed during the optimisation of each machine learning algorithm to preserve the sequential integrity of each drill hole. Balanced accuracy or macro-average recall was the primary evaluation metric for each algorithm.

The analysis revealed that while most algorithms can easily achieve over 90 per cent balanced accuracy within the same drill hole, their performance declined by approximately 30 per cent when applied to independent and previously unseen drill holes. A small increase in performance was observed in the same context when the models were trained with an expanded set of four drill holes. Among the tested algorithms, RF and SVC consistently outperformed the field on average across all testing scenarios. These results imply that access to a more extensive data set, featuring consistent geological assessments with higher confidence intervals, may significantly improve each algorithm's capability to classify lithology accurately and rapidly.

REFERENCES

Horrocks, T, Holden, E and Wedge, D, 2015. Evaluation of automated lithology classification architectures using highly-sampled wireline logs for coal exploration, *Computers and Geosciences*, 83:209–218.