A Machine-Learning Workflow for the Automated Stratigraphic Interpretation of Pilbara Iron Ore Stratigraphy

Elisabeth Scibiorski[1], Jack Maughan[1], Yasin Dagasan[1], Rian Dutch[1], Bhanu Voutharoja[1], Leam Gatehouse[2], Eric Hein[2], Patric Wilson[2]
[1]Datarock, Australia, [2]Rio Tinto Iron Ore, Australia

Accurately modelling subsurface stratigraphy is crucial for iron ore mining, but manual stratigraphic interpretation is a time-consuming process with varying consistency. We developed a machine learning (ML) workflow to automate interpretation at an iron ore deposit in the Pilbara, Western Australia, using a combination of conventional downhole wireline geophysics (gamma, density, magnetic susceptibility) and geochemical assay data.

The workflow comprises a nested XGBoost model architecture across five levels, designed to predict: detrital/bedrock boundary; mineralised/waste classification; weathering status; stratigraphic unit; and strand unit (a subdivision of stratigraphic unit). Predictions are generated for each 2m geological interval in turn, starting with the detrital/bedrock model. Selection and application of the next model in the sequence depends on the result of the previous model.

Post-prediction processing is used to reduce noise and improve geological continuity. Because the XGBoost models treat each interval in spatial isolation, the initial sequence of predictions may not be geologically reasonable; e.g., multiple alternating detrital and bedrock intervals within a single drillhole. The softmax class probabilities are used to smooth the initial predictions, with algorithms ranging from the enforcement of a single detrital/bedrock boundary at the most probable location, to the use of a Hidden Markov Model for the stratigraphy and strand models.

A robust flagging system is implemented to detect anomalies in ML predictions, both within individual drillholes and in a spatial context. Downhole flagging checks for internal consistency against constraints such as minimum unit thicknesses, and spatial flagging evaluates prediction continuity with neighbouring drillholes.

Overall, the automated workflow is able to produce geologically reasonable stratigraphic predictions, and the models perform well when training data is sufficiently representative of the target geology. These results demonstrate that automation via ML is a promising route to improve the consistency, reproducibility, and efficiency of stratigraphic predictions in iron ore deposits.