Use of Machine Learning to predict recovery of alumina involving multiple analytical methods at Worsley Alumina

J Dangin[1], P Soodi Shoar[2,] T Wilson[3], J Harvey[4] and C Cavelius[5]

1.Data Scientist, DeepLime, Perth WA 6000, johann.dangin@deeplime.io
2.Principal Resource Geology, South32, Perth WA 6000. payam.soodishoar@south32.net
3.Data Scientist, DeepLime, Perth WA 6000, tom.wilson@deeplime.io
4.Lead Resource Geology, South32, Perth WA 6000. joshua.d.harvey@south32.net
5.Lead Developer, DeepLime, Pau France 64000, claude.cavelius@deeplime.io

Keywords: Orebody Knowledge, Machine Learning, Analytical method, Python, Geometallurgy

ABSTRACT

The bauxite deposit at South32's Worsley Alumina in Western Australia extends over ~4000 km^2 from Brookton in the north to Collie in the south. The deposit has been explored and mined for over 40 years and contains more than 2.5 million analytical results. Over time, the analytical method to determine Available Alumina (AAl_2O_3) and Reactive Silica ($RxSiO_2$) has been revised to better represent recovery of alumina through the refinery in response to the change in underlying geology. Linear regressions between different methods are currently being used to convert all data into the latest analytical method (Worsley Laboratory Available Alumina - WLAA) to adjust for differences between the analytical methods. A significant amount of effort is currently needed to maintain this process as new data is added to the database, including evaluating sensitivities and impacts of updating the regressions.

An approach using Machine Learning (ML) has been developed to infer WLAA values for AAl_2O_3 and $RxSiO_2$ where the measurements are based on legacy analytical methods. The automated workflow selects samples which have been analysed with multiple methods, to derive predictive models for each analyte in WLAA. These models are used to predict the missing WLAA results.

Additional workflows have been created to automate geological interpretation, using ML classified geological domains and implicitly modelled surfaces for each domain. This enables rapid assessment and validation of any changes resulting from updating the predicted WLAA values.