

ABSTRAK

Eksplorasi mineral merupakan proses kompleks dalam industri pertambangan yang memerlukan analisis mendalam terhadap data geokimia dan geospasial. PT Citra Palu Minerals (CPM) sebagai perusahaan pertambangan belum menginterpretasikan kandungan mineral pada spasi jarak di antara lubang bor yang telah ada. Keterbatasan ini menyebabkan sulitnya memprediksi potensi mineral di area yang belum dieksplorasi, berdampak pada ketidakpastian strategi pengeboran dan alokasi sumber daya yang tidak optimal. Penelitian ini bertujuan mengembangkan model prediksi kandungan mineral menggunakan algoritma *Random Forest Regression* dengan memanfaatkan data geokimia berupa data *fire assay* yaitu *Au_FA* dan *alteration* serta data geospasial berupa kedalaman pengeboran dan *lithology* dari 187 sampel pengeboran.

Metodologi penelitian meliputi tahapan *preprocessing* data, pemodelan, dan evaluasi. *Preprocessing* data mencakup *feature construction* untuk membuat kolom *hole_id*, penggabungan data dari berbagai file, penanganan *missing value* menggunakan metode *mean* dan interpolasi, serta *feature encoding* dengan *One-Hot Encoding*. Augmentasi data dilakukan menggunakan *Gaussian Copula Synthesizer* untuk mengatasi keterbatasan jumlah sampel dan meningkatkan kemampuan generalisasi model. Model dibangun menggunakan *Random Forest Regression* dengan *hyperparameter tuning* melalui *RandomizedSearchCV* untuk mencari kombinasi parameter optimal. Data dibagi menjadi 80% data latih dan 20% data uji. Evaluasi model menggunakan metrik *Mean Absolute Error (MAE)*, *Mean Squared Error (MSE)*, *Root Mean Squared Error (RMSE)*, dan koefisien determinasi (R^2).

Hasil penelitian menunjukkan model dengan *hyperparameter tuning* menghasilkan performa lebih baik dibandingkan tanpa tuning, dengan nilai *MAE* sebesar 6,762, *MSE* sebesar 79,418, *RMSE* sebesar 8,912, dan R^2 sebesar 0,055. Model berhasil memprediksi lima mineral target yaitu *Quartz*, *Muscovite*, *Montmorillonite/Smectite*, *Illite*, dan *Laumonite*. Sistem prediksi kandungan mineral berbasis *web* berhasil diimplementasikan menggunakan framework *Streamlit* dengan fitur evaluasi model, visualisasi pengeboran 2D dan 3D, serta prediksi interaktif. Sistem telah melalui pengujian *black box testing* dengan hasil sukses pada semua skenario. Penelitian ini memberikan kontribusi dalam penerapan *machine learning* untuk eksplorasi mineral di Indonesia dan menyediakan solusi praktis bagi PT CPM dalam mengoptimalkan strategi pengeboran berdasarkan data geokimia dan geospasial.

Kata Kunci: *Random Forest Regression*, Prediksi Kandungan Mineral, Data Geokimia, Data Geospasial, Eksplorasi Mineral

ABSTRACT

Mineral exploration is a complex process in the mining industry that requires in-depth analysis of geochemical and geospatial data. PT Citra Palu Minerals (CPM) as a mining company has not yet interpreted mineral content in the spacing between existing boreholes. This limitation makes it difficult to predict mineral potential in unexplored areas, resulting in uncertainty in drilling strategies and suboptimal resource allocation. This research aims to develop a mineral content prediction model using the Random Forest Regression algorithm by utilizing geochemical data in the form of fire assay data namely Au_FA and alteration, as well as geospatial data in the form of drilling depth and lithology from 187 drilling samples.

The research methodology includes data preprocessing, modeling, and evaluation stages. Data preprocessing includes feature construction to create the hole_id column, data merging from various files, missing value handling using mean and interpolation methods, and feature encoding with One-Hot Encoding. Data augmentation was performed using Gaussian Copula Synthesizer to overcome the limitation of sample size and improve the model's generalization capability. The model was built using Random Forest Regression with hyperparameter tuning through RandomizedSearchCV to find the optimal parameter combination. The data was divided into 80% training data and 20% testing data. Model evaluation uses Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and coefficient of determination (R^2) metrics.

The research results show that the model with hyperparameter tuning produces better performance compared to without tuning, with MAE value of 6.762, MSE of 79.418, RMSE of 8.912, and R^2 of 0.055. The model successfully predicts five target minerals namely Quartz, Muscovite, Montmorillonite/Smectite, Illite, and Laumontite. A web-based mineral content prediction system was successfully implemented using the Streamlit framework with features including model evaluation, 2D and 3D drilling visualization, and interactive prediction. The system has passed black box testing with successful results in all scenarios. This research contributes to the application of machine learning for mineral exploration in Indonesia and provides a practical solution for PT CPM in optimizing drilling strategies based on geochemical and geospatial data.

Keywords: *Random Forest Regression, Mineral Content Prediction, Geochemical Data, Geospatial Data, Mineral Exploration*