

## INTISARI

### Penerapan Algoritma *XGBoost* Dan *Random Forest* Untuk Estimasi Porositas Efektif Data Log Sumur

Oleh

Muhammad Fajrul Haqqi  
17/411322/PA/17842

Pemodelan regresi menggunakan pembelajaran mesin (*machine learning*) dapat diaplikasikan untuk estimasi data log sumur. Estimasi porositas efektif dari data log sumur perlu untuk dilakukan karena data *core* biasanya tidak tersedia untuk beberapa sumur. Pada penelitian ini dilakukan penerapan *machine learning* menggunakan algoritma *eXtreme Gradient Boosting (XGBoost)* dan *Random Forest* yang berbasis model non-linear *decision tree* untuk estimasi porositas efektif. Model dioptimalkan dengan metode *GridSearchCV* (GS) dan dievaluasi dengan metrik skor  $R^2$ , RMSE, dan MAE. Model diterapkan pada data lapangan “Fajrul” dengan 7 sumur serta variasi pemisahan data pelatihan dan pengujian berdasarkan jumlah sumur.

Hasil penelitian menunjukkan algoritma *Random Forest* lebih akurat dan stabil dalam mengestimasi porositas efektif. Algoritma *XGBoost* lebih cepat dalam proses pembelajaran data dengan waktu terlama 0,72 detik, sedangkan *Random Forest* 6,13 detik. Peningkatan jumlah data pelatihan beriringan dengan peningkatan akurasi dan kestabilan model. Hasil pengujian model terbaik diperoleh saat menghilangkan data porositas efektif yang bernilai 0 dengan optimisasi GS pada pelatihan di 6 sumur, dan pengujian di 1 sumur.

**Kata kunci:** pembelajaran mesin, log sumur, porositas efektif, *GridSearchCV*, *Random Forest*, *XGBoost*.

## **ABSTRACT**

### **Application of XGBoost and Random Forest algorithm for Effective Porosity Estimation from Well Log Data**

By

Muhammad Fajrul Haqqi  
17/411322/PA/17842

Regression model using machine learning can be applied to estimate well log data. Estimation of effective porosity from well log data is necessary because core data are not available for some wells. In this study, machine learning is implemented using eXtreme Gradient Boosting (XGBoost) and Random Forest algorithm based on non-linear decision tree model for effective porosity estimation. Model is optimized by GridSearchCV (GS) and evaluated by metrics  $R^2$  score, RMSE, and MAE. Model is applied to from “Fajrul” field with 7 wells and variations of training and testing data split based on amount of wells.

The result of this study shows that Random Forest algorithm is more accurate and stable in estimating effective porosity. XGBoost is faster in data learning process with the longest time 0,72 seconds, while Random Forest is 6,13 seconds. The increasing of amount of the training data is in tandem with the increasing of accuracy and stability of the model. The best results of model test are obtained when eliminating the zero value of effective porosity with GS optimization on training in 6 wells, and testing in one well.

**Keywords:** machine learning, well log, effective porosity, GridSearchCV, Random Forest, XGBoost.