

SISTEM PREDIKSI ZONA POTENSIAL HIDROKARBON BERDASARKAN DATA *WELL-LOG* MENGGUNAKAN METODE *RANDOM FOREST*

Aditya Nur'ahya

ABSTRAK

Kegiatan interpretasi data geofisika merupakan aktivitas analisis yang dilakukan oleh ahli geofisika dalam menginterpretasikan data hasil akuisisi dengan data – data yang berkaitan. Sehingga, mendapatkan pendugaan awal untuk menjelaskan keadaan bawah permukaan. Penelitian ini dilakukan untuk membuat prediksi yang dapat membantu mempercepat dugaan awal kandungan hidrokarbon yang dilakukan oleh seorang ahli geofisika sebagai kegiatan interpretasi. Kegiatan yang interpretasi membutuhkan kecepatan dan ketelitian yang tinggi sehingga dapat menghasilkan hasil analisis yang tepat. Penerapan konsep *data mining* digunakan untuk mengklasifikasikan kandungan hidrokarbon pada lapisan tanah yang di prediksi menggunakan *data well-log*. Salah satu algoritma klasifikasi data mining yang digunakan pada penelitian ini yaitu algoritma *Random Forest*. Algoritma ini digunakan untuk mengklasifikasikan kandungan hidrokarbon. Pada penelitian ini digunakan *dataset* yang didapatkan dari situs resmi *National Offshore Petroleum Information Management System* (NOPIMS) dengan *dataset* awal berjumlah 28.378 data dan setelah dibersihkan data yang digunakan berjumlah 4.556 data, kemudian data dilakukan pembagian menjadi data latih dan data uji untuk pelatihan model. Pada penelitian ini dilakukan evaluasi model yang mencakup pengukuran akurasi, presisi, nilai *recall*, nilai *f1-measure*, dan nilai AUC score yang dilakukan dengan mengukur setiap perbandingan rasio pembagian data. Dengan perbandingan rasio 80- 20 nilai akurasi yang didapatkan sebesar 0,9923 atau 99,23%, nilai presisi 0,97 atau 97%, nilai *recall* 0,98 atau 98%, nilai *f1-measure* 0,97 atau 97%, dan nilai AUC Score 0,9919. Diharapkan implementasi model pada sistem dapat membantu ahli geofisika dalam menginterpretasikan *data well-log* dalam menentukan zona potensial hidrokarbon.

Kata kunci: Prediksi, Hidrokarbon, *Random Forest*

**SISTEM PREDIKSI ZONA POTENSIAL HIDROKARBON
BERDASARKAN DATA *WELL-LOG* MENGGUNAKAN METODE
*RANDOM FOREST***

Aditya Nur'ahya

ABSTRACT

The activity of interpreting geophysical data is an analytical activity carried out by geophysicists in interpreting the acquired data with related data. Thus, obtaining initial estimates to explain the subsurface conditions. This research was conducted to make predictions that could help speed up the initial prediction of the hydrocarbon content made by a geophysicist as an interpretation activity. Interpretation activities require high speed and accuracy so as to produce the right analysis results. The application of the concept of data mining is used to classify the hydrocarbon content in the soil layer which is predicted using well-log data. One of the data mining classification algorithms used in this study is the Random Forest algorithm. This algorithm is used to classify the hydrocarbon content. In this study, datasets obtained from the official website of the National Offshore Petroleum Information Management System (NOPIMS) were used with an initial dataset of 28,378 data and after cleaning the data used amounted to 4,556 data, then the data was divided into training data and test data for model training. In this study, an evaluation of the model was carried out which included measurements of accuracy, precision, recall value, f1-measure value, and AUC score value which was carried out by measuring each ratio of data sharing ratios. With a ratio of 80-20, the accuracy value obtained is 0.9923 or 99.23%, the precision value is 0.97 or 97%, the recall value is 0.98 or 98%, the f1-measure value is 0.97 or 97%, and AUC Score value of 0.9919. It is hoped that the implementation of the model in the system can assist geophysicists in interpreting well-log data in determining potential hydrocarbon zones.

Keyword: Prediction, Hydrocarbon, Random Forests