PROCEEDINGS
JOINT CONVENTION BANDUNG (JCB) 2021
November 23rd – 25th 2021

Software Development using Machine Learning to Predict PPFG and 3D Geomechanics

Pipit Irfani Amorokhman[1], Allen Haryanto Lukmana[2], Pertamina Research Technology Innovation[1] UPN Veteran Yogyakarta[2]

Abstract

Pressure prediction plays a fundamental role to design mud weight and well trajectory for wellbore stability and prevents stuck pipe. Some manual calculations are only able to calculate on certain condition such as clean-shale formation and under-compaction mechanism formation. Unfortunately, real formation can be very heterogenic. A method to produce independent formation type shall be developed solve the issue. Therefore, software using machine learning (ML) were developed to generate scrupulous pressure prediction.

Logging data (e.g., Density, Sonic, Gamma Ray) and drilling parameter (e.g., ROP, RPM, WOB) from 2 wells (MLC-01, TTA-01) were used as machine learning input. In this research, 3 methods which are Artificial Neural Network (ANN) Feedforward type, Random Forest (RF), and Support Vector Machine (SVM) were applied.

The result exhibits (1) ANN showed the least Root Square Mean Error (RSME) of 0.11401 in comparison to the other 3 methods, Determination Coefficient (R^2) 0.9789. Thus, ANN will be used for the rest of the analysis. (2) 4 data (Density, Sonic, Gamma Ray, Depth) together achieve the most precise with actual condition with RSME 0.0714 and R^2 0.9826. (3) After plotting the result in one graph, pore pressure prediction from ANN method is closer to actual pore pressure rather than manual calculation result.

It is to conclude that this software gives a promising result to predict Pore Pressure, Fracture Gradient, and Shear Failure Gradient. The comparative analysis results show that ANN Feedforward type has the feature estimation by its shorter time prediction and high accuracy (a coefficient of determination of 0.99 and RSME 0.08 – 0.23. The overpressure prediction, XRD and Geomechanics can be analysis in one integrated software.

Introduction

Basic geomechanical components are include pore pressure (PP), unconfined compression strength (UCS), Overburden or Vertical stress (Sv), Minimum horizontal stress (Shmin), and maximum horizontal stress (SHmax). These component gained from core measurement. However, core measurement cannot performed in the whole interval regarding of core limitation. Prediction calculation can be alternative, but it require many data and only specific condition to use the method
Pore pressure at depth is equivalent to hydraulic potential measurement respect of earth surface. Assumed to be uniform in small volume of interconnected pores. Therefore, pore pressure can be variated refer to geological events. Overpressure occurs as several factors such as Disequilibrium compaction, tectonic compaction, hydrocarbon column height, Aqua-thermal compaction, mineral diagenesis, and hydrocarbon maturation.

.
Artificial Intelligence (AI) and Machine Learning (ML) grown its popularity among various industry. Its supported by development of technology in computer service makes big data acquisition is more easier than before. However, energy sector is still left behind.

In this paper we rise a case study using data from 2 exploration wells located in North Sumatera. Depth of well reached to 9000 ft, TTA-01 and MLC-01. Overpressure occurred for both of wells during drilling activity. Team developing a software using AI and ML to predict PP from existing log and drilling data. The result of AI prediction will be compared to manual calculations and recorded data.

Data and Method

Geological and Stratigraphical Setting

The fold system is dominated by the WNW-ESE trending anticlinoria. Generally forming an en-echelon pattern explanation:

- Wrench movement along the NW-SE basement faults (related to the Paleogene graben system) (Harding,1988) .
-Draping over uplifted blocks due to compressive regime of the subduction (Moulds, 1989).

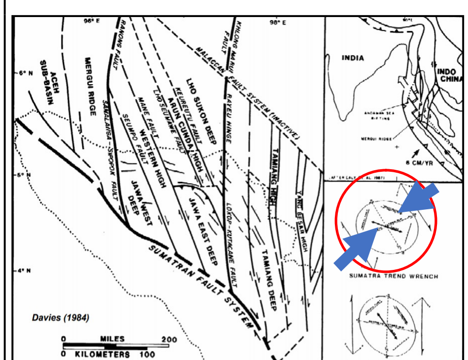

Figure 1. North Sumatera Basin where MLC-001 and TTA-001 Location

The Present Stress was dominantly from Plio-Pleistocene Stress. Well MLC-1 dan TTA-1 are under U-S pattern with stress direction NE-SW (N 020 E – N 060 E) & (N 200 E – N 240 E).

Log and Drilling Data

Data from well TTA-01 will be input as learning data, consist of RHOB, DTCO, GR, ECD, RPM, WOB, ROP, and recorded pore pressure. The final model will be used to predict pore pressure for well MLC-001

DEPTH	ECD	RPM	WOB	ROP	PorePressure
xx5.724	10.912	20.133	2.778	132.929	9.753
xx5.877	10.912	20.134	2.778	132.935	9.754
xx6.029	10.912	20.135	2.777	132.942	9.755
xx6.182	10.912	20.137	2.777	132.949	9.756
xx6.334	10.912	20.138	2.779	132.956	9.757
xx6.486	10.912	20.139	2.782	132.962	9.758
xx6.639	10.912	20.14	2.783	132.969	9.759
xx6.791	10.912	20.142	2.783	132.976	9.76
xx6.944	10.912	20.143	2.787	132.983	9.761

Figure 2. Drilling Parameter of Welll TTA-01

Method:
Artificial Neural Network (ANN)

ANN is a method which inspired by "neuron" system in human brain. It is using connectivity between "neuron" to find the right model. The target can be iterations limit or maximum error value determine by researchers.

$$y = ax + by + cz + \quad \ldots \ldots .. \quad + \; mn$$

$a, b, c, \ldots, m$ is input variable

$x, y, z, \ldots, n$ is training weight

Analysis Steps:
Preprocessing

The data cleaning process required raw data preprocessing. Then it will be divided into 2 types of data input, training and testing.

- Data 1: in this process, all of data output (Pore Pressure) smaller than 0 (negative) will be deleted.
- Data 2: in this process, all of data output (Pore Pressure) smaller than 0 (negative) will be convert 0.

Testing
Machine Learning Method

Testing process will determine the most suitable machine learning methods and final validation will use 10-fold cross. The method with result R^2 closest to 1, RSME closest to 0, and optimum training time.

Metode	*Best* RMSE	*Best* R^2	10-*fold* RMSE	10-*fold* R
ANN	0.11401	0.9789	0.1240	0.9644
RF	0.36512	0.9213	0.841	0.918
SVM	0.4771	0.946	0.5112	0.933

Table 1. Accuracy of Machine Learning Methods

Based on table 1, ANN show the best result with the lowest RMSE 0.11401 and R^2 0.9644. ANN will be use as ML method in this research.

Result and Discussion
Learning Rate

Learning Rate	Best RMSE (nilai LR)	Best R^2 (nilai LR)
0.1 – 0.5	3.274 (0.1)	0.847 (0.1)
0.05 – 0.1	1.172 (0.05)	0.912 (0.05)
0.01 – 0.05	0.1121 (0.01)	0.978 (0.01)
0.0005 – 0.01	0.110 (0.00025)	0.979 (0.00025)

Table 2. Determination of Optimum Learning Rate

We need to find the correct learning rate for well TTA-001. Normally, the smaller learning rate the smaller error rate will gain but the longer convergent time needed. So the optimum value should be obtained by trial and error.

Learning rate 0.01 shown as optimum value as RSME had reached 0.1121 and R^2 0.978. Researcher tried to decrease learning rate to 0.00025 as RSME and R^2 improved insignificantly yet had increased learning time significantly.

Error Threshold

Error threshold is limit of training process time. If error threshold reached its value, then the training will stop. If value of error threshold is too large the training will stop faster causing the optimal condition has not been achieved (error percentage may still high). However, if error threshold value is too small then training will last longer, and most likely stop because the maximum iteration has been reached before error threshold value is reached. In this paper, there were 4 error threshold values tested with result below.

Error Threshold	RMSE	R^2
0.001	0.4016	0.9013
0.0005	0.1104	0.9781
0.00025	0.1121	0.9671
0.0001	N/A	N/A

Table 3. Determination of Error Threshold

Error threshold 0.0001 cannot be reach because training process stopped first (500.000 iterations). Error threshold value 0.0005 result RSME 0.1104 and R^2 0.9781. However, error threshold value 0.00025 give insignificant improved result but it took a longer training process time.

Data Combination

The quantity of neuron in hidden layer determine complexity of ANN. The greater amount of neuron, the more complex and time consuming the process will be. There will be 4 data combinations as the input training parameters with minimum 3 input variable. The result is combinations of depth, DTCO, GR, and RHOB got the lowest RMSE with R^2 closest to 1.

Kombinasi Variabel	RMSE	R^2
DEPTH + DTCO + GR	0.3124	0.8901
DEPTH + DTCO + RHOB	0.1123	0.9223
DEPTH + GR + RHOB	0.412	0.901
DTCO + GR + RHOB	0.1819	0.9633
DEPTH + DTCO + GR + RHOB	0.0714	0.9826

Table 4. Determination of Optimal Data Combinations

Comparison Between Machine Learning Result vs Calculation Result

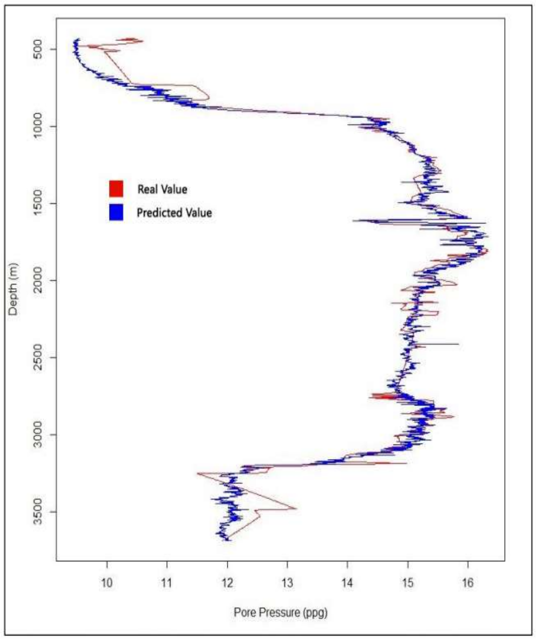

Figure 3. Comparison Between Manual Calculation and Predicted Pressure by Machine Learning

The picture shows that machine learning is having the same capacity to predict pore pressure as manual calculation. The input data used in this calculation is wireline log data.

Conclusions

1. Artificial Neural Network (ANN) is the most suitable machine learning method with R^2 0.9789, RSME 0.11401, 10-Fold RSME 0.1240, 10-Fold R^2 0.9644.
2. In this case study using learning rate 0.01, error threshold 0.1 as the optimum training variable.
3. Machine learning are able to predict pressure with a more efficient and faster result compared to manual calculation.

References

Moulds, P.J 1989, Ann. Conv. Indonesia Petroleum Association (IPA), **18**, 217-245.
Wilcox, R.E., T.P. Harding & D.R. Seeley 1973. American Assoc. Petrol. Geol. (AAPG) Bull. 57, 74-96.