Automatic machine learning (AutoML) for petrophysical evaluation: Case study in Sirikit Field Thailand

Sujintara Muenban [1*], Pongthep Thongsang [1]

[1]Petroleum Geoscience Program, Department of Geology, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand
*Corresponding author e-mail: sujintaram@hotmail.com

Received: 14 Jun 2021
Revised : 2 Jul 2021
Accepted: 11 Jul 2021

Abstract

The petrophysical interpretation is critical for assessing the economic justification. However, the existing workflow of petrophysical assessment is time-consuming. This study aims to investigate the petrophysical interpretation utilizing the machine learning algorithms in the determination of lithology classification and reservoir identification from well log data of Sirikit field. This experiment is based on data from 50 deviated wells located throughout the Sirikit main area, which is the main production area of the Sirikit field, containing oil and gas reservoirs from the Yom, Pratu Tao, and Lan Krabue formations. The programming will concentrate on four well log data types: gamma ray, resistivity, density, and neutron log, as well as two interpretation logs: lithology interpretation and fluid interpretation log. The approach is separated into two basic phases, the first of which is to develop an artificial architecture of neuron networks capable of categorizing lithology, namely sandstone and shale. The lithology will then lead to the secondary goal of reservoir categorization, which includes gas-, oil- and water-saturated-sandstones and shale. This research will focus on the extreme gradient boosting (XGBoost) technique developed as a result of automated machine learning (AutoML). The mean squared error (MSE) and customized error measurement (CEM) accuracy on prediction is the main accuracy metrics used to assess the model score. The best lithology prediction receives an average MSE of 2.76 percent and average CEM of 4.27 percent. Furthermore, the best reservoir classification prediction receives an average MSE of 0.17 percent and average CEM of 1.90 percent. Consequently, the algorithm developed in this work help shorten the time required for petrophysical interpretation.

Keywords: Petrophysical interpretation, Automatic machine learning, Extreme gradient boosting, Sirikit field, Phitsanulok basin

1. Introduction

PTTEP (PTT Exploration and Production Public Company Limited)'s Sirikit field is the largest onshore oilfield in the Phitsanulok Basin, central part of Thailand (Chantraprasert and Utitsan, 2021). More than 2,000 wells have been drilled and produced over 300 million barrels of oil with approximately 30,000 barrels of oil a day. A large number of development wells have been drilled in order to sustain the production rate.

The wireline log and formation pressure testing procedures are used to assess the wells' potential. The petrophysical interpretation is critical for assessing the economic justification. However, the current petrophysical evaluation methodology is time-consuming (Halotel, Demyanov and Gardiner, 2020). If petrophysical interpretation takes less time, leading to cost savings.

The automated machine learning (AutoML) technique is used to design an

algorithm that can aid in analyzing petrophysics, with the goal of creating AI-based programming that can automate forecasting reservoir characteristics.

This study focuses on one technique called extreme gradient boosting (XGBoost) that was developed as a preliminary output of AutoML. This improved approach is a success of random forest, which was applied with well data from using the Sirikit field.

The major goal is to develop an artificial architecture of neuron networks capable of categorizing geological facies, including sandstone and shale. The accomplishment of this lithology classification leads to the secondary goal of reservoir identification, which includes gas-, oil- and water-saturated-sandstones and shale. This result aims to lower the timing of interpretation in the petrophysical interpretation processes.

2. Study area and geological background

The study area is located in Sirikit field, which is the largest onshore oil field in the Phitsanulok Basin. The Phitsanulok Basin, a continental rift basin, located in central Thailand. It is part of the N-S trending belt of Cenozoic basins. The N-S trending belt extends from northern Thailand to the Gulf of Thailand (Chantraprasert and Utitsan, 2021).

The northern half of the basin is a deep half-graben (the Sukhothai Depression). The basin is bounded on the west by an east-dipping normal fault (the Western Boundary Fault). The southern half composes of a smaller sub-basin series separated by basement highs.

Based on the structural evolution studies of Chantraprasert and Utitsan, 2021, the result suggest that the Phitsanulok Basin has been experienced three main evolution periods.

● A rifting period during 30-18 Ma: The Phitsanulok Basin was initially opened by NE-SW extension at the junction between two existing Mae Ping and Uttaradit faults (30-18 Ma).

● A transtention and inversion period during 18-10 Ma: The stress regime changed from vertical to horizontal direction causing the lateral movement of strike-slip faults, which are a sinistral movement of the Uttaradit fault and a dextral movement of the Mae Ping fault.

● A post-rift period from 10 Ma to the present day: The tectonic setting reverted to an extension regime with a minor episodic varying between E-W & WNW-ESE during 10-0 Ma. There were fewer activities on strike-slip faults with a minor episodic inversion.

The depositional environment principally focuses on the three main stratigraphic intervals. The succession is described from the older age to the younger age in the following account:

● Sarabop – NongBue – Khom Formations

During the early syn-rift period, the Phitsanulok Basin was primarily filled by clastic sediments of Sarabop – Nong Bue – Khom Formation overlying above the Pre-Tertiary Basement.

● Nam Nan – Lan Krabue (LKU) – Chum Saeng (CS) Formations

This formation contains the lateral facies variation due to the changing in depositional environments from a fluvial dominated Nam Nam Formation to the fluvial-deltaic LKU Formation The LKU Formation, can be divided into four reservoir units (D, K, L and M) (Lawwongngam and Philp, 1991) and the open lacustrine Chum Saeng Formation.

● Pratu Tao (PTO) – Yom Formations

These formations are mainly developed by fluvial channels and floodplain deposits in an alluvial plain setting.

3. Dataset

This study focuses on the well log data. The available data include 50 wells that located though out the Sirikit main area which is the main production area of Sirikit field. The experiment used four well log data including Gamma ray (GR), Deep Resistivity (RT), Bulk Density (RHOB) and Neutron (NPHI) and two interpretation log including lithology interpretation (SAND) and fluid interpretation log (FCOL). The well log data distributions are illustrated in Figure 1 and 2.

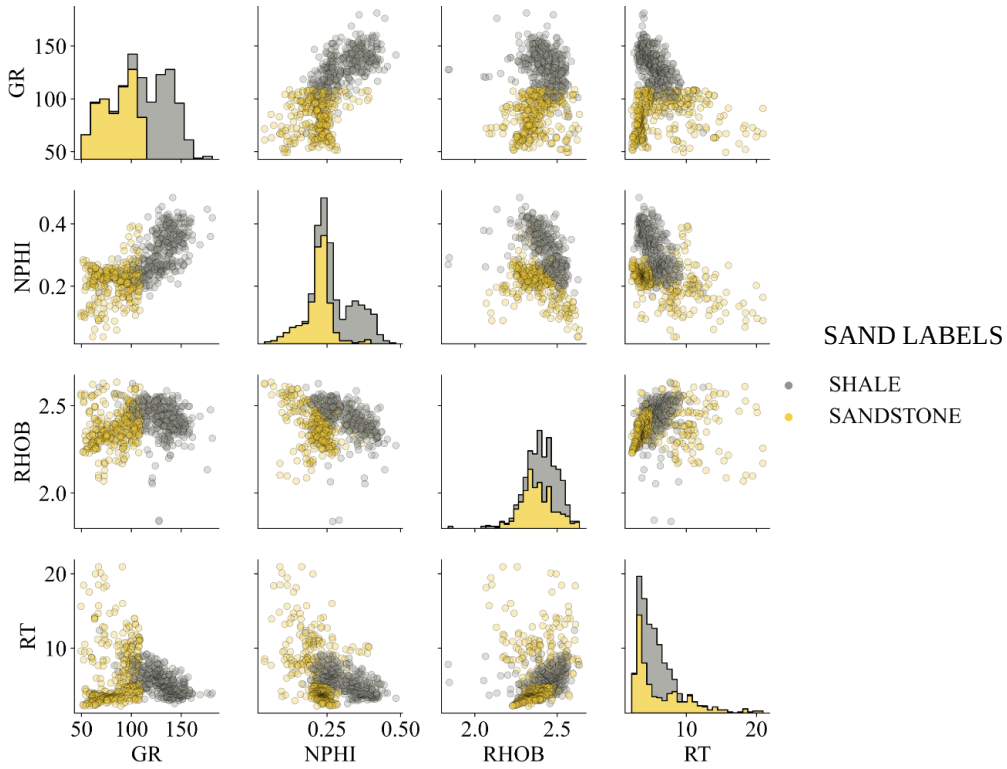

Figure 1. Pairwise scatterplots and grouped histograms of well logs by lithology interpretation log (show in different colors)

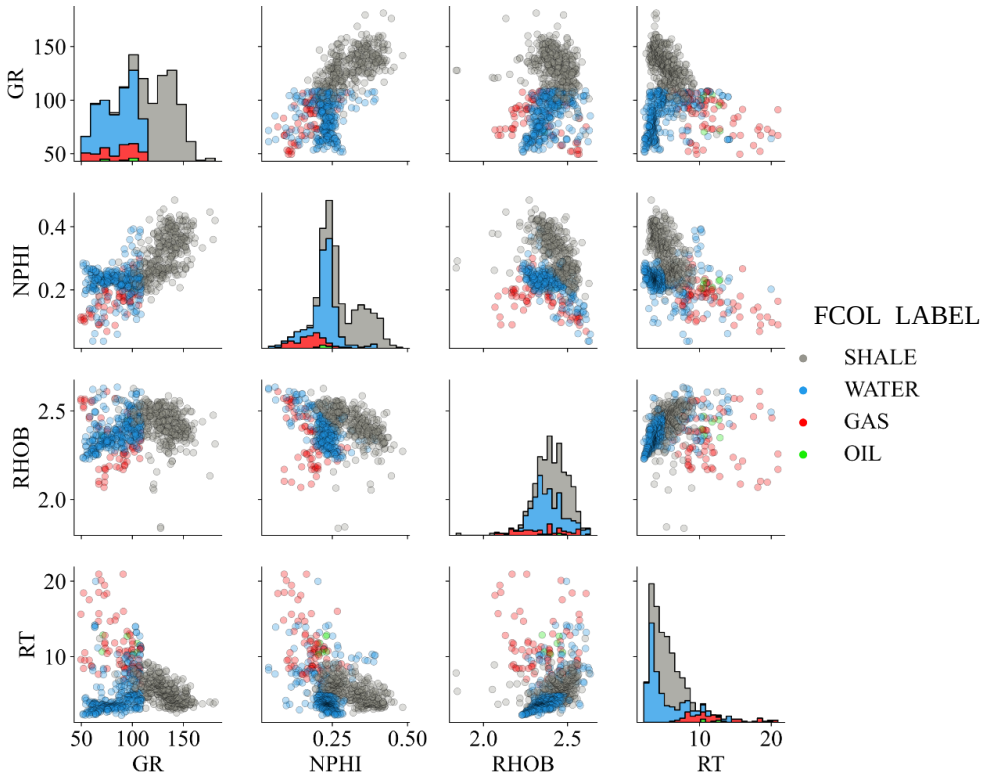

Figure 2. Pairwise scatterplots and grouped histograms of well logs by reservoir classification log (show in different colors)

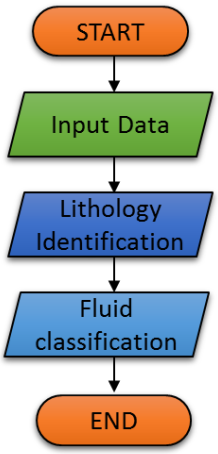

Figure 3. The experiment workflow

4. Methodology and experiment setup

The goal of this study is to create artificial intelligence based (AI-based) programming that can aid in the evaluation of petrophysics through the machine learning methods, also known as automatic machine learning (AutoML).

The methodology is divided into two main processes, which include creating an artificial architecture of neuron networks that can categorize geological facies, including sand and shale. Then the achievement of this clustered sandstone leads to the secondary objective that is reservoir classification, i.e., gas-, oil- and water-saturated-sandstones and shale (Fig. 3).

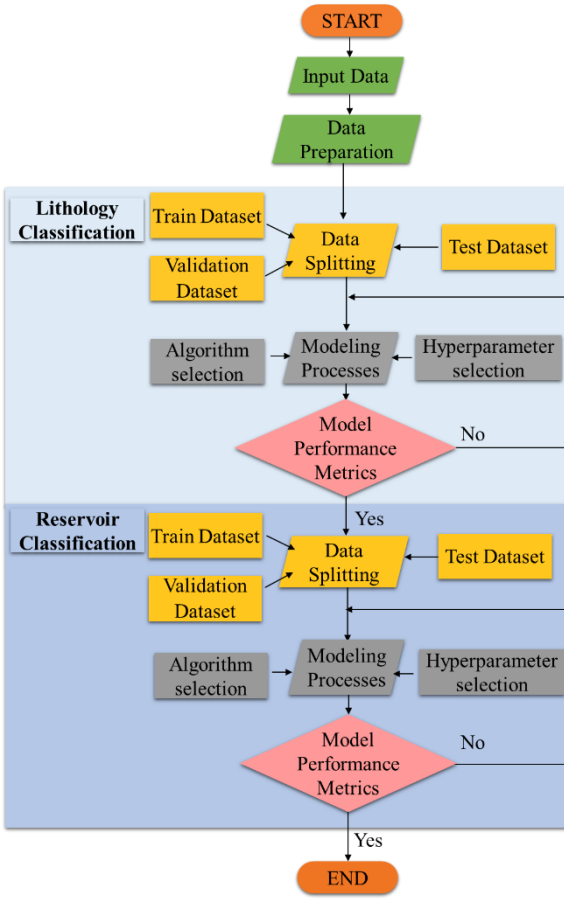

Figure 4. Flow of methodology work chart

4.1 Methodology

The first step consists of preparing the data. This involves loading and cleaning the data for use in the modeling process. The second step involves selecting features to be used in creating the model. The third stage involve dividing the available data are divided into train, test and validation data. (Fig. 4).

4.1.1 Data Preparation

It is regarded as an important step in the use of machine learning**.** The data preparation processes aim to improve data quality (Simeone *et al.*, 2019). Data cleaning removes obsolete records.

4.1.2 Data Splitting

The study, learn and development of algorithms from data are typical jobs in machine learning. The algorithms make data-driven predictions or judgments by constructing a mathematical model from input data.

The data are separated into training, validation, and test to prevent overfitting, model over-learning from training data, and to adequately evaluate your model.

In our experiment keeps ten wells to be a test dataset, and the rest of the data are split as a fraction of 0.75 for training and 0.25 for validation (Fig. 5).

- **Training Dataset**

The training dataset is a sample of data utilized throughout the learning process. For classification tasks, a supervised learning algorithm examines the training dataset to discover the optimal variables that will provide a successful prediction model.

- **Validation Dataset**

The validation dataset is a subset of data used to offer an unbiased evaluation of a model fit on the training dataset while fine-tuning the

model parameters. The validation set is also known as the development set since it is used throughout the model's development stage (Torlay *et al.*, 2017).

● **Test Dataset**

A test dataset is distinct from the training dataset but has the same probability distribution as the training dataset. A test dataset is utilized to provide an unbiased assessment of the final model fit on the training dataset. The test dataset in this study are chosen based on criteria from the complete classes on lithology and fluid.

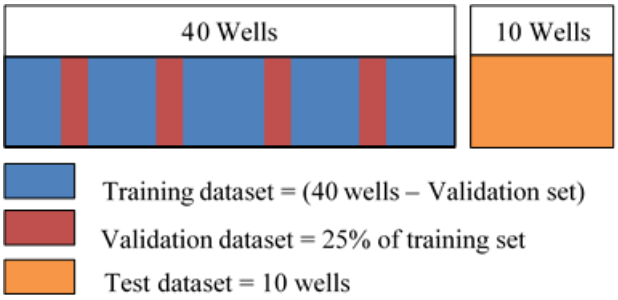

Figure 5. Visualization of the splits

4.1.3 Modeling Processes

Following completion of the data preparation processes, the model is run through the algorithms with the appropriate parameters.

The model is built based on a decision tree algorithm. The decision tree technique is a supervised machine learning algorithm that is widely utilized. It is a typical machine learning algorithm that makes predictions using a "tree structure." (Chang *et al.*, 2019). The decision tree solves the challenge of mapping non-linear relationship by employing the whole training set as the root, tree corresponds to a class label and the leaf node are expressed on the interior characteristic. Tree-based algorithms provide great accuracy, stability, and interpretability to prediction models (Dhaliwal, Nahid and Abbas, 2018).

To deal with enormous amounts of data, such as more than 2,100,000 data points from 7 well logs, a huge number of branches must be created. Extreme gradient boosting (XgBoost) (Chen and Guestrin, 2016) which is one of the implementations of gradient boosting machines (Friedman, 2001). XgBoost, which is known as one of the best performing algorithms for supervised learning (Giglou et al., 2017), is used for model prediction.

Automated machine learning (AutoML) is used in this project to construct algorisms and forecast prediction. AutoML is a developing area that attempts to automatically choose, build, and parameterize machine learning models to achieve optimal performance on a given dataset, in order to make machine learning techniques more accessible and decrease the need for human expertise.(Waring, Lindvall and Umeton, 2020) (Fig. 6).

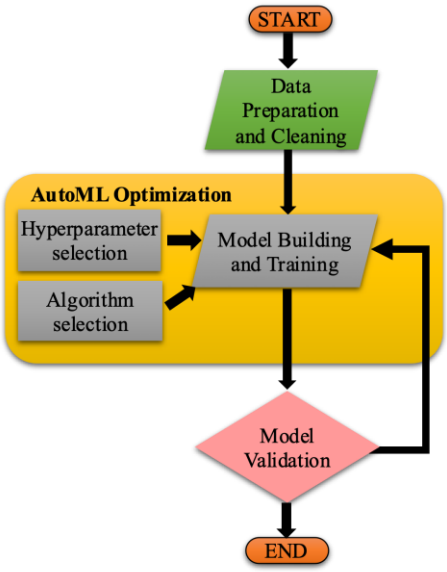

Figure 6. AutoML Optimization pipeline

This study will work on an open source H2O's AutoML platform. H2O's AutoML is a machine learning platform designed to scale to very big datasets. It is one of the highest qualities among AutoML benchmarks. It is widely used in business and academics, and has many advocates in the open-source machine learning community (Ledell and Poirier, 2020).

4.1.4 Model parameters

Two of the most difficult jobs in machine learning are parameter tuning and algorithm selection. This procedure involves the selection of input variables for the models. The selection of appropriate input variables for machine learning models is essential to the model's performance.

The input variables are chosen to guarantee that the model captures the relationship between the inputs and the target variable during the training phase in order to identify the optimal set of algorithm parameters that provides the best model.

Five parameters were chosen for this study, including

● Max runtime

This parameter is used when constructing a model to set the maximum number of seconds spent to training each individual model in the AutoML run. This parameter determines the maximum runtime in seconds for the whole grid while doing a grid search.

● Max models

This parameter specifies the maximum number of models that will be built during the AutoML run.

● nfolds

When nfolds is specified, the method will generate models with nfolds +1. If you give nfolds = 5, for example, 6 models are created. The first five models (cross-validation models) are created using 80% of the training data, with a different 20% kept out for each of the five models. The primary model is then created using all of the training data. For the whole training dataset, the 5 holdout/validation data predictions are merged into a single prediction. This "validation prediction" is then compared to the real labels, and the total cross-validation metrics are calculated.

● Seed

The random sample order can be controlled using the seed function. The basic goal of seeding is to make reproducible the result of random.

● Balance classes

If it is determined during model training that the majority of data fit into a single category. To balance the class distribution, use the balance classes option. When activated, H2O will either undersample the majority or oversample the minority.

4.1.5 Model Performance Metrics

Validation the algorithm and obtain the best predicted model is to determine which algorithm is most suited to solving the problem.

The model's performance will be evaluated using both the standard accuracy measurement and the customized error measurement designed for this study.

This study's accuracy measurements include

● Mean Squared Error (MSE)

The MSE metric calculates the average of the square root of the errors. MSE squares the distances between the points and the regression line (the "errors") to remove any negative signs.

MSE equation:

$$MSE = \frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \quad (1)$$

Where:
N is the total number of rows (observations) of your corresponding data frame.
y is the actual target value.
$\hat{y}$ is the predicted target value.

● Root Mean Squared Error (RMSE)

The RMSE metric assesses how well a model can forecast a value.

RMSE equation:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y}_i)^2} \quad (2)$$

Where:
N is the total number of rows (observations) of your corresponding data frame
y is the actual target value
$\hat{y}$ is the predicted target value

● Mean Per Class Error

Mean Per Class Error is the average of the errors of training, validation, and testing class in the dataset. This metric shows toward misclassification of the data across these classes. The lower of this metric represent the better prediction.

● Customized Error Measurement (CEM)

CEM assesses the error in the depth domain. The evaluation metric derived from the comparison of predicted and real data results. If the forecast is right, the measurement will be considered accurate. However, if the forecast is not equal to true, the measurement will be considered erroneous. The right and wrong answers are then summarized as an accuracy percentage (Fig. 7).

CEM equation:

$$CEM = 100\left(\frac{n}{c+n}\right) \qquad (3)$$

Where:

CEM = Customized error measurement

c = number of correct predictions counted

n = number of incorrect predictions counted

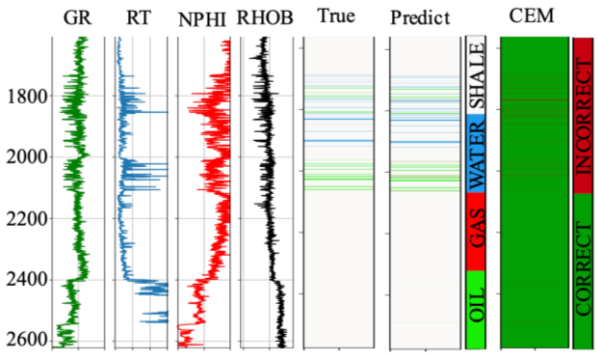

Figure 7. Customized error measurement

5. Results

5.1 Lithology Classification

To determine the optimal method for predicting the model. Three sets of parameter tuning are carried out. The parameter settings used are displayed in the table below (Table 1).

Table 1. The parameter for lithology classification

Parameters	Model 1	Model 2	Model 3
Max runtime (sec)	45	55	120
Max models	25	25	25
nfolds	2	2	3
seed	1	3	3
Balance classes	TRUE	TRUE	TRUE

Each model produces a distinct forecast and accuracy result. After obtaining the results for each model, a model comparison will be performed to choose the best prediction for use in the subsequent stages.

According to the results of the experiment, when using the XGBoost technique, each models get the best outcomes, though varied number of decision trees (Table 2).

Table 2. The number of decision trees and average number of decision trees

Parameters	Model 1	Model 2	Model 3
Number of decision trees	3 - 12	8 - 25	23 - 41
Average number of decision trees	9	15	31

According to the various in the parameter, model provides a distinct forecast with a varied accuracy (Fig. 8).

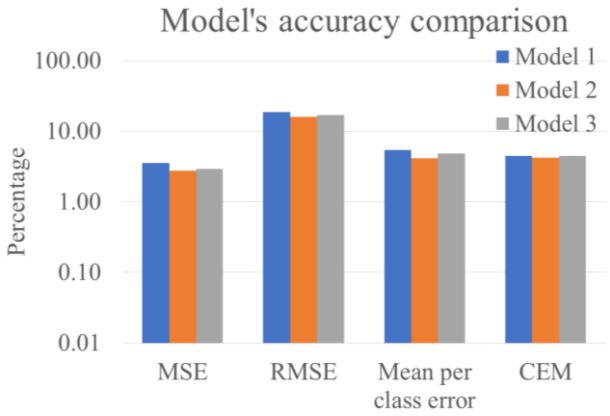

Figure 8. The lithology classification model's accuracy comparison results

In the processes of parameter tuning and performance evaluation approaches, the training goal is to find a sweet spot between overfitting and underfitting to ensure optimal training. Figures 9 - 11 represent the fit ability of each model. The orange color denotes sand forecast, whereas the green color indicates shale prediction. Model 1 (Fig. 9) shows an underfitting situation. It means a model can neither learn the training data nor generalize to new data because it is not powerful enough. Fig. 9 shows that the model 1 is unable to capture the

relationship between sand and shale data points. It has plenty of both sandstone data points predicted as shale and shale data points predicted as sandstone. Model 2 (Fig. 10) shows an appropriate fitting situation. Figure 10 shows model is able to capture the relationship between sand and shale data points. Model 3 (Fig. 11) is an overfitting model. It shows model works well on the training data but not on the evaluation data (Fig. 8). This is due to the model remembering the data it has seen and being unable to generalize to unseen examples.

Combining the highest accuracy with appropriate fitting reveals that model 2 is the best lithology prediction algorithm.

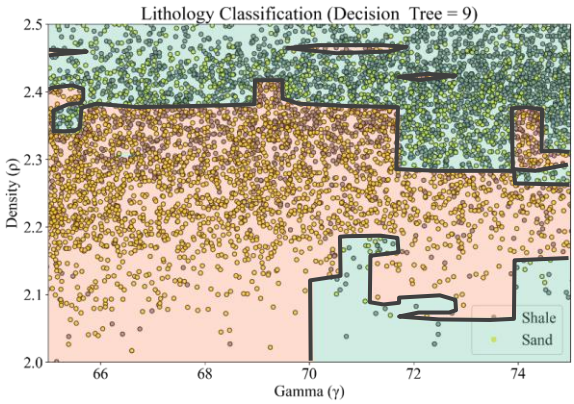

Figure 9. Model 1's lithology classification (Decision Trees = 9) underfit situation

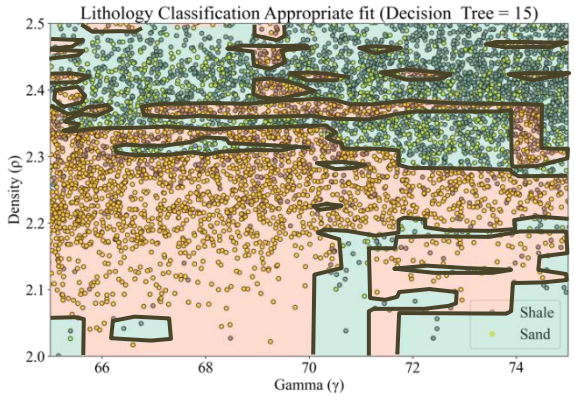

Figure 10. Model'2 lithology classification (Decision Trees = 15) appropriate fit case

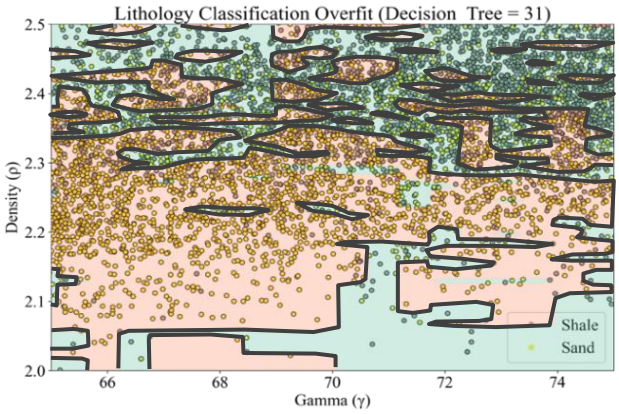

Figure 11. Model'3 lithology classification (Decision Trees = 31) overfit circumstance

5.2 Reservoir Classification

After completing the lithology prediction, the reservoir classification will be proceeded as the following procedure.

The reservoir classification prediction algorithm will be modelled using the same approach as the lithology prediction in the previous stage.

Unlike lithology prediction, reservoir classification will be predicted into four classes: gas-, oil- and water-saturated-sandstones and shale.

To determine the optimal method for predicting the model, three sets of parameters are carried out. The parameter settings used are shown in the table below (Table 3).

Table 3. Parameter using on each model

Parameters	Model 1	Model 2	Model 3
Max runtime (sec)	60	140	180
Max models	25	25	25
nfolds	2	2	2
seed	1	1	1
Balance classes	TRUE	TRUE	TRUE

According to the results of the experiment, when using the XGBoost technique, each models get the best outcomes, though varied number of decision trees. Table 4 summarizes the number of decision trees and average number of decision trees for each model.

Table 4. The number of decision trees and average number of decision trees

Parameters	Model 1	Model 2	Model 3
Number of decision trees	3 - 11	9 - 25	23 - 61
Average number of decision trees	3	15	44

Each model produces a distinct forecast and accuracy result. After obtaining the results for each model, a model comparison was performed to determine the best forecast. The model accuracy is critical in determining the optimum method (Fig. 12).

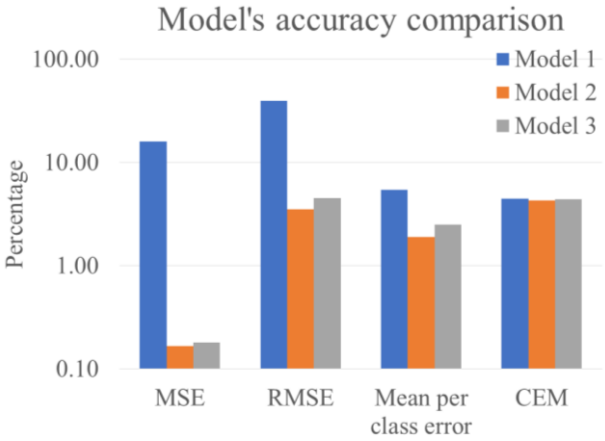

Figure 12. The reservoir identification model's accuracy comparison results

Similar to lithology prediction, the training aim to identify a sweet spot between overfitting and underfitting to ensure optimal training. Figures 13–15 demonstrate how well each model fits the data. The white color represents shale forecasts, the yellow color represents gas-, the pink color represents oil-, and the grey color represents water-saturated-sandstones forecasts. Model 1 (Fig. 13) shows an underfitting situation. It shows the model is unable to capture the relationship between gas-, oil- and water-saturated-sandstones and shale data points. It has plenty of misclassification of data points. Model 2 (Fig. 14) shows an appropriate fitting situation. It shows that model 10 is enable to capture the relationship between gas, oil, water-saturated-sandstone and shale data points. Model 3 (Fig. 15) is an overfitting model. It shows model works well on the training data but not on the evaluation data (Fig. 12). This is because the model remembers the data it has seen and is unable to generalize to unseen examples.

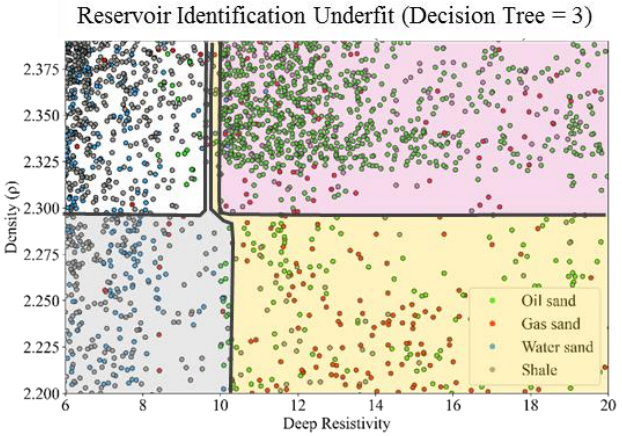

Figure 13. Model 1's reservoir identification (Decision Trees = 3) underfit situation

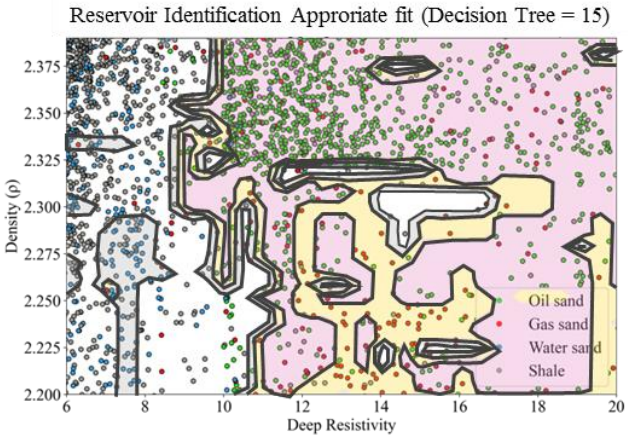

Figure 14. Model'2 reservoir identification (Decision Trees = 15) appropriate fit case

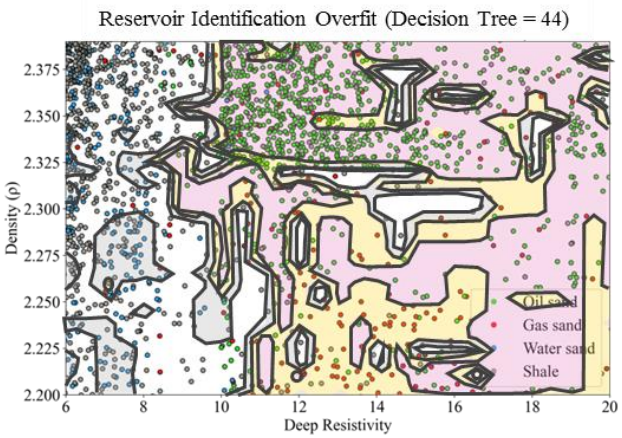

Figure 15. Model'3 reservoir identification (Decision Trees = 44) overfit circumstance

Model 2 is the best reservoir identification prediction algorithm when the highest accuracy and proper fitting are combined.

6. Discussion

Figure 16 represents the graph of model 2's testing well lithology classification accuracy result. The graph shows three trends in accuracy based on CEM and MSE metrics. The first trend includes the wells from A to C. The second trend includes D and E wells. This trend has the highest error of the three groups. The third trend includes the wells F to J. This trend has the minimum CEM error.

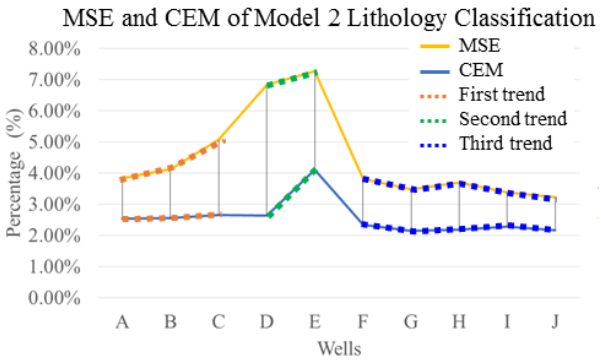

Figure 16. Evaluation metric using MSE and CEM. The measurement indicates the separations of lithological trends as three zones.

The graph of model 2's testing well on reservoir identification accuracy result is shown in Figure 17. The graph depicts three trends in accuracy based on CEM and MSE metrics. The first trend includes all of the wells from A to C. The second trend includes D and E wells. This trend has the highest error of the three trends. The third trend includes well F through J. This trend has the lowest error.

To determine the cause of the trend in accuracy, a well correlation was performed to examine the variation in well log characters to see whether it may disclose the underlying causes. The correlation illustrates that the log characters vary in PTO, LKU-D, and LKU-M formations.

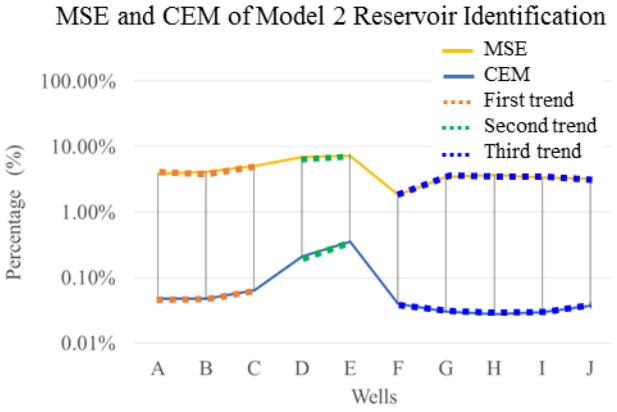

Figure 17. Evaluation metric using MSE and CEM. The measurement indicates the separations of reservoir identification trends as three zones.

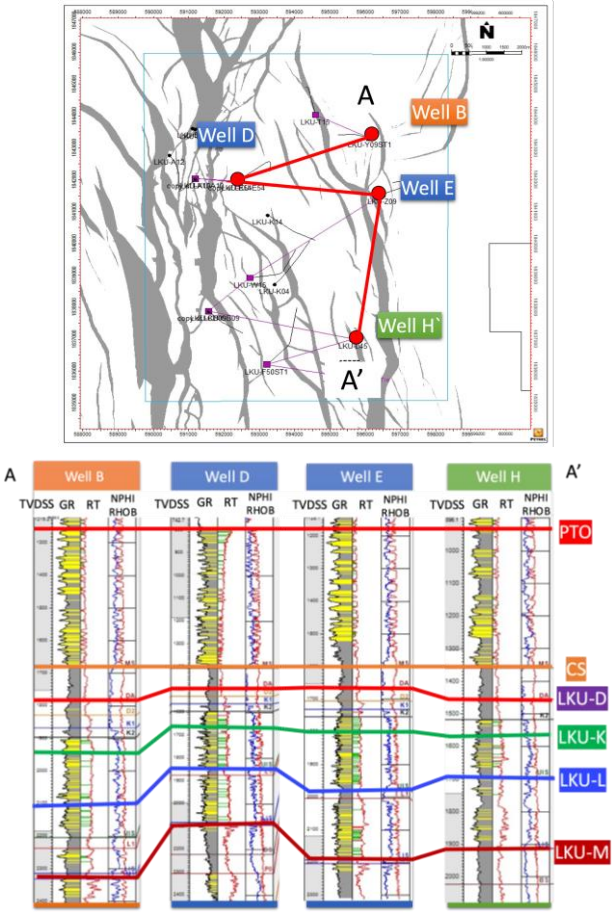

Figure 18. The well correlation of model 2 testing wells.

The gamma ray log in the first trend of the PTO formation shows a greater sandstone layer content. The second trend shows a moderate amount of sandstone layer content. The third trend indicates increased shale content.

In addition, Gamma ray log in the first trend of the LKU-D and LKU-M formations reveal high sandstone layer composition. In the second trend, the characters combination of sandstone and shale. The third trend indicates increased shale content. (Fig. 18).

After combining the accuracy with the depositional environment map, it was discovered that in the PTO, LKU-D, and LKU-M formations, the second trend including D and E wells, is situated in the depositional environment's transition zone (Fig. 19 and 20).

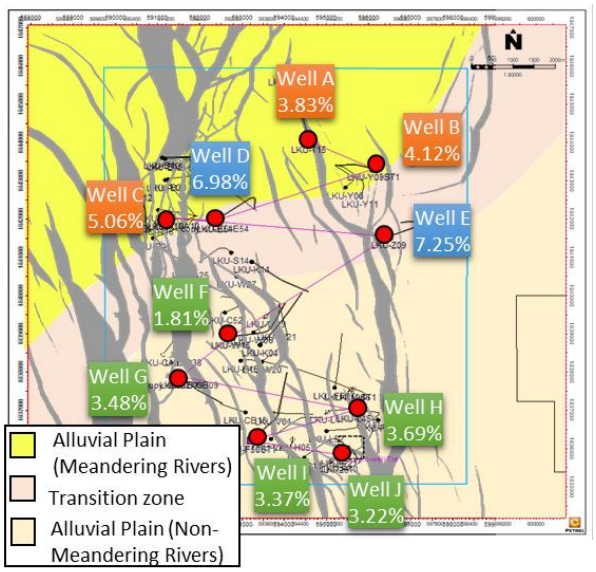

Figure 19. CEM accuracy of prediction posting on a depositional environment map of the PTO formation.

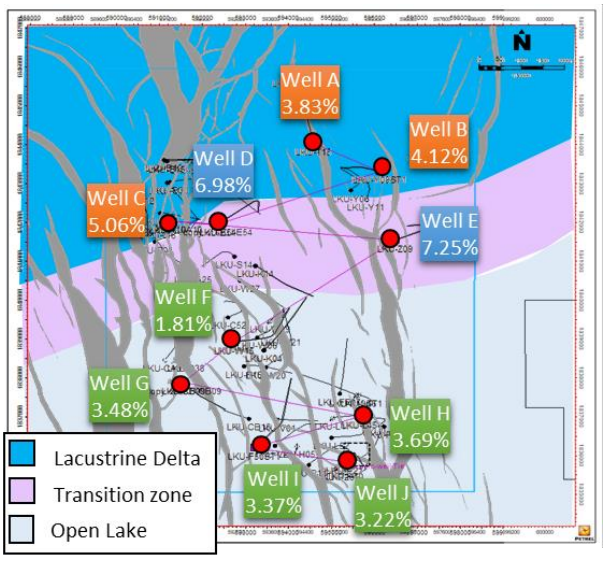

Figure 20. CEM accuracy posting on depositional environment map of LKU-D and LKU-M formation.

7. Conclusions

The main objective of this research is to develop AI-based programming that can help in the evaluation of petrophysics using machine learning approaches.

The algorithm modeling working on four well log data, including Gamma ray and Resistivity. Density and Neutron, as well as two interpretation logs, comprising lithology interpretation and reservoir classification, were obtained from 50 deviated wells located across the Sirikit main region, the Sirikit field.

Automated machine learning (AutoML) is used in this project to construct algorisms and forecast prediction. The model algorithm is created based on supervised machine learning algorithm of Extreme gradient boosting (XGBoost).

The methodology is divided into two main processes, which include create an artificial architecture of neuron networks that can categorize geological facies including sandstone and shale then the achievement of this clustered will lead to the secondary objective that is reservoir classification including gas-, oil- and water-saturated-sandstones and shale

The findings of the experiment demonstrate that the average forecast time for lithology identification per well is 62 seconds. The average time for reservoir categorization estimation per well is 364 seconds. When all prediction procedures are combined, the average prediction time per well is 426 seconds, or 7 minutes and 6 seconds.

Comparing with traditional petrophysical evaluation, which required more than 30 minutes per well, this study demonstrates that machine learning algorithms can reduce timing on petrophysical interpretation.

8. Acknowledgments

The first author would like to thank PTTEP for a M.Sc. scholarship.

Last but not the least, the authors are grateful to all those, who helped in the entire study directly and indirectly.

9. References

Chang, W. *et al.* (2019) 'A machine-learning-based prediction method for hypertension outcomes based on medical data', *Diagnostics*, 9(4). doi: 10.3390/diagnostics9040178.

Chantraprasert, S. and Utitsan, S. (2021) 'Origin of synchronous extension and inversion in a rift basin: The Phitsanulok Basin, central Thailand', *Journal of Asian Earth Sciences*, 213(April). doi: 10.1016/j.jseaes.2021.104774.

Chen, T. and Guestrin, C. (2016) 'XGBoost: A scalable tree boosting system', *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 13-17-Augu, pp. 785–794. doi: 10.1145/2939672.2939785.

Dhaliwal, S. S., Nahid, A. Al and Abbas, R. (2018) 'Effective intrusion detection system using XGBoost', *Information (Switzerland)*, 9(7). doi: 10.3390/info9070149.

Giglou, A. N., Mccorquodale, J. A. and Solari, L. (2017) 'Ain Shams Engineering Journal'. Available at: https://www.researchgate.net/profile/Abolfazl%7B_%7DNazari%7B_%7DGiglou2/publication/314686506%7B_%7DNumerical%7B_%7Dstudy%7B_%7Don%7B_%7Dthe%7B_%7Deffect%7B_%7Dof%7B_%7Dthe%7B_%7Dspur%7B_%7Ddikes%7B_%7Don%7B_%7Dsedimentation%7B_%7Dpattern/links/58f7ac2a0.

Friedman, J. Greedy function approximation: A gradient boosting machine. Ann. Stat. 2001, 29, 1189–1232.

Halotel, J., Demyanov, V. and Gardiner, A. (2020) 'Value of Geologically Derived Features in Machine Learning Facies Classification', *Mathematical Geosciences*, 52(1), pp. 5–29. doi: 10.1007/s11004-019-09838-0.

Lawwongngam, K. and Philp, R. P. (1991) 'Geochemical characteristics of oils from the Sirikit Oilfield, Phisanulok Basin, Thailand', *Chemical Geology*, 93(1–2), pp. 129–146. doi: 10.1016/0009-2541(91)90068-3.

Ledell, E. and Poirier, S. (2020) 'H2O AutoML: Scalable Automatic Machine Learning', *Automl.Org*. Available at: https://scinet.usda.gov/user/geospatial/#tools-and-software%0Ahttps://www.slideshare.net/0xdata/intro-to-automl-handson-lab-erin-ledell-machine-learning-scientist-h2oai%0Ahttps://docs.h2o.ai/h2o/latest-stable/h2o-docs/automl.html.

Simeone, A. *et al.* (2019) 'ScienceDirect ScienceDirect A new methodology to analyze the functional and physical architecture of existing products for an assembly oriented product family identification service reliability assessment in cloud', *12th CIRP Conference on Intelligent Computation in Manufacturing Engineering*, 99, pp. 233–238. Available at: https://doi.org/10.1016/j.procir.2018.03.074.

Torlay, L. *et al.* (2017) 'Machine learning–XGBoost analysis of language networks to classify patients with epilepsy', *Brain Informatics*, 4(3), pp. 159–169. doi: 10.1007/s40708-017-0065-7.

Waring, J., Lindvall, C. and Umeton, R. (2020) 'Automated machine learning: Review of the state-of-the-art and opportunities for healthcare', *Artificial Intelligence in Medicine*, 104 (October 2019). doi: 10.1016/j.artmed.2020.101822.