Adaptive boosting of random forest algorithm for automatic petrophysical interpretation of well logs

V. Srivardhan

Received: 5 March 2022 / Accepted: 1 June 2022 / Published online: 5 September 2022
© Akadémiai Kiadó 2022

Abstract
The power of Machine Learning is demonstrated for automatic interpretation of well logs and determining reservoir properties for volume of shale, porosity, and water saturation respectively for tight clastic sequences. Random Forest algorithms are reputed for their efficiency as they belong to a class of algorithms called ensemble methods, which are traditionally seen as weak learners, but can be transformed into strong performers and they promise to deliver highly accurate results. The study area is located offshore Australia in the Poseidon and Crown fields situated in the Browse Basin, which are gas fields in tight complex clastic reservoirs. There are 5 wells used in this study with one well manually interpreted which is subsequently used in developing a machine learning model which predicts the output for the other 4 wells. The basic open hole logs namely Natural gamma ray, Resistivity, Neutron Porosity, Bulk Density, P-wave and S-wave sonic travel-time, are used in interpretation. One of the wells has a missing S-wave travel-time log which was also predicted by developing a Random Forest Machine Learning model. The results indicate a very robust improvement in performance when Random Forest algorithm was combined with Adaptive Boosting when interpreting the well logs. The training accuracy using Random Forest alone was 98.21%, but testing was 77.62% which suggested over-fitting by the Random Forest model. The Adaptive Boosting of the Random Forest algorithm resulted in the overall training accuracy of 99.40% and an overall testing accuracy of 97.03%, indicating a drastic improvement in performance. S-wave travel-time log was predicted by preparing a training set consisting of Natural gamma ray, Resistivity, Neutron Porosity, Bulk Density, and P-wave travel-time logs for the 4 wells using Random Forest which gave a training accuracy of 99.79% and a testing accuracy of 98.54%. Machine learning algorithms can be successfully applied for interpreting well log data in complex sedimentary environment and their performance can be drastically improved using Adaptive Boosting.

Keywords Random forest · Machine learning · Adaptive boosting · Well log interpretation and construction · Supervised learning
1 Introduction

The idea of developing an automated methodology towards well log interpretation has been a research theme for many years as it offers the tremendous advantage of reduced time and computing power. The usage of Fourier and Wavelet Transform in well logs using multi-scale analysis for finding hidden correlations between open-hole logs and lithologies through sequence stratigraphy for various wells has been attempted in the past (Mukherjee et al. 2016; Perez-Muñoz et al. 2013; Srivardhan 2016; Panda et al. 2015). The ideas associated help in identifying lithofacies with well log signatures. Deep learning associated with neural networks have been used in the recent past in identifying lithologies and creating neural network models in order to predict clay volume, effective porosity, water saturation, and permeability (Peyret et al. 2019; Gupta and Soumya 2020). These methods have demonstrated good accuracy and the prediction models developed have been specific to geologic formations. The Deep Learning and Neural Network models though require more computing as compared to Machine Learning models. The application of Machine Learning (ML) algorithms in well logs has demonstrated tremendous utility. They have been used extensively in order to interpret facies (Bestagini et al. 2017; Pratama 2018; Alexandre et al. 2017), synthetic well log generation (Akinnikawe et al. 2018), Rock Physics modelling (Jiang et al. 2020), and predicting lithologies ahead of drill bit using LWD logs (Zhong et al. 2019). Well log interpretation has in the past been carried out deterministically (Senosy et al. 2020) using relationships between reservoir properties and well log responses, and comparatively recent methods like interval inversion (Szabó et al. 2022; Dobróka et al. 2016) in which the number of unknown lithological and reservoir properties are determined as an overdetermined problem by developing suitable relationships between well log responses and their causative reservoir properties which are to be determined. In this paper the performance of Random Forest algorithm is tremendously enhanced using Adaptive-Boosting algorithm which gives automatic near perfect interpretation of well logs in tight gas bearing clastic reservoirs.

The Random Forest is a Machine Learning Algorithm and an ensemble learning method which can be used for providing classification and regression solutions. It is a supervised learning algorithm and is inspired by Decision Trees, but generally found to give superior results as Decision Trees generally are sometimes found to overfit during training (Hastie et al. 2008; Piryonesi et al. 2020, 2021). It uses the concept of Bagging or Bootstrap Aggregating when finalizing the ensemble Regressor or Classifier over simple averaging which is done in Decision Trees. Random Forest has been applied successfully in many areas like Insurance (Lin et al. 2017), Image Classification (Akar and Güngör 2012), Land Cover and Ecological studies (Kulkarni and Lowe 2016; Mutanga et al. 2012), Medical Science (Sarica et al. 2017), and Consumer Behaviour (Valecha et al. 2018). Adaboost or short form for Adaptive Boosting is a statistical technique for improving the performance of weak learners and can be applied to other machine learning algorithms for improving the learning rate and avoid over fitting of the data. In regression problems the algorithm tends to fit the dataset through a regressor and then further fits additional instances of the regressor on the same dataset, but using different weights in accordance with the error in the prediction. It is commonly used in classification and regression problems. It has been found to improve the performance of Random Forest algorithm and has also been used with other Machine Learning algorithms to improve their performance and has been demonstrated in banking (Sanjaya et al. 2020), structural engineering (Feng et al. 2020), cybersecurity (Sornsuwit and Jaiyen 2019), plant species identification (Kumar et al. 2019) and many other avenues. In this study the application of Random Forest
algorithm is demonstrated to predict reservoir properties of 4 wells using 1 well as a training dataset, and then using Adaptive Boosting to increase the performance of the Random Forest algorithm. The predictions are then compared with manual interpretation of the wells and there is very good accuracy in results. Random Forest is also used to predict missing Shear-Sonic log in one of the wells which also gives a good performance.

2 Geology and description of study area

The Browse Basin is located in NW Shelf of Australia with water depths reaching up to 4000 m. The Browse Basin was formed during Early Carboniferous-Early Permian as a series of intracratonic half graben systems due to the breakup of Gondawana and formation of the Neo-Tethys Ocean along the northwest Australian margin (Symonds et al. 1994; Struckmeyer et al. 1998). Structural developments during extension established the sub-basin architecture and compartmentalization, which would later control the rate of sedimentation and tectonism up to Miocene (Struckmeyer et al. 1998). The evolution of the Browse Basin reflects the breakup of the Gondwana supercontinent and the creation of the Westralian Superbasin and also includes other basins in NW Australia including Carnarvon, offshore Canning, and Bonaparte Basins (Stephenson and Cadman 1994; Struckmeyer et al. 1998). During the Late Triassic—Early Jurassic, an inversion event correlated with onset of rifting on the NW Shelf which reactivated Paleozoic faults resulting in partial inversion of the Paleozoic half-grabens and formation of large-scale anticlinal and synclinal features within their hanging walls (AGSO Browse Basin Project Team 1997). Continued extension in Early Jurassic resulted in the collapse of numerous Triassic anticlines (Blevin et al. 1998), which culminated in Callovian-Oxfordian. Post rift thermal subsidence commenced in Mid-Callovian and the basin transitioned into passive margin. Sediment supply was controlled by eustasy levels, sediment supply from hinterland, small scale growth faults, and reactivation of pre-existing faults. Carbonate deposition increased in Upper Cretaceous with many fluvio-deltaic sediment systems in the Cenozoic (Poidevin et al. 2015).

The wells Kronos-1, Boreas-1, Poseidon-2 are located in the Poseidon discovery, while wells Pharos-1 and Proteus-1 are on the adjacent Crown discovery (Fig. 1). Both the Poseidon and Crown discoveries are located in the Browse Basin. The main reservoir for these discoveries are the Plover formation in Upper-Middle Jurassic which are syn-rift fluvio-deltaic clastic sequences. The gas is sourced from the Middle-Lower Jurassic sequences of the Plover formation which are fluvio-deltaic claystones. The hydrocarbons are sealed by intraformational shales of the Plover formation (Rollet et al. 2018). Volcanic activity was prevalent in nearby areas of the discovery in the Browse basin in the Jurassic and possible deposition of volcano-clastics is reported along with fluvio-deltaic sediments. In this analysis based on the regional geological history, the petrophysical interpretation for the wells has been performed in order to determine the reservoir properties including Vshale, Porosity, and Water Saturation.

3 Random forest theory

Random Forest algorithm is an ensemble technique for creating regression models using Bootstrap Aggregating or Bagging. Consider a training dataset of size \( m \) with \( TR = \{TR_1, TR_2, \ldots, TR_m\} \), with its corresponding output \( OP = \{OP_1, OP_2, \ldots, OP_m\} \), the random forest algorithm aims to find out a loss function \( f(a) \) to predict output \( b \), where \( b \in OP \).
In the concept of Bagging, at each given iteration $i$ which is repeated for say $B$ times where $b = \{1,2,3,...B\}$ and $i \in B$, there are $m$ pair of corresponding input and output samples randomly selected from $TR$ and $OP$ respectively and a function $f_b$ is created for the regression operation using the samples created which is the best fit decision tree for the given input and output samples. At each iteration input and output samples which are not part of boostraping are considered separately as unseen samples or out-of-bag samples. The operation is repeated $B$ times and the overall regressor model $f'$ is created as the aggregation of all trees applied on the unseen $x'$ samples at each iteration and averaging them which can be represented in Eq. 1:

$$f' = \frac{1}{B} \sum_{b=1}^{B} (f_b(x''))$$  \hspace{1cm} (1)

The mean accuracy of the model $f'$ during training for the predicted output $OPred = \{OPred_1,OPred_2,...OPred_m\}$ is calculated as $r^2$ (Coefficient of Determination) and determined as in Eq. 2:

$$r^2 = 1 - \frac{\sum_{i=1}^{m} (OP_i - OPred_i)^2}{\sum_{i=1}^{m} (OP_i - OPmean)^2}$$  \hspace{1cm} (2)

In the above equation the $OPmean$ is the mean of the output data $TR$. For testing the model the same $r^2$ can be determined for the testing samples of the input data. In the present analysis a ratio of 60% to 40% has been considered for dividing the input data for Training and Testing purposes. The $r^2$ is used to check the overall performance of the algorithm and is a measure of how good the model fits all the data points. It varies between 0 and 100% with 100% having the best possible fit for all the data points. It is also a measure of accuracy of the Machine Learning model.
4 Adaptive boosting

The AdaBoost or Adaptive Boosting algorithm was introduced by Freund and Schapire (1997) and discussed first in 1995. The performance of weak machine learning algorithms can be enhanced without knowing any prior knowledge through multiplicative-weight update technique. Considering the output of a weak learning algorithm \( f_i \) as \( O_{P_i} \), \( O_{P_2} \), ..., \( O_{P_m} \) where the objective of the weak learner is to fit a function \( f \) between \( TR \) and \( OP \) through least square error, that is \((OP-f(x'))^2\) and \( x' \in TR \), the error function in adaptive boosting is \( e^{-OPf(x')} \) which takes into account that only the sign of the final result is considered and the final error is the multiplicative addition from each stage, that is \( e^{-\sum_i OPf(x')} \).

At each stage and segment of the iteration the weights are updated by the algorithm so that segments which tends to increase the error are identified and the weights are adjusted so that the error is brought down.

In the case of training of Random Forest there is a chance for the model to overfit as in nodes which are seen are strong learners the weights assigned are left unaltered and only the nodes which are weak learners are further iteratively progressed and their weights altered to arrive at a state where the overall error is reduced during training. When this model is applied on the testing data, it may lead to overfitting since the weights associated with strong learning nodes are not altered and the nodes and leaves derived from it tend to become biased. In case of adaptive boosting the weights from both strong learners and weak learners are changed based on performance, such that iteratively the weights of strong learning nodes are reduced and the weights of weak learning nodes are increased so that the nodes and leaves which are further derived from these do not have any bias. This process is done iteratively so that the overall model has different weights at different nodes depending on its learning rate and the overall error of the model is also reduced.

5 Methodology

There are 5 wells in the study area namely Boreas-1, Kronos-1, Pharos-1, Poseidon-2, and Proteus-1 as shown in Fig. 1. All the wells are recorded with Natural gamma ray (GR), Resistivity (RD), Neutron Porosity (TNPH), Bulk Density (RHOB), P-wave travel-time (DTP), and S-wave travel-time (DTS) logs. The wells have all intersected Gas and Water in the Early-Middle Jurassic clastic sequences comprising of fluvio-deltaic to marine siliclastics which are interspersed with syn-sedimentary volcanics. The log motifs of the 5 wells are shown in Fig. 2.

The S-wave travel-time log available only partly in Proteus-1 as seen in Fig. 3a. A training dataset was made using the input curves Depth, GR, RD, TNPH, RHOB, and DTP from the other 4 wells with output being DTS log. A Machine Learning model relating input and output was made using Random Forest algorithm. The following Table 1, indicates the input parameters and the results which were obtained. A plot comparing the predicted DTS log and the partially recorded DTS log is shown in Fig. 3b. There is an accuracy of 97.04% between predicted log and actual available log in Fig. 3b.

The Poseidon-2 well was subsequently used as a training well which was manually interpreted using the input log curves and reservoir properties such as Vshale, Porosity, and Water Saturation were deduced as shown in Fig. 4. The green shade shows the amount of shale present in the reservoir and the yellow shade shows the fraction of available reservoir.
The same well and the interpretation results was used for training the Machine Learning (ML) model using Random Forest. The relationship between the various input curves are shown in Fig. 5, for reservoir and non-reservoir facies using litho-facies (Fac) interpreted from the well logs. The colour codes green (facies 0) denotes non-reservoir and consist of dominantly shales and siltstones, orange (facies 1) denote reservoir facies which are sandstones, and blue (facies 2) denotes coal which is very occasionally present. There were 3 different ML models created for determination of Vshale, Porosity, and Water Saturation respectively using Random Forest. During training of the algorithm, cross-validation and stratification (Ojala and Garriga 2010) on the input dataset was performed with different k-fold permutations (50–500) for better training of the algorithm. The results of the training and testing on the dataset are shown in Table 2. The difference in training scores points out that the model is overfitting the dataset and the learning rate does not improve even if the number of trees are increased. The overall training and testing scores are arrived using averages for the three cases.

The adaptive-boosting of the Random Forest algorithm was performed with the same parameters and the results tabulated in Table 3. There has been a massive improvement in performance with testing accuracy improving to 97.03%. The model was used on the other 4 wells and manual petrophysical interpretation was also subsequently performed on the 4 wells individually in order to compare the performance. The comparison between the ML model results and the manual interpretation is shown in Fig. 6a, b.
Fig. 3  a The partially available Shear-Sonic log is depicted for Proteus-1. b The predicted Shear-Sonic log is compared with the partially recorded shear log which shows a very good prediction.
6 Discussion and conclusion

The study successfully demonstrates the ability of making weak learning Random Forest algorithms into strong performing algorithms through Adaptive Boosting. The study also demonstrates the utility of using Random Forest algorithms for automatic petrophysical interpretation, which reduces a lot of time and effort. Typically the manual interpretation of the well logs for interpreting Vshale, Porosity, and Water Saturation requires concerted effort in understanding various data which includes, understanding the geological conditions for deposition, collating and studying well cuttings data, understanding reservoir and non-reservoir characteristics on recorded logs, grain densities from core data, electrical properties of reservoirs, and salinity information of the formation water. The manual interpretation of wells including Poseidon-2 well as shown in Fig. 4 and Fig. 6b respectively, was done using information from core data from some of the wells which included a matrix density of reservoir ~ 2.67 g/cm³. The densities of shale varied between ~2.55 and 2.70 g/cm³. The shale volume and porosity was estimated using Gamma and Neutron-Density logs. The Indonesian Equation was used to estimate Water Saturation (Poupon and Lev-\makebox{eaux 1971}) with electrical rock properties considered as a = 1, m = 2, and n = 2. The Archie parameters or electrical rock properties (Archie 1942) namely ‘a’, ‘m’, and ‘n’ are called the tortuosity factor, cementation factor, and saturation exponent respectively. The tortuosity factor relates to the tortuous pathway of the pore spaces available in the rock. The cementation factor relates to the hindrance in the pathways available in the pore spaces of the rock for fluid migration. The saturation exponent relates to the presence of non-conductive fluid in relation to conducting fluid water available in the pore spaces of the rock. A formation water salinity of ~ 12,000 to 14,000 ppm was used in the study as per reports from the drilled wells.

The petrophysical interpretation of well logs using ML was done based on the input logs which are required to be continuous for the entire well or the zone of interpretation. The log curves needs to be corrected for inaccuracies which may arise during acquisition or due to borehole conditions. As shown in Fig. 5 the relationship between input curves after interpretation of Poseidon-2 well is shown. The reservoir section can be differentiated from the non-reservoir shale and coal in the RHOB vs DTSM and RHOB vs DTCO plots. The GR, TNPH, and Resistivity logs are also good indicators for differentiating hydrocarbon bearing reservoirs from other non-reservoirs. The RD log in Fig. 5 is plotted in Log₁₀ scale. The weak machine learning algorithms tend to overfit the model especially during training when due to limited sampling of the data the model tends towards the more dominant relationship amongst input curves. This is evident during testing of the data by the model as the minority points in the training become more dominant in the testing dataset which gives a lesser accuracy as evident in Table 2. The adaptive boosting of the Random Forest ML algorithm becomes necessary to weigh all the points equally and adjust the relative weights assigned depending on how they fit in the model in order to prevent it from overfitting and make them strong performers as evident from Table 3 and Fig. 6a, b.

| Table 1 | The parameters used and accuracy achieved by the ML algorithm for DTS log prediction |
|-----------------|--------------------------------------|
| Number of trees | 1500                                 |
| Percentage of training data | 60%                               |
| Minimum split | 2                                    |
| Error criteria | Mean square error                     |
| Training score | 99.79%                               |
| Testing score | 98.54%                               |
Fig. 4 The manual interpretation results for Poseidon-2 well
Fig. 5  The relationship between various input logs for reservoir and non-reservoir facies are shown

Table 2  The parameters used and accuracy achieved by random forest algorithm on Poseidon-2 well

| Parameter                  | Model for Vshale | Model for porosity | Model for water saturation |
|----------------------------|------------------|--------------------|---------------------------|
| Number of trees            | 2000             |                    |                           |
| Percentage of training data| 60%              |                    |                           |
| Minimum split              | 2                |                    |                           |
| Training score             | 98.65%           | 97.85%             | 98.12%                    |
| Testing score              | 78.15%           | 69.54%             | 85.17%                    |
| Overall training score     | 98.21%           |                    |                           |
| Overall testing score      | 77.62%           |                    |                           |
Random Forest ML algorithms were also good performers at predicting shear logs from the input dataset as seen in Fig. 3 and Table 1, which gave good results and subsequently used in the interpretation.

| Parameter                  | Model for Vshale | Model for porosity | Model for water saturation |
|----------------------------|------------------|--------------------|---------------------------|
| Number of trees            | 2000             |                    |                           |
| Percentage of training data| 60%              |                    |                           |
| Minimum split              | 2                |                    |                           |
| Training score             | 99.48%           | 98.85%             | 99.87%                    |
| Testing score              | 97.72%           | 97.12%             | 96.24%                    |
| Overall training score     | 99.40%           |                    |                           |
| Overall testing score      | 97.03%           |                    |                           |

**Table 3** The parameters used and accuracy achieved by adaptive-boosted random forest algorithm on Poseidon-2 well

**Fig. 6**

(a) The ML interpretation of wells after adaptive boosting of random forest algorithm. (b) The figure comparing the manual interpretation (Black) and ML output (Red) after adaptive boosting is shown. A very good match in results are observed between both outputs for the 5 wells. (Color figure online)
Acknowledgements The author would like to thank TerraNubis (www.terranubis.com) for making the Poseidon Field dataset publicly available for the purpose of Research and Development along with its associated reports and supporting material under the Creative Commons Licence.

Author contributions Not applicable as single author.

Funding The research was completely undertaken by the author during the COVID-19 Pandemic utilising the time during the Lockdown without any external funding.

Availability of data and material The complete research was done based on dataset available online under the Creative Commons license as acknowledged.

Code availability The demonstration of the methodology was done using Python and Jupyter Notebook which is available in the public domain along with codes as part of the Creative Commons licence. The manual interpretation was done using Petrel and Techlog which are proprietary softwares licensed by my company.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethical approval The author has undertaken this research following the highest standards of ethics.

Consent to participate Not Applicable as there is only single Author who is consenting to participate.

Consent for publication The author gives consent for publication.

References

AGSO Browse Basin Project Team (1997) Browse Basin high resolution study, North West Shelf, Australia. Int Rep Record 1997(38):1–123
Akınnikawa O, Lyne S, Roberts J (2018) Synthetic well log generation using machine learning techniques. In: Unconventional resources technology conference 2018, doi: https://doi.org/10.15530/artec-2018-2877021
Alexandro GC, da P. Carlos AC, Geraldo GN (2017) Facies classification in well logs of the Namorado oilfield using Support Vector Machine algorithm. In: 15th International congress of the Brazilian geophysical society & EXPOGEF 2017 https://doi.org/10.1190/sgbf2017-365
Archie GE (1942) The electrical resistivity log as an aid in determining some reservoir characteristics. Trans AIME 146:54–67
Bestagini P, Lipari V, Tubaro S (2017) A machine learning approach to facies classification using well logs. SEG Tech Prog Expand Abstr. https://doi.org/10.1190/segam2017-17729805.1
Blevin JE, Boreham CJ, Summons RE, Struckmeyer HIM, Loutit TS (1998) An effective Lower Cretaceous petroleum system on the North West Shelf: evidence from the Browse Basin. In: Purcell, P.G., Purcell, R.R. (ed.), The sedimentary basins of Western Australia 2: Proceedings of the Petroleum Exploration Society of Australia Symposium, Perth, WA, 1998, 397–420
Dobróka M, Szabó NP, Tóth J, Vass P (2016) Interval inversion approach for an improved interpretation of well logs. Geophysics 81:D155–D167
Feng DC, Liu ZT, Wang XD, Chen Y, Chang JQ, Wei DF, Jiang ZM (2020) Machine learning-based compressive strength prediction for concrete: an adaptive boosting approach. Constr Build Mater. https://doi.org/10.1016/j.conbuildmat.2019.117000
Freund Y, Schapire RE (1997) A decision-theoretic generalization of on-line learning and an application to boosting. J Comput Syst Sci 55(1):119–139
Gupta A, Soumya U (2020) Well log interpretation using deep learning neural networks. In: International petroleum technology conference, Dhahran, Kingdom of Saudi Arabia, January 2020. https://doi.org/10.2523/IPTC-19678-Abstract

Hastie T, Tibshirani R, Friedman J (2008) The elements of statistical learning (2nd ed.). Springer. ISBN 0-387-95284-5

Jiang L, Castagna JP, Russell B, Guillian P (2020) Rock physics modeling using machine learning. SEG Tech Progr Expand Abstr. https://doi.org/10.1190/segam2020-3427097.1

Kulkarni, Arun D, Lowe B (2016) random forest algorithm for land cover classification, Computer Science Faculty Publications and Presentations. Paper 1. http://hdl.handle.net/10950/341

Kumar M, Gupta S, Gao X, Singh A (2019) Plant Species Recognition Using Morphological Features And Adaptive Boosting Methodology. IEEE Access 7:163912–163918. https://doi.org/10.1109/ACCESS.2019.2952176

Lin W, Wu Z, Lin L, Wen A, Li J (2017) An ensemble random forest algorithm for insurance big data analysis. IEEE Access 5:16568–16575. https://doi.org/10.1109/ACCESS.2017.2738069

Madeh PS, El-Diraby TE (2021) Using machine learning to examine impact of type of performance indicator on flexible pavement deterioration modeling. J Infrastructure Syst 27(2):04021005

Mukherjee B, Srivardhan V, Roy PNS (2016) Identification of formation interfaces by using wavelet and Fourier transforms. J Appl Geophys 128(2016):140–149. https://doi.org/10.1016/j.jappgeo.2016.03.025

Mutanga O, Adam E, Cho MA (2012) High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm. Int J Appl Earth Observ Geoinform 18:399–406. https://doi.org/10.1016/j.jag.2012.03.012

Ojala M, Garriga GC (2010) Permutation tests for studying classifier performance. J Mach Learn Res 11:1833–1863

Panda SC, Srivardhan V, Chatterjee R (2015) Lithological characteristics analysis in straton field oil using wavelet transform. In: 77th EAGE conference and exhibition. 2015: 1–3. https://doi.org/10.3997/2214-4609.201412493

Perez-Muñoz T, Velasco-Hernandez J, Hernandez-Martinez E (2013) Wavelet transform analysis for lithological characteristics identification in siliciclastic oil fields. J Appl Geophys 98(2013):298–308. https://doi.org/10.1016/j.jappgeo.2013.09.010

Peyret AP, Ambia J, Torres-Verdin C, Strobe J (2019). Automatic interpretation of well logs with lithology-specific deep-learning methods. In: SPWLA 60th annual logging symposium, 2019. https://doi.org/10.30632/T60ALS-2019.SSSS

Piryonesi SM, El-Diraby TE (2020) Role of data analytics in infrastructure asset management: overcoming data size and quality problems. J Transp Eng Part B Pavements 146(2):04020022

Poidevin SR, Kuske T, Edwards D, Temple R (2015) Australian petroleum accumulations report 7 Browse Basin. Record 2015/10. Geoscience Australia, Canberra, Australia, pp 1–109

Poupon A, Leveaux J (1971) Evaluation of water saturation in shaly formations. In: SPWLA 12th annual logging symposium, Society of Petrophysicists and Well-Log Analysts

Pratama H (2018) Machine learning: using optimized KNN (K-Nearest Neighbors) to predict the facies classifications. In: The 13th SEGJ international symposium 2018, doi: https://doi.org/10.1190/SEGJ2018-139.1

Purcell PG, Purcell RR (ed.), The sedimentary basins of Western Australia 2: Proceedings of the petroleum exploration society of australia symposium, Perth, WA, 1998, 347–367

Rollet N, Edwards D, Grosjean E, Palu TJ, Hall L, Torterdell JM, Borreham C, Murray A (2018) Regional Jurassic sediment depositional architecture, Browse Basin: implications for petroleum systems. In: Australasian exploration geoscience conference, 18–21 February 2018, Sydney, Australia. Page(s) 1–8

Sanjaya J, Renata E, Budiman V, Anderson F, Ayub M (2020) Prediksi Kelaianan Pinjaman bank Menggunakan random forest Dan adaptive boosting. Jurnal Teknik Informatika Dan Sistem Informasi. https://doi.org/10.28932/jutisi.v6i1.2313

Sarica A, Cerasa A, Quattrone A (2017) Random forest algorithm for the classification of neuroimaging data in alzheimer’s disease: a systematic review. Front Aging Neurosci 9:329. https://doi.org/10.3389/fnagi.2017.00329

Senosy AH, Ewida HF, Soliman HA, Ebraheem MO (2020) Petrophysical analysis of well logs data for identification and characterization of the main reservoir of Al Baraka Oil Field, Komombo Basin, Upper Egypt. SN Appl Sci 2:129. https://doi.org/10.1007/s42452-020-3100-x

Sornsuwit P, Jaiyen S (2019) A new hybrid machine learning for cybersecurity threat detection based on adaptive boosting. Appl Artif Intell 33(5):462–482. https://doi.org/10.1080/08839514.2019.1582861
Srivardhan V (2016) Stratigraphic correlation of wells using discrete wavelet transform with fourier transform and multi-scale analysis. Geomech Geophys Geo-Ener Geo-Resour 2:137–150. https://doi.org/10.1007/s40948-016-0027-1

Stephenson AE, Cadman SJ (1994) Browse Basin, Northwest Australia: the evolution, palaeogeography and petroleum potential of a passive continental margin. Palaeogeography Palaeoclimatol Palaeoecol. https://doi.org/10.1016/0031-0182(94)90071-X

Struckmeyer HIM, Blevin JE, Sayers J, Totterdell JM, Baxter K, Cathro DL (1998) Structural evolution of the Browse Basin, North West Shelf: new concepts from deep-seismic data

Symonds PA, Collins CDN, Bradshaw J (1994) Deep structure of the Browse Basin: implications for basin development and petroleum exploration. In: Purcell PG, Purcell RR (ed.), The Sedimentary Basins of Western Australia: Proceedings of petroleum exploration society of australia symposium, Perth, WA, 1994, 315–332

Szabó NP, Remeczki F, Jobbik A, Kiss K, Dobróka M (2022) Interval inversion based well log analysis assisted by petrophysical laboratory measurements for evaluating tight gas formations in Derecske through, Pannonian basin, east Hungary. J Pet Sci Eng. https://doi.org/10.1016/j.petrol.2021.109607

Valecha H, Varma A, Khare I, Sachdeva A, Goyal M (2018) Prediction of consumer behaviour using random forest algorithm, 2018 5th IEEE Uttar Pradesh Section international conference on electrical, electronics and computer engineering (UPCON), pp. 1–6 https://doi.org/10.1109/UPCON.2018.859707

Zhong R, Johnson RL, Chen Z (2019). Using machine learning methods to identify coals from drilling and logging-while-drilling LWD data. In: Asia Pacific Unconventional Resources Technology Conference 2019. https://doi.org/10.15530/AP-URTEC-2019-198288